Université Bordeaux Segalen

Année 2011 Thèse n°1899

THÈSE

Pour le

DOCTORAT DE L'UNIVERSITÉ BORDEAUX 2

Mention : Sciences Cognitives

Manuscrit confidentiel

Présentée et soutenue publiquement

Le 20 Décembre 2011

Par Benoît PAPAÏX

Né le 12 Avril 1984 Castres (81)

Outiller la conception centrée utilisateur en milieu industriel complexe :

Des techniques de traitement automatique de textes pour la conception des cockpits

Membres du Jury

M. Francis Lowenthal, Professeur, Université de Mons .....(Rapporteur)
M. Farid El Massioui, Professeur, Université Paris VIII .....(Rapporteur)
M. Bernard N'Kaoua, Professeur, Université Bordeaux II.....(Directeur de thèse)
Mme Florence Reuzeau, Docteur, Airbus Operations SAS .....(Examinatrice)
Mme Hélène Sauzéon, Maître de conférences, Univ. Bordeaux II.....(Examinatrice)

UNIVERSITÉ BORDEAUX SEGALEN

AIRBUS AN EADS COMPANY

[…]

« Jugeant qu'il n'y a pas péril en la demeure
Allons vers l'autre monde en flânant en chemin
Car, à forcer l'allure, il arrive qu'on meure
Pour des idées n'ayant plus cours le lendemain
Or, s'il est une chose amère, désolante
En rendant l'âme à Dieu c'est bien de constater
Qu'on a fait fausse route, qu'on s'est trompé d'idée
Mourrons pour des idées, d'accord, mais de mort lente
D'accord, mais de mort lente »
[…]

Georges Brassens

A ma famille…

REMERCIEMENTS

Mes remerciements vont tout d'abord à Florence Reuzeau pour la qualité de son encadrement, son écoute et la confiance qu'elle a su m'accorder pour réaliser ce travail. Toutes ces qualités, dont je lui suis reconnaissant, ont largement contribué à faire de ces années de thèse, une expérience dont je garderai un très bon souvenir. Je remercie également mon directeur Bernard N'Kaoua pour ses précieuses relectures qui ont su arriver à temps !

Je souhaite remercier chaleureusement Florence Buratto pour sa grande qualité d'écoute, pour son aide dans la réalisation de la présentation orale mais aussi pour les qualités humaines dont elle fait preuve dans la gestion de son équipe. Pour tout cela (et bien plus) je souhaite lui exprimer ma grande amitié.

Mes plus vifs remerciements vont aussi à Francis Lowenthal et Farid El Massioui pour avoir accepté d'être rapporteurs de cette thèse ainsi qu'à Hélène Sauzéon pour son soutien et qui m'a fait le plaisir de participer au jury.

Je souhaite exprimer ma gratitude à toutes les personnes qui ont pris le temps de m'aider ou pris part aux réflexions sur ce sujet. Je pense particulièrement à Jérôme Barbé, Florence Beaujard, Christine Bézard, Marie-christine Bressolle, Emmanuelle Cannesson, Laure Christophe, Isabelle Dardenne, Olivier Durou, Charlotte Heulin, Barbara Mawhin, Marielle Plat-Robain, Laurent Spaggiari, Sonja Straussberger, Frédéric Thérond, Cecile Vollard pour leurs remarques judicieuses et les critiques constructives (parfois piquantes) qui m'ont souvent permis d'accomplir des avancées dans ce travail et garanti la bonne humeur dans le bureau.

Je tiens enfin exprimer ma reconnaissance envers mon entourage personnel. Mes parents et mes frères pour leur soutien indéfectible tout au long de mon parcours universitaire. Je pense également à ma compagne, amie et confidente, Jéjé pour ses encouragements, son soutien, sa patience et sa mignonne tête de mule ! Je remercie aussi mes amis, en particulier Florian, compagnon de galère, qui comme il le sait est vraiment le plus beau !

RESUME

Dans le milieu aéronautique professionnel (un des secteurs professionnels les plus sûr au monde), la gestion des conséquences des erreurs humaines doit être améliorée pour garantir une sécurité maximum. Pour ce faire, il est nécessaire de mettre en place des techniques de conception centrées sur l'utilisateur. Cependant, la mise en place de ces techniques est rendue difficile par les particularités des systèmes sociotechniques complexes (la certification, la complexité des systèmes conçus, le nombre de personnes impliquées…). Notre étude a pour but de développer et de valider des outils d'aide à la conception centrée sur l'utilisateur, notamment pour le traitement automatique de grande quantité de données. Pour ce faire, nous avons, dans un premier temps, réalisé une étude basée sur le jugement d'expert visant à identifier, dans une base de données, des scenarii susceptibles de contenir une erreur de l'équipage. Les résultats de cette méthode par jugement d'expert ont été comparés à ceux obtenus à l'aide d'outils de traitement automatique. Cette comparaison a permis :

1/ D'identifier des algorithmes pertinents pour l'extraction d'information dans des bases de données (algorithme des plus proches voisins et de filtrage bayesien) ;

2/ De proposer une méthodologie permettant l'extraction automatique de situations à risque pouvant donner lieu à des études plus approfondies, sur simulateur par exemple. Cette étape est primordiale dans cadre de la conception centrée utilisateur.

Les liens établis avec les études des incidents/accidents laissent envisager des impacts positifs sur la sécurité aérienne.

Mots clés : Conception centrée utilisateur, Traitement automatique de textes, Erreur Humaine, Sûreté des Systèmes, Aviation

ABSTRACT

In the professional aeronautical field (one of the safest in the world), human error management must be improved to reach a better safety level. To do this, a user centred design process has to be implemented. However, due to the complexity of socio technical systems, the implementation of an efficient user centred design process could be challenging. To ease this process, our study aims to develop and validate specific tools, particularly for processing large amounts of textual data. In our study, we will firstly undertake an extraction of scenarios that can contain human errors in a specific database. This extraction will be based on expert judgment (control condition). Secondly, we will confront the control condition with results obtained automatically. The results of this comparison allow:

1/ The identification of relevant algorithms for automatic information extraction within large textual databases (Nearest Neighbour, Bayesian filtering);

2/ The identification of a methodology to extract risk situations that could be included in specific studies. This step is very important for the user centred design process.

Links that we have established between our results and incident/accident studies allow us to consider positive impacts on aviation safety.

Keywords: User Centred Design, Automatic Text Processing, Human Error, System Safety, Aviation

SOMMAIRE

Introduction *12*

Partie I : Cadre théorique **19**

CHAPITRE. 1 La conception centrée utilisateur : cadre général 21

I La conception centrée utilisateur : Les normes 21

I.1 La norme ISO 13407 23

I.2 Etapes du processus de conception centrée utilisateur 23

I.2.1 Planification du processus de conception 24

I.2.2 Spécification du contexte d'utilisation 24

I.2.3 Spécification des besoins utilisateurs 26

I.2.4 Production des solutions 27

I.2.5 Evaluation des solutions 27

I.3 Discussion sur la normalisation de la conception centrée utilisateur 28

II Mise en œuvre de la conception centrée utilisateur 28

II.1 Techniques de recueil de données 28

II.1.1 L'étude documentaire 28

II.1.2 L'interview 29

II.1.3 Le questionnaire 29

II.1.4 Le maquettage 30

II.1.5 La conception collaborative 30

II.2 Les théories et méthodes de formalisation 30

II.2.1 Les modèles d'explication de l'erreur humaine 30

II.2.2 Les modèles de compréhension des causes d'accident 35

II.2.3 Conclusion sur les théories et méthodes de formalisation 41

II.3 Les facteurs particulier structurant la conception 42
II.3.1 Automatisation et Charge de travail : 43
II.3.2 Automatisation et compréhensibilité des systèmes embarqués 44
II.3.3 Automatisation et partage des tâches entre l'Homme et la machine 44
II.3.4 Augmentation de la fiabilité des systèmes 45
II.3.5 Conclusion sur les facteurs influençant la conception 45
II.4 Tracer les solutions de conception 46
II.4.1 La documentation comme support de réflexion. 46
II.4.2 La documentation comme support à la gestion des liens entre les problématiques de conception 46
II.4.3 La documentation comme support de réutilisation, conception, re-conception 47
II.4.4 La documentation comme support à la gestion de projet 47
III Conclusion 48

CHAPITRE. 2 La conception centrée utilisateur en milieu sociotechnique complexe : Cas de l'aéronautique 49
I Un processus et un avion certifié 50
I.1 Le règlement CS25-1302 52
I.2 Les méthodes d'application AMC25-1302 53
I.2.1 La méthode recommandée 53
I.2.2 Les objets de conception à considérer dans la certification 55
I.2.3 Les moyens de démonstration de conformité 57
I.3 Conclusion sur le règlement de certification 58
II Processus de conception : Mise en œuvre en aéronautique 59
II.1 Les objets d'études 61
II.2 Des méthodes pour expliciter, observer et analyser les difficultés et erreurs humaines pertinentes 62
II.2.1 Les méthodes prédictives 63
II.2.2 Les méthodes expérimentales 63
II.3 Conception : prise en compte de la complexité et évaluation des nouveautés 64
II.3.1 Prise en compte de la complexité pour l'utilisateur 64
II.3.2 Gérer les nouveautés 65
II.4 Les expérimentations 65
III Conclusion : Difficultés du cycle de conception centrée utilisateur 67
CHAPITRE. 3 Optimiser la conception centrée utilisateur : Les analyses de sûretés des systèmes, une source d'hypothèses à mieux utiliser 70

I Les analyses de sûreté de fonctionnement 71
I.1 La sûreté de fonctionnement des systèmes dans le monde de l'aéronautique 71
I.2 Les FHA : les scénarii de pannes 75
I.2.1 FHA de niveau avion 76
I.2.2 FHA de niveau système 76
I.3 L'utilisation des FHA dans la conception d'un avion 77
I.4 Les hypothèses de comportement de l'utilisateur dans les scénarii des FHA 77
II Les caractéristiques des scénarii à rechercher dans les analyses de sûreté 78
II.1 Les caractéristiques des scénarii : Ce que cherchent les experts 78
II.2 Les difficultés de sélection des scénarii 79
III Les méthodes de recherche des scénarii 80
III.1 Les méthodes fiabilistes 81
III.1.1 SHERPA : Systematic Human Error Reduction & Prediction 82
III.1.2 HEART : Human Error Assessment & Reduction Technique 83
III.1.3 THERP : Technique for Human Error Rate Prediction 84
III.1.4 TAFEI : Task Analysis for Error Identification 86
III.1.5 Conclusion sur les modèles fiabilistes 87
III.2 Une méthode basée sur le jugement d'expert 88
III.2.1 L'expertise 88
III.2.2 L'utilisation du jugement d'expert en sûreté de fonctionnement 91
III.2.3 Intérêts et inconvénients de l'utilisation du jugement 92
III.3 Une méthode optimisée pour répondre aux besoins de l'industrie 93
III.3.1 Les outils pour répondre aux besoins des experts 94
III.3.2 Les choix de développement 95
III.3.3 Le test des outils 103
IV Conclusion : outils et processus industriel 104

CHAPITRE. 4 Modèle d'analyse 106
I Problématique 106
II Schéma du modèle d'analyse : articulation des notions du cadre théorique 108
III Conclusion 110*

Partie II : Démarche expérimentale 111

CHAPITRE. 1 Sélection des scénarii pertinents par un groupe d'experts 112
I Méthode pour la sélection des scénarii pertinents (Objectif 1) : 115
II Méthode d'étude des scénarii pertinents : le modèle de Reason (Objectif 2) 116
III Résultats : 117
III.1 La base de données initiale 117
III.2 Sélection de scénarii par des experts 118
III.3 Résultats de l'étude des scénarii grâce au processus de prise de décision : 120
III.3.1 Etude des scénarii choisis selon les étapes globales de compréhension et de correction de la défaillance 121
III.3.2 Étude des scénarii choisis selon les 4 étapes du processus de prise de décision 122
IV Interprétation des résultats 124
V Conclusion 126

CHAPITRE. 2 Des descripteurs pour des requêtes pertinentes 128
I Matériel et méthode 129
II Résultats 130
II.1 Résultats 1 : Les descripteurs qui augmentent avec la classification 130
II.2 Comparaison des proportions de descripteurs entre les scénarii sélectionnés et de la base de données initiale 132
III Interprétation des résultats 134
IV Conclusion : 137

CHAPITRE. 3 Développement et test des outils de traitement automatique de textes 139
I Méthode 140
I.1 Etape préliminaire : Formalisation sous forme de vecteurs 140
I.2 Méthode pour les algorithmes de recherche (Les plus proches voisins, le filtrage bayésien) 142
II Résultats 143
II.1 Résultats pour l'algorithme des « plus proches voisins » 143
II.1.1 Résultats pour l'analyse du CDS 143
II.1.2 Résultats pour l'analyse du PA 145
II.2 Résultats pour l'algorithme du filtrage bayesien 145
II.2.1 Résultats pour l'analyse du CDS 146
II.2.2 Résultats pour l'analyse du PA 147
III Interprétation des résultats 149

IV Conclusion 152

Discussion Générale et Perspectives ***153***

I Synthèses des résultats obtenus 153

I.1 Sélection des scénarii pertinents par un groupe d'experts. 154

I.2 Des descripteurs pour des requêtes pertinentes 155

I.3 Développement et test des outils de traitement automatique de textes 155

II Inscription de l'étude dans l'environnement industriel et la littérature existante 156

II.1 Réponses aux attentes industrielles 156

II.2 Inscription dans la littérature des incidents/accidents 159

III Apports de l'étude à la démarche de conception 160

IV Discussion autour des méthodes utilisées et perspectives 162

V Epilogue 164

Références bibliographiques ***165***

Table des figures ***173***

Liste des Tableaux ***176***

Liste des Abréviations ***177***

Annexes ***179***

CHAPITRE. 1 Des descripteurs pour des requêtes pertinentes : Analyses Statistiques 180

CHAPITRE. 2 Précisions sur le « basic Airmanship » 188

I L'Airmanship 188

II Conception et « Basic Airmanship »: 190

III Systèmes et environnement contraignent le « Basic Airmanship » : 191

IV Conclusion 193

CHAPITRE. 3 Architecture des programmes informatiques 194

I Préparation des données 195

II Algorithme des plus proches voisins 196

III Algorithme de filtrage bayésien 197

INTRODUCTION

Aujourd'hui, le milieu aéronautique professionnel est un des secteurs professionnels et industriels les plus sûrs au monde, avec un niveau de sûreté de l'ordre de 10^{-5} à 10^{-6} (Amalberti, 2001): il se produit seulement un incident majeur ou un accident pour 100 000 à 1 000 000 vols complets. Comme le montre la Figure 1 (extraite de Boeing Company (2010)), le niveau de sécurité s'est très largement amélioré ces trente dernières années grâce aux progrès effectués sur les matériaux, les automatismes, la formation des équipages, etc. Cependant, la part de l'erreur humaine dans ces incidents ou ces accidents stagne, et se place dans une fourchette de 65% à 80% (Amalberti, 1996) et logiquement, elles apparaissent toujours comme la dernière frontière à franchir pour une sécurité maximum.

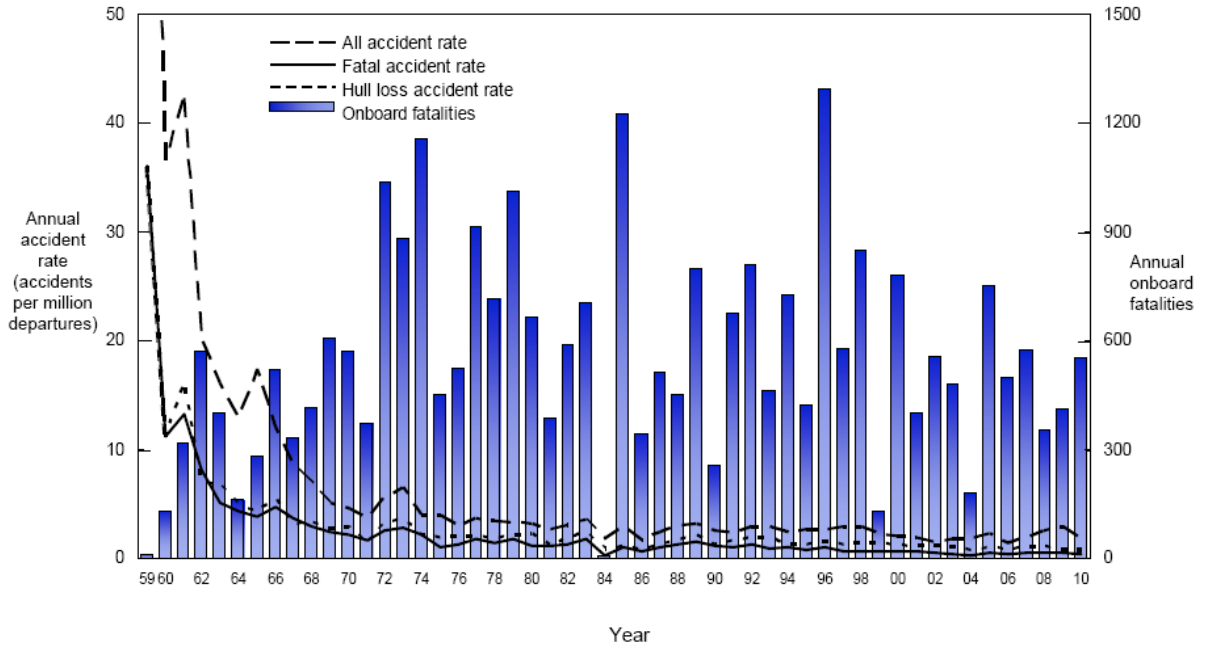


Figure 1 : Evolution du nombre d'accidents d'avion sur les 50 dernières années.

Toutefois, comme le dit Reason (1990), les erreurs humaines sont le « prix à payer » à l'intelligence humaine. De ce fait, les solutions qui espéraient supprimer les erreurs sont en perte de vitesse : automatisation augmentée, procéduralisation poussée, surveillance accrue, et même certaines logiques de formation recherchant l'absence d'erreur chez l'élève comme résultat idéal. Les études les plus actuelles, sans renier les bénéfices des précédentes prennent l'erreur comme un composant structurel du comportement humain dont nous devons tenir compte en conception. Ces approches de l'erreur humaine, auxquelles appartient notre étude, cherchent à maîtriser ses conséquences. Nous savons donc, qu'il faut accepter l'erreur humaine, expliciter le plus fidèlement possible ses occurrences et mettre en place des solutions de conception permettant la gestion de ses conséquences pour la sécurité.

La conception d'un avion (ou de tout autre système) est fondée sur des hypothèses. Lorsque l'on s'intéresse à ces hypothèses pour la conception de postes de pilotage, nous remarquons qu'elles appartiennent à trois grandes classes. Les premières sont des **hypothèses de fonctionnement de systèmes**, les secondes des **hypothèses d'architecture des systèmes** et les dernières, les **hypothèses d'utilisation de ces systèmes par un opérateur** (pilotes, personnels navigant cabine, agents de maintenance...). Ces grandes classes d'hypothèses sont étroitement liées, et faire une hypothèse sur l'une des classes implique souvent de faire une hypothèse sur une autre. Le travail de conception accorde ces différentes hypothèses entre elles. Une hypothèse d'architecture ou de fonctionnement système entraînant une hypothèse de comportement utilisateur[1] erronée, pourrait entraîner la survenue d'incidents ou d'accidents. Optimiser l'accord des hypothèses les unes avec les autres pourrait potentiellement augmenter la sécurité.

Pour mieux comprendre ces hypothèses, prenons un exemple. Pour accomplir un vol, le pilote a besoin de connaitre son altitude. L'avion doit donc posséder une fonction donnant au pilote cette information. Pour cela, les concepteurs font les hypothèses suivantes :

- **Hypothèse d'architecture :** Le système devra posséder des sondes barométriques, reliées à des calculateurs qui transmettront l'information d'altitude à un système d'affichage chargé de la rendre disponible pour le pilote.
- **Hypothèse de fonctionnement :** Le système devra donner l'information au pilote selon différents formats nécessaires en opérations, c'est-à-dire : hauteur (calage QFE : l'altitude 0 est prise au niveau de l'aéroport), altitude (calage QNH : l'altitude 0 est prise au niveau de la

[1] *NB : Dans la suite de la thèse, le mot « utilisateur » pourra avoir un sens général ne dépendant pas de la conception d'un système particulier. Ou bien, il sera précisé la nature de l'utilisateur comme : pilote de ligne.*

mer), niveaux de vol (calage standard : l'altitude 0 est prise par rapport au niveau moyen de la mer c'est-à-dire : 1013,25 hectopascal).

- **Hypothèse de comportement pilote :** Le pilote sera en mesure de retrouver cette information dans les informations présentes dans le cockpit et saura différencier les calages QFE, QNH et Standard.

Ces hypothèses sont prises pour les situations normales et anormales. Par exemple, dans le cas de la perte de l'information d'altitude, l'hypothèse de fonctionnement du système pourrait être que celui-ci ne détecte pas la perte de l'information d'altitude. Dans ce cas, le pilote doit détecter la perte de son altitude et nous aurions donc une hypothèse de comportement pilote correspondante.

La prise en compte des Facteurs Humains centre la conception sur l'utilisateur. Les processus de *conception centrés utilisateur* alors mis en place, permettent de **concevoir et d'évaluer les hypothèses d'utilisation de ces systèmes par un opérateur**. Bien entendu, ceci pourra aboutir à des modifications affectant toutes les classes d'hypothèses. Ce travail représente une tâche immense, surtout pour concevoir un système aussi complexe qu'un avion de ligne. Le nombre d'hypothèses n'est pas quantifiable de manière absolue. Elles sont plus ou moins formalisées dans de nombreux documents issus des processus de conception, des procédures d'utilisation de l'avion et des programmes de formation des équipages. De plus, la réalité explorera infiniment plus d'horizons que ces hypothèses ne pourront couvrir. Cependant, puisqu'elles conduisent la conception cela reste, de ce point de vue, un excellent point d'entrée à la gestion du risque lié à l'erreur humaine.

Il existe un document reflétant la majorité des hypothèses de conception : *les analyses de sûreté des systèmes*. Ces analyses sont rédigées en réponse à une exigence de certification et sont présentes chez toutes les industries à risques : transport aérien, ferroviaire et maritime, industrie nucléaire, industrie spatiale…

Dans le cas de la conception aéronautique, les analyses de sûreté des systèmes contiennent, pour chaque groupe fonctionnel de l'avion (ATA[2]), les scénarii décrivant les conséquences prévues de toutes les défaillances fonctionnelles. Les conséquences ainsi décrites sont alors classées sur une

[2] L'ATA 100 défini de manière empirique les groupes fonctionnels qui servent de référence lors de la conception et de l'utilisation d'un avion de ligne. Ce document a été rédigé par « *l'Air Transportation Association* » d'où il tient son nom.

échelle de gravité donnée par le document de certification CS 25[3] 1309 (European Aviation Safety Agency, 2011) dont le Tableau 1 présente un extrait. Cette échelle de gravité commence à « *No Safety Effects* » pour les scénarii ayant les conséquences les plus faibles à « *Catastrophic* » pour les scénarii aux conséquences les plus graves.

Effect on Aeroplane	No effect on operational capabilities or safety	Slight reduction in functional capabilities or safety margins	Significant reduction in functional capabilities or safety margins	Large reduction in functional capabilities or safety margins	Normally with hull loss
Effect on Occupants excluding Flight Crew	Inconvenience	Physical discomfort	Physical distress, possibly including injuries	Serious or fatal injury to a small number of passengers or cabin crew	Multiple fatalities
Effect on Flight Crew	No effect on flight crew	Slight increase in workload	Physical discomfort or a significant increase in workload	Physical distress or excessive workload impairs ability to perform tasks	Fatalities or incapacitation
Allowable Qualitative Probability	No Probability Requirement	<----Probable---->	<-----Remote----->	Extremely <----------> Remote	Extremely Improbable
Allowable Quantitative Probability: Average Probability per Flight Hour on the Order of:	No Probability Requirement	<----------> $<10^{-3}$ Note 1	<----------> $<10^{-5}$	<----------> $<10^{-7}$	$<10^{-9}$
Classification of Failure Conditions	No Safety Effect	<-----Minor----->	<-----Major----->	<---Hazardous--->	Catastrophic
Note 1: A numerical probability range is provided here as a reference. The applicant is not required to perform a quantitative analysis, nor substantiate by such an analysis, that this numerical criteria has been met for Minor Failure Conditions. Current transport category aeroplane products are regarded as meeting this standard simply by using current commonly-accepted industry practice.					

Tableau 1 : Extrait du CS 25 1309

Ce document contient la majorité des hypothèses de conception pour les raisons suivantes :

- Pour les **hypothèses d'architecture des systèmes et hypothèses de fonctionnement des systèmes :** A chaque classification correspond une probabilité à respecter, les scénarii aux conséquences les plus graves étant, bien entendu, ceux les moins probables. La classification présente dans les analyses de sécurité, induit l'architecture des systèmes. L'architecture des systèmes, induit les logiques de leur fonctionnement.
- Pour les **hypothèses de comportement des utilisateurs :** Dans ces analyses, les spécialistes en sécurité prévoient les comportements de l'équipage confronté à la défaillance d'un système. Cet ensemble de « comportements attendus » fait partie du **« Basic Airmanship ».** Cette notion fédère les comportements que les concepteurs peuvent raisonnablement attendre d'un

[3] CS 25 est le règlement de certification auquel doivent se soumettre tous les constructeurs d'avion de ligne. Edité par l'EASA, les « *Certification Specifications* » contiennent plusieurs chapitres : le 1309 correspondant au règlement concernant la Sécurité et le 1302 au règlement concernant les Facteurs Humains.

équipage, c'est-à-dire des hypothèses de comportements des utilisateurs. *NB : Des précisions sur cette notion sont présentes en annexe à ce document.*

Pour donner un aperçu de ce que représentent les analyses de sécurité, l'ensemble des ATA d'un avion tel que l'Airbus A380, contiennent environ 100 000 scénarii de pannes. Chaque scénario présente :

- Un titre donnant une description de la panne fonctionnelle et l'évaluation de ces conséquences rédigée dans un style permettant une intégration assez aisée dans une base de données.
- Une description des conséquences de la panne sur l'avion et les comportements attendus de l'équipage. Cette dernière partie est rédigée en langage naturel (Anglais). L'information contenue dans cette partie est donc plus difficilement accessible pour des outils de recherche informatique.

Dans ce cadre, les concepteurs chez Airbus souhaiteraient renforcer la confiance dans les hypothèses de comportement pilote. Pour cela, ils demandent **une méthode ou des outils permettant de trouver rapidement dans ces grands ensembles de scénarii, impliquant des hypothèses d'architecture des systèmes et/ou des hypothèses de fonctionnement des systèmes, des hypothèses de comportement des pilotes « discutables », « incertaines » qui pourraient être étudiées de manière approfondie.** Ceci permettant de les éprouver en condition de test afin d'assurer leur robustesse.

Afin de répondre à cette demande, notre démarche a été tout d'abord de nous intéresser à la conception centrée utilisateur, car celle-ci doit théoriquement prévenir l'apparition de ces hypothèses discutables de conception.

Dans le **chapitre 1,** nous verrons comment se fait la mise en place de la conception *centrée sur utilisateur*. Notamment, nous aborderons sa normalisation qui a été mise en place à la fin des années 1990 pour répondre aux besoins grandissant des concepteurs d'interfaces Homme/machine. Dans ce chapitre l'utilisateur sera pris au sens général.

Ensuite dans le **chapitre 2**, nous verrons comment se fait la *conception centrée utilisateur* lorsque l'on se place dans le cadre d'un système sociotechnique complexe comme celui de la conception en aéronautique. Ce chapitre permettra aussi de donner le contexte applicatif des résultats de cette étude.

Après avoir indiqué le contexte, nous présenterons les outils théoriques à disposition pouvant répondre à la demande industrielle. Le **chapitre 3** présentera ces outils dans le but d'en choisir un qui permette

de répondre à la demande industrielle. Afin d'atteindre cet objectif ce chapitre précisera tout d'abord le besoin industriel puis s'intéressera aux outils à disposition permettant d'y répondre.

Enfin le **chapitre 4** présentera le modèle d'analyse de ce travail. En faisant la synthèse des 3 premiers chapitres, il liera les différentes notions qui auront été vues afin de préparer la méthode et son opérationnalisation.

Ces quatre premiers chapitres formeront le cadre théorique de cette thèse (**Partie I**).

Nous présenterons ensuite les chapitres relatifs à la démarche expérimentale (**Partie II**). Pour développer, tester et évaluer les outils définis dans la partie précédente, nous verrons qu'il est nécessaire de connaitre 1/Un groupe de scénarii référents pour développer, tester et évaluer les outils conçus, 2/Des relations sur des critères de recherche permettant aux experts de filtrer rapidement la base de données en fonction de leurs besoins ; 3/Des prototypes d'outils afin de tester leur efficacité et la réponse qu'ils apportent aux besoins des experts. Afin d'obtenir ces informations, nous avons mis en place une démarche expérimentale avec les étapes suivantes :

- Etape 1 : Des experts avaient pour tâche de sélectionner les scenarii qu'ils considéraient comme étant discutables du point de vue des comportements attendus de l'équipage dans un ensemble réduit de la base de données initiale (analyse de sûretés). Cet ensemble réduit contenait environ 2200 scénarii.
- Etape 2 : Ensuite, ces scenarii ont été étudiés afin d'identifier des thématiques récurrentes, des mots clés, des descripteurs, des relations entre descripteurs, etc. L'objectif à terme est de pouvoir utiliser ces informations pour réaliser des requêtes, c'est-à-dire des extractions automatisées des scénarii discutables à partir de mots clés.).
- Etape 3 : Toujours dans le but de créer des outils d'extractions automatiques, nous avons utilisé les scénarii sélectionnés par les experts pour tester la pertinence de 2 méthodes existantes d'extraction d'information dans une base de données : la méthode des plus proches voisins et la méthode du filtrage bayésien.

D'un point de vue général, les étapes 1 et 2 correspondent à des étapes nécessaires au développement de règles pour la sélection des scénarii. Elles serviront 1/ pour donner aux experts des exemples de requêtes pertinentes pour cibler les scénarii pertinents. Elles serviront aussi 2/ à la calibration des outils développés, mis en pratique et testés lors de l'étape 3.

Dans le **chapitre 1,** nous présenterons la mise en place, les résultats et leur interprétation pour l'étape 1. Après avoir formé des groupes d'experts, nous sélectionnerons avec eux les scénarii pertinents dans 3 analyses de sûreté des systèmes. Nous verrons, dans les scénarii sélectionnés, quelles sont les hypothèses de comportement de l'équipage remises en doute par les experts.

Ensuite dans le **chapitre 2**, nous verrons qu'il existe des descripteurs présents dans les titres des scénarii qu'il est possible d'utiliser pour réaliser des requêtes sur l'ensemble de la base de données. Ces requêtes pourraient permettre aux experts de faciliter leur démarche de recherche et d'initialiser les outils d'extraction automatique de textes. Pour comprendre quelles sont les meilleures requêtes à réaliser, nous chercherons des relations entre pertinence d'un scénario et présence de certains critères.

Enfin, dans le **chapitre 3**, après avoir expliqué comment ont été développés les outils de traitement automatique de textes, nous évaluerons leur efficacité. Ces outils sont basés sur des algorithmes éprouvés en traitement automatique de textes (Algorithme des plus proches voisins, Filtrage Bayésien). Ils devraient permettre aux concepteurs 1/de sélectionner des scénarii en fonction de leur proximité avec un scénario déjà sélectionné (Plus proches voisins). Mais aussi de donner une probabilité de sélection, *a priori*, grâce au filtrage Bayésien. Ces choix d'algorithmes seront expliqués dans la suite de cette thèse. Dans ce chapitre, nous les testerons et évaluerons afin de conclure sur la réponse aux besoins des concepteurs.

Ce chapitre conclura la partie expérimentale.

Pour finir, la dernière partie donnera **la discussion générale et les perspectives** ouvertes par cette thèse. Dans cette partie, nous reprendrons la problématique ainsi que les différents champs théoriques auxquels cette thèse appartient, puis nous donnerons les apports ainsi que les perspectives de recherche ouvertes par cette étude tant en matière de recherche fondamentale qu'appliquée.

Partie I : CADRE THEORIQUE

Lors que l'on se place dans un contexte de conception, d'autant plus lorsque l'on cherche à concevoir en tenant compte de l'utilisateur final, les différentes théories en lien avec les Facteurs Humains, se structurent autour du processus de conception centré sur l'utilisateur. La *conception centrée utilisateur* considère les utilisateurs et leurs besoins tout au long du processus de développement d'un produit. Ce processus ne se contente pas de demander aux utilisateurs ce qu'ils désirent, mais consiste à appliquer des méthodes rigoureuses de recueil de données concernant l'utilisabilité, l'utilité, l'acceptabilité dans l'utilisation d'un produit existant ou d'un prototype. La mise en œuvre de ce processus doit permettre en conception de limiter le risque lié à l'erreur humaine. Un processus de *conception centrée utilisateur* typique comprend trois phases principales mises en œuvre de façon itérative :

- Analyse.
- Conception.
- Evaluation.

Notre étude trouve son cadre dans les théories relatives à la conception dans les milieux socio-techniques complexes notamment celui de la *conception centrée utilisateur*. L'objectif de cette partie est de remettre dans son contexte le problème soulevé dans l'introduction : Comment garantir de meilleures hypothèses de conception du point de vue des Facteurs Humains afin de limiter le risque lié aux erreurs humaines ?

Pour garantir une bonne performance aux systèmes sociotechniques complexes, ceux-ci doivent mettre en place des outils et/ou des méthodes adaptés à leurs besoins mais aussi à la situation dans lesquels ils vont être mis en place.

Cette partie abordera la mise en place d'un outil dans le cadre de l'aéronautique et sera organisée comme suit :

- Chapitre 1 : La *conception centrée utilisateur* : cadre général
- Chapitre 2 : La *conception centrée utilisateur* en milieu sociotechnique complexe : Cas de l'aéronautique
- Chapitre 3 : Optimiser la *conception centrée utilisateur* : Les analyses de sûretés des systèmes, une source d'hypothèses à mieux utiliser
- Chapitre 4 : Modèle d'analyse

Le déroulement de ces chapitres partant du cadre le plus large vers la problématique, permettra de positionner ce travail dans le cadre théorique correspondant.

CHAPITRE. 1 LA CONCEPTION CENTREE UTILISATEUR : CADRE GENERAL

L'approche de *conception centrée utilisateur* considère les utilisateurs et leurs besoins tout au long du processus de développement de produit. Cette démarche est normalisée par la norme ISO 13407 (International Organization for Standardization, 1999). Selon cette norme, les utilisateurs finaux doivent être des acteurs pour l'évaluation et le développement d'un produit. Cette méthode de conception impose que le développement soit guidé par les besoins des utilisateurs plutôt que par les possibilités technologiques. Un processus de *conception centrée utilisateur* d'un poste de pilotage doit donc mettre en œuvre des moyens pour adapter le produit aux utilisateurs cibles. Cette démarche, inclut un ensemble de méthodes spécifiques destinées à recueillir des données et à les convertir en choix de conception. Ces méthodes permettent de collecter des données sur les tâches et besoins des utilisateurs puis de valider la pertinence de ces données. Cette implication des utilisateurs doit être à la fois précoce (elle est nécessaire dès le début du projet) et itérative (elle doit se répéter tout au long des étapes clés du projet) (Norman, et al., 1986) (Vredenburg, et al., 2001). Concevoir une application performante selon la *conception centrée utilisateur*, nécessite de se demander à chaque étape critique de la conception si le produit correspond aux besoins des utilisateurs finaux. Cette méthode de conception a été définie par plusieurs normes internationales notamment l'ISO 13407, qui sera présentée en détail dans le paragraphe suivant.

I La conception centrée utilisateur : Les normes

La normalisation dans le domaine de la prise en compte de l'Homme pour conception d'interfaces en général, couvre deux champs distincts : le premier se place au niveau du produit et le second au niveau des processus.

Si l'on prend par exemple les normes développées en matière d'ergonomie informatique, celles-ci ont d'abord été dédiées aux préconisations sur la qualité ergonomique finale du produit. Tout d'abord il y a eu les normes ISO 9241 (International Organization for Standardization, 1998). Celles-ci sont relatives aux « Exigences ergonomiques pour le travail de bureau avec terminaux à écrans de visualisation » comprennent 17 parties. Les parties 1 à 9 concernent les équipements, les environnements et les postes de travail. Les parties 10 à 17 traitent de l'ergonomie du logiciel.

Puis la norme ISO 13407 (International Organization for Standardization, 1999) a été créée, celle-ci vise le champ plus large du cycle de conception et détermine les exigences auxquelles un projet doit répondre pour être considéré comme "centré sur l'humain". Elle concerne la méthodologie de conception et l'intégration de la démarche ergonomique dans un cycle de développement. La norme ISO 13407 définit les conditions de mise en œuvre d'un processus de *conception centrée utilisateur*.

La norme ISO 13407 a été complétée par un rapport technique l'ISO/TR 16982 (International Organization for Standardization, 2002). Ce rapport indique des méthodes préconisées pour la *conception centrée utilisateur*. Celui-ci présente une liste de méthodes ergonomiques pouvant être appliquées aux différentes étapes du cycle de conception tout en précisant leurs avantages et inconvénients. On y présente les méthodes impliquant directement les utilisateurs finaux (ex., l'observation, la mesure des performances, la technique des incidents critiques, les questionnaires, les entretiens, les verbalisations au cours des tâches, les techniques de conception et d'évaluation participatives, etc.) et les méthodes n'impliquant pas directement les utilisateurs finaux (ex., l'analyse de documents, guide style, grilles d'évaluation, critères ergonomiques, les méthodes formelles {ex., KLM, GOMS, MAD*}, etc.). Bref, il s'agit ici d'une présentation rapide de méthodes ayant fait leurs preuves dans le domaine de la *conception centrée utilisateur*.

Aujourd'hui, ce type de conception entre de plus en plus en considération lors de la fabrication d'un produit. On distingue plusieurs niveaux de prise en compte des processus centrés sur l'utilisateur dans les organisations :

- Non reconnus.
- Reconnus.
- Considérés.
- Implémentés.
- Intégrés.
- Institutionnalisés.

Ces niveaux ont été établis dans le cadre d'un modèle sur l'utilisabilité : UMM. Ils sont certifiés par la norme ISO 15504 (International Organization for Standardization, 2011). Cette norme définit une approche pour estimer la maturité d'un processus. Dans le paragraphe suivant, nous allons nous intéresser à la norme ISO 13407 car celle-ci précise la méthode pour mettre en place une *conception centrée utilisateur*.

I.1 La norme ISO 13407

La norme ISO 13407 concerne le cycle de conception et détermine les exigences auxquelles un projet doit répondre pour être considéré comme « *centré utilisateur* ». Elle concerne la méthodologie de conception et l'intégration de la démarche ergonomique dans le cycle développement.
Cette norme a été complétée une première fois en 2000 par le rapport technique ISO/TR 18529 (International Organization for Standardization, 2000) qui décrit sept ensembles de pratiques de base pour mettre en œuvre le processus de conception centrée sur l'humain. Ce rapport ISO/TR 18529 permet d'identifier quelles méthodes d'utilisabilité sont recommandées à chacun des stades du cycle de conception. Douze méthodes génériques ont été identifiées et décrites succinctement avec leurs principaux avantages et inconvénients selon les caractéristiques du projet.

La norme ISO 13407 définit les conditions de mise en œuvre d'un processus centré sur l'opérateur humain à travers cinq principes nécessaires :

Une préoccupation en amont des utilisateurs, de leurs tâches et de leur environnement ;

- La participation active de ces utilisateurs, ainsi que la compréhension claire de leurs besoins et des exigences liées à leurs tâches ;
- Une répartition appropriée des fonctions entre les utilisateurs et la technologie ;
- L'itération des solutions de conception ;
- L'intervention d'une équipe de conception multidisciplinaire.

I.2 Etapes du processus de conception centrée utilisateur

La norme ISO 13407 (ISO 13407) définit les étapes du cycle de *conception centrée utilisateur* tel que décrit dans la Figure 2 (source ISO 13407). Le processus est d'abord planifié. Dans cette étape, un consensus est recherché au sein du projet afin de préparer la mise en place du processus de conception. Cette étape préliminaire précède la boucle itérative de conception. La première station du cycle consiste à définir le contexte d'utilisation. S'ensuit l'étape de captation des besoins utilisateurs. Une fois ces deux étapes réalisées, l'équipe doit être capable d'émettre des propositions de solutions qui seront évaluées au regard des exigences posées au préalable.

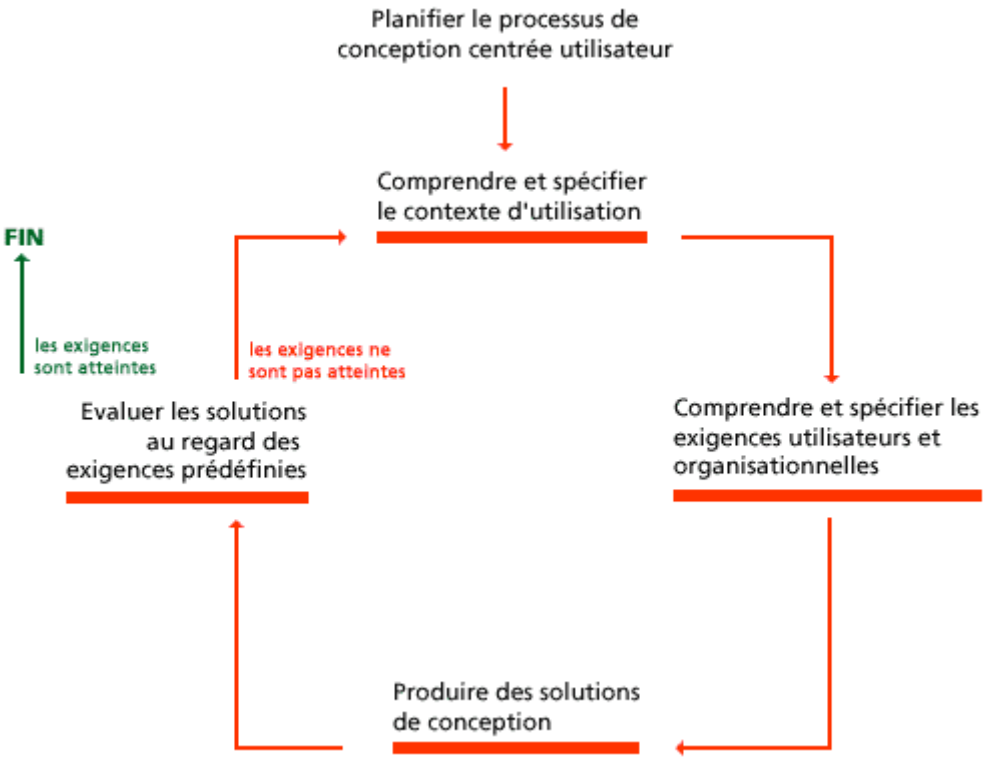


Figure 2 : Processus de conception centré utilisateur (ISO 13407)

Dans la suite, nous détaillerons ces étapes de conception.

I.2.1 Planification du processus de conception

Le pré-requis essentiel à une démarche *conception centrée utilisateur* consiste à planifier sa mise en œuvre. L'équipe de conception doit donc avoir atteint un consensus concernant la recherche de la satisfaction de la norme ISO 13407 et donc de ses implications sur les plans techniques, méthodologiques et de conduite de projet. Cette pré étape consiste à planifier les activités de développement dans une optique de *conception centrée utilisateur*. L'adaptation des outils et méthodes se basera sur la consultation de documentation et des discussions autour des pratiques et des contraintes de la conception. Ces discussions sont notamment menées avec la participation des futurs utilisateurs du système constitués en groupes de travail.

I.2.2 Spécification du contexte d'utilisation

Divers auteurs ont souligné l'importance du contexte d'utilisation d'un artefact. Ce contexte participe au partage d'une structure de connaissances et de représentation partagées par l'ensemble des acteurs d'une situation (Grice, 1975) ; (Marwel, et al., 1975). De Terssac et Chabaud (De Terssac, et al., 1990) y voient : « *La mise en commun des compétences pour préparer et réaliser une action ; cette mise en commun des compétences, en même temps qu'elle complète la représentation que chacun se fait de la*

tâche à réaliser, constitue un référentiel commun permettant d'ajuster les décisions de chacun en fonction des connaissances des autres. » (Pages 128-129). Il s'agit donc à la fois de connaissances et de représentations partagées pour pouvoir tenir compte des autres acteurs dans ses décisions. Ce référentiel commun crée un espace de coopération. Piaget (Piaget, 1965), ne donne pas particulièrement de nom à cette structure bien qu'il définisse une échelle de valeurs fondamentales auxquelles se référer en cas de discussion. Ceci permettant aux acteurs de s'accorder sur des valeurs auxquelles se référer en cas de besoin : *« un langage, un système de notions et des propositions fondamentales »*(Piaget, 1965). De ce point de vu, ce référentiel joue un rôle crucial dans la compréhension des communications et les explications. Ce référentiel pouvant être modifié lors de l'intégration d'une information nouvelle à une information ancienne (Karsenty, 1996) (Karsenty, 2000). La construction et le maintien de ce référentiel est une tâche importante dans le déroulement d'une coopération. Certaines expériences (Hoc, 1998) (Carlier, et al., 1999) impliquant la coopération entre deux contrôleurs radar pour gérer un même secteur montre d'ailleurs que l'essentiel des communications portaient sur le maintien d'un référentiel commun. Ce référentiel commun est une structure complexe qui peut être prise à plusieurs niveaux. Il existe des référentiels locaux ou globaux, par exemple, un référentiel commun pour les contrôleurs aériens, un référentiel commun pour un équipage d'avion de ligne et bien entendu un référentiel commun pour la coopération des contrôleurs et de l'équipage. Hoc (Hoc, 2003) définit des niveaux en fonction de la conscience qu'ont les acteurs de ce référentiel commun (Niveaux Action, Plan, Méta) Dans un référentiel commun de niveau Méta, l'opérateur est amené à prendre en compte se propres caractéristiques de fonctionnement et les connaissances de son activité pour la coopération. C'est le cas des pilotes de lignes.

C'est pourquoi, la première étape proprement dite du cycle de *conception centrée utilisateur* vise à comprendre et spécifier le contexte d'utilisation. Il s'agit donc de saisir les caractéristiques, buts et tâches des utilisateurs ainsi que leur environnement d'utilisation. A cet effet, il s'agit d'abord de décrire l'environnement des points de vue technique, matériel, social, organisationnel et législatif. Par ailleurs, l'identification des profils utilisateurs est une base essentielle de cette première étape du cycle. La connaissance de ces profils permettra de choisir les méthodes d'évaluation et de sélectionner des participants pour mener des tests utilisateurs. Les utilisateurs finaux sont qualifiés par des caractéristiques données : connaissances, compétences, fonctions, langage, éducation, formation, caractéristiques physiques, psychologiques, habitudes…

Des méthodes fondées sur l'observation et le recours aux questionnaires sont proposées dans la norme. Pour certains projets, le recueil des données peut bénéficier d'une analyse des usages des outils existants. Dans le domaine des interactions homme-machine, l'analyse des tâches de l'utilisateur

permet de comprendre comment ceux-ci vont utiliser le système, quelles sont les tâches critiques, les plus fréquentes, leur importance, leur durée et leurs niveaux de difficulté. Le choix des méthodes à utiliser est fonction du contexte technique, des marges temporelles et financières et aussi de la disponibilité des utilisateurs. Les différentes techniques nécessitent en effet plus ou moins de matériel, de temps et de compétences de la part de l'expérimentateur.

I.2.3 Spécification des besoins utilisateurs

Lorsque l'on s'intéresse à la conception centrée sur l'Homme, l'ensemble des auteurs s'accordent à dire qu'il est primordial de **bien connaitre l'utilisateur et son besoin** pour concevoir de bonnes interfaces. Cette connaissance possède plusieurs niveaux qui débutent à la connaissance d'un utilisateur « typique » en allant vers le contexte le plus intégré dans lequel une multitude d'utilisateurs vont interagir dans un environnement dynamique. Ce message est véhiculé tant par des travaux sur le processus de conception (Nielsen, 1993) que des travaux sur les lignes directrices de conception (Mayhew, 1992). A première vue, nous ne pouvons être qu'en accord avec ce point de vue, cependant dés lorsqu'il s'agit de le mettre en pratique, la situation se corse. Au-delà de toutes les questions basiques que se posera le concepteur : Qui sont-ils ? Que font-ils ? Quelles connaissances ont-ils du système que je conçois ? ; Il faut organiser et optimiser les connaissances de l'utilisateur nécessaires à notre conception :

- Que faut-il savoir au juste des utilisateurs et pour quoi faire ?
- Quelle démarche adopter pour connaitre les utilisateurs pour la conception d'interface ?
- Quel rôle joue l'utilisateur dans la conception ?

Nous défendons ici le même positionnement que Robert (Robert, 2003) : il est nécessaire *« d'avoir un cadre large et inclusif pour bien comprendre les utilisateurs et fournir des informations pertinentes et utiles à leur sujet en vue de la conception de systèmes interactifs ».*

Robert (Robert, 2003) définit les propriétés de ce cadre de référence de la façon suivante :

- Il ne doit pas se limiter à la seule utilisabilité des systèmes mais tenir compte de l'utilité, de l'accessibilité, etc.
- Il ne doit pas seulement tenir compte des études de sciences cognitives mais inclure des études sociologiques, démographique, de marché…
- Il ne doit pas se limiter à un seul type de tâche, d'interface ou de contexte d'utilisation.
- Il ne doit pas seulement être satisfaisant du point de vue théorique mais aussi être pertinent et utile pour la pratique du concepteur.

A travers ce cadre de référence donné pour les connaissances des utilisateurs, apparaissent des notions importantes sur les connaissances préliminaires à la conception. En plus des connaissances sur l'utilisateur, il est nécessaire de connaitre et comprendre les variabilités du contexte et de tous les facteurs externes pouvant influencer l'interaction Utilisateur/Produit. Enfin, l'ensemble des connaissances sur l'utilisateur, les tâches qu'il réalisera, les connaissances sur le contexte et les facteurs externes doivent être pertinents et utiles pour la pratique des concepteurs. Lors de cette étape, il faut donc spécifier les exigences liées à l'utilisateur et à l'organisation. Il s'agit de prendre en compte les besoins, compétences et l'environnement de travail de tous les intervenants sur le système. L'identification des buts et tâches des utilisateurs est la base du travail de spécification d'exigences. Cette caractérisation des tâches est documentée par l'étape précédente. Les documents de spécification consistent en des descriptions précises des profils d'utilisateurs et des cas d'utilisation. Les exigences doivent ensuite être ordonnées selon leur importance. A ce niveau, les objectifs d'utilisabilité vont guider et justifier les choix de conception. Ils fourniront par la suite des critères de validation lors des tests utilisateurs.

I.2.4 Production des solutions

L'étape de production de solutions de conception utilise les connaissances acquises lors des étapes précédentes pour matérialiser les solutions sous forme de prototypes. Le choix de solutions potentielles se fait en deux grandes étapes. L'ergonome se base d'abord sur son expertise et ses connaissances pour déterminer un éventail de choix possibles. Il teste ensuite ces options avec les utilisateurs pour définir la plus adaptée.

I.2.5 Evaluation des solutions

Les prototypes créés au stade précédent sont utilisés pour évaluer les solutions conçues en fonction des exigences. Le pilotage de tests utilisateurs selon un protocole d'évaluation précis permet de détecter les défauts. On peut ordonner les défauts de conception selon leur importance en fonction des objectifs d'utilisabilité définis précédemment. Diverses méthodes sont indiquées pour l'évaluation telle que les tests d'évaluation individuels ou en groupe et/ou les questionnaires de satisfaction. Par ailleurs, des observations en situation sont recommandées afin de valider les hypothèses des phases précédentes. L'objectif de l'évaluation des solutions est de recueillir un retour sur le prototype développé. Il permettra d'améliorer la conception. C'est une évaluation de la satisfaction des objectifs utilisateurs et organisationnels.

I.3 Discussion sur la normalisation de la conception centrée utilisateur

Cette base normative peut servir à la conception ou à la refonte de toute interface Homme/machine. L'idée est d'adapter le système aux caractéristiques des utilisateurs, des tâches et du contexte d'utilisation. Cependant, la norme prescrit des méthodes pour les différentes phases sans spécifier précisément la forme des résultats. Il est précisé qu'une étape préliminaire consiste à trouver un consensus sur la pertinence et les modalités de l'application de la norme dans un projet. Or, elle ne précise pas les modalités pratiques de cette négociation, tous les acteurs pourraient avoir une vision différente de la réalisation de ce cycle de conception. De plus, un décalage peut apparaître avec les situations réelles de conception. En effet, les fortes contraintes pesant sur un projet de développement, les pratiques de conception habituelles et les possibilités technologiques peuvent rendre difficile l'application stricte et entière d'une démarche centrée sur l'utilisateur. Il s'agit alors de garder les lignes directrices de la norme telles que le cycle itératif et le retour utilisateur nécessaire. Pour parer à ces éventuelles difficultés, la mise à œuvre de la *conception centrée utilisateur* a été outillée. Sans reprendre les outils du rapport ISO/TR 16982, nous nous attâcherons dans le paragraphe II à présenter les méthodes en référence à la bibliographie sur le sujet.

II Mise en œuvre de la conception centrée utilisateur

Dans le paragraphe précédent, nous nous sommes attachés à définir les étapes de la *conception centrée utilisateur*, qu'est ce que l'on doit faire et pourquoi le faire ? Nous allons maintenant définir comment réaliser la *conception centrée utilisateur*. Ce paragraphe présentera les méthodes à disposition pour recueillir des données, qu'elles soient pour la connaissance de l'utilisateur ou pour l'évaluation d'un prototype. Ensuite, afin que le concepteur structure sa connaissance du fonctionnement de l'utilisateur, il est nécessaire qu'il utilise des modèles d'abstractions et de représentation des connaissances. Nous verrons ensuite qu'une étude bibliographique sur le domaine de conception est nécessaire afin de formaliser l'impact des facteurs externes sur l'activité de l'utilisateur. Pour finir nous verrons l'importance pour la *conception centrée utilisateur* de tracer les décisions de conception, afin d'intégrer au mieux à chaque itération les sources de changement de conception.

II.1 Techniques de recueil de données

II.1.1 L'étude documentaire

Cette ressource est largement utilisée afin de capturer des besoins en se familiarisant avec le contexte d'application. L'étude concerne des documents techniques mais aussi organisationnels et statistiques. Les rapports techniques concernent les rapports de tests de systèmes existants, manuels,

documentation techniques, des captures d'écran, etc. Les documents organisationnels comprennent les procédures, les guidelines, les descriptions métier, etc. Enfin, différents documents spécifiques du métier peuvent constituer une ressource d'information comme les revues du domaine d'application par exemple. Ce mode de recueil d'information, malgré sa richesse, reste une ressource incomplète car dépourvue du contexte de l'usage du système. Le recueil hors contexte induit des informations sur l'utilisation prescrite ou théorique et non des pratiques réelles d'usage. Nous verrons que le même problème se pose avec la technique des questionnaires.

II.1.2 L'interview

C'est une technique fondamentale et constitutive de la plupart des processus d'identification des besoins en conception d'IHM. Le contact direct favorise une étude dynamique, conviviale et permet les itérations et validations citées plus haut. L'interviewé peut être non seulement l'utilisateur final mais aussi un expert métier afin de collecter des informations sur le domaine d'application. L'un des freins à une bonne évaluation des besoins recueillis par interview est la complexité d'analyse des données en l'absence de méthodologie structurante (Kotonya, et al., 1998) (Sommerville, et al., 1997). Les informations collectées peuvent paraître vagues, hors du sujet, ou encore inopportunes et inutilisables. Cet état de fait s'explique souvent par l'absence d'une grille de lecture structurée des données d'interview.

II.1.3 Le questionnaire

Parfois utilisé en complément de l'interview, le questionnaire suppose une vision claire des objectifs poursuivis par le processus de développement. Le risque apparaît donc de mettre en avant les « a priori » des concepteurs sur les besoins des utilisateurs du système. Lorsqu'il n'est pas passé durant une interview mais envoyé directement à l'utilisateur, le questionnaire est un outil passif. Aucune opportunité d'approfondir ou de clarifier les données n'est possible. Cependant, le questionnaire permet des études quantitatives sur des échantillons plus larges ainsi qu'un traitement statistique des données. Dans ce dernier but, les questionnaires de besoins utilisateurs contiennent un minimum d'interrogations ouvertes au profit des Questionnaires à Choix Multiples (QCM). Dans son étude intitulée « *What we say / What we do* » (Deutcher, 1974) le sociologue américain Irwin Deutcher présente un certain nombre d'exemples de résultats contradictoires obtenus selon que l'enquête est conduite par observation ou par questionnaire. Cette analyse met en lumière le biais possible entre les déclarations des personnes et leurs actes réels. Deutcher démontre que c'est la relation directe et particulière avec les individus qui induit des comportements pouvant être opposés aux attitudes exprimées.

II.1.4 Le maquettage

Le maquettage est de plus en plus utilisé dans le développement d'interface Homme/machine. Les maquettes sont conçues de façon à visualiser tout ou partie du système afin que les utilisateurs puissent l'évaluer. Une maquette d'IHM est essentiellement une GUI (Graphical User Interface) qui simule le comportement du système dans différentes situations d'usage. Les réponses de simulation sont souvent simplement codées dans une base de données, l'objectif étant d'obtenir l'évaluation par les utilisateurs de tel ou tel comportement système. Le maquettage devient indispensable lorsque les fonctionnalités développées sont très innovantes ou encore dans les phases intermédiaires d'évaluation (dans le contexte de conceptions itératives).

II.1.5 La conception collaborative

La technique consiste à réunir utilisateurs, concepteurs et parfois développeurs dans des ateliers d'explicitation. Les acteurs du projet sont constitués en groupes avec un leader, un secrétaire, des utilisateurs et des développeurs. Le groupe doit sortir avec un capital de problèmes et de solutions grâce à la synergie qu'il aura réussi à créer (Caelen, et al., 2006).

II.2 Les théories et méthodes de formalisation

Pour réaliser le cycle de *conception centré utilisateur*, il est nécessaire d'expliciter les modes de fonctionnement du système Homme/machines. Pour cela, il faut utiliser des méthodes d'abstraction. Ces méthodes sont basées sur des théories de formalisation et d'explication des comportements, des modes de fonctionnement et des erreurs de l'Homme. Les modèles les plus utilisés en conception sont principalement de deux sortes, les modèles d'explication de l'erreur humaine et les modèles de compréhension des causes des accidents.

II.2.1 Les modèles d'explication de l'erreur humaine

Les principaux modèles d'erreur humaine seront exposés dans ce paragraphe :

- La théorie SRK de Rasmussen (Rasmussen, 1983)
- La théorie des « *Slips, Lapses, Mistakes, Violations* » de Norman (Norman, 1988)
- Le modèle GEMS de Reason (Reason, 1990)
- La théorie des génotypes et des phénotypes d'Hollnagel (Hollnagel, 1991)

II.2.1.1 La théorie SRK de Rasmussen

Rasmussen (Rasmussen, 1983) distingue trois niveaux de processus cognitifs de prise de "décision/action". Le niveau « habiletés » (*Skill*) pour les activités réalisées automatiquement. Le niveau « règles » (*Rule*) pour lequel le processus d'action est basé sur des règles construites lors de situations précédente. Pour finir le niveau « connaissances » (*Knowledge*) correspond aux actions pour

lesquelles la réponse a du être construite entièrement à partir des connaissances, c'est-à-dire lors de situations totalement nouvelles et pour lesquelles aucune règle n'a précédemment été construite. A chacun de ces niveaux, Rasmussen associe des erreurs. La classification des erreurs se fait donc en fonction des mécanismes cognitifs mis en jeux.

Situations	**Mode de contrôle**		
	Conscient	Mixte	Automatique
Routines			Basé sur des habiletés
Connues		Basé sur des règles	
Nouvelles	Basé sur des connaissances		

Tableau 2 : Le modèle SRK

II.2.1.2 La théorie Slips, Lapses, Mistakes, Violations de Norman

Don Norman (Norman, 1988) propose, lui aussi, une classification des erreurs humaines en fonction des mécanismes cognitifs mis en jeu. Les *« Slips »* résultent d'un comportement automatique, c'est-à-dire lorsque les actions inconscientes mises en place pour réaliser un objectif sont stoppées avant de l'avoir atteint. Les *« Mistakes »* proviennent de mécanismes conscients. Les *« Slips »* sont souvent des petites erreurs comme des actions mal placées, les *« mistakes »* peuvent être des évènements plus importants et moins facilement détectés puisqu'ils sont en accord avec le processus d'actions mis en place. A cette classification, Reason ajoute en 1997 (Reason, 1997) les *« Lapses »* et les *« Violations »*. Une erreur de type *« Lapse »* est observable à la différence d'une erreur de type *« Slip »*. Les erreurs de type *« violation »* sont des actions conscientes et correspondent à des transgressions volontaires d'une règle. La Figure 3 reprend ces différentes catégories.

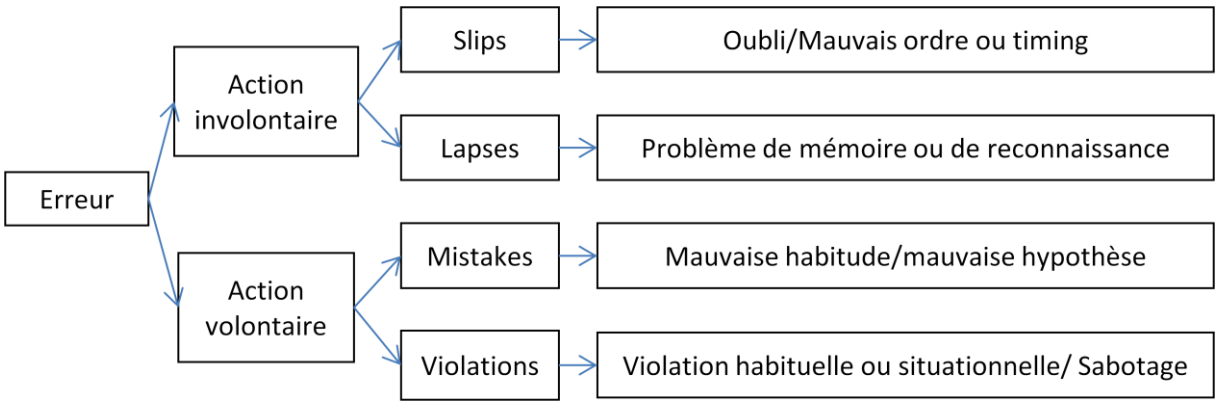


Figure 3 : Le modèle Slips, Lapses, Mistakes, Violations

II.2.1.3 Le modèle GEMS de Reason

Le modèle GEMS (*Generic Error Modelling System*) de Reason (Reason, 1990) est basé sur la théorie du modèle SRK. Il permet de montrer comment est fait le passage entre les différents processus cognitifs du modèle SRK définis par Rasmussen. L'approche de ce modèle ne prend pas en compte les facteurs environnementaux, ou d'autres facteurs liés au contexte ou aux autres processus cognitifs humains. Le modèle GEMS reprend en partie la classification de *« Slips », « Lapse », « Mistake », « Violation »*. Les erreurs de type *« Slips »* et *« Lapses »* correspondant à des erreurs au niveau des habiletés (*Skill*). Les erreurs de type *« Mistakes »* se situant au niveau des règles et des connaissances (*Rule, Knowledge*). Les erreurs de type *« Violations »* ne sont pas reprises. La Figure 4 présente le modèle GEMS et reprend les liens dans avec les modèles SRK et *« Slips », « Lapse », « Mistake »*.

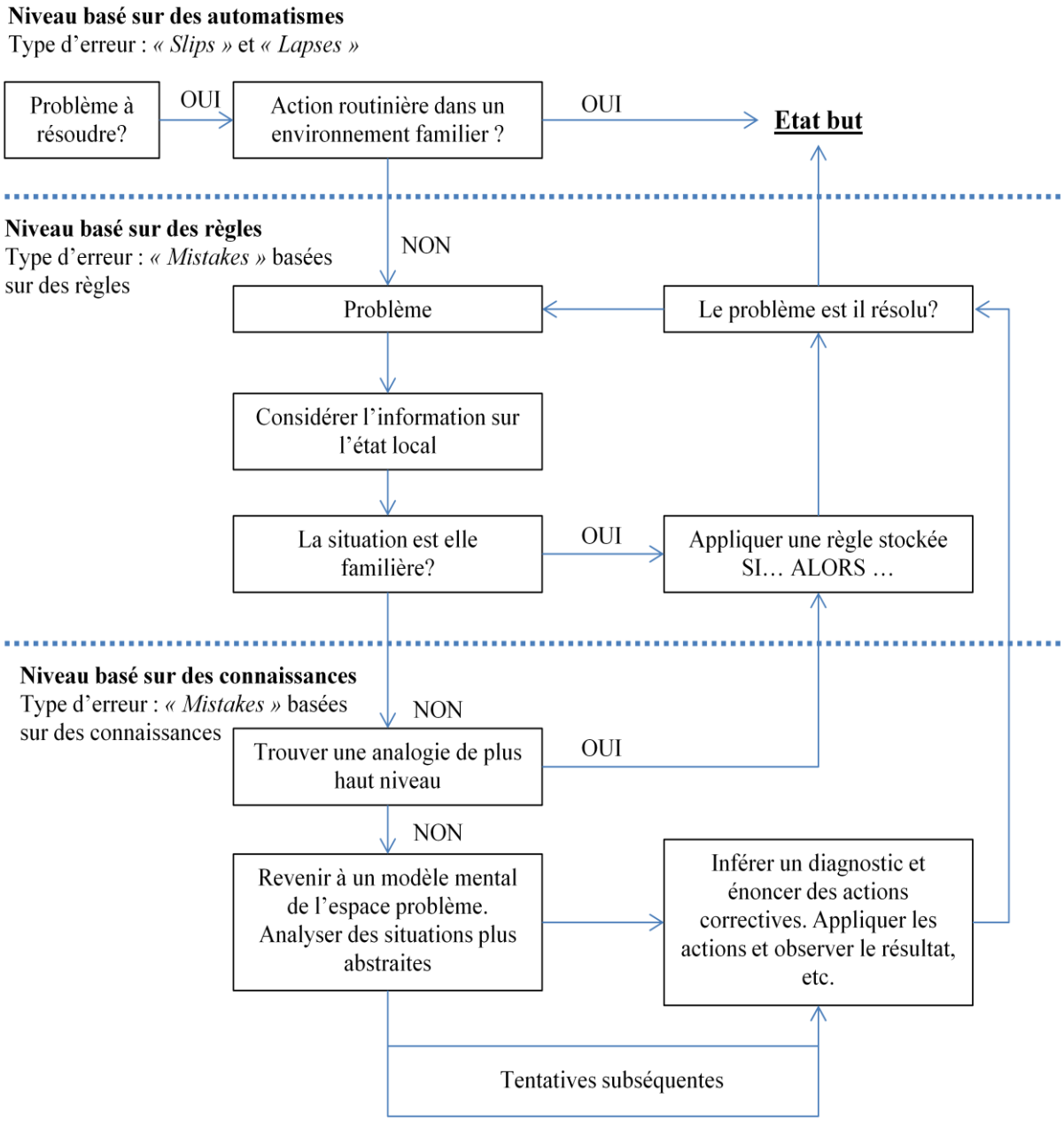


Figure 4 : Le modèle GEMs

II.2.1.4 Les phénotypes et génotypes d'Hollnagel

La classification de Hollnagel (Hollnagel, 1991) est basée sur les termes phénotypes et génotypes. Ces termes proviennent de la biologie et sont définis comme suit :

Phénotype : C'est la partie observable d'un organisme, les parties externes, les manifestations physiques. C'est-à-dire la somme des atomes, des molécules, des macromolécules, des cellules, des structures, des tissus, des organes ; tout ce qui fait partie de la structure observable d'un être vivant.

Génotype : Correspond à la partie « codée » en interne, l'information héréditaire possédée par tout organisme vivant. Cette information stockée est utilisée comme instruction pour construire et maintenir une créature vivante.

En appliquant ces définitions à l'erreur humaine, le phénotype correspond à la partie observable de l'erreur c'est à dire l'activité de l'opérateur. Le génotype correspond à la partie non observable c'est à dire aux mécanismes cognitifs qui ont conduit à l'erreur. Ce qui est intéressant avec ce modèle, c'est que nous n'avons pas besoin de connaître l'activité cognitive de l'opérateur pour classifier ces erreurs. Hollnagel a défini huit phénotypes simples et sept complexes.

Phénotypes simples	Phénotypes complexes
Répétition d'une action	Recommencement de l'action
Mélange dans l'ordre des actions	Saut d'une action à l'autre
Oubli d'action	Dépassement du temps pour une action
Action en retard	Action réalisée trop tôt
Action trop précoce	Rapprochement incorrect de deux actions dû à la proximité de ces actions
Remplacement d'une action par une autre	Rapprochement incorrect de deux actions dû à la proximité de ces actions dans l'arbre de décision
Insertion d'une action présente ailleurs dans la tâche	Capture de l'attention par une autre action
Intrusion d'une action non présente dans la tâche	

Tableau 3 Le modèle des Phénotypes et Génotypes

Basé sur les phénotypes définis ci-dessus, Hollnagel a présenté une taxonomie des phénotypes des erreurs. Cette taxonomie est présentée dans la Figure 5 ci-dessous.

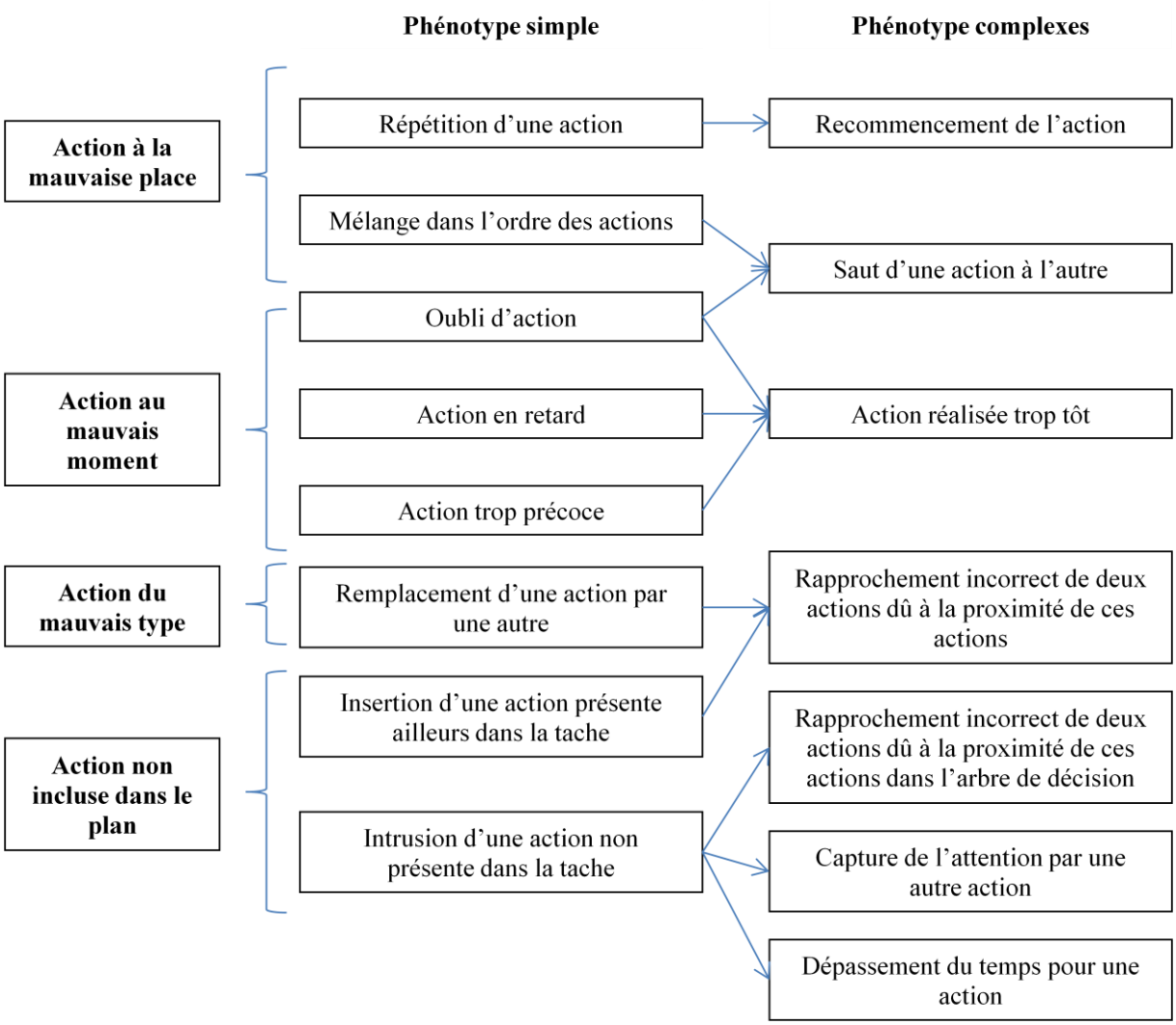


Figure 5 : Taxonomie des phénotypes d'Hollnagel.

Nous venons de présenter les méthodes d'abstraction basées sur des théories de formalisation et d'explication des comportements et des erreurs humaines. Ces théories permettent au concepteur d'avoir l'abstraction nécessaire à la compréhension des liens entre les comportements qu'il pourra observer chez l'utilisateur et les logiques d'interaction qu'il aura développées dans le produit de sa conception. Ces modèles sont à compléter par les modèles de compréhension des causes des accidents.

II.2.2 Les modèles de compréhension des causes d'accident

Nous allons présenter, dans ce paragraphe, les principaux modèles qui aident à la compréhension des causes des accidents :

- Le modèle des dominos (Heinrich, et al., 1980)
- Le Swiss Cheese Model (Reason, 1990)

- Le modèle HFACS (Shappell, et al., 2000)
- Le modèle FRAM (Hollnagel, 2004)

Nous allons suivre la chronologie de ces modèles, commençant par le modèle de plus séquentiel : le modèle des dominos, en allant vers le modèle dit de résonnance le modèle FRAM.

II.2.2.1 Le modèle des dominos Heinrich et al, 1980

Le modèle des Dominos d'Heinrich (Heinrich, et al., 1980) est un exemple classique du modèle d'accident séquentiel ; les dominos tombent du fait d'un unique évènement. Heinrich propose que l'accident puisse être causé par un acte et/ou une condition non sûr. Le modèle est considéré comme déterministe puisque les conséquences sont l'issue d'un évènement spécifique. En traçant la chaine des causes, l'enquêteur d'un accident doit pouvoir trouver la racine de l'accident et retracer l'enchainement des évènements.

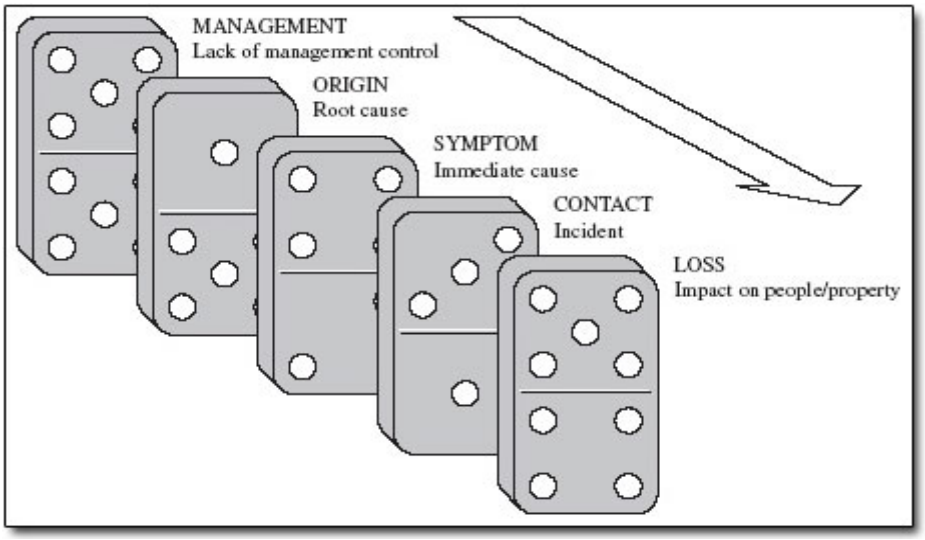


Figure 6 : Le modèle des dominos

II.2.2.2 Le Swiss Cheese (« Gruyère ») Model de Reason

Le modèle proposé par Reason (Reason, 1990) et (Reason, 1997) est l'un des plus influents modèles de compréhension des accidents. Il est basé sur le concept de barrières qui permettent de séparer les sources d'un accident. Les "trous" dans les barrières sont les failles d'un niveau de défense d'un système. Ainsi, un accident a lieu lorsque les trous du « gruyère » s'alignent, c'est à dire lorsque toutes les conditions sont réunies. Ce qui explique le fait que plusieurs événements sont nécessaires pour provoquer un accident (ceci est un plus, par rapport au modèle des dominos). Selon Reason, pour éviter un accident, il suffit qu'une seule de ces faiblesses situées sur l'une des barrières soit supprimée. Au lieu de blâmer les opérateurs qui contrôlent directement les systèmes ou processus complexes, il

faudrait rechercher les causes latentes et plus ou moins distantes des évènements redoutés, comme une mauvaise politique de gestion du risque (Reason, 1990). La théorie de Reason montre que le précurseur d'un accident est le résultat d'imperfections dans l'organisation et l'environnement du système complexe.

Ce modèle présente 4 niveaux de barrières, chacun s'influençant mutuellement. Le niveau 1 correspond aux *actes dangereux* (*Unsafe Act*). C'est ce que Reason appelle les erreurs actives. Cela correspond à une combinaison de facteurs du système (humains ou machines) qui peuvent déclencher un accident.

Les autres niveaux sont des erreurs latentes. Le niveau 2 contient *les précurseurs des actes dangereux* (*Preconditions for Unsafe Acts*). Il contient les conditions de l'opérateur et du système (fatigue ou panne par exemple) et la façon dont ces conditions impactent les performances de ces composantes du système. Le niveau 3 correspond à des défauts de supervision (*Unsafe Supervision*). Un exemple de niveau 3 pourrait être une mauvaise communication entre les acteurs du système, notamment entre le niveau décisionnel et le niveau opérationnel. Le dernier niveau montre que des erreurs latentes peuvent aussi se trouver dans l'organisation même du système, par exemple si l'organisation ne prend pas en compte des facteurs comme la sécurité des systèmes.

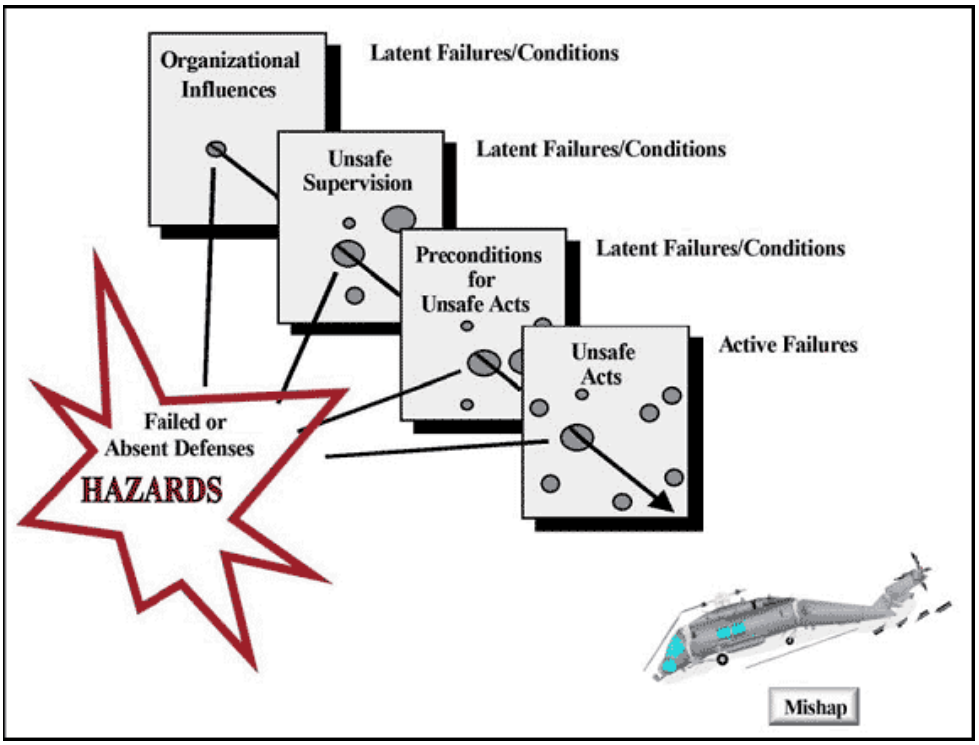


Figure 7 : Le Swiss Cheese Model

II.2.2.3 Le modèle HFACS de Shappell & Wiegman

Le problème avec le modèle du Swiss Cheese est sa mise en œuvre, dans la mesure où Reason donne peu d'éléments pour étudier et améliorer un processus de conception avec son modèle. C'est sur ce point que le modèle HFACS (Shappell, et al., 2000) (Shappell, et al., 2007) amène une amélioration. Le modèle *Human Factors Analysis and Classification System*, *HFACS* décrit les « trous dans le gruyère » du modèle de Reason en faisant le lien entre la théorie et les observations qui ont pu avoir lieu en pratique, principalement pour l'industrie aéronautique. Les figures Figure 8, Figure 9, Figure 10, Figure 11 présentent ce modèle. Elles reprennent les différents niveaux du modèle « *swiss cheese* » de Reason et rajoutent les éléments donnés par HFACS.

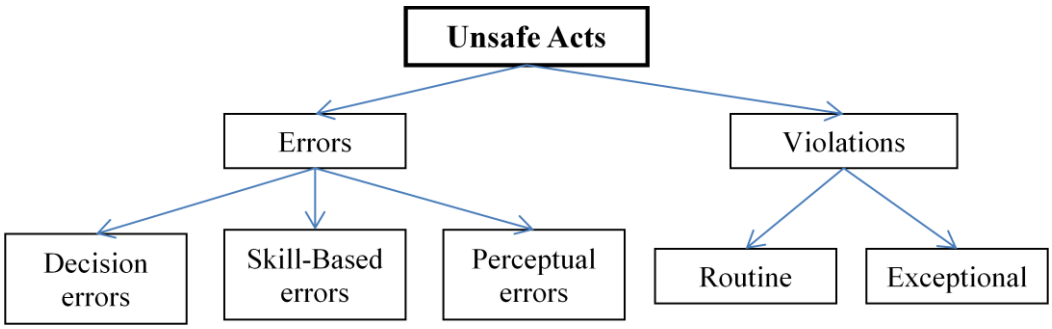


Figure 8 : HFACS, Unsafe act.

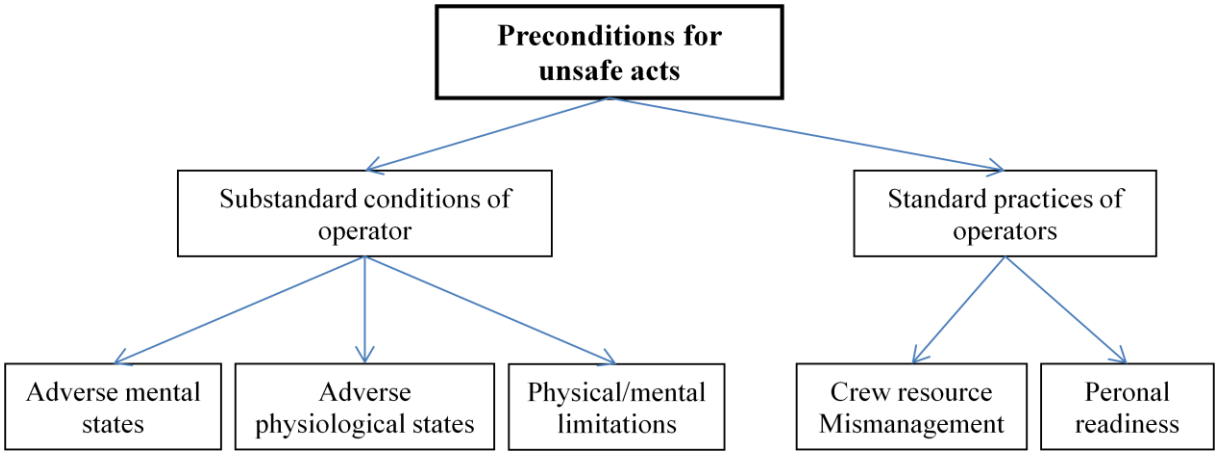


Figure 9 : HFACS, preconditions for unsafe acts

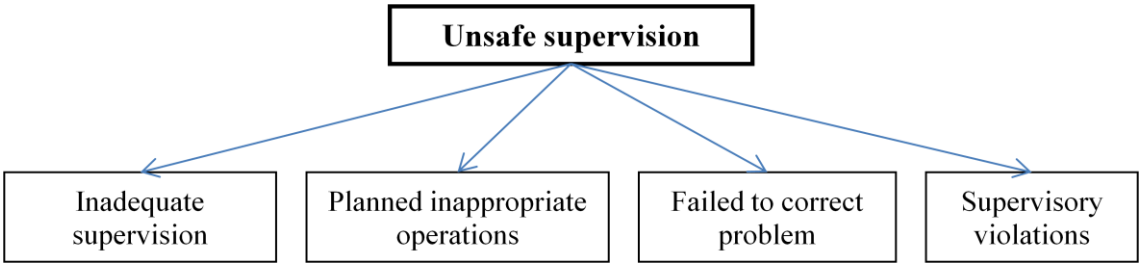


Figure 10 : HFACS, Unsafe supervision

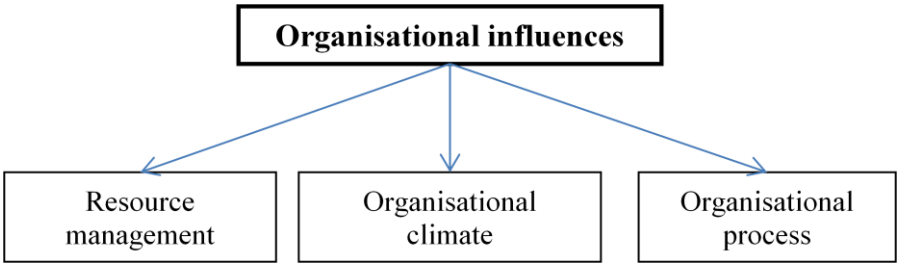


Figure 11 : HFACS, Organisational influences

II.2.2.4 Le modèle FRAM

FRAM (*Functional Resonance Analysis Method*) (Hollnagel, 2004) est un modèle qui appartient au courant de l'ingénierie de la résilience (Hollnagel, et al., 2006). Ce courant propose que chaque système possède une capacité propre à se rétablir après l'erreur d'un de ces composants. Cette capacité doit pouvoir se mesurer et s'optimiser pour garantir une meilleure sécurité. Contrairement aux autres modèles qui structurent les systèmes selon leur organisation, FRAM est un modèle qui caractérise les systèmes socio techniques par les fonctions qu'ils remplissent. C'est-à-dire, au lieu de représenter les organisations de façon hiérarchique de l'opérateur vers l'organisation, FRAM va les représenter à travers leurs fonctions (Conception, Achats, Maintenance, Commerciaux, etc.). Les frontières des systèmes sont donc définies à travers la description de leurs fonctions (Hollnagel, et al., 2005). FRAM capture la dynamique des systèmes à travers les dépendances non linéaires entre les différentes fonctions (Par exemple, la conception ne peut se faire que si les commerciaux ont vendu des avions pour payer les concepteurs).

FRAM est basé sur 4 principes (Hollnagel, et al., 2008). **Premièrement,** les succès ou les échecs sont le résultat des adaptations réalisées par des organisations, des groupes et des individus afin de répondre à la complexité. Les succès dépendent de leurs capacités de reconnaitre et d'anticiper les risques. Les échecs sont dus à une absence (permanente ou temporaire) de cette capacité plutôt qu'à

l'échec d'un de ces composants (système ou machine). **Deuxièmement,** les systèmes sociotechniques complexes sont par nécessité sous spécifiés et partiellement prédictibles. Les procédures et les outils sont adaptés à la situation pour répondre à une grande variabilité de buts potentiellement contradictoires. La variabilité de la performance est donc normale et nécessaire. La variation suffisamment grande d'une seule fonction pourrait suffire à provoquer un accident. Cependant, **le troisième principe** indique que la variabilité des différentes fonctions peut se combiner de manière inattendue, menant à un résultat aux conséquences disproportionnées. Les performances normales ou accidentelles sont donc un phénomène émergeant qui ne peut être expliqué seulement en regardant les performances d'un composant du système. Enfin, pour **le quatrième principe**, la variabilité d'un nombre de fonctions peut entrer en résonnance et impliquer la sortie des limites de fonctionnement du système, avec pour conséquences : l'émergence d'un accident.

FRAM propose une méthode afin d'implémenter ces principes et supporter l'analyse et la prédiction des accidents. Cette méthode se déroule en 4 étapes :

Etape 1 : Identifier les fonctions du système et caractériser pour toutes les fonctions les 6 paramètres de base composant la fonction. Ces paramètres sont :

- L'entrée que la fonction est chargée de modifier *(I : Input)*.
- La sortie produite, c'est à dire ce que la fonction doit apporter au système *(O : Output)*
- Les conditions préalables pour réaliser la fonction *(P : Preconditions)*.
- Les ressources que la fonction devra consommer pour fonctionner *(R : Resources)*.
- Le temps nécessaire pour la réalisation de la fonction *(T : Time)*.
- Le contrôle qui est chargé de superviser et d'ajuster l'activité de la fonction *(C : Control)*.

Etape 2 : Caractériser la variabilité potentielle grâce aux *conditions communes de performance (Common Performance Conditions : CPCs*). Il existe 11 CPCs identifiées dans la méthode FRAM. 1/Disponibilité du personnel et des équipements. 2/entraînement, préparation, compétences. 3/Qualité de communication. 4/Interaction Homme/machine, et support opérationnel. 5/Disponibilité de procédures. 6/ Conditions de travail. 7/ Nombre de buts à atteindre et leurs conflits potentiels. 8/Temps disponible. 9/Rythmes circadien, Stress, constantes biologiques. 10/Collaboration au sein des équipes. 11/Qualité organisationnelle. Ces CPCs adressent la combinaison des aspects humains, technologiques et organisationnels de chaque fonction.

Etape 3 : Définir les processus de résonnances possibles entre les fonctions. La résonnance est définie selon les dépendances et les couplages possibles entre les fonctions et les variabilités potentielles. Par

exemple, la sortie d'une fonction peut être l'entrée d'une autre fonction, ou lui fournir des ressources ou encore être une condition préalable à sa réalisation. Si la sortie de cette fonction varie de façon non prédictible, les fonctions suivantes seront impactées.

Etape 4 : Identifier la zone de variabilité opérationnelle et spécifier les barrières permettant de garantir un niveau de variabilité de performance acceptable pour le fonctionnement global du système.

La Figure 12 est une illustration de la représentation graphique obtenue avec le modèle FRAM.

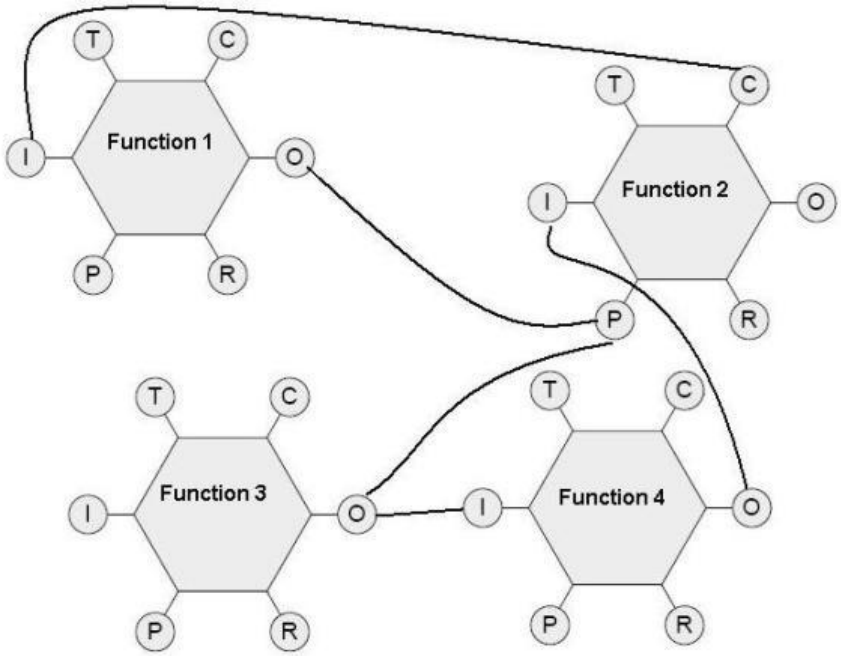


Figure 12 : Illustration du modèle FRAM

Ce modèle a comme avantage de proposer une nouvelle approche, reposant sur une analyse fonctionnelle, pour décrire la performance des systèmes socio techniques complexes (Mawhin, 2011). L'opérationnalisation de ce modèle est en développement pour l'analyse d'accidents et en est encore aux prémisses pour l'analyse et la prévention des risques.

II.2.3 Conclusion sur les théories et méthodes de formalisation

Pour conclure, il me semble important de montrer quel point de vue doit avoir le concepteur, sur les théories et méthodes de formalisation. Les approches précédentes, fournissent un cadre théorique descriptif et opérationnel de l'erreur humaine et des causes d'accidents. Cependant, ils peuvent être délicats à mettre en place dans un environnement sociotechnique complexe. Enfin de fournir un cadre général aux concepteurs, Reuzeau (Reuzeau, 2000) s'appuie sur le modèle de Rabardel (Rabardel,

1995) et Boy (Boy, 1998) qui offre une représentation des champs de connaissances que les concepteurs doivent disposer dans leur activité de conception sous la forme de la pyramide AUTO. Cette représentation définit les composantes structurantes de la solution «humain-machine».

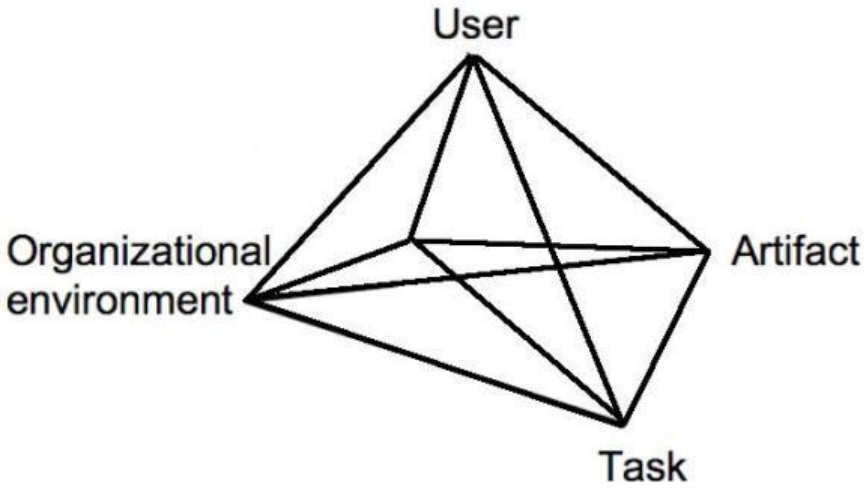


Figure 13 : La pyramide AUTO

- La Tâche : réfère aux buts des opérateurs et aux moyens de réalisation de ces buts. Elle prend en compte les procédures et les méthodes de travail.
- L'utilisateur : savoirs et savoir-faire existants ou futurs, pratiques, régulations individuelles et collectives, déviations d'usage…
- L'artefact : adéquation entre l'outil proposé et les besoins de l'utilisateur, son utilisabilité et tous les critères basiques de l'évaluation d'une IHM.
- L'organisation : les dimensions du contexte sont des facteurs d'influence premiers sur l'utilisation d'un produit (contexte social, organisationnel, économique).

Ce modèle reprend à un niveau d'abstraction plus élevé les différentes notions qui ont été abordées dans les modèles précédents, cela permet de poser un référentiel simplifié dans lequel le concepteur pourra plus facilement situer les chalenges nécessaires à la *conception centrée utilisateur*.

II.3 Les facteurs particulier structurant la conception

Il existe une multitude de facteurs connus influençant le comportement des utilisateurs et qui servent de guides pour la conception. Ces facteurs peuvent être l'environnement d'utilisation (par exemple dans le cas d'un pilote : la météo, le trafic,…), mais aussi ceux dépendant des systèmes embarqués (le degré d'automatisation par exemple). Il existe de nombreuses études sur ces facteurs et, afin d'illustrer leur importance pour la conception, nous allons prendre l'exemple d'un thème central pour la conception en aéronautique : l'automatisation et la fiabilité des systèmes.

L'automatisation des cockpits lors des vingt dernières années a modifié les compétences nécessaires au métier de pilote. Le rôle de l'équipage a évolué, partant de celui d'acteur du vol à celui de manageur du vol. Dans les paragraphes suivants, sans minimiser les avantages liés à l'automatisation, nous préfèrerons appuyer sur les inconvénients et les problèmes qu'elle crée ou révèle pour l'établissement des hypothèses de comportement de l'équipage.

II.3.1 Automatisation et Charge de travail :

L'automatisation change l'activité humaine (Parasuraman, et al., 1997) (Sarter, et al., 1992) (Wiener, et al., 1980) de manière parfois inattendue pour les concepteurs (Parasuraman, et al., 2000). L'objectif premier de l'automatisation des avions de ligne est d'aider l'équipage à remplir sa mission. Le concept d'aide sous-tend un allégement de la charge de travail des membres de l'équipage. Hughes et Dornheim (Hughes, et al., 1995) ont d'ailleurs montré que l'instauration de systèmes automatisés dans l'aviation commerciale a réduit la charge de travail des pilotes en diminuant la quantité d'affichages et en automatisant certaines procédures par exemple. Il semble néanmoins qu'elle ne soit pas seulement réduite mais que **sa répartition soit modifiée au cours du vol**. L'automatisation a augmenté la charge de travail lors de courtes phases critiques comme la (re)programmation du *« Flight Management System »* lors du décollage et atterrissage, par exemple. C'est ce que Bainbridge appelle «l'ironie de l'automatisation » (Bainbridge, 1987). Les systèmes automatiques ont encouragé l'apparition de nouvelles tâches (comme les tâches de manipulation, de gestion ou de surveillance). Ces nouvelles tâches peuvent être des sources possibles d'erreurs et peuvent être consommatrices de temps. Au contraire, durant la phase de croisière, qui est par nature moins chargée, l'automatisation réduit la charge de travail de l'équipage. Or, cette diminution de l'activité et de la charge mentale engendre des difficultés de maintien de l'attention des pilotes. En fait, le problème se pose en ces termes : l'automatisme a tendance à réduire la charge de travail de façon globale, créant par moment des points de faible charge de travail. Un manque éventuel de charge peut produire des pertes de concentration, des pertes d'efficacité en situation critique (Roscoe, 1989). L'automatisation produira aussi par moment des points de charge de travail élevés avec un accroissement de la consommation des ressources dans des phases difficiles et/ou en situation de crise (Kantowitz, 1994) (Hoc, 2003). En résumé, la sous charge de travail par l'automatisation est donc une préoccupation aussi importante que la surcharge(Young, et al., 2002). De plus, l'Homme régule efficacement sa charge de travail lorsqu'il peut anticiper à condition qu'il soit véritablement dans la boucle de l'action. Or, les systèmes automatisés tendent à mettre les pilotes hors de la boucle de pilotage, c'est ce que soulignent les travaux de Sarter et al. (Sarter, et al., 1995). Cependant, cette éventualité issue de l'automatisation semble être nuancée par le niveau de contrôle des opérateurs agissant comme un facteur modérant

cette perte de conscience de la situation (Endsley, et al., 1995). De plus, Bowers (Bowers, et al., 1995), soutient que les systèmes automatisés redistribuent la charge de travail plus qu'ils ne l'augmentent. Cela se traduit par des communications plus fréquentes dans le cockpit entre les membres d'équipage.

II.3.2 Automatisation et compréhensibilité des systèmes embarqués

De nombreuses études dans le domaine de l'aviation (notamment celles de Sarter & al (1995), (1992) (2000a), (2000b)) ont montré que l'automatisation informatisée des cockpits est une variable importante de la complexité de la tâche de pilotage. En même temps, ces automatismes sont des instruments de pilotage auxquels les pilotes peuvent déléguer des tâches de conduite, et qui de ce fait facilitent le pilotage (Plat, et al., 2000). Il y a donc a priori une tension entre deux effets : complexité et aide et, face à un dysfonctionnement d'automatisme, les pilotes vont à fois chercher à comprendre le fonctionnement du système et chercher à ne pas perdre l'appui d'un tel instrument. Tout d'abord, la complexité d'un système peut amener le pilote à se questionner sur ce que fait le système et pourquoi il agit (Wiener, 1989). Les automatismes, pour le pilote sont régis par une « logique opaque » (Woods, 1996). Ce phénomène est d'ailleurs augmenté par le fait que les automatismes sont devenus très autonomes et peuvent modifier leur comportement sans l'intervention du pilote et sans apporter toujours les informations suffisantes pour qu'il comprenne l'action engagée (Sarter, et al., 2000b). Enfin, Bainbridge, Amalberti et Hoc (Bainbridge, 1987) ; (Hoc, 2003) ; (Amalberti, 2003) soulignent que le choix entre contrôle automatique ou contrôle manuel dépend de ce que l'équipage peut raisonnablement imaginer du comportement du système. Le choix de la reprise en manuel dans le cas d'une panne dépendra alors de la compréhension qu'a le pilote sur l'automatisme.

Ces constats amènent donc directement des informations sur les liens qui se créent entre conception des systèmes et comportements attendus des équipages.

II.3.3 Automatisation et partage des tâches entre l'Homme et la machine

Lorsque nous parlons de systèmes d'aide au pilotage, comme l'expression le désigne, nous sous entendons que le pilote réalise son activité en étant aidé par le système. Néanmoins, il s'avère que les pilotes sont déchargés de nombreuses tâches au profit des systèmes qui les effectuent de manière automatisée c'est-à-dire sans l'intervention du pilote. La raréfaction de certaines actions conduit les pilotes à perdre des habiletés nécessaires à la bonne réalisation des tâches (Bainbridge, 1987) (Hopkin, 1994). De la même façon, Rasmussen & Golstein (Rasmussen, et al., 1985) expliquent qu'il y aura une perte de compétences dans tous les domaines qui sont généralement pris en charge par l'automatisme. Or cela donne un paradoxe bien connu dans le domaine des activités automatisées : c'est lorsque le

système ne pourra plus gérer seul la situation que le pilote devra intervenir et mettre en œuvre ses compétences perdu. Finalement, plus le pilote est assisté et plus la sécurité en situation normale est assurée, mais plus **certaines compétences s'érodent.** Même rare, la survenue de l'imprévu pourra le mettre alors en défaut, au moment où il a le plus besoin des dites compétences. L'érosion des compétences par l'aide apporté par des systèmes automatisés doit être prise en compte lors de la conception. Les hypothèses sur le comportement pilote, qui sont valides pour les anciens modèles ne peuvent pas être forcement réutilisées pour les nouveaux.

II.3.4 Augmentation de la fiabilité des systèmes

L'amélioration sans cesse croissante de la fiabilité des systèmes et des automatismes, a abouti à un très faible taux de confrontation des pilotes face à une défaillance technique. Il ne s'agit pas ici de remettre en cause cette fiabilité qui est nécessaire à la sécurité du transport aérien, mais plutôt de mettre en avant les effets pervers que cela peut entraîner. En effet, en cas de problème le pilote peut avoir tendance à remettre en cause son activité plutôt que celle du système. Que cela soit pour un novice ou un expert, le fait d'être moins confronté à des pannes (la formation ne pouvant pas couvrir la grande diversité des cas réels) peut amener les pilotes à perdre confiance dans leur capacité à gérer les pannes (Amalberti, 1996). Le pilote peut avoir tendance, en cas de problème, à remettre son travail en cause avant de penser à une erreur de l'automatisme.

Le faible taux de confrontation aux pannes accentue d'autant plus l'érosion des compétences, vue dans le paragraphe précédent, due au manque de pratique dans tous les domaines que l'automatisme prend habituellement en charge, (Rasmussen, et al., 1985). La capacité à faire un « retour en manuel » peut poser problème et, les pilotes auront tendance à reprogrammer les automatismes au lieu de passer directement en manuel.

Plus généralement, les pilotes ont tendance à croire que l'on doit passer toujours par l'automatisme, (Amalberti, 2003).

II.3.5 Conclusion sur les facteurs influençant la conception

Nous avons repris, dans ce paragraphe, un exemple de facteur qui influence la conception : l'automatisation. Aujourd'hui toutes les études de conception d'un avion de ligne prennent en compte la place de l'Homme et celle de la machine dans la tâche à réaliser afin de garantir une automatisation optimisée. Bien attendu l'effet de l'automatisation n'est pas le seul facteur à prendre en compte. Mais cet exemple illustre bien la nécessité, durant la phase de conception, de s'intéresser aux études faites sur ces facteurs particuliers.

II.4 Tracer les solutions de conception

Concevoir un objet nécessite de nombreuses prises de décision. Ainsi, retrouver l'explication de ces prises de décision est souvent une tâche ardue, car ainsi que l'indique Moran et Carroll (Moran, et al., 1996) : *« l'artefact n'exprime pas ou de manière très subtile, le processus de conception. En effet, l'artefact ne reflète pas les objectifs initiaux de sa conception, le cahier des charges, les conditions de mise en œuvre, les luttes, les discussions et les négociations, les essais et les réflexions, l'équilibrage et le compromis entre divers facteurs, le pourquoi de ces caractéristiques et le pourquoi des caractéristiques qu'il n'a pas, etc. »*

Pour Moran et Carroll (Moran, et al., 1996), la logique de conception représentant les raisons et les raisonnements sous jacents à la conception et à la spécification de l'artefact doit être documentée. Cette documentation doit rassembler les éléments suivants :

- La documentation de l'histoire de la conception (intentions du concepteur, contraintes contextuelles, objectifs, besoin de l'utilisateur, etc.)
- La documentation des justifications de conception.
- La documentation des étapes de conception.

La logique de conception constitue une aide si elle **aide la réflexion du concepteur**, si elle lui permet de **gérer les dépendances entre les problèmes de conception**, si elle peut **favoriser la réutilisation de connaissances pour la conception ou la re-conception** et si **elle permet aux différentes personnes impliquées dans le projet de communiquer et de coopérer.** Ces différents points serviront de structure à la fin de ce paragraphe.

II.4.1 La documentation comme support de réflexion.

La mise en place de la logique de conception nécessite une certaine discipline (Grudin, 1996) et une certaine réflexion critique (Fischer, et al., 1996) qui facilite le raisonnement et donc favorise la conception. Ainsi, la logique de conception est utile si elle aide les concepteurs à créer, représenter et réfléchir leurs décisions. Cela permet notamment la représentation et la réflexion des alternatives possibles et de poser les arguments pour et contre (MacLean, et al., 1996).

II.4.2 La documentation comme support à la gestion des liens entre les problématiques de conception.

La conception peut être assimilée, selon Lee (Lee, 1997), à un processus de gestion de dépendances entre les besoins exprimés et leur mise en œuvre. De ce point de vue, la logique de conception explicite les relations de dépendances, les arguments et les alternatives possibles afin de garantir une

meilleure cohérence. La gestion des dépendances peut être effectuée simplement en affichant les problèmes qui dépendent d'un problème donné ou de manière plus complexe pour les systèmes détectant les conflits entre différentes contraintes.

II.4.3 La documentation comme support de réutilisation, conception, re-conception

La documentation de la logique de conception doit permettre la réutilisation de la connaissance, au sein des mêmes équipes mais bien entendu au sein d'équipes différentes. Ormerod, et al (Ormerod, et al., 1999), distinguent deux catégories d'informations réutilisables : la première orientée produit correspond à la réutilisation d'une solution ou d'une information. La seconde, orientée processus, porte sur la réutilisation des idées, des conversations, des arguments, des décisions, des critiques… Dans le même ordre d'idée, Lee et Lai(Lee, et al., 1996) distinguent, d'une part, la description et la justification fonctionnelle par rapport aux besoins et exigences de conception et d'autre part la description de « l'espace » dans lequel c'est produit la conception. La logique de conception peut contribuer à la réutilisation des connaissances, lorsqu'elle permet de concrétiser les connaissances issues de projets antérieurs et, également, lorsqu'elle peut être exploitée par des concepteurs en tant que telle pour, par exemple, rechercher les alternatives de conception étudiées, rechercher les critères d'évaluation déterminant, la sélection d'une alternatives de conception… (Lee, 1997).

II.4.4 La documentation comme support à la gestion de projet

La conception est un processus social qui, en général, implique de nombreuses personnes, chacune d'elles ayant différents buts et différentes compétences (MacLean, et al., 1996). La documentation en conception est un support de communication nécessaire au sein d'une équipe de conception, mais aussi entre les équipes de conception, entre les concepteurs et l'utilisateur et pour finir entre les équipes de conception actuelles et les futures équipes qui auront pour objectif la re-conception ou l'utilisation de partie de la conception présente. La logique de conception peut être utile aux utilisateurs finaux s'ils sont, par exemples, amenés à participer à l'affinage du système afin que celui-ci soit plus adapté à leurs besoins. Ils ont ainsi les moyens de comprendre comment le système fonctionne et comment y apporter des modifications en préservant sa cohérence (MacLean, et al., 1989). Cette documentation permet également de retrouver les problèmes non résolus, former les nouveaux participants au projet, reprendre des conceptions interrompues et favoriser la répartition des tâches (Lee, 1997).

III Conclusion

Dans ce chapitre nous avons présenté la *conception centrée utilisateur* principalement à travers ISO 13407 et les méthodes qui peuvent soutenir cette conception. Cela nous a permis de percevoir l'intérêt reconnu des processus itératifs pour une conception répondant aux exigences des utilisateurs. L'intégration de phases de prototypage et de validation successives favorise l'implication des utilisateurs au cours de la conception. La *conception centrée utilisateur*, propose aussi un grand panel d'outils permettant de supporter sa mise en place. Même si ces outils n'ont pas une opérationnalisation facile, cela permet un support théorique aux questions que vont se poser les concepteurs et une participation active à la formalisation des connaissances nécessaires à la conception. Ces méthodes, qu'elles permettent de recueillir des données, de formaliser les connaissances, de prévoir les effets du contexte, servent aux concepteurs pour acquérir le niveau d'abstraction nécessaire à la conception. Pour finir, bien que consommatrice en temps, cette démarche est pour le moment la seule garante de l'intégration des besoins de l'utilisateur dans le futur produit.

Cependant, la difficulté pour mettre en place ce cycle de conception s'accroit avec la complexité du système sociotechnique. Certaines méthodes qui fonctionnent parfaitement à une échelle réduite se révèlent inutilisables dès que la complexité du système s'avère trop grande. Cela correspond au deuxième principe du modèle FRAM pour lequel, dans un milieu sociotechnique complexe, les procédures et les outils sont adaptés à la situation pour répondre à une grande variabilité de buts potentiellement contradictoires. Nous allons voir dans le chapitre suivant comment est mise en place la *conception centrée utilisateur* dans le cas d'un système sociotechnique complexe : La conception en aéronautique.

CHAPITRE. 2 LA CONCEPTION CENTREE UTILISATEUR EN MILIEU SOCIOTECHNIQUE COMPLEXE : CAS DE L'AERONAUTIQUE

L'activité créatrice d'une industrie en charge de la fabrication de systèmes complexes est régie par des **objectifs** à atteindre (performances, bénéfices, sécurité…) compte tenu d'un ensemble de **contraintes** à respecter (type d'utilisation, type d'utilisateurs, besoins des utilisateurs, technologies à disposition, règles…). L'accord entre ces objectifs et ces contraintes donnera un ensemble divers d'exigences qui mèneront au produit final. De ce fait, quelle que soit l'industrie, **le travail de conception mené par les spécialistes en Facteurs Humains revient donc à harmoniser les contraintes en présence et les hypothèses prises sur le comportement des utilisateurs.** Cette harmonisation est faite dans le cadre de la *conception centrée utilisateur* dont nous avons parlé dans le chapitre précédent. Dans le cas des systèmes complexes, compte tenue de leur taille, de leur complexité et du nombre d'acteurs, l'harmonisation des hypothèses de conception devient rapidement un vrai « casse tête » et une activité souvent extrêmement consommatrice de temps. Par exemple, dans l'environnement réel d'utilisation d'un système complexe, il existe une infinité de comportements utilisateurs observables, compte tenu du contexte du jour, de l'utilisateur, de l'état du système… Cette combinatoire est un obstacle majeur à la réalisation des objectifs des Facteurs Humains. Ces éléments vont à l'encontre du bon déroulement du processus de *conception centrée utilisateur* tel que nous l'avons défini dans le chapitre 1. Dans le cas où ce processus serait imparfait, l'utilisation de mauvaises hypothèses de conception peut mener à des erreurs d'utilisation et *de facto* aura un impact sur la sécurité de ces systèmes.

Avec la prise de conscience sur l'importance de l'Homme dans la sécurité des systèmes complexes et l'avènement des Facteurs Humains des 15 dernières années, l'implication des utilisateurs finaux et des spécialistes en Facteurs Humains dans la conception du produit s'ancre aujourd'hui dans les processus industriels et ce malgré la difficulté que cela représente. Dans le cas des industries à risques tel que

l'aviation par exemple, ce processus est certifié afin de garantir légalement la prise en compte des problématiques Facteurs Humains dans la conception. De ce fait le processus se doit d'être rationalisé et formalisé afin de garantir la visibilité et la traçabilité nécessaire à la certification.

Dans ce chapitre, nous resterons dans le cadre de l'industrie aéronautique et nous présenterons tout d'abord la certification à laquelle doit répondre Airbus dans le cadre des Facteurs Humains. Puis nous verrons le processus mis en place afin d'y répondre et pour finir nous présenterons une des difficultés présente dans la mise en place de ce processus de conception.

I Un processus et un avion certifié

Afin de procéder à la mise en service d'un avion, tout avionneur doit répondre à un règlement de certification. C'est ce règlement qui permet de justifier et démontrer la conformité de l'avion aux exigences minimales requises pour le niveau de sécurité actuel (par exemple moins de 10^{-7} pannes catastrophiques par heure de vol). L'autorité de certification est aussi en charge de délivrer le certificat de navigabilité qui le suivra pendant son cycle de vie. Même si chaque pays de l'Union Européenne possède encore sa propre autorité aérienne (*Direction Générale de l'Aviation Civile : DGAC,* pour la France), l'harmonisation au niveau européen a abouti à la création de l'Agence Européenne de la Sécurité Aérienne (*AESA,* en anglais *EASA*). C'est une agence de la Communauté européenne qui traite la sécurité aérienne, basée à Cologne (Allemagne) et qui est entrée en fonctions en septembre 2003. Elle a remplacé *l'European Joint Aviation Authorities* qui représentait la majorité des organisations européennes de certification et de règlementation aéronautique civile. L'AESA a des partenariats avec les autorités d'autres pays notamment la FAA aux Etats Unis *(Federal Aviation Administration).* Ces partenariats facilitent la certification des avions au niveau mondial.

La création de l'AESA à ouvert la voie à une nouvelle réglementation communautaire en matière de sécurité et de compatibilité environnementale de l'aviation civile. Le nouveau règlement de base (EC) No 216/2008 (European Union, 2008) est entré en vigueur le 8 avril 2008 et définit les missions de l'Agence.
L'AESA a pour mission d'aider la Communauté à :

- Promouvoir le plus haut niveau possible de sécurité et de protection environnementale de l'aviation civile ;
- Faciliter la libre circulation des biens, des personnes et des services ;
- Favoriser la rentabilisation des processus réglementaires et de certification ;

- Aider les États membres à remplir, sur une base commune, les obligations que leur impose l'OACI ;
- Promouvoir, au niveau mondial, les vues qu'elle défend quant aux normes de sécurité à appliquer dans l'aviation civile.

Le règlement auquel doit se conformer Airbus est le CS 25 (*Certification Specifications for Large Aeroplanes*) (European Aviation Safety Agency, 2011). Ce document contient des règles aussi bien relatives à la structure, qu'aux moteurs, au système électrique… Un effort est exercé depuis plusieurs années par les autorités de l'aviation civile afin d'établir une réglementation formalisant mieux la prise en compte des Facteurs Humains dans le processus de certification. La FAA a notamment édité deux documents temporaires :

- *FAA Policy memo ANM-99-2. Guidance for Reviewing Certification Plans to Address Human Factors for Certification of Transport Airplane Flight Decks. September 29, 1999.*

Ce document proposait une revue pratique des méthodes assurant la prise en compte des Facteurs Humains dans le plan de certification d'un avion de ligne. Ce document a été révisé en 2001 :

- *FAA Policy memo ANM-01-03. Factors to Consider When Reviewing an Applicant's Proposed Human Factors Methods of Compliance for Flight Deck Certification, Feb 14, 2001.*

Au niveau européen, les JAA ont publié en 2001 le texte suivant :

- *JAA Project of interim policy. Human Factors Aspects of Flight Deck Design, Issue 2, 15-03-2001.*

Ce document donnait les points liés aux Facteurs Humains, principalement sur les nouveautés présentes dans le cockpit, qui n'étaient que partiellement couverts par le document de certification existant (*JAR25*)

Ce texte n'est plus d'actualité en Europe depuis le transfert d'autorité à l'AESA et la parution du CS25. Le paragraphe CS 25.1302 établi sous l'impulsion du groupe d'harmonisation Facteurs Humains contient le règlement qui leurs est consacré. Celui-ci est complété par des textes d'interprétation et des moyens de conformité AMC25-1302 (*Acceptable Means of Compliance*). Nous allons maintenant présenter les idées principales du règlement CS25-1302 ainsi que des moyens de conformité AMC25-1302 afin de mieux comprendre les contraintes auxquelles Airbus doit se conformer.

I.1 Le règlement CS25-1302

Le CS25-1302 s'applique aux équipements qui seront utilisés par les membres d'équipage (Pilote et Copilote uniquement) lors de l'utilisation de l'appareil. Les démonstrations sont demandées pour une position normale du pilote dans le cockpit. Airbus doit montrer que chaque équipement pris individuellement mais aussi en combinaison avec d'autres, peut être utilisé de manière sûre pour réaliser des tâches demandées par les équipages. Voici un extrait de ce règlement reprenant les exigences principales de l'AESA :

*« a) Flight deck **controls** must be installed to allow accomplishment of these tasks and **information** necessary **to accomplish these tasks** must be provided.*
*b) Flight deck **controls and information** intended for flight crew use must :*
*1) Be presented in a **clear and unambiguous form**, at resolution and precision appropriate to the task.*
*2) **Be accessible and usable** by the flight crew **in a manner consistent with the urgency, frequency, and duration** of their tasks.*
*3) Enable flight **crew awareness**, if awareness is required for safe operation, of the effects on the aircraft or systems resulting from flight crew actions.*
*c) **Operationally-relevant behaviour** of the installed equipment must be:*
*1) **Predictable and unambiguous**.*
*2) Designed to enable the flight crew **to intervene in a manner appropriate to the task**.*
*d) […] installed equipment must **enable the flight crew to manage errors** resulting from the kinds of flight crew interactions with the equipment **that can be reasonably expected** in service, assuming the flight crew is acting in good faith […] ».*

En résumé, ce règlement a deux volets principaux.
Les parties a), b), c) forment le premier volet relatif aux comportements des systèmes (informations et contrôles) permettant aux équipages de réaliser leurs tâches. Les notions principales se trouvent dans les 3 sous parties de ce volet. Premièrement, l'installation des équipements dans le poste de pilotage doit être réalisée avec des règles anthropométriques et physiologiques afin de garantir l'accessibilité des commandes et la lisibilité des informations par exemple. Deuxièmement, les contrôles et les informations doivent être présentés de manière claire et non ambigüe afin d'aider l'équipage à réaliser sa tâche. La présentation de ces informations doit prendre en compte le contexte et peut être modifiée en fonction de l'état de l'avion et de sa phase de vol afin d'optimiser la réalisation des tâches de l'équipage. Troisièmement, le comportement des systèmes doit être prévisible et non ambigu afin que

l'équipage se construise une bonne représentation de la situation et prenne les bonnes décisions pour accomplir sa tâche.

Le deuxième volet (d) est relatif à la gestion et la récupération par l'équipage de ses erreurs. En excluant les comportements intentionnels (suicides par exemple), le règlement impose que les erreurs « raisonnablement attendues en service » soit prévues afin d'aider l'équipage à les gérer et les récupérer.

Ce règlement permet de donner un cadre à l'activité en Facteurs Humains, mais nécessite, pour être appliqué, des guides afin que le constructeur puisse si conformer, c'est objectif de l'AMC25-1302.

I.2 Les méthodes d'application AMC25-1302

L'objectif de l'AMC25-1302 est de fournir une approche méthodologique pour la certification afin de guider les concepteurs aéronautiques et leur permettre de certifier leurs avions. Nous allons reprendre dans ce paragraphe les principaux points méthodologiques apportés par l'AMC :

- La méthode recommandée
- Les objets de conception à considérer dans la certification
- Les moyens appropriés pour démontrer la conformité au règlement

I.2.1 La méthode recommandée

L'approche recommandée consiste en 4 étapes :

Etape 1 : Identifier les systèmes en interaction avec des tâches de l'équipage. Cette étape doit permettre au constructeur de cibler les systèmes sur lesquels s'appliquera la certification. Ces systèmes sont tous ceux en interaction avec l'équipage. Par exemple, le système CDS, (*ATA 31 : Control and Display System*) qui correspond à l'ensemble des écrans ainsi qu'une partie des commandes présentent dans le cockpit, fait partie des systèmes pour lesquels la certification s'applique. En revanche pour le train d'atterrissage, seul les parties en interaction avec à l'équipage sont prises en compte pour la certification Facteurs Humains (commande d'entrée et sortie du train, information d'état du système…)

Etape 2 : Identifier les équipements, les fonctions, les tâches selon leur nouveauté, leur complexité et leur niveau d'intégration. Après avoir défini les systèmes sur lesquels s'applique la réglementation, il

faut les décrire. Chaque système est composé d'équipements et doit remplir certaines fonctions. Les avions d'ancienne génération n'avaient pas à se conformer à ce règlement. Le constructeur peut s'appuyer sur ses acquis, en ciblant en priorité les nouveautés apportées par ce système par rapport à la génération d'avion précédente. Pour l'A380 (premier avion à devoir remplir cette certification) il existait énormément de changement par rapport aux générations précédentes, la certification a donc touché la majorité des systèmes en interaction avec l'équipage.

Etape 3 : Identifier les objets de conception à considérer dans la certification. Cette étape ce fait à l'aide des chapitres 5.2 à 5.6 de l'AMC, qui seront présentés dans le paragraphe suivant. L'idée de cette étape est à nouveau de décomposer le système en composants élémentaires qui pourront être alors évalués.

Etape 4 : Sélectionner les moyens de démonstration de conformité appropriés. Une fois le système décomposé, il faut trouver un moyen de l'évaluer qui est autorisé par la réglementation. Nous verrons rapidement ces moyens dans le chapitre I.2.3.

L'approche recommandée par les autorités de certification est représentée sur la Figure 14 extraite de l'AMC.

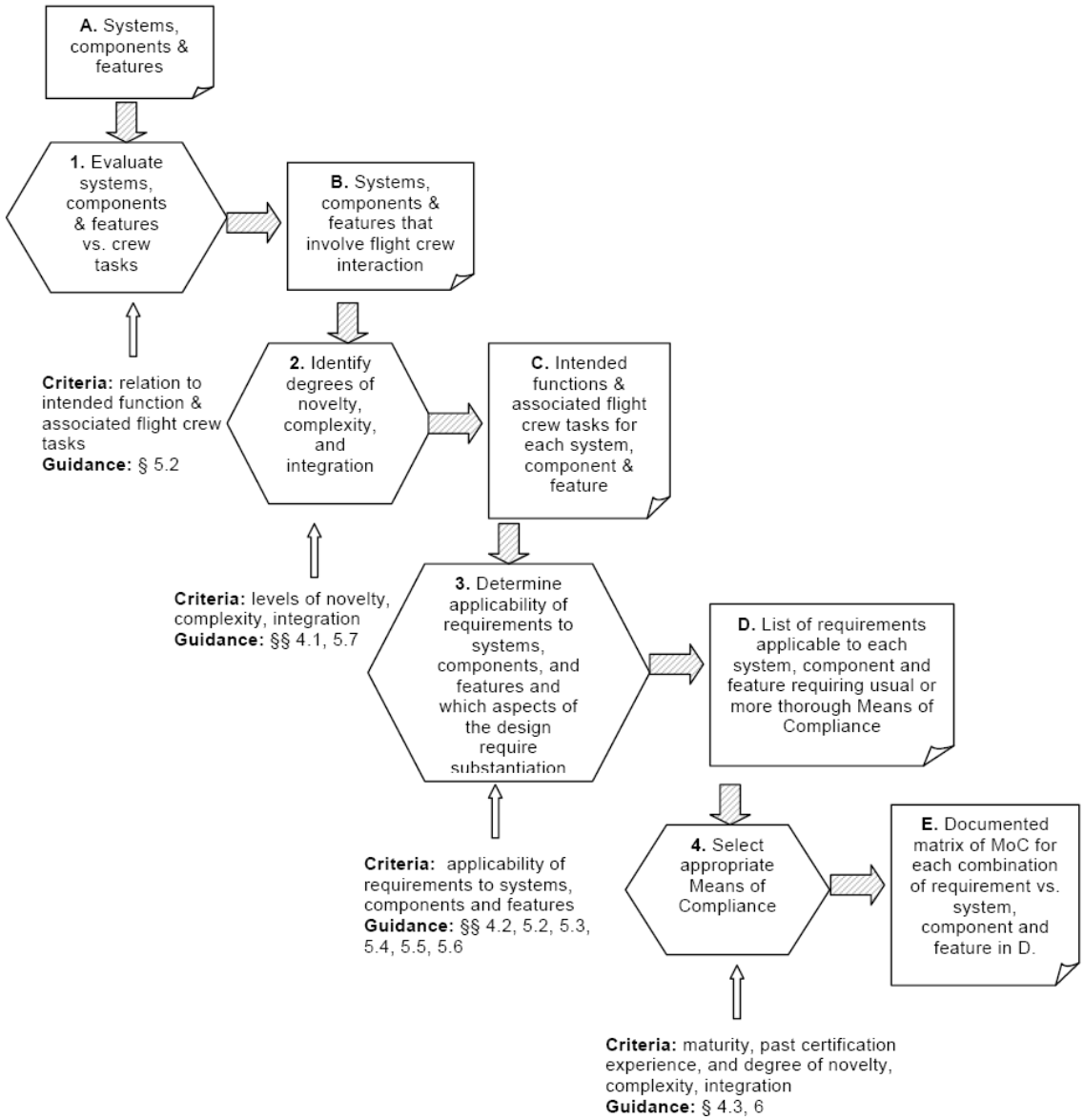


Figure 14 : Approche recommandée, extrait de l'AMC25-1302

I.2.2 Les objets de conception à considérer dans la certification

Comme évoqué lors de l'étape 3 du processus recommandé par l'AESA, l'AMC donne les objets de conception à évaluer lors du processus de certification. Ils sont de 6 grandes familles :

1/Fonction attendue et tâche (paragraphe 5.2 de l'AMC) :

Pour chaque composant du système, la fonction et la tâche qu'il doit remplir sont des aspects à certifier. Il faut savoir, par exemple, comment la fonction d'un élément particulier s'insère dans la tâche du pilote, qu'est ce qu'elle va apporter au pilote, quel était le besoin initial de cette fonction…

2/Contrôles (paragraphe 5.3 de l'AMC) :

Chaque fonction doit avoir un processus de contrôle de son activité. Ce contrôle peut être automatique et transparent du point de vue du pilote, ou bien nécessiter un monitoring de la part de l'équipage. La part automatisée et la part humaine doivent être définies ici en fonction des besoins de l'utilisateur.

3/Présentation des informations (paragraphe 5.4 de l'AMC) :

Les fonctions du système vont envoyer des informations à l'utilisateur. La façon dont se fera cette présentation est à certifier afin de garantir la meilleure compréhension de ces informations par le pilote.

4/Comportement du système (paragraphe 5.5 de l'AMC) :

Les tâches réalisées par l'équipage de l'avion vont varier en fonction de comportements prévus du système. D'un système à l'autre, les comportements sont différents, la réponse peut être directe pour un système mécanique ou indirecte lorsqu'il y a intervention d'un ou plusieurs sous-système. Dans ce contexte, il est nécessaire, pour aider la réalisation des tâches du pilote, d'assurer que le comportement du système soit prédictible et non ambigu.

5/Gestion des erreurs de l'équipage (paragraphe 5.6 de l'AMC) :

Tout équipage bien entraîné, avec de l'expérience et en bon état physique peut faire des erreurs. Ce fait reconnu est présent dans la certification. Cependant, il faut s'assurer que le système proposera des moyens adaptés de récupération des erreurs à l'équipage.

6/Intégration (paragraphe 5.7 de l'AMC) :

Il y a beaucoup de systèmes intégrés physiquement et/ou fonctionnellement dans le cockpit. Ces systèmes peuvent interagir les uns avec les autres. Il est donc important de considérer la conception de ces systèmes dans ce contexte et non pas de manière isolée. La certification de l'intégration des systèmes inclue non seulement la façon dont est installé un écran ou une commande mais aussi l'homogénéité des systèmes les uns avec les autres.

Après avoir présenté les points à certifier pour un équipement, nous allons voir les moyens de démonstration autorisés par l'AMC.

I.2.3 Les moyens de démonstration de conformité

Comme pour les objets à considérer dans la conception, l'AMC donne les moyens d'évaluation de la conformité. Pour chaque moyen, l'AESA pose certaines conditions d'utilisation. Nous allons présenter rapidement les 5 moyens utilisables pour montrer la conformité.

1/ Déclaration de similarité (paragraphe 6.3.1 de l'AMC)

Une déclaration de similarité contient la description d'un système qui doit être approuvé ainsi que la description d'un système précédemment approuvé. La description doit contenir les détails physiques, logiques et opérationnels de chaque système et montrer la similitude entre les deux systèmes. Il existe des limitations dans l'utilisation de la déclaration de similarité car le cockpit doit être pris dans son ensemble et non comme la somme de plusieurs systèmes. La déclaration de similarité s'applique, par exemple, pour des avions aux pare-brises similaires.

2/ Description de la conception (paragraphe 6.3.2 de l'AMC)

Une description de la conception à l'aide de schémas ou de dessins industriels montrant l'arrangement et la position du système peut servir à la certification. Les codes couleurs et les philosophies d'interaction montrant l'homogénéité dans le cockpit sont à inclure dans ce document. Globalement, il faut inclure, toutes les descriptions pouvant être utilisées pour démontrer une philosophie centrée sur la sécurité des occupants de l'avion.

3/ Analyses et calcul (paragraphe 6.3.3 de l'AMC)

Les analyses et les calculs peuvent être des moyens de démonstration. Ce sont des analyses qui ne demandent pas de participation d'utilisateurs et qui s'appuient sur une expertise ou l'utilisation d'outils mathématiques. Par exemple les cônes de vision extérieure calculés grâce à des mannequins numériques suffisent à démontrer la forme prévue des vitres du cockpit.

4/ Evaluations (paragraphe 6.3.4 de l'AMC)

Les évaluations s'effectuent sur des maquettes, des bancs d'essais, en laboratoire ou sur avion réel. Le constructeur fait passer sur un de ces moyens d'essais des utilisateurs et recueille des informations sur différents aspects de la conception à considérer dans la certification, comme par exemple, la récupération des erreurs de l'équipage. Les rapports d'évaluations qui reprennent les objectifs d'évaluation et les résultats servent à la certification.

5/ Tests (paragraphe 6.3.5 de l'AMC)

Les tests sont des évaluations auxquelles les autorités de certification participent. Cela permet de garantir une meilleure prise en compte du règlement. Les moyens utilisés pour réaliser des tests doivent être conformes à la réalité des systèmes.

Le moyen approprié peut être sélectionné en fonction du système et des moyens d'étude à disposition. Cependant l'AESA propose une série de questions permettant d'aider au choix du moyen de test de conformité.

- Quels sont les moyens de démonstration qu'il est possible d'utiliser pour rassembler les données de certification ?
- Est-ce qu'un seul moyen est possible, ou est-il possible d'en utiliser plusieurs en parallèle ?
- Quel est le degré de fidélité nécessaire à avoir pour collecter les données ?
- Qui doit participer aux moyens de démonstration ?
- Quel est le niveau d'entraînement et de connaissances nécessaires pour participer à la démonstration ?
- Comment doivent être présentées les données d'évaluation pour montrer le respect du règlement ?
- Est-ce que les résultats doivent provenir d'une démonstration pour être intégrés dans la certification ?
- Si un test est nécessaire, quel est le moyen conforme à utiliser ?

I.3 Conclusion sur le règlement de certification

Parmi les textes consacrés à l'évaluation des Facteurs Humains, le nouveau paragraphe 25.1302 du règlement de certification et son AMC représentent une avancée considérable par rapport aux textes précédents difficilement applicables. Fait remarquable, il est entré en vigueur d'abord au niveau européen. Les Facteurs Humains sont une discipline de l'ingénierie qui a la maturité suffisante pour réaliser cette certification. Les processus et les méthodes mis en place, complètent ceux de l'ingénierie générale.

Cependant, la mise en place de ces méthodes n'est pas triviale. La réponse aux exigences de certification mais aussi aux exigences de pilotage du projet global, oblige la réalisation de compromis. Ces compromis sont d'ordres techniques mais aussi industriels. Dans le prochain paragraphe, nous allons présenter comment Airbus a mis en place les méthodes pour répondre à ces exigences. Durant cette présentation, après avoir rapidement indiqué le processus général, nous tâcherons de nous positionner plus à un niveau opérationnel. En effet, nous voulons pointer les difficultés de mise en

place de la *conception centrée utilisateur* et ces difficultés ne sont pas toujours visibles au niveau des processus.

II Processus de conception : Mise en œuvre en aéronautique

Airbus a acquis une large expertise pour la prise en compte des Facteurs Humains durant la conception des cockpits et la certification des avions de ligne (Reuzeau, 2005). Toutes les études entreprises sur les cockpits d'avion sont organisées comme une séquence d'expériences scientifiques : création d'hypothèses à tester selon un protocole expérimental ; collection des données ; étude des données. Le processus de *conception centrée utilisateur* (présenté sur la Figure 15) mis en place chez Airbus reprend les étapes de la norme ISO13407, c'est-à-dire une étape de recueil des besoins, une étape de conception, une étape d'évaluation. Ces étapes entrent dans un processus itératif de conception.

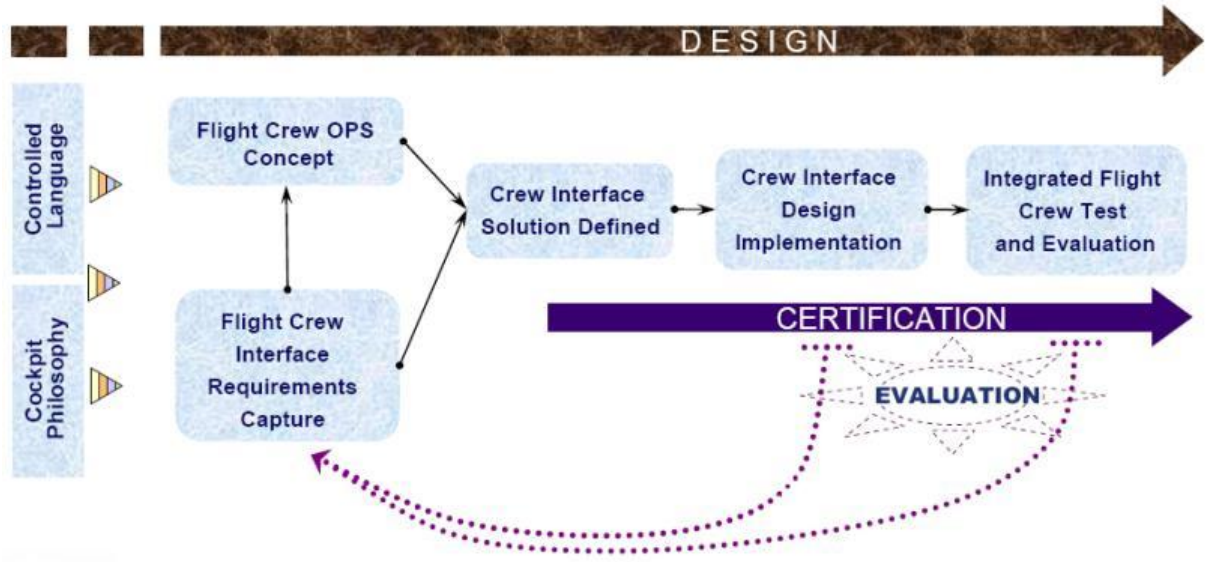


Figure 15 : Processus de conception centrée utilisateur mis en place chez Airbus.

Du point de vue des processus industriels, le constructeur doit aussi garantir que la maturité de l'avion est atteinte à la mise en service de l'appareil. L'activité de mise au point qui pouvait exister trente ans auparavant pendant la première année d'exploitation doit donc être incluse dans le cycle de conception/validation de l'avion. Ce cycle est relativement court : d'une durée moyenne de 6 ans, alors que l'exploitation d'un avion peut courir sur plus de 20 ans. Nous verrons donc que les actions de conception/validation du cockpit en Facteurs Humains peuvent être assimilées à une activité de conception (incubation et matérialisation de concepts) conjuguée à une simulation « en accéléré » de la vie future de l'avion dans une compagnie aérienne. Cette période intègre l'exploration et la

compréhension des comportements futurs raisonnables pour résoudre les problèmes de conception liés à l'utilisation (Reuzeau, 2005). Les activités de validation portent en elles le devoir d'acquérir un niveau de confiance suffisant sur la performance Homme/Machine, sur le respect des objectifs de sécurité et de l'efficacité opérationnelle.

L'objet de la conception, le cockpit, est une infrastructure d'accueil dans laquelle les objectifs fonctionnels du système aérien sont déclinés en fonctions abstraites liées au vol et à la circulation des aéronefs puis en fonctions générales du couplage équipage-avion. Les fonctions générales arrangées en une architecture multi-système (allocations de fonctions générales entre humains, ressources matérielles et immatérielles) sont structurées par des logiques de fonctionnement et des logiques d'utilisation. Le processus de conception et de validation se met notamment en quête des difficultés d'utilisation et des erreurs humaines pour trouver les meilleures parades (Reuzeau, 2005) (Reuzeau, et al., 2010). En réalité, l'ampleur de la recherche est sans limite mais d'après Reuzeau (Reuzeau, 2005) nous pouvons le résumer aux comportements individuels et collectifs qui conduisent à une conséquence dommageable pour la performance de l'avion ou de l'équipage (performance opérationnelle et impact sur la sécurité) ou pour la santé de l'équipage. Par exemple, un briefing d'approche non réalisé a un impact potentiel sur la sécurité s'il existe plusieurs procédures d'approche possibles. Les pilotes pourraient diverger dans la gestion de l'approche s'ils ont en tête des plans d'actions différents. Nous appellerons l'ensemble de ces comportements dommageables, les difficultés d'utilisation ou les erreurs humaines, définies et distinguées comme suit :

- **Les erreurs humaines** sont des manifestations physiques ou cognitives, intentionnelles qui échouent dans la réalisation de l'intention ou des échecs dans la mise en place du plan d'action.
- **Les difficultés d'utilisation** diffèrent par le fait que l'utilisateur réussit le plan d'action mais d'une manière indirecte et coûteuse à son encontre.

Les difficultés d'utilisation et les erreurs humaines traquées sont celles qui pourraient se produire «raisonnablement en service» dans des vols quotidiens même si ce quotidien peut être source d'inconfort pour les équipages en raison de gestions de contraintes et d'imprévus imposés par le contrôle aérien, les retards passagers et toutes contraintes posées par le transport commercial. La source de ces erreurs ou difficultés est multiple, l'intérêt des évaluations est de voir comment la conception pourrait réduire leur nombre d'occurrence.

De plus, lorsque l'on s'intéresse aux systèmes complexes, il faut posséder des outils descriptifs capables d'expliciter les modes de fonctionnement du système Homme/Avion, système extrêmement intégré avec énormément d'interdépendances entre ses différents composants (Reuzeau, et al., 2010). Nous avons consacré une partie du chapitre 1, à une présentation de ces modèles d'abstraction. Ces modèles sont utilisés notamment pour poser les objectifs de tests et d'évaluation. Cela permet d'assurer la cohérence du processus envers les théories du fonctionnement humain.

Ce paragraphe présentera tout d'abord les objets de l'étude Facteurs Humains, c'est-à-dire au niveau général, les cockpits d'avion. Nous verrons quelles sont les méthodes mises en place pour leur évaluation mais aussi la façon dont elles sont mises en place. Pour finir nous présenterons les expérimentations mises en place lors de la certification de l'Airbus A380.

II.1 Les objets d'études

La quête des erreurs et des difficultés se fait tout d'abord, sur des primo-solutions (voir l'illustration Figure 16 du simulateur de l'A380 en fin de conception). Ces premiers développements « humain-machine » émergent de la rencontre entre des concepts techniques disponibles et connus du collectif de conception, des modes d'usages envisageables a priori, des hypothèses sur les compétences des pilotes qui seront qualifiés sur la future machine, d'une organisation probable du futur travail des pilotes, des rôles et tâches futures probables des utilisateurs (en vol et au sol) que ces tâches soient individuelles ou collectives. L'organisation de la conception comprend les structures sociales cadrant le travail des opérateurs, les conditions de travail (i.e. horaires de travail, environnement de travail), les politiques des sociétés (i.e. politique de profit maximum, politique salariale, politique sécuritaire), l'appartenance historique de la société (i.e. les transports ferroviaires, le nucléaire, l'aviation, l'imprimerie) avec toutes les contraintes implicites ou explicites, véhiculées par l'image de ces domaines.



Figure 16 : Simulateur pour l'A380

La solution de conception est donc un objet multiforme. Chacun des pôles ainsi que les interactions entre pôles doivent être validés. Suivant les types de nouveautés introduites dans un cockpit (innovation fonctionnelle ou modification d'un système existant) un ou plusieurs pôles de connaissances peuvent être impactés. Tout comme les niveaux d'abstraction choisis, l'ensemble des pôles doit être évalué pour conclure sur l'utilisabilité de l'avion.

Les solutions de conception sont testées de manière itérative grâce à des méthodes agrées par les autorités que nous allons présenter dans le paragraphe suivant.

II.2 Des méthodes pour expliciter, observer et analyser les difficultés et erreurs humaines pertinentes

D'après Reuzeau (Reuzeau, et al., 2010), deux classes de méthodes sont envisagées dans le domaine de la conception aéronautique pour des contraintes de temps de cycle de conception, de coût et d'accessibilité à des données valides : **les méthodes prédictives et les méthodes expérimentales.** Ces méthodes sont agrées par les autorités puisqu'elles correspondent 1/aux méthodes d'analyses et calcul (paragraphe 6.3.3 de l'AMC) pour les méthodes prédictives, 2/aux méthodes d'évaluations et aux tests (paragraphe 6.3.4 et 6.3.5 de l'AMC) pour les méthodes expérimentales.

II.2.1 Les méthodes prédictives

Les méthodes prédictives consistent à inventorier une liste d'erreurs théoriques (i) par application de modèles d'erreurs humaines (voir chapitre 1) sur les nouveautés dans des conditions normales et dégradées, (ii) par jugement d'experts (sécurité, Facteurs Humains, concepteurs). Les avantages des méthodes prédictives à base de modèles résident dans l'exhaustivité des hypothèses de comportement testées. Cependant, l'expérience industrielle a montré que cette exhaustivité est tellement coûteuse en temps pour les concepteurs qui les réalisent, qu'elles ne peuvent se faire qu'au détriment d'analyses approfondies. Par ailleurs, l'exhaustivité n'est qu'apparente car la qualité du résultat relève de la compétence des personnes qui déclinent les modèles d'erreurs. Un modèle aussi juste et précis soit-il ne garantit pas la prédiction des erreurs humaines sur des solutions innovantes pour un système aussi complexe que l'aéronautique. Les méthodes prédictives efficaces reposent donc sur des experts rompus aux modèles d'erreurs humaines, seuls à même de choisir 1/où stopper l'analyse et 2/ les points à analyser en profondeur.

II.2.2 Les méthodes expérimentales

Les méthodes expérimentales consistent à établir des scénarii opérationnels représentant une ou plusieurs activités singulières sur la base de l'ensemble des hypothèses de test des nouveautés et des contextes opérationnels donnés. Puis de conduire des évaluations pilotées (en simulateur), avec la participation d'un échantillon représentatif de pilotes futurs. Ces scénarii (Reuzeau, et al., 2010) ressemblent à la réalité mais font l'objet d'un ingénieux jeu de montage d'événements, propres à stimuler les pilotes et faire émerger le plus banal, le meilleur ou le pire des comportements « humain-machine ». La réalisation du vol simulé doit laisser des espaces d'improvisation aux pilotes où ils exerceront librement des choix de stratégies et exploreront des comportements tout en restant sur un schéma de vol contrôlable par l'expérimentateur. Les pilotes sont tout à la fois naïfs, manipulés et maîtres du jeu. Un avantage notable de cette méthode est d'explorer des territoires à la fois connus par le concepteur du scénario mais aussi des espaces inconnus puisque les pilotes ont une liberté d'action et d'initiative. Elle permet donc de se doter de la capacité de découvrir et d'apprendre des comportements pilotes nouveaux.

En Aéronautique, nous sommes en présence d'un système extrêmement complexe avec des interdépendances qui peuvent être difficilement envisageables à priori. De plus, la réglementation demande l'étude des nouveautés apportées à la famille d'avion en conception. C'est pourquoi, il nous semble important de préciser comment ces deux notions sont prises en compte dans les méthodes de conception en aéronautique.

II.3 Conception : prise en compte de la complexité et évaluation des nouveautés

II.3.1 Prise en compte de la complexité pour l'utilisateur

Du point de vue de l'utilisateur, les difficultés d'utilisation émanent de plusieurs facteurs. Notamment des imperfections :

- De chacun des nombreux d'éléments du système « humain-machine » (plus d'une trentaine de systèmes techniques sont accessibles aux pilotes dans un cockpit) ;
- De l'intrication entre les différents éléments, (les échanges de données entre systèmes par exemple) ;
- De l'immersion de ce système dans l'immense diversité des situations opérationnelles (les types de vol, les conditions météorologiques, les pays survolés et leurs particularités de contrôle aérien, les structures et les contraintes organisationnelles des compagnies aériennes, les règlements nationaux ou internationaux).

La complexité perçue par les utilisateurs dépend de leur capacité individuelle mais aussi de leur performance instantanée, et donc des capacités qu'ils peuvent mettre en jeu «là-haut et maintenant » (Jouanneaux, 1999). Ces capacités sont mobilisées pour gérer la situation en fonction de leur état physique (fatigue, vigilance), de leur ressources cognitives mobilisées et de leur état psychique. Les difficultés d'utilisation et les erreurs humaines sont à rechercher dans les composantes de cette complexité. Ces composantes naissant de la composition du couplage «humain-machine» et des facteurs agissant sur la complexité perçue.

C'est dans ce contexte que la conception des cockpits et des systèmes avion s'attâche à organiser les informations, les commandes et les logiques suivant des principes structurants de manière à réduire la complexité perçue. Ces principes sont :

- Les quatre tâches opérationnelles pivots (Voler, Naviguer, Communiquer, Gérer les systèmes) ;
- Les phases de vol (Roulage, Décollage, Montée, Croisière, Préparation de Descente, Descente, Approche, Atterrissage, Roulage) ;
- Une organisation différente entre les conditions normales du point de vue de l'avion et les conditions dégradées ;
- Les fonctions générales principales qui structurent aussi la formation initiale du pilote de ligne (Carburant, hydraulique, commandes de vol, pilote automatique, conditionnement d'air.).

Ces principes, complétés par des règles de conception d'IHM classiques, sont les garants de la cohérence du cockpit, condition nécessaire à sa bonne compréhension et à sa facilité d'utilisation. Ces principes d'organisation sont agréés avec les autorités de certification.

II.3.2 Gérer les nouveautés

Pour chaque nouveau programme d'avion ou pour chacune des évolutions majeures d'un programme existant, par règlement, le constructeur doit définir le niveau de nouveauté et de complexité de l'évolution proposée. Pour le programme A380, une douzaine de nouveautés furent décrites. Certaines de ces nouveautés concernaient la fonction physique comme l'introduction d'écrans interactifs pour la manipulation du plan de vol, d'autres se plaçaient sur le plan de la nouveauté fonctionnelle comme la mise à disposition d'une nouvelle manière de conduire une approche de non-précision, d'autres encore se rapportaient à des modifications de forme physique comme le codage de l'information ou des détails de manipulation.

En plus de la nouveauté elle-même, une nouvelle fonction peut amener un pilote à négliger une autre fonction. Les processus de conception et de validation doivent donc inclure des études et des objectifs d'évaluation capables d'adresser plusieurs niveaux d'abstraction. Tout d'abord, pour aider la compréhension des nouveautés mais aussi sur un périmètre plus large que son propre niveau d'abstraction parce qu'une nouveauté peut créer des perturbations sur d'autres fonctions.

Ces deux niveaux d'abstraction sont donc nécessaires pour évaluer les nouveautés. Les concepteurs devront donc décrire à l'aide des modèles d'abstraction à la fois le système nouveau en lui-même mais aussi les liens et les influences que celui-ci aura sur les autres systèmes. Ce raisonnement est applicable aussi en cas de nouveautés au niveau fonctionnel[4].

Les méthodes précédentes quelles soient expérimentales ou prédictives, mises en place pour gérer la complexité ou la nouveauté alimentent les entrées des évaluations. Les sommes d'hypothèses de test servent donc à produire des scénarii dans lesquels des panels sélectionnés d'utilisateurs se confronteront au design. C'est ce qui est présenté dans le prochain paragraphe.

II.4 Les expérimentations

A l'aide des méthodes décrites ci-dessus, les spécialistes en Facteurs Humains et les concepteurs décrivent une liste limitée de difficultés d'utilisation en fonction de leurs connaissances sur les

[4] *NB : Dans l'aéronautique, comme dans d'autres industries, l'objet de la conception possède des fonctions. Ces fonctions sont implémentées à l'aide d'un ou plusieurs systèmes.*

nouveautés, de la maturité de la solution et des connaissances sur les compétences des pilotes. Cette liste est transformée en hypothèses de tests pour être incluse dans un scénario de vol.
Les scénarii de vol comportent nombre d'autres événements qui instancient une multiplicité de conditions normales et dégradées dans la perspective d'exposer les pilotes à une diversité optimum de situations. Le choix des situations rencontrées dans les scénarii d'évaluation ne se fait pas en fonction d'une situation moyenne de vol. En effet, afin de gérer au mieux la diversité de situations présentes il est préférable de choisir des situations structurantes, limites. Le choix de ces situations doit permettre à la fois de confronter le design à une utilisation classique de vol mais surtout de voir la stabilité du design dans des conditions marginales.

Une expérimentation peut être locale ou globale en fonction de ce que l'on veut adresser. Les évaluations locales permettent de tester les logiques d'un système mais avec un environnement contextuel réduit. Pour les évaluations globales le contexte est, là, équivalent à un cas réel. Pour le programme A380 il y a eu 3 évaluations globales à différents niveau de maturité du cockpit. Dans une expérimentation donnée, chaque pilote joue en équipage 3 à 5 scénarii d'environ 2 heures. Pour chacune de ces évaluations, 8 équipages (1 pilote et 1 copilote) ont participé. Les campagnes d'évaluations globales font appels tant à des pilotes d'essais, qu'à des pilotes instructeurs (plus à même d'avoir du recul sur les difficultés des pilotes en formation), que des pilotes des compagnies aériennes. Ce protocole est agréé par les autorités de certification. Au total plus de 1000 heures de simulation ont été réalisées pour le programme A380. Durant les campagnes globales plus de 617 erreurs et difficultés ont été observées. Au total plus de 302 erreurs et difficultés uniques ont été tracées, étudiées et prises en compte dans la conception (Reuzeau, et al., 2010).
La prise en compte des erreurs et difficultés passe tout d'abord par la classification de leurs conséquences sur la sécurité des vols. Ce classement se base sur la classification, « *NSE* », « *MIN* », « *MAJ* », « *HAZ* », « *CAT* » provenant du règlement de certification CS25-1309 relatif à la sûreté de systèmes et présenté succinctement dans l'introduction. Ces termes signifient :

- *NSE, « No Safety Effect » :* Dans cette classification, l'évènement (erreur/difficulté ou panne système) n'a pas d'impact sur les capacités opérationnelles de l'avion et sur sa sécurité.
- *MIN, « Minor » :* **Effets sur l'avion :** Faible baisse des marges de sécurité ; **Effets sur les passagers :** Inconfort. **Effets sur l'équipage :** Augmentation légère de la charge de travail. (Probabilité d'occurrence : inferieure à 10-3 par heures de vol).
- *MAJ, « Major » :* **Effets sur l'avion :** Baisse significative des capacités opérationnelles ou des marges de sécurité. **Effets sur les passagers :** Douleur physique incluant possiblement de graves blessures. **Effets sur l'équipage :** Inconfort physique et augmentation significative de la charge de travail. (probabilité d'occurrence < 10^{-5} par heures de vol).

- *HAZ, « Hazardous » :* **Effets sur l'avion :** Baisse très importante des capacités opérationnelles ou des marges de sécurité. **Effets sur les passagers :** Très graves blessures incluant possiblement des morts dans les passagers et dans les membres d'équipage en cabine. **Effets sur l'équipage :** Douleur physique ou très grande charge de travail avec diminution des capacités de réalisation des tâches. (probabilité d'occurrence : <10-7 par heures de vol).
- *CAT, « Catastrophic » :* **Effets sur l'avion :** Perte de l'avion. **Effets sur les passagers** : Nombreux morts. **Effets sur l'équipage :** Morts ou blessures très graves. (probabilité d'occurrence : inferieure à 10-9 par heures de vol).

En fonction de la nature de l'erreur et des conséquences de celle-ci, des recommandations de conception ou de formation sont émises. Les recommandations de formation ne pouvant être émises seulement que pour les erreurs aux conséquences inférieures ou égales à *« MAJ »*. Ces recommandations sont injectées pour être prises en compte lors de la prochaine itération du cycle de conception.

Avant de passer au chapitre 3 qui montrera comment optimiser ce cycle de conception, nous allons dans la conclusion déceler les difficultés qui se posent au concepteur pour la production et l'évaluation des cockpits.

III Conclusion : Difficultés du cycle de conception centrée utilisateur

Nous venons de voir qu'en conception aéronautique, le processus de *conception centrée utilisateur* doit répondre à un règlement qui pose les objectifs à tenir mais propose aussi les points à étudier et les méthodes. Pour y répondre Airbus a mis en place un processus d'élaboration et de test d'hypothèses. Celui-ci demandant aux concepteurs non seulement d'analyser et prévoir comment les utilisateurs doivent utiliser une interface, mais aussi de tester la validité de leurs hypothèses par l'analyse du comportement des utilisateurs lors de tests en environnement réel. Les hypothèses testées sont basées à la fois sur les connaissances relatives aux domaines qui rentrent dans le processus de conception. Mais elles peuvent être aussi le fruit de recherches fondamentales ou de l'expertise acquise au cours des développements des précédents avions. Le test de ces hypothèses est nécessaire car il est souvent très difficile pour des concepteurs d'une interface de prévoir ce que va vivre un utilisateur qui utilisera pour la première fois son interface (adapté de : ISO 13407 ; ISO/TR 16982 ; Robert, J. M. (International Organization for Standardization, 1999) ; (International Organization for Standardization, 2002) ; (Robert, 2003)).

Dans un monde idéal, tout concepteur souhaiterait imaginer le fonctionnement du produit qu'il conçoit et que chaque utilisation réelle ne soit que la « répétition » ou l'imitation d'un fonctionnement déjà éprouvé en conception. Cependant la réalité des systèmes complexes (tels que l'aviation commerciale) ne permet que l'étude d'une faible partie de scénarii. Même si ceux ci sont pensés, discutés, tracés tout au long du processus de conception, la réalité explore d'autres horizons et des situations si diverses qu'elles rompent avec l'utopie d'une prévision simple et exhaustive.
Cependant, sans chercher le mythe de l'exhaustivité, il est nécessaire définir les scénarii de test afin qu'ils forment la meilleure représentation du domaine d'utilisation du système. Afin de garantir la meilleure représentation possible, il faut utiliser au maximum les sources d'hypothèses de conception. Cependant certaines de ces sources sont difficilement intégrables dans le processus facteur humain, notamment les analyses de sûretés des systèmes que nous avons évoquées dans l'introduction. Actuellement ces analyses sont prises partiellement en compte en fonction des demandes des experts en sécurité. Le fait que ces analyses ne contiennent que des cas de pannes est certainement à l'origine de cette prise en compte partielle. En effet, certaines de ces pannes sont des évènements extrêmement improbables et donc des conditions très marginales. Même des pannes aux conséquences relativement faibles (majeures par exemple) ont des objectifs de conception devant assurer des probabilités d'occurrences très faible (10^{-5} dans le cas d'une panne majeure, cela correspond à une occurrence de panne sur toute la carrière d'un pilote de ligne).

Cependant, pour aller vers une représentation la plus complète possible, il est nécessaire de garantir une prise en compte générale de ces analyses dans les évaluations Facteurs Humains. De plus, le niveau de sûreté du transport aérien actuel est extrêmement haut, avec un accident tous les millions de décollages. C'est pourquoi afin d'améliorer ce niveau de sécurité, il faut nous intéresser à des situations de plus en plus marginales.

Cet objectif n'est pas simple à réaliser, car il y existe plus de 100 000 scénarii de pannes pour un Airbus A380. Toutes ces pannes ne sont pas intéressantes pour dimensionner les études Facteurs Humains, il faut donc trouver le moyen de récupérer les situations présentant un intérêt pour ces évaluations.

Dans le prochain chapitre, nous allons présenter plus en détail ces analyses de sûreté, notamment les processus mis en place pour parvenir à leur rédaction. Puis nous définirons les caractéristiques des scénarii qu'il est utile d'intégrer dans le processus de conception centré sur l'utilisateur. Enfin nous présenterons les méthodes existantes pour extraire les scénarii intéressants avant de choisir la méthode la plus adaptée à la situation.

CHAPITRE. 3 OPTIMISER LA CONCEPTION CENTREE UTILISATEUR : LES ANALYSES DE SURETES DES SYSTEMES, UNE SOURCE D'HYPOTHESES A MIEUX UTILISER

Comme nous l'avons précédent expliqué, l'avion est considéré comme l'un des moyens de transport les plus sûrs. Cette réputation perdure car la sécurité est une préoccupation majeure du monde aérien :

- tout avion qui décolle d'un aéroport répond à des exigences nationales, voire internationales (suivant son périmètre d'exploitation), garantissant l'intégrité des personnes, tant à bord qu'autour de l'appareil.
- le contrôle aérien applique des procédures précises pour guider chaque avion afin qu'il effectue un vol en évitant toute collision.

Nous avons vu dans le chapitre 2, qu'Airbus, en tant qu'avionneur européen, se doit de certifier, auprès des autorités européennes, chaque avion pour que la compagnie aérienne cliente puisse en prendre possession et l'exploite sur ses lignes. Airbus est concerné par toutes les réglementations liées aux aéronefs, durant chaque phase du cycle de vie d'un avion :

- **La conception :** avant d'être produit et assemblé à grande échelle, chaque nouvel appareil doit obtenir une certification de type ;
- **La production :** chaque avion produit est réalisé spécialement pour une compagnie cliente ; des écarts par rapport aux spécifications certifiées durant la conception sont possibles mais ils doivent également répondre aux exigences des autorités européennes. Si la compagnie cliente réside en dehors d'Europe, il est également nécessaire de fournir des garanties aux autorités locales, chaque avion obtient donc une certification individuelle.
- **L'exploitation :** chaque avion est suivi durant toute sa "vie" afin qu'il reste conforme à son état de certification. Chaque incident est signalé puis analysé pour démontrer la préservation de la conformité, toute intervention (maintenance, réparation, modification) est enregistrée et contrôlée.

Durant la conception, en vue d'obtenir la certification de type, des analyses sont menées afin de chercher tout ce qui pourrait conduire l'avion dans une situation dégradée pouvant éventuellement porter atteinte à l'équipage, aux passagers ou à toute personne dans le périmètre de l'appareil : on parle d'analyses de sûreté de fonctionnement des systèmes. Comme la prise en compte du facteur humain dans le cycle de conception, les analyses de sécurité sont rédigées en réponse à une exigence de certification. Le paragraphe I de ce chapitre présentera donc les analyses de sécurité et le contexte réglementaire dont elles dépendent. Afin, puisque nous n'avons pas perdu de vue l'objectif de rechercher dans ces analyses les scénarii à prendre en compte pour les évaluations Facteurs Humains, nous présenterons dans le paragraphe II les caractéristiques des scénarii à rechercher. Nous verrons qu'il n'est pas facile de les repérer malgré avoir défini leurs caractéristiques. Puis pour terminer, nous verrons quelles sont les méthodes à notre disposition pour récupérer de façon automatique ces scénarii dans les analyses.

I Les analyses de sûreté de fonctionnement

I.1 La sûreté de fonctionnement des systèmes dans le monde de l'aéronautique

La sûreté de fonctionnement des systèmes constitue une partie de la sécurité d'un avion. Les acteurs de la sûreté de fonctionnement des systèmes aériens peuvent être classés en trois catégories :

- Les autorités de certification, qui définissent des règles et s'assurent de leur application ;
- Les avionneurs, les équipementiers et les motoristes, qui doivent appliquer ces règles et démontrer qu'elles sont correctement appliquées ;
- Des groupes de travail, constitués à la fois de représentants des avionneurs et des autorités, qui s'entendent sur des pratiques permettant de satisfaire les règles et de répondre aux exigences.

Chaque catégorie d'acteurs élabore ses propres documents décrivant les pratiques permettant de satisfaire les règles.

D'une manière similaire au règlement pour les Facteurs Humains, l'AESA élabore un règlement pour la sûreté de fonctionnement. Comme ce que nous avons présenté dans le chapitre 2, ce règlement s'appelle le *CS25-1309* et possède une *AMC* correspondante (*AMC25-1309*). Dans ce cadre afin d'améliorer constamment le niveau de sécurité des passagers, les principaux acteurs du monde aéronautique (avionneurs, motoristes, équipementiers) collaborent au sein de deux principaux groupes de travail :

- **le groupe WG-63** (*Complex Aircraft Systems*) de l'European *Organisation of Civil Aviation Equipment* (EUROCAE)
- **le groupe S-18** (*Aircraft and System Development and Safety Assessment Committee*) de la *Society of Automotive Engineers* (SAE)

Ces groupes élaborent des recommandations, applicables aux processus, méthodes et outils, pour répondre aux exigences des autorités précédemment citées. La présence de représentants des autorités garantit la bonne interprétation des réglementations.

Le groupe S-18 travaille principalement à l'élaboration de deux documents dits Aerospace Recommended Practice (ARP) :

- *l'ARP 4754* (*Certification considerations for highly-integrated or complex aircraft systems*) (S-18, Aircraft And System Developement And Safety Assessment Committee, 2010) fournit des recommandations pour démontrer que des systèmes complexes (i.e. dont le niveau de sécurité ne peut pas être démontré par test ou dont la compréhension du fonctionnement nécessite le support d'outils) ou hautement intégrés (i.e. qui réalise ou contribue à plusieurs fonctions) satisfont les exigences de navigabilité ;
- *l'ARP 4761* (*Guidelines and methods for conducting the safety assessment process on civil airborne systems and equipment*) (S-18, Aircraft and System Development and Safety Assessment Committee, 1996) propose des méthodes possibles pour évaluer la sûreté de fonctionnement des systèmes d'un avion en vue de sa certification (les différentes étapes sont présentées dans la suite de ce chapitre).

Le groupe WG-63 élabore des documents équivalents baptisés respectivement ED-79(WG-63, Complex Aircraft Systems, 2007) et ED-135 (WG-63, Complex Aircraft Systems, 2007). Des réunions communes entre les deux groupes permettent une évolution cohérente des documents et une meilleure réflexion.

Conformément aux ARP 4754 et 4761, les avionneurs réalisent différentes analyses pour démontrer la tenue des exigences imposées par les réglementations internationales en matière de navigabilité aérienne. Le processus global se décompose selon les grandes étapes suivantes :

- Identifier les fonctions souhaitées du système à étudier.
- Rechercher les conditions défaillantes liées à chaque fonction pour déterminer les possibles configurations de panne, en anglais *Failure Condition* (*FC*).

- Définir une criticité pour chaque *FC* en fonction de ses effets sur l'avion et sur ces occupants (en se référant au Tableau 1 extrait du CS25-1309).
- Déduire des exigences (principes de conception et probabilité d'occurrence maximum tolérée) pour prévenir chaque FC ou maintenir leurs conséquences à un niveau réglementaire.
- Démontrer que l'étude est complète et que les résultats sont en accord avec les réglementations en vigueur : la probabilité d'occurrence de chaque FC est inférieure ou égale à l'objectif déduit de la criticité.

Une fonction est définie par la réglementation comme un service (abstrait) que doit rendre un objet physique ou un ensemble d'objets physiques (concrets). Les programmes aéronautiques considèrent la décomposition physique suivante :

- un avion est constitué d'une structure, appelée fuselage, de moteurs et de systèmes ;
- un système est composé d'équipements.

Plusieurs systèmes peuvent fonctionner conjointement afin de remplir une fonction, ces systèmes sont alors dits interfacés ou en interface. Par exemple, le système CDS (*Control and Display System*) constitué par la majeure partie des interfaces avec le pilote, est en relation avec les calculateurs de vol car c'est lui qui affichera les informations de ces systèmes, mais aussi avec le système de génération électrique car lui fourni l'énergie nécessaire à son fonctionnement.

Cette décomposition physique est transposée en décomposition fonctionnelle hiérarchique : les fonctions de niveau supérieur sont reliées à un ensemble de fonctions de niveaux inférieurs. Les relations entre niveaux signifient que les fonctions de niveau inférieur contribuent à la réalisation de la fonction de niveau supérieur. Cette décomposition est utile pour, guider les analyses de sûreté de fonctionnement des systèmes et allouer des exigences aux concepteurs des fonctions. Les trois principaux niveaux de décomposition utilisés dans les programmes aéronautiques sont :

- le niveau "avion", le plus haut dans la hiérarchie, qui regroupe les fonctionnalités principales de l'avion comme "commander l'avion", "naviguer" ;
- le niveau "système", qui regroupe des fonctions au sein de systèmes tels que le système des commandes de vol ;
- le niveau "équipement", dernier niveau dans la hiérarchie, qui regroupe des fonctions réalisées par un même équipement.

Les analyses à conduire concernent initialement l'avion dans son ensemble, puis les systèmes embarqués (considérés individuellement, liés fonctionnellement ou liés par leur placement à bord) et

enfin les équipements qui composent un système. En nous référant aux *ARP*, nous avons résumé ci-après le processus actuel que nous décomposons en trois catégories d'analyses :

- **Evaluation des risques fonctionnels**, en anglais *Functional Hazard Assessment* (*FHA*) : étude des fonctions pour déterminer les défaillances fonctionnelles possibles et classifier les risques associés à des configurations de panne spécifiques.
- **Evaluation de la sécurité des système**s, en anglais *System Safety Assessment* (*SSA*) : évaluation de la conformité de l'architecture d'un système avec les exigences de sécurité déduites des *FHA*. La première étape dans le processus de création des évaluations de la sécurité des systèmes, amène à la rédaction d'une évaluation préliminaire, en anglais *Preliminary System Safety Assessment* (*PSSA*)
- **Analyses des causes communes**, en anglais *Common Cause Analysis* (*CCA*) : études complémentaires des risques liés à des causes communes à plusieurs systèmes.

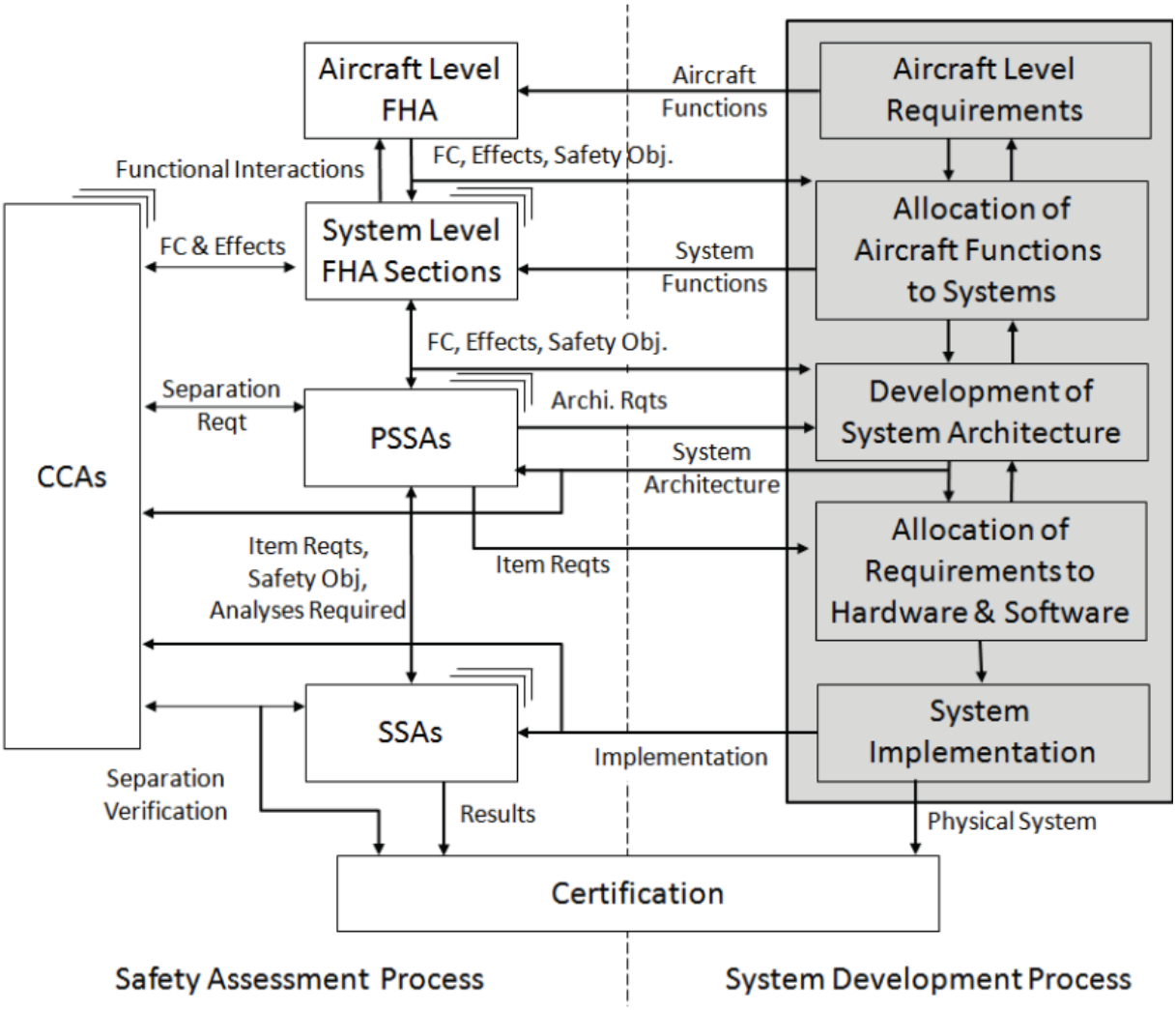


Figure 17 : Processus d'évaluation de la sécurité (extrait de l'ARP 4754)

I.2 Les FHA : les scénarii de pannes

Chaque nouvel avion est conçu pour une utilisation spécifique : transport de passagers, acheminement de marchandises, utilisation militaire. Chaque utilisation ajoute des fonctions spécifiques aux fonctions communes à tout avion (décoller, se mouvoir dans le ciel...). Or toute fonction peut être perdue ou remplie de façon incorrecte, partiellement, de façon intempestive, en retard... L'examen systématique des fonctions permet de recenser toutes les *Failure Conditions* (*FC*) possibles. Généralement quatre grands modes de défaillances sont donnés :

- La perte totale de la fonction
- La perte partielle
- Un déclenchement non commandé (valable généralement pour des fonctions à deux états, *« allumé/éteint »* par exemple)
- Un comportement erroné (valable généralement pour des fonctions qui ont une infinité d'état, l'affichage de la vitesse de l'avion par exemple)

En croisant chaque fonction et chaque défaillance fonctionnelle, une recherche systématique des scénarii opérationnels de panne est réalisée. Ensuite, l'étude des répercussions de chaque scénario (sur l'avion, son équipage, ses occupants et les personnes évoluant autour de l'avion) permet d'établir sa criticité, en se référant au Tableau 1 présenté dans l'introduction (Le tableau présente les niveaux de criticité : *No safety effect (NSE), Minor (MIN), Major (MAJ), Hazardous (HAZ), Catastrophic (CAT)*). Les scénarii sont ensuite regroupés par effets équivalents en *Failure Conditions*. Cela permet d'associer à chaque FC :

- un objectif quantitatif : une probabilité par heure de vol à atteindre ;
- un niveau de confiance du développement, en anglais Design Assurance Level (DAL) : chaque catégorie de DAL impose des règles de conception tant sur les architectures physiques que sur les techniques d'implémentation logicielle.

La liste des fonctions, des FC et leur criticité constituent l'évaluation des risques fonctionnels (FHA). Outre les exigences quantitatives, des exigences qualitatives sont associées à chaque FC en fonction de sa criticité. Ces exigences sont de la forme « *la FC de criticité C ne peut se produire en moins de N défaillances* ». Le tableau présenté ci-dessous (Tableau 4), présente la relation entre le nombre de défaillances nécessaires et la criticité.

Classification	Catastrophic	Hazardous	Major	Minor	No Safety Effect
Qualitative requirement	3	2	2	1	0

Tableau 4 : Criticité et exigences qualitatives associées

Durant la conception d'un avion, une première FHA de niveau avion considère l'avion dans sa globalité puis différentes FHA sont réalisées pour traiter chaque système embarqué individuellement.

I.2.1 FHA de niveau avion

A partir de la liste des fonctions attendues pour un avion et des choix de conception initiaux une évaluation qualitative de haut niveau recense les risques potentiels liés aux fonctions basiques de l'avion. Une fonction est réalisée par un système ou par une combinaison de systèmes. La FHA niveau avion permet d'identifier des FC impliquant plusieurs systèmes à la fois. Les exigences de sécurité, qualitatives (règles de ségrégation, DAL...) et quantitatives (objectif de probabilité) déduites de la FHA niveau avion, sont allouées aux systèmes contribuant aux différentes fonctions et doivent être prises en compte durant leur conception.

I.2.2 FHA de niveau système

La FHA système est également une analyse qualitative et itérative. Elle s'appuie sur les résultats de la FHA niveau avion et traite les fonctions du système concerné en étudiant les conséquences d'une défaillance ou d'une combinaison de défaillances tant sur le système que sur l'avion. Lorsque les fonctions avion ont été affectées à des systèmes par le processus de conception, chaque système fait l'objet d'une FHA système. Nous donnons un exemple d'un scénario de panne issu d'une FHA niveau système sur la Figure 18.

S-01.01.01.A	Loss of attitude (roll / pitch) information display	
01	Permanent loss of attitude, one side Repercussion : Attitude information display is lost on one side in spite of reconfiguration (automatic, manual or reset of DU) . It remains back up information (ISIS) and other side information to continue the flight. Effect are similar to a total loss of PFD image on one side. Next flight is cancelled. Repercussions are considered : - MAJOR in T/O (refer to S/R parameter, MAJOR classification in T/O due to possible RTO between 100kt and V1 if the failure occurs on flying pilot's side) - MINOR in other flight phases (crew task reversion possible). Crew detection : Blank zone or flagged parameters.	MAJ

Figure 18 : Exemple d'un scénario extrait de la FHA du CDS

I.3 L'utilisation des FHA dans la conception d'un avion

Les concepteurs de systèmes doivent élaborer une architecture pour réaliser les fonctions mentionnées précédemment. Compte tenu des FHA déjà réalisées, il est nécessaire d'évaluer chaque architecture pour s'assurer qu'elle répond aux exigences définies au préalable. Ces évaluations sont menées dans la continuité des FHA en suivant un processus itératif : chaque évaluation peut détecter une correction à effectuer qui nécessite une nouvelle évaluation. Bien que l'évaluation d'une architecture s'intéresse principalement à un système, la forte interconnexion des systèmes dans un avion implique de considérer les systèmes en interface.

Dans un premier temps, des évaluations dites "préliminaires" (*PSSA*), sont réalisées. Une PSSA permet, tout d'abord, de compléter la liste des FC et donc éventuellement d'ajouter de nouvelles exigences de sécurité. D'autre part, une PSSA permet de déterminer la liste préliminaire des actions pilotes à accomplir et/ou des opérations de maintenance à effectuer lorsqu'une FC se produit. La PSSA vérifie également que l'architecture du système est cohérente avec les modes de défaillance envisagés. Pour cela, une recherche des combinaisons minimales de défaillances est effectuée : c'est ce que l'on appelle une coupe minimale. La probabilité d'occurrence d'une FC est calculée à partir de ces coupes minimales, de la probabilité individuelle de chaque défaillance ainsi que de paramètres complémentaires (temps de vol moyen, intervalle de temps entre 2 opérations de maintenance...). Enfin, cette évaluation permet d'allouer au niveau équipement les exigences de sécurité du niveau système. Ces exigences sont soumises aux fournisseurs d'équipements.

I.4 Les hypothèses de comportement de l'utilisateur dans les scénarii des FHA

Lors de la rédaction des analyses de sécurité pour les systèmes en interaction avec l'équipage, il est nécessaire d'imaginer quel comportement la défaillance du système va entraîner chez les pilotes. Dans l'exemple donné par la Figure 18, par conception nous faisons l'hypothèse que la perte des informations d'attitude de l'avions (roulis ou tangage) sur l'un des instruments primaires de pilotage de l'avion (*Primary Flight Display : PFD*), entraînera l'utilisation par l'équipage des informations restantes sur le second PFD. Ces hypothèses sont écrites par des groupes d'experts en sûreté. Ils sont constitués d'experts en sûreté des systèmes, d'experts en sûreté au niveau opérationnel, de pilotes d'essais et ponctuellement d'experts en Facteurs Humains. Compte tenu du nombre de scénarii présents dans les FHA systèmes pour un avion entier (environ 100 000) l'étude systématique de chaque hypothèse de comportement est extrêmement difficile. De plus, ces hypothèses sont contenues dans de petits paragraphes de textes et donc non accessibles directement sauf à lire le scénario. De

plus, toutes les hypothèses ne sont pas pertinentes à étudier, seule les plus structurantes devraient être extraites pour une étude en facteur humain.
Dans le prochain paragraphe, nous allons nous intéresser aux caractéristiques des scénarii à rechercher, nous verrons que ces caractéristiques ne suffisent pas forcement à rechercher facilement les scénarii dans les FHA.

II Les caractéristiques des scénarii à rechercher dans les analyses de sûreté

Afin de compléter les travaux réalisés par les experts Facteurs Humains, nous souhaitons permettre l'évaluation des hypothèses de comportement de l'équipage prises après la défaillance d'un système que nous trouvons dans les scénarii des FHA. Dans ce paragraphe, nous allons définir les caractéristiques des scénarii à sélectionner. Les points que nous allons aborder ici sont le fruit du travail de définition du besoin qui a eu lieu en début de thèse. Il y a eu plusieurs rencontres entre les experts en Facteurs Humains et les experts en sûreté de fonctionnement des systèmes qui ont permis de préciser ce besoin.

II.1 Les caractéristiques des scénarii : Ce que cherchent les experts

Les caractéristiques des scénarii à rechercher ont été construites par une démarche itérative d'affinement du besoin. En partant d'un but, idéal absolu, qui sera affiné au cours du temps pour répondre aux contraintes de la conception.

L'objectif principal des experts est : « d'acquérir la plus grande confiance possible dans les hypothèses de comportement de l'équipage prises dans les analyses de sûreté de fonctionnement des systèmes ». Pour ce faire il faudrait mettre en place un processus Facteurs Humains de test de ces hypothèses (par évaluation experte, simulation…). Cependant le nombre scénarii et donc d'hypothèses potentielles est bien trop grand pour permettre leur test exhaustif.

Le premier pas de l'itération, s'est donc centré sur la sûreté des systèmes. Ce qui intéressait les experts ce sont les hypothèses qui pourraient, si elles sont fausses, dégrader le niveau de sûreté. Pour cela, ils sont partis de la classification des analyses de sûreté. « Les hypothèses à rechercher sont celles pour lesquelles l'équipage joue un rôle dans le maintien de la classification. » Une erreur de l'équipage, c'est-à-dire lorsque celui-ci ne réagit pas comme prévu, entraînerait alors la dégradation de la classification du scénario. Il y aurait donc des « sauts de classifications » lors de la survenue d'erreurs

de l'équipage. (*NB :* Nous parlons ici d'erreur bien que cela soit, à mon avis, un abus de langage. Il s'agit plutôt d'une mauvaise hypothèse de conception que d'une erreur à proprement parler). La Figure 19 illustre ces sauts de classifications. Nous pouvons déjà remarquer que les scénarii aux conséquences les plus graves (*CAT*) ne sont pas à prendre en compte, puisqu'ils sont déjà au niveau le plus haut de la classification, la survenue d'une erreur ne peut donc pas aggraver la situation.

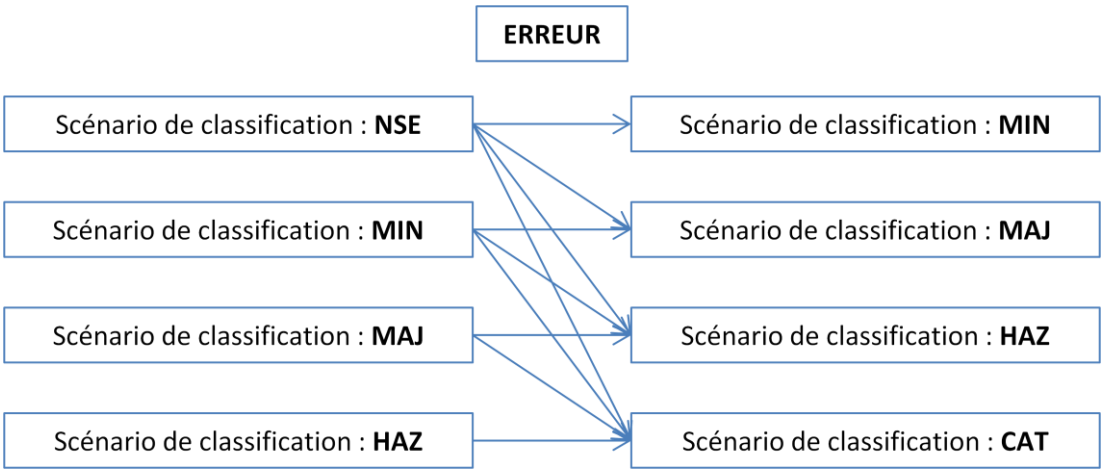


Figure 19 : Illustration des sauts de classifications

Bien que cette itération permette de réduire le scope des scénarii à rechercher, beaucoup trop de scénarii restent encore. Ce sont maintenant des expertises au niveau Facteurs Humains qui rentrent en jeu. En effet, beaucoup d'hypothèses sont justes et il est difficilement envisageable qu'un pilote qui perd son écran primaire de vol par exemple, n'utilise pas les moyens de reconfiguration du cockpit pour afficher ces informations sur un autre écran, ou plus simplement qu'il se reconfigure avec le copilote qui lui a son écran qui fonctionne. Les scénarii à rechercher sont donc :

« Ceux pour lesquels une *erreur* peut être raisonnablement envisageable et qui, si elle a lieu, entraînerait une dégradation de la classification ».

Cette dernière boucle de précision des besoins, cible au mieux les scénarii à rechercher. Cependant cette heuristique, ne permet pas une sélection facile des scénarii. De plus, les experts demandent une extraction rapide de ces scénarii. Cette extraction doit se faire durant le cycle d'écriture des scénarii afin de permettre la meilleure anticipation des modifications de design. Le processus d'extraction doit donc correspondre aux processus de design actuels.

II.2 Les difficultés de sélection des scénarii

Bien que le besoin des experts soit très bien défini, les scénarii correspondant aux caractéristiques demandées sont toujours difficilement extractibles. Nous avons vu dans l'introduction, que le nombre scénarii de départ est très grand (environ 100 000). Si nous nous limitons aux systèmes en interface avec l'équipage (environ 30) cela nous laisse encore environ 30 000 scénarii à étudier.
De plus, la nature de l'information présente dans les scénarii est aussi un problème. Comme nous l'avons vu dans l'introduction chaque scénario présente : 1/ une description de la panne fonctionnelle et l'évaluation de ces conséquences rédigée dans un style permettant une intégration assez aisée dans une base de données. 2/ une description des conséquences de la panne sur l'avion et des comportements attendus de l'équipage. Cette dernière partie est rédigée en langage naturel (Anglais). Il est donc difficile d'imaginer des requêtes simples permettant de récupérer les hypothèses de comportement équipage.
Enfin, les caractéristiques à rechercher dans les scénarii demandent une analyse de chaque scénario afin de savoir, si la classification va être modifiée et si l'erreur est raisonnablement envisageable.

En conclusion, il faut donc trouver un moyen permettant, 1/ de traiter rapidement de grande quantité de données, 2/ afin de garantir l'intégration dans le processus de conception existant, 3/ avec des données non ou partiellement formalisées, 4/ pouvant supporter des critères experts.

Nous allons, dans le paragraphe suivant, présenter des méthodes existantes afin de trouver ou d'élaborer une méthode répondant à ces critères.

III Les méthodes de recherche des scénarii

Le besoin industriel, exposé dans le chapitre précédent, est en rapport avec les théories de prévision de l'erreur humaine qui forment le cadre théorique principal de ce travail d'étude. Nous avons vu par exemple, dans le chapitre 1, que l'application des modèles décrivant les types d'erreurs (Hollnagel, 1991), (Norman, et al., 1986), (Rasmussen, 1983), (Reason, 1990) peut entrer dans le cycle de conception et participer à la prévision et l'évaluation des risques liés à l'erreur humaine. Mais, en nous rapprochant du besoin industriel qui consiste à évaluer, dans une situation décrite, le risque lié à une erreur humaine, nous nous rapprochons du domaine des méthodes fiabilistes.

Les méthodes fiabilistes, dont les plus utilisées seront exposées dans le paragraphe III.1**,** permettent, à partir d'une situation, de définir les comportements possibles de l'utilisateur, d'évaluer ceux considérés comme non souhaités afin de faire en sorte que ceux-ci n'apparaissent plus ou qu'ils

n'engendrent pas des situations aux conséquences pouvant être graves. La conclusion du paragraphe III.1 donnera les avantages et les inconvénients de ces méthodes.

Dans le paragraphe III.2, nous verrons comment la sélection des scénarii pourrait se faire par jugement d'expert. Cette évaluation permettra, tout comme les méthodes fiabilistes, de définir au mieux le comportement de l'utilisateur et de détecter les cas où ce comportement ne serait pas en accord avec une utilisation sûre du système conçu. La conclusion du paragraphe III.2 donnera les avantages et les inconvénients d'une méthode basée sur le jugement d'expert.

Enfin, dans le paragraphe III.3, nous présenterons une méthode optimisée pour répondre aux besoins des experts et adaptée au contexte industriel.

III.1 Les méthodes fiabilistes

Les méthodes fiabilistes reposent sur l'identification des erreurs possibles et sont de plus en plus proches du terrain pour mieux cerner les erreurs à redouter et mieux rendre compte de leurs conséquences. Plus généralement, les techniques d'évaluation de la fiabilité humaine (*HRA : Human Relability Assesment*), centrées classiquement sur la quantification du risque d'erreur humaine, évoluent vers des techniques de détection de l'erreur (*HEI : Human Error Identification*). Il s'agit d'échanger une vision tournée vers l'analyse causale de l'erreur (l'arbre des causes en est un exemple achevé) pour une vision tournée vers la détection des erreurs et la mesure du risque d'accidentalité associé à chaque erreur. Ces méthodes voudraient incorporer à terme, une évaluation des chances de récupération de l'erreur par l'opérateur ou pas le système ou par toute combinaison des deux.

Dans le but d'extraire des scénarii de la FHA, nous nous sommes penchés sur l'application de ces méthodes. Puisque chaque scénario décrit un contexte opérationnel, il est théoriquement possible d'appliquer ces modèles prévisionnels de l'erreur humaine. Dans les prochains paragraphes, nous allons présenter rapidement les principaux modèles fiabilistes :

- *SHERPA : Systematic Human Error Reduction & Prediction Approach* (Stanton, et al., 2000)
- *HEART : Human Error Assessment & Reduction Technique* (Williams, 1985)
- *THERP : Technique for Human Error Rate Prediction* (Swain, et al., 1983)
- *TAFEI : Task Analysis for Error Identification* (Baber, et al., 2002)

III.1.1 SHERPA : Systematic Human Error Reduction & Prediction

La méthode SHERPA (Stanton, et al., 2000) (Stanton, 2003) commence par une analyse hiérarchique de la tâche (en anglais *HTA Hierarchical Task Analysis*). A l'origine, la HTA avait été conçue pour identifier les besoins de formation (Annett, et al., 1967). La HTA demande une décomposition en tâches et sous tâches… Le point de départ étant donné par le but à atteindre en utilisation. SHERPA classifie ensuite chaque tâche selon cinq catégories : *Checking*, *Selection*, *Action*, *Communication*, *Information retrieval*. Nous appliquons alors, une taxonomie des types d'erreurs (celle d'Hollnagel par exemple présentée dans le chapitre 1) sur chaque tâche. Pour chaque type d'erreur, nous faisons une évaluation experte de la probabilité d'occurrence et de ces conséquences. Cette décomposition permet d'être exhaustif sur les erreurs humaines possibles et donc de mieux les prévoir.

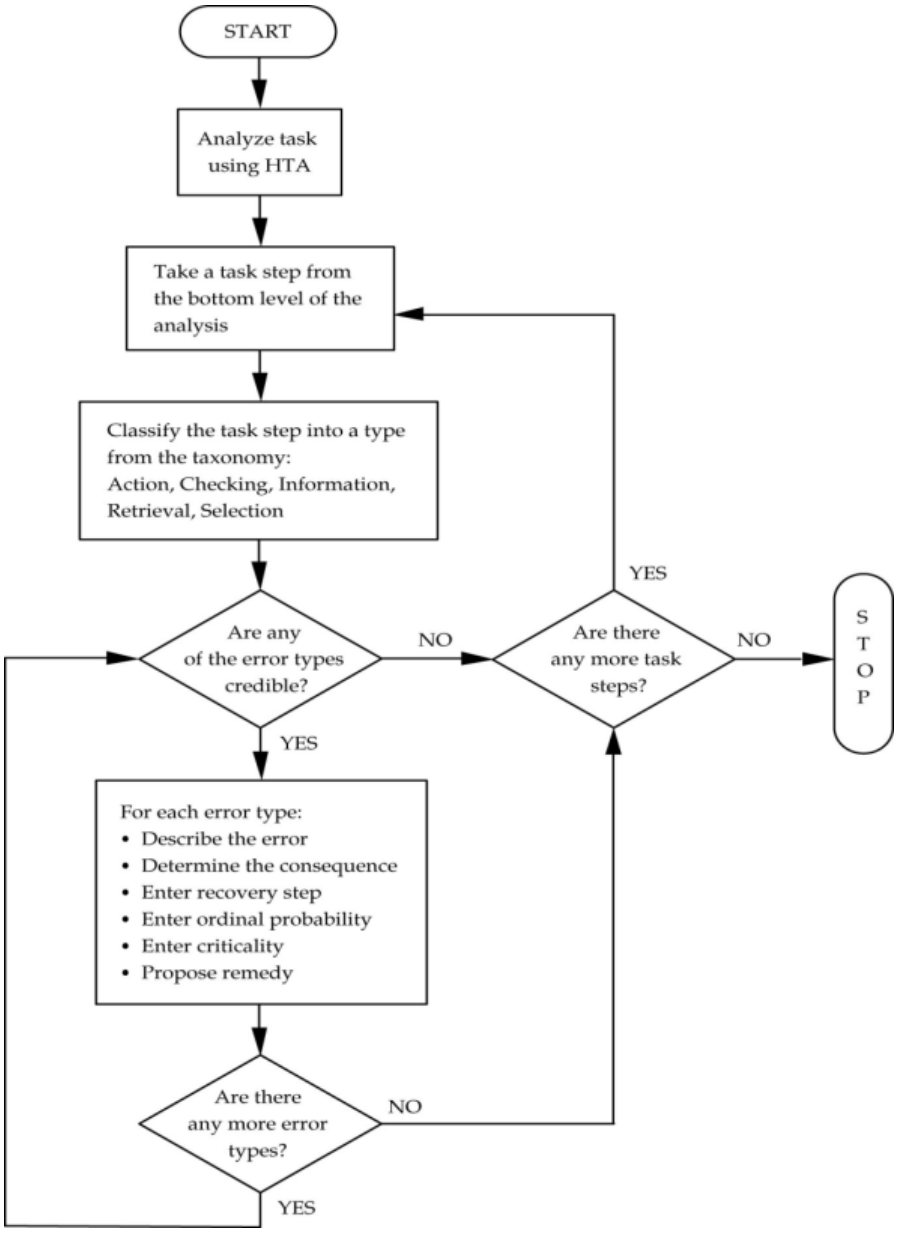


Figure 20 : The Systematic Human Error Reduction and Prediction Approach (SHERPA)

III.1.2 HEART : Human Error Assessment & Reduction Technique

Tout comme SHERPA, HEART (Williams, 1985) commence par une HTA. Par jugement d'expert, chaque tâche est ensuite classifiée selon les descriptions du tableau ci-dessous (Tableau 5). Une probabilité d'erreur humaine empirique (« r ») est associée à chaque classe. En complément, il faut donner les conditions favorisant l'erreur humaine avec des facteurs empiriques associés (EPC_i) (voir Tableau 6). Ensuite, il y a un facteur « p » (p compris entre 0 et 1) donné par jugement de l'analyste selon qu'il pense que l'erreur et probable ou non. Selon le scénario de tâche à analyser, nous appliquons alors les probabilités d'erreur humaine revues par les facteurs des conditions favorisant l'erreur. Nous obtenons une probabilité d'erreur humaine *(HEP : Human Error Probability)* pour le scénario étudié à partir d'une formule empire :

$$HEP = r \times \prod_i \{p_i \times (EPC_i - 1) + 1\}$$

Classe de la méthode	Description des classes	Probabilité d'erreur humaine (r)
A	*Totally familiar, performed at speed with no idea of likely consequences*	0,55
B	*Shift or restore system to new or original state on a single attempt without supervision or procedures*	0,26
C	*Complex task requiring high level of comprehension and skill*	0,16
D	*Fairly routine task performed rapidly or given scant attention*	0,09
E	*Routine highly-practised, rapid task involving relatively low level of skill*	0,02
F	*Restore or shift a system to original or new state following procedures with some checking*	0,003
G	*Completely familiar, well designed, highly practised routine task occurring several times per hour*	0,0004
H	*Respond correctly to system command even when there is an augmented or automated supervisory system*	0,00002
M	*None of the above*	--

Tableau 5 : Table des classes de la méthode HEART

Identifiant de la condition	Conditions favorisants l'erreur humaine	Facteurs associés (EPC_i)
1	*Unfamiliarity*	17
2	*Shortage of Time*	11
3	*Low signal to noise ratio*	10
4	*Ease of information suppression*	9
5	*Ease of information assimilation*	8
6	*Model mismatch (operator / designer)*	8
7	*Reversing unintended actions*	8
8	*Channel capacity overload*	6
9	*Technique unlearning*	6
10	*Transfer of knowledge*	5,5
11	*Performance standard ambiguity*	5
12	*Mismatch between perceived / real risk*	4

Tableau 6 : Table des conditions favorisant l'erreur humaine

III.1.3 THERP : Technique for Human Error Rate Prediction

THERP (Technique for Human Error Rate Prediction) est une méthode développée par Swain (Swain, et al., 1983) pour l'autorité de sûreté nucléaire américaine. Appartenant au domaine des HRA *(Human reliability Assessment)*, elle a comme objectif d'évaluer la probabilité d'occurrence d'une erreur humaine lors de la réalisation d'une tâche spécifiée. La méthode ce déroule en 5 étapes :

Etape 1 : Définir les tâches du système Homme/machine possédant un intérêt d'étude.

A partir de la situation à étudier, il faut définir les tâches d'intérêt. Une erreur dans la réalisation de la tâche entraînant des conséquences sur la sécurité. En effet, il n'est pas intéressant d'investiguer des erreurs pour lesquels des moyens de mitigation ont déjà été donnés ou pour lesquels il n'y a pas de conséquences sur la sécurité opérationnelle.

Etape 2 : Analyser les tâches sélectionnées et donner les possibilités d'erreurs.

Les tâches sélectionnées lors de l'étape 1 sont analysées grâce à une analyse hiérarchique de la tâche (*HTA*) représentée sous forme d'arbres d'évènements. Ces arbres d'évènements sont structurés avec des opérateurs logiques qui lient entre elles les tâches extraites de l'activité humaine étudiée. En plus de cette HTA, une décomposition fonctionnelle du système peut être réalisée et couplée à la HTA. L'arbre pourra donc avoir des « branches » liées au fonctionnement du système technique et des « branches » liées au fonctionnement humain. Pour chaque étape humaine de l'arbre, les erreurs

possibles doivent être considérées par l'analyste et définies précisément. La taxonomie pour décrire les erreurs est la suivante :

- Les erreurs d'omission, ce sont des oublis d'une partie de la tâche ou de la tâche en entier.
- Les erreurs de commission, ces erreurs peuvent être des types suivant :
    - Erreurs de sélection – Erreurs dans l'utilisation des moyens de commande.
    - Erreurs de séquence – La séquence d'action pour réaliser la tâche est faite dans un ordre incorrect.
    - Erreurs de timing – La tâche est réalisée avant ou après le moment où elle est nécessaire.
    - Erreurs de quantité – La tâche est réalisée un nombre de fois inadéquat (répétition par exemple)

L'arbre final prendra donc en considération la réalisation normale de la tâche mais il possèdera aussi des branches reflétant les conséquences des erreurs. Les dernières branches de l'arbre correspondront à des états du système Homme/machine.

Etape 3 : Estimer les probabilités d'erreur

Pour chaque branche de l'arbre, il faut donner une estimation de la probabilité d'erreur du système. En fonction de comment est construit l'arbre, ce peut être des erreurs humaines, des erreurs du système ou les deux. (*NB :* De ce point de vue, l'arbre des défaillances donné dans la FHA est la partie correspondant au système dans l'arbre à réaliser pour la méthode THERP). L'estimation de la probabilité d'erreur humaine peut provenir d'un large panel de sources, comme : des données de simulation, de la base de données THERP, du jugement d'expert…

Etape 4 : Estimer les effets des erreurs humaines

Grâce à la réalisation d'un arbre contenant à la fois les possibles défaillances humaines et les défaillances du système technique, il est possible de mesurer la part sur la sécurité des erreurs humaines. Pour les erreurs humaines, une échelle similaire aux pannes systèmes est réalisée et les conséquences des erreurs humaines sont codifiées. (NB : Avec la méthode THERP, le système est pris comme ayant des composants à la fois humains et techniques. Ces deux types de composants sont traités de la même façon).

Etape 5 : Proposer des modifications du système et recalculer les probabilités d'erreurs du système complet.

Une fois que la contribution de l'erreur humaine aux défaillances du système est connue, il est possible de déterminer comment gérer son risque (suppression de la possibilité d'erreur, mise en place de moyens de mitigation…).

III.1.4 TAFEI : Task Analysis for Error Identification

TAFEI (*Task Analysis for Error Identification*) est une méthode mise au point par Barber et Stanton (Baber, et al., 2002), l'idée de cette méthode est qu'il est possible, si l'on a une vision des tâches qu'un utilisateur doit réaliser avec une interface, de savoir où sont les possibles sources d'erreurs. Pour réaliser l'analyse de tâches les auteurs expliquent que cela peut être fait à partir de n'importe quel modèle de tâche. La Figure 21 ci-dessous montre quelles sont les étapes de la méthode.

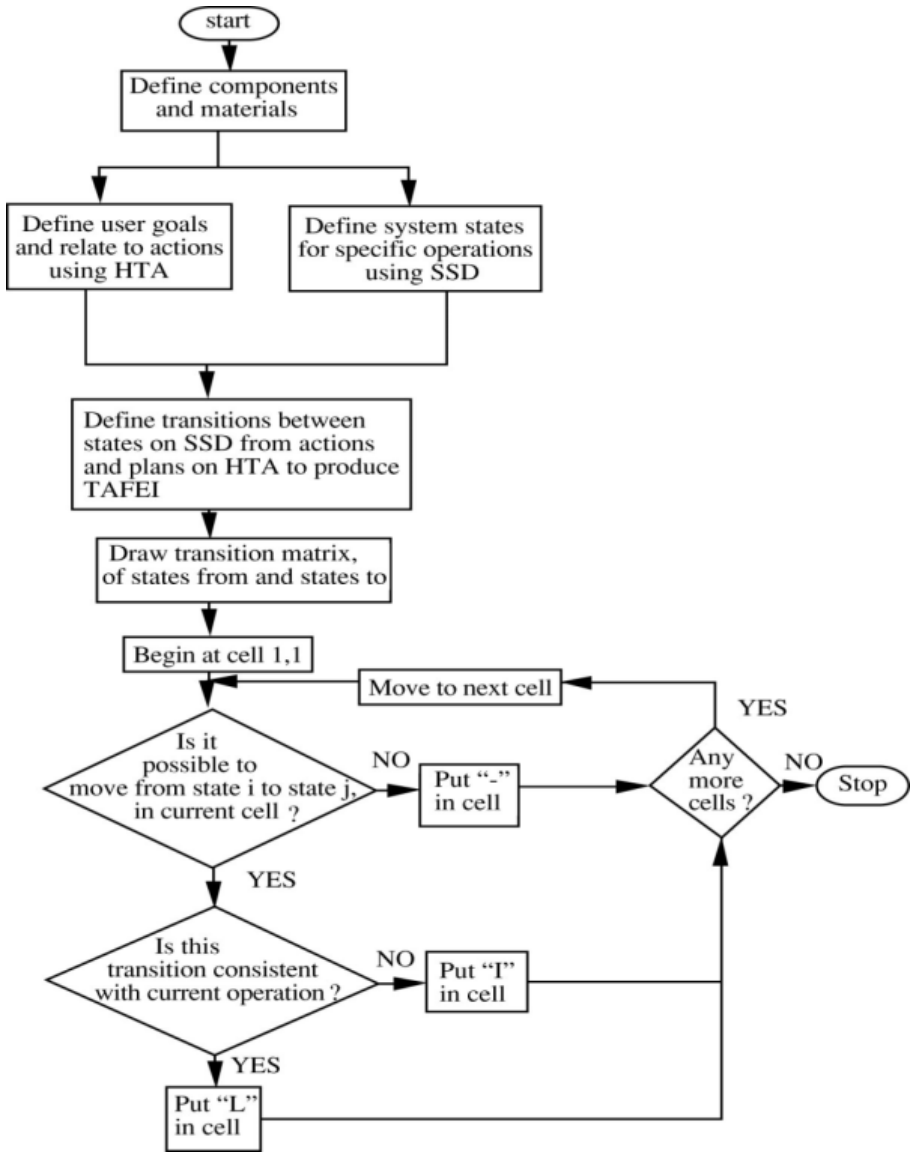


Figure 21 : Processus d'application de la méthode TAFEI

La **première étape** de la méthode TAFEI consiste dans la construction de deux arbres, le premier correspondant à l'analyse de la tâche par l'opérateur grâce à une HTA, le second aux différents enchainements d'états du système c'est-à-dire un diagramme d'états du système *(SSD : State Space Diagrams).*

La **deuxième étape** de cette méthode consiste dans la création de liens entre l'analyse des tâches de l'opérateur et le diagramme d'état du système. Ces liens se trouvent, par exemple, lorsqu'une action de l'opérateur entraîne un changement d'états du système. Cette étape permet de donner tous les états possibles du système.

La **troisième et dernière étape** compare tous les états possibles du système Homme/machine. A l'aide des liens présents dans le système (créés lors du rassemblement entre la HTA et le SSD), il est possible de savoir si l'on peut passer et comment passer d'un état à l'autre. La façon de passer d'un état à l'autre sera noté « -- » s'il n'est pas possible de passer d'un état à un autre, « L » pour « *Legal* » lorsque le passage d'un état à un autre correspond à un fonctionnement normal du système. Le changement d'un état à un autre sera noté « I » pour « *Illegal* » lorsque cela correspond à un mauvais fonctionnement du système. Ce passage « illégal » d'un état à l'autre peut être le fait d'une panne technique ou d'une erreur humaine. En comprenant les chaines pouvant amener d'un état à l'autre, il est possible d'identifier des erreurs humaines possibles.

III.1.5 Conclusion sur les modèles fiabilistes

L'application de modèles fiabilistes présente l'avantage d'avoir une approche systématique de l'erreur humaine, approche identique à celle mise en place dans les industries à risques pour les pannes des systèmes techniques. Il pourrait être envisageable de les utiliser pour formaliser les scénarii des FHA, notamment en intégrant les classes du modèles HEART, les catégories de la méthode SHERPA et en utilisant des HTA en complément des formalisations des systèmes comme le font les méthodes THERP ou TAFEI.

Cependant, outre des interrogations théoriques notamment sur l'utilisation de probabilités sur des erreurs humaines (HEART et THERP), l'utilisation des modèles fiabilistes est basée d'une manière ou d'une autre sur une HTA. Cette étape peut être rapide à réaliser pour des systèmes simples. Mais pour des systèmes extrêmement complexes pour lesquels la combinatoire de l'interactivité est presque sans limite cette tâche est difficilement réalisable. Les HTA sont utilisées chez Airbus en phase de conception soit 1/ à haut niveau afin de guider la conception des nouvelles philosophies d'interaction de l'équipage mais aussi 2/ au niveau d'un équipement particulier afin de guider les règles de

conception de celui-ci. Sans renier l'apport de ces modèles à la prévision des erreurs, nous pensons qu'ils ne sont pas adaptés au traitement de plusieurs milliers de scénarii surtout pour une industrie dont les cycles de conception sont réduits (5 ans pour développer un A380). Il est donc nécessaire de trouver une méthode dont le temps d'application soit en accord avec les contraintes de la conception.

III.2 Une méthode basée sur le jugement d'expert

Nous avons vu dans le paragraphe I de ce chapitre, que la rédaction des scénarii des analyses de sûreté est faite par des experts en sûreté des systèmes étudiés. Ce sont ces experts qui ont aussi définis les caractéristiques des scénarii à rechercher. Ce sont donc certainement eux les plus à même de sélectionner les scénarii qu'ils souhaitent étudier.
Dans ce paragraphe, nous allons tout d'abord faire une rapide présentation de l'expertise. Cette description de l'expertise permettra de soutenir notamment les idées qui seront développées dans une deuxième partie sur l'utilisation du jugement d'expert en sûreté de fonctionnement. Pour terminer ce paragraphe nous donnerons les avantages et les inconvénients de l'utilisation du jugement d'expert pour sélectionner les scénarii à étudier dans les FHA systèmes.

III.2.1 L'expertise

Avant de nous intéresser à l'utilisation du jugement d'expert, nous commencerons par définir l'expertise. En effet l'expertise, bien que le terme apparaisse aller de soi ou relever du sens commun, possède une définition et une manière d'opérationnalisation dans les études empiriques, loin d'être homogènes.

A un niveau très général et pour lequel un accord est dégagé, l'expert est quelqu'un qui réalise des performances supérieures stables à celles d'autres individus. Il y a accord également sur le fait que cette supériorité est acquise à travers la pratique, l'entraînement et la confrontation répétée au domaine de tâches de son champ d'expertise. La terminologie habituellement utilisée est celle de novice vs expert. Plus rarement on trouvera la terminologie débutant vs expérimenté qui met davantage l'accent sur la durée de la pratique. Malgré cette homogénéité de surface, dans le vocabulaire et le nombre extraordinaire d'études qui sont publiées dans ce domaine, il n'existe pas de consensus véritable.

Au plan des définitions théoriques, un auteur comme Shanteau (Shanteau, 1992) définit le novice comme quelqu'un qui possède non seulement des connaissances déclaratives dans son domaine mais

aussi un certain nombre d'habiletés acquises par la pratique. En revanche, pour Dreyfus (Dreyfus, et al., 1980) le novice possède uniquement des connaissances déclaratives et décontextualisées. De même pour l'expert, alors que le terme est employé généralement de façon globale, Patel et Groen (Patel, et al., 1991) (Patel, et al., 1986), dans le domaine médical, réservent le terme d'expert à celui qui possède des connaissances spécialisées (en cardiologie par exemple) et désignent par sub-expert celui qui possède les connaissances générique du domaine (par exemple médecin généraliste). On pourra noter que dans ce cas experts et sub-experts peuvent avoir la même quantité de pratique. Enfin, il existe une distinction selon que l'on définisse l'expert comme quelqu'un qui a acquis des habiletés et des routines et qui est donc très performant dans les situations et tâches habituellement rencontrées dans son domaine ; ou au contraire, définir l'expert comme quelqu'un qui a un fonctionnement flexible, qui est à même de traiter efficacement des situations nouvelles non rencontrées précédemment au cours de son expérience.

Ce flou dans la définition se retrouve au niveau de l'opérationnalisation, c'est à dire dans les choix des critères ou indicateurs qui vont conduire à distinguer un échantillon de participants qui seront qualifiés de novice vs un échantillon de participants qualifiés d'experts.

Un **premier mode d'opérationnalisation** de l'expertise consiste à distinguer les individus sur la base de leurs performances à une (ou plusieurs) tâche(s). La première implication de ce choix est qu'il faut pouvoir comparer des groupes d'individus différents dans la même situation. Or ceci est impossible dans les situations naturelles. D'où le recours à des tâches contrôlées en laboratoire. La question est alors celle du rapport entre tâche réelle et tâche expérimentale et la représentativité de la tâche expérimentale du point de vue de l'expertise. Une première façon est de construire une simulation qui soit la plus proche possible de la situation naturelle (fréquent dans l'étude de l'expertise professionnelle par exemple dans les expériences d'Ericsson (Ericsson, et al., 1988) et de Bennett (Bennett, 1983) sur les serveurs de café/restaurant). Une seconde façon de traiter cette question est d'identifier une composante de la tâche qui apparaît cruciale du point de vue de la performance ; ainsi, dans les études initiales sur les jeux d'échec, De Groot (De Groot, 1978) a pris comme critère le choix du meilleur mouvement à venir à un moment donné de la partie.

En l'absence d'une tâche permettant d'évaluer précisément la performance, un **deuxième mode d'opérationnalisation** va partir d'un classement des individus. Dans certains domaines, le classement est le reflet d'une performance répétée ; c'est le cas dans le classement des joueurs d'échec ou dans certains sports de compétition. Dans d'autres recherches, le classement sera le fruit d'une évaluation par les pairs ; c'est une procédure moins souvent utilisée et qui peut être sujette à distorsions. Dans le

domaine scolaire académique, on pourra évaluer la performance par des résultats à des examens. Plus récemment, et notamment chez l'enfant, on a vu apparaître des évaluations fondées sur un test de connaissances (ex. pour distinguer des novices des experts en football, Schneider & al (Schneider, et al., 1998) font passer un test de connaissances sur le football et séparent le groupe en deux à partir de la médiane).

Enfin, un **troisième mode d'opérationnalisation** de l'expertise s'appuie sur le fait que celle-ci se construit avec la pratique. On distinguera alors les notices et experts par la longueur de leur expérience : soit leur avancée dans le domaine académique (ex. des étudiants de licence et des doctorants) ; soit la présence ou non d'une pratique (ex. des étudiants en fin d'étude et des professionnels) ; soit le nombre d'années de pratique. Ce dernier critère ne va pas sans poser de problème. D'une part, il est fondé sur un fait implicite, à savoir que plus on pratique quelque chose, plus la performance s'accroît. Or, chacun d'entre nous doit posséder en tête un exemple d'une personne qui malgré une longue pratique, ne pourrait être qualifiée d'experte. On peut pallier ce problème en contrôlant que la durée de pratique s'accompagne effectivement d'un gain de performance. Le deuxième problème est que l'on ne dispose pas de modèles qui nous permettent de dire que, dans tel domaine de tâches, il faut tant d'années de pratique pour être expert. Il est clair pourtant qu'il ne faut pas le même nombre d'années pour être expert aux échecs ou dans la sécurité aérienne. En l'absence d'un tel modèle, les chercheurs définissent empiriquement novices et experts en fonction de l'échantillon de sujets dont ils disposent, ce qui aboutit à des groupes pour le moins disparates. Ainsi sur plusieurs études recensées, les novices allaient de personnes sortant juste de la formation jusqu'à des personnes ayant 7 ans d'expérience professionnelle ; les experts quant à eux, allaient de 1 an et demi de pratique à 27 ans de pratique. Pour une part non négligeable, ceux qui étaient considérés comme novices dans une étude étaient considérés comme experts dans une autre.

Globalement, même s'il est difficile de placer de façon absolue une limite entre experts et novices, de façon relative et pour une situation et un problème donnés (espace problème) :

- Un novice aura plutôt un raisonnement basé majoritairement sur des règles concoctées par des *expériences* étrangères à l'espace problème donné.
- Un expert aura plutôt un raisonnement basé sur des cas (*Pattern matching*) et des habilités (Skills), construits à partir de ces expériences sur cet espace problème.

Nous conclurons en donnant une définition appliquée de l'expertise par la *norme AFNOR NFX-50-110* (AFNOR : Association française de Normalisation, 2003) : *« l'expertise est un ensemble d'activités ayant pour objet de fournir à un client, en réponse à la question posée, une interprétation,*

un avis ou une recommandation, aussi objectivement fondé que possible, élaboré à partir des connaissances disponibles et de démonstrations accompagnées d'un jugement professionnel : les démonstrations incluent essais, analyses, inspections, simulations, etc. »

La norme NFX-50-110 fait la différence entre 3 types d'expertises :

- **Expertise individuelle :** « expertise réalisée par un seul expert sous sa propre responsabilité »
- **Expertise institutionnelle / Expertise collective :** « expertise conduite sous la responsabilité propre d'une institution, et réalisée par un ou plusieurs experts habilités par elle même, expertise réalisée par plusieurs experts appartenant à une même entité (organisme, entreprise…) »
- **Expertise collégiale :** « expertise réalisée par un collège d'experts choisis par chacune des parties clientes à une question déterminée ou par l'autorité compétente, l'expertise étant conduite sous la responsabilité collective des experts. L'un des experts est désigné pour coordonner l'expertise ».

Cette définition de l'expertise collégiale correspond à ce que l'on appelle le jugement d'expert.

III.2.2 L'utilisation du jugement d'expert en sûreté de fonctionnement

En sûreté de fonctionnement comme dans d'autres domaines, le jugement d'expert peut être considéré comme une des réponses possibles à un problème technique, un appoint là où un retour d'expérience et déficient ou un support à l'aide à la décision. Il est en effet utilisé pour compléter ou suppléer les données objectives éventuellement disponibles sur un problème considéré. C'est aussi, d'après Lannoy et Proccacia (Lannoy, et al., 2001), *« souvent la seule source d'information disponible dans un processus décisionnel ou lorsqu'il s'agit de parier sur l'occurrence d'évènement futurs et faire des prévisions où des prédictions »*. Plus généralement, la sûreté de fonctionnement peut faire appel à l'expertise pour les raisons suivantes :

- Compléter, valider, interpréter des données existantes, apprécier l'impact d'une modification (de conception, d'exploitation, de maintenance, d'environnement),
- Estimer l'occurrence d'événements futurs, prévoir les conséquences d'une décision
- Déterminer un état de connaissances d'un domaine,
- Fournir les éléments de décision nécessaires au décideur, lorsque plusieurs alternatives se présentent.

Toutefois, même si l'expertise peut être considérée comme une source de données, les données de l'expertise d'après Lannoy et Proccacia (Lannoy, et al., 2001) sont une « *représentation imparfaite et*

incertaine de ce qu'elles sont censées représenter ». La qualité des données recueillies par jugement d'expert varie en fonction :

- De l'information disponible et de l'expert, en particulier, de sa formation : son éducation, sa culture, sa profession, sa formation théorique et pratique sur le sujet, son expérience, son environnement, sa personnalité, d'une façon générale de la qualité intrinsèque de l'expert.
- De sa motivation et de son sérieux et de sa façon de recueillir, appréhender, transmettre, corriger et agréger ses connaissances,
- De la complexité relative du problème et de l'état général des connaissances sur ce problème, et de la manière dont cette complexité est modélisée dans l'esprit de l'expert,
- De la façon dont va s'instaurer le dialogue entre l'expert et l'analyste, chargé de recueillir puis d'interpréter son expertise.

Ainsi la qualité d'une expertise dépend non seulement de l'expert mais aussi de ceux qui vont la recueillir, l'analyser et l'utiliser.

III.2.3 Intérêts et inconvénients de l'utilisation du jugement

Comme nous l'avons vu en conclusion du paragraphe II de ce chapitre, la méthode doit permettre le traitement rapide de grandes quantités de données. L'utilisation du jugement d'expert peut aider à cela. En effet, comme nous l'avons vu dans les sous parties précédentes, le jugement d'expert est une source importante d'information dans le cadre de la sûreté de fonctionnement. De part leurs fonctions, les experts en Facteurs Humains et les experts en sûreté ont connaissance des évènements qui peuvent avoir lieu en service, des argumentaires et des résultats d'évaluation précédentes qui leur permettent de poser un avis argumenté sur la criticité et la survenue d'erreurs humaines sur des scénarii notamment ceux de la FHA. Nous pouvons donc imaginer une méthode basée sur le jugement d'expert afin de sélectionner les scénarii pertinent de la FHA.

Les experts en sûreté de fonctionnement des systèmes produisent les scénarii des FHA, par conséquence, ils les passent tous en revue. Ils pourraient donc potentiellement marquer les scénarii à étudier du point de vue des Facteurs Humains. Cependant, cette expertise est incomplète, elle se doit d'être couplée à une expertise en Facteurs Humains. Les experts Facteurs Humains ne participent pas à la rédaction des scénarii mais sont là en support de la rédaction. De ce fait, on pourrait imaginer que les experts en Facteurs Humains ne regardent que les scénarii extraits par les spécialistes en sûreté des systèmes. Mais, cette méthode est à éviter puisqu'elle risque de laisser passer des scénarii qui devraient être étudiés.

D'ailleurs, le principal inconvénient du jugement d'expert est qu'il ne permet pas la certitude absolue sur la sélection des scénarii. Cet inconvénient peut être contrôlé :

- En guidant le recueil de l'expertise – Par exemple, en utilisant des modèles théoriques aptes à aider l'expert à structurer sa réflexion.
- En utilisant plusieurs expertises – Par exemple, en recoupant des sensibilités expertes différentes. Dans ce cas, la sélection des scénarii serait réalisée en utilisant des connaissances expertes provenant de champs d'études différents augmentant de fait la précision de la sélection.
- En utilisant le retour d'expérience – Par exemple, les scénarii sélectionnés sur la FHA d'un système d'un avion particulier pourraient servir pour sélectionner les scénarii d'une FHA d'un système d'un autre avion. Cela afin de ne pas perdre les apprentissages d'un programme à l'autre. Cette étape pourrait permettre de gagner du temps qui pourrait être alors consacré à la sélection d'autres scénarii et donc optimiser la sélection.

En comparaison avec les méthodes fiabilistes, il faut savoir que malgré tout le soin apporté au recueil de ces expertises, il n'est pas certain d'égaler en termes de précision les résultats de ces méthodes qui traitent les scénarii de manière absolue. Par contre, l'utilisation du jugement d'expert permet une extraction beaucoup plus rapide des scénarii en comparaison du temps nécessaire à la réalisation d'une analyse hiérarchique des tâches.

En conclusion, l'utilisation d'un jugement d'expert simple ne permet pas de répondre aux besoins formulés par les experts. En effet, **la précision** qui pourrait s'avérer inadéquate pour un processus certifié, **la rapidité** qui malgré le gain par rapport aux méthodes fiabilistes pourrait ne pas suffire pour être en accord avec le processus industriel, le fait que **les spécialistes en Facteurs Humains doivent passer en revue l'intégralité des scénarii** pour avoir une expertise collégiale, tous ces points sont autant d'arguments pour ne pas choisir le jugement d'expert simple.

C'est pourquoi dans le paragraphe suivant nous allons proposer une méthode, s'appuyant sur un jugement d'expert outillé de façon à contrôler les faiblesses du jugement d'expert simple.

III.3 Une méthode optimisée pour répondre aux besoins de l'industrie

Après avoir présenté les conclusions précédentes aux experts qui avaient participé à la définition des besoins de l'industrie (voir paragraphe II de ce chapitre), nous avons défini les fonctions que devraient

remplir les outils ou les méthodes afin de palier les inconvénients du jugement d'expert simple à savoir :

- La rapidité
- La précision

Dans le premier paragraphe, nous présenterons le résultat de ce recueil de besoin. A la sortie de ce chapitre nous aurons les fonctions que remplirons les outils. Dans le paragraphe suivant, nous verrons quelles sont les possibilités de développements de ces outils, nous ferons aussi nos choix quant à ces possibilités de développements. Pour finir, nous présentation la méthode mise en place pour tester les outils.

III.3.1 Les outils pour répondre aux besoins des experts

De la même façon qu'ont été construites les caractéristiques des scénarii à rechercher, les fonctions des outils ont été construites par une démarche itérative d'affinement du besoin.

La **première itération** a permis de définir les besoins de haut niveau. Ces besoins pourraient être récapitulés ainsi :

- Donner les moyens aux experts en sûreté de fonctionnement des systèmes, de consolider l'ensemble des scénarii qu'ils auront préalablement sélectionnés.
- Puisque les experts en Facteurs Humains ne sont pas directement dans le processus d'écriture des FHA, donner les moyens aux experts en Facteurs Humains d'interroger de la façon la plus efficace possible la base de données. Cela doit permettre l'expertise collégiale.
- Garantir le transfert des connaissances acquises d'un programme à l'autre.

La **seconde itération** a permis de se rapprocher des fonctions que devront remplir les outils. Les résultats de cette itération pourraient se résumer dans le développement d'outils informatiques (pour garantir la rapidité maximale) permettant :

- La recherche et l'extraction de scénarii basés sur des critères experts.
- La recherche et extraction de scénarii basés sur les ensembles de scénarii de précédents programmes avions.

Deux autres points ont aussi été émis mais ils ne sont pas directement en lien avec les outils, ils se situent plutôt au niveau des processus. De ce point de vue, il faudra :

- Assurer le suivi des résultats de l'évaluation des scénarii sélectionnés.
- Assurer le suivi les modifications possibles résultant de l'évaluation des scénarii.

Enfin la **dernière itération** a permis de définir les fonctionnalités de ces outils.
Les deux premières fonctionnalités pourraient s'énoncer de la façon suivante :

- Permettre la sélection de scénarii en fonction de filtres particuliers. Par exemple, les experts voudraient pouvoir sélectionner tous les scénarii de catégorie MAJ, se passant à l'atterrissage, avec une panne où le système a un comportement erroné.
- Permettre la sélection de scénarii en fonction de leur « proximité » avec un scénario existant. Par exemple, les experts voudraient lorsqu'ils ont sélectionné un scénario, vérifier les scénarii qui lui sont proches afin d'évaluer s'il faut les sélectionner ou non. Cette fonction permet de donner d'apporter de la cohérence à la sélection et ouvrir le champs de validité des scénarii retenus.

Ces deux premières fonctionnalités doivent permettre d'aider l'application des expertises. C'est-à-dire, pour les experts en sûreté de plus facilement gérer la cohérence de leur sélection durant l'écriture. Pour les experts en Facteurs Humains cela permet d'interroger plus facilement la large base de données.

Les outils devraient aussi avoir la fonctionnalité suivante :

- Permettre la sélection de scénarii à partir d'un ensemble de scénarii. Par exemple, après une première sélection, les experts ont une base de scénarii sélectionnés. Afin de compléter et consolider cet ensemble, il faudrait que l'outil permette de sélectionner les scénarii qui auraient des caractéristiques similaires aux scénarii sélectionnés mais ne l'ayant pas été.

Cette fonctionnalité permettrait tout d'abord d'apporter de la cohérence à l'ensemble final de scénarii et fur et à mesure des itérations de sélection permettre une plus grande maturité de l'ensemble sélectionné. De plus, cette fonctionnalité sera nécessaire pour faciliter le transfert d'un programme avion à l'autre. En effet, l'ensemble final de scénarii pour un avion qui a fini son développement pourrait servir pour d'interroger la base de scénarii d'un avion en cours de développement.

Dans le prochain paragraphe, nous allons détailler les choix de développement des outils répondant aux fonctionnalités précédentes.

III.3.2 Les choix de développement

Pour remplir ces fonctions, nous nous sommes tournés vers les méthodes de Traitement Automatique de la Langue *(TAL),* les méthodes de filtrage automatique de textes et l'utilisation de bases de données.
Compte tenu des contraintes de temps nécessaires à l'implémentation des méthodes présentent dans la littérature nous n'avons testé qu'une méthode par fonctionnalité. Ce paragraphe sera organisé selon les

fonctionnalités décrites précédemment. Pour chaque fonctionnalité, une rapide revue des méthodes à disposition sera réalisée. Puis, de manière argumentée, nous en choisirons une et nous décrirons son mode de fonctionnement, ses avantages et ses inconvénients.

III.3.2.1 Permettre la sélection de scénarii en fonction de filtres particuliers

Pour cette fonctionnalité, le choix du moyen de réalisation est assez simple puisqu'il n'existe que deux méthodes principales de filtrage par rapport à des mots clefs.

La première utilise des requêtes dans une base de données, la deuxième similaire dans son fonctionnement mais plus limité dans la complexité des requêtes est l'utilisation de requêtes dans un tableur. L'utilisation d'outils de base de données existants permettra l'envoi de requêtes faites par l'expert (mots clefs, filtres, requêtes plus complexes type *« SELECT… FROM… WHERE… »*). Nous avons donc choisi d'utiliser des bases de données. Cependant, la difficulté ne réside pas tant dans le choix de la méthode de filtrage, mais plutôt dans la nature des requêtes à réaliser. Par exemple, nous allons voir plusieurs descripteurs tels que la phase de vol, le type de panne, la classification *(« MIN », « MAJ », « HAZ », « CAT »)*, la présence d'action de l'équipage,… qui sont des données importantes pour ces requêtes. Mais comment préciser ces requêtes ?

Pour répondre à cette question, nous avons eu deux angles d'approche :

1/Tout d'abord, nous avons pris l'hypothèse suivante :

Puisque les scénarii à sélectionner sont ceux pour lesquels une erreur d'un membre de l'équipage entraînera une modification de la classification, alors les données (ou descripteurs) : classification, type de panne, phase de vol... qui voient leur répartition modifiée en fonction de la classification, sont à prendre en compte.

Pour illustrer ce postulat, un scénario classé « MAJ », qui possède des descripteurs statistiquement plus associés à la catégorie « CAT » (par exemple : Atterrissage, Comportement erroné du système) est potentiellement intéressant. En effet, dans ce cas, l'équipage pourrait participer au maintien de la classification « MAJ ». Les requêtes visant à sélectionner les scénarii MAJ, à l'atterrissage avec un comportement erroné du système sont donc à privilégier.

2/Ensuite, nous avons utilisé des scénarii sélectionnés par les experts et étudié la répartition des descripteurs entre l'ensemble des scénarii sélectionnés et l'ensemble global des scénarii. Cela permet de retrouver les critères pertinents pour les choix des scénarii.

En comparant les résultats de ces deux méthodes, nous pourrons alors valider ou infirmer l'hypothèse précédente.

Cette étude sera réalisée dans le second chapitre de la partie II sur l'étude expérimentale. Voyons maintenant comment remplir la seconde fonctionnalité.

III.3.2.2 Permettre la sélection de scénarii en fonction de leur « proximité » avec un scénario existant

Cette fonctionnalité devra permettre, à partir d'un exemple de scénario de la base de données, de trouver des scénarii qui auront des propriétés similaires. Pour y répondre, il existe une suite d'algorithmes nommés : algorithmes des « plus proches voisins » (*Nearestneighbor* : Cover et Dasarathy (Cover, et al., 1967) et (Dasarathy, 1991)). Ils sont utilisés notamment pour les systèmes de raisonnement à partir de cas (*Case Based Reasoning : CBR ;* (Kolodner, 1992), (Aamodt, et al., 1994), (Maher, et al., 1995))

Algorithmes des « plus proches voisins » :

Le problème de la recherche des plus proches voisins est un problème courant en algorithmique et de nombreux auteurs ont proposé des algorithmes efficaces pour le résoudre rapidement notamment Dasarathy (Dasarathy, 1991).

La méthode s'applique sur un espace E de dimension D (dans notre cas E est l'ensemble des scénarii de la base de données). Le cardinal de E est :

Card E = Nombre de scénarii dans la base de données

La dimension de D de E dépend de la structure de E. Définir la dimension de E va dépendre des études que nous allons faire sur cet espace. Pour cette étude nous voulons trouver les scénarii les plus proches d'un scénario S_1 appartenant à E. La notion de proximité induit la notion de distance mathématique. La distance peut être mesurée entre deux vecteurs. Pour Robert M. Losee (Losee, 1998), tout texte peut se transformer en vecteur de la façon suivante. L'ensemble des mots unitaires du texte (le lexique) forme la base de l'ensemble. C'est sur cette base que vont pouvoir ce décomposer tous les extraits de ce texte.

Par exemple, voici un ensemble E de 3 textes (*Card E = 3)*,

Texte 1 : « je je vais ».

Texte 2 : « je je je je vais vais ».

Texte 3 : « je vais vais ».

Il y a deux mots présents dans le lexique, « je » et « vais ». Ces deux mots forment la base d'un ensemble à deux dimensions. Les vecteurs correspondants sont : [1,0] et [0,1]. « Je » correspond au vecteur de base [1,0] et « vais » à [0,1]. Dans ce cas là, les vecteurs correspondant aux trois textes précédents sont :

V_1 = [2,1]

V_2 = [4, 2]

V_3 = [1, 2]

Grace à cette décomposition en vecteurs, il est alors possible de mesurer la distante entre deux de ces vecteurs. Les mesures de distance utilisées peuvent être :

- La distance de Manhattan
- La distance euclidienne
- La distance angulaire

Soit deux vecteurs $V_1 = (x_1, x_2,\ldots, x_n)$ et $V_2 = (y_1, y_2, \ldots, y_3)$, voici les formules pour calculer ces distances :

La **distance Manhattan** se calcule de la façon suivante :

$$d = \sum_{i=1}^{n} |xi - yi|$$

La **distance euclidienne** se calcule de la façon suivante :

$$d = \sqrt{\sum_{i=1}^{n} (xi - yi)^2}$$

La **distance angulaire**, qui est la mesure de l'angle entre les deux vecteurs, se calcule de la façon suivante :

$$\alpha = \cos^{-1}\left\{\frac{V1.V2}{||V1|| \times ||V2||}\right\}$$

Pour les ensembles où la base est de petite dimension (D égal à 1, 2 ou 3) il est possible de représenter les textes graphiquement comme le présente la Figure 22.

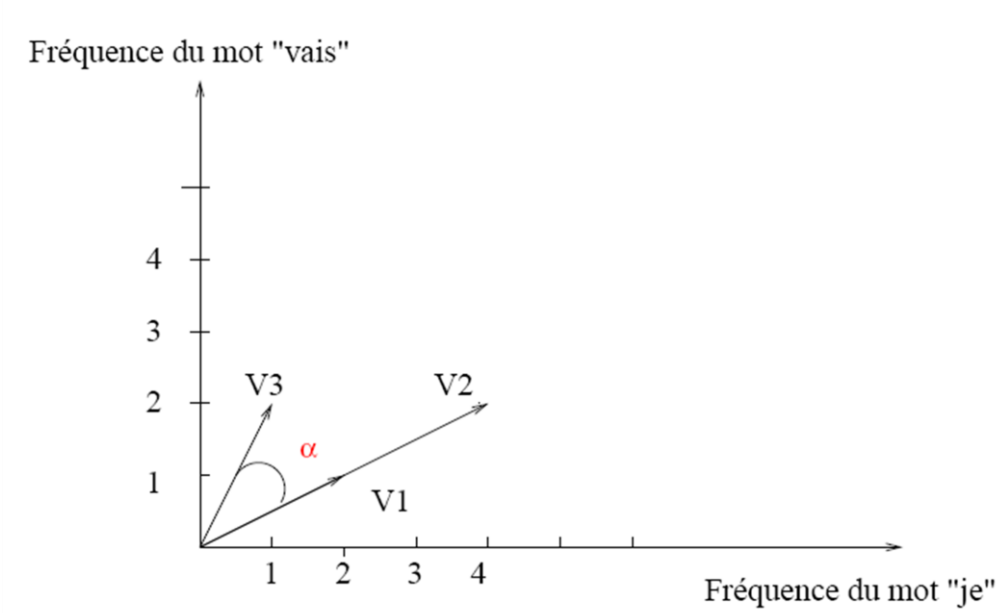


Figure 22 : Représentation graphique des textes précédents

La distance euclidienne est la plus utilisée, c'est celle que nous utiliserons.

L'utilisation de la méthode des proches voisins est une méthode puissante de mesure de similarité entre deux textes. Néanmoins, la décomposition en vecteurs part d'un postulat remis en cause au niveau de la linguistique. Ce postulat est le suivant : *« Le message véhiculé par un texte est entièrement dépendant des mots qui le constituent »*. Ce postulat a des limites puisque le passage en vecteurs même s'il garde les mots constituant le texte, ne prend pas en compte les éléments de syntaxe, de grammaire, qui sont pourtant des éléments porteurs de sens dans le texte. Cependant, de nombreuses études ont montré son intérêt, nous utiliserons donc cette méthode tout en gardant ses défauts à l'esprit afin de mettre en place de moyens pour évaluer leur impacts.

L'utilisation de la méthode des proches voisins sera développée dans le chapitre 3 de la partie expérimentale

III.3.2.3 Permettre la sélection de scénarii à partir d'un ensemble de scénarii

Cette fonctionnalité devra permettre, à partir d'un ensemble de scénarii de la base de données, de trouver des scénarii qui auront des propriétés similaires. Il existe beaucoup d'algorithmes permettant de réaliser cela. Par exemple : le filtrage bayesien (filtrage du pourriel : (Sahami, et al., 1998)), les réseaux neuronaux (Le perceptron : (Rosenblatt, 1958) ; Perceptron Multicouches : (Rumelhart, et al., 1986)), et les machines à vecteurs de support (Vapnik, et al., 1963), (Duda, et al., 1973), (Duda, et al., 2001).

Nous nous sommes intéressés au filtrage bayésien puisque son application actuelle est la plus proche de notre besoin. En effet, cette méthode est utilisée quotidiennement par des millions d'utilisateurs à travers les filtres du courrier indésirable de leur messagerie électronique. Par analogie, nous avons des scénarii « indésirables » que nous souhaiterions extraire. Nous allons maintenant voir les principes de la méthode de filtrage bayésien.

Le filtrage bayésien :

Le principe général du filtrage bayésien est basé sur l'utilisation de probabilité d'occurrence de mots. Les mots ont des probabilités d'apparaître dans un « scénario à sélectionner (S1)» et dans un « scénarii à ne pas sélectionner (S2)». Le filtre ne connaît pas à l'avance ces probabilités, c'est pourquoi il lui faut un temps d'apprentissage pour les évaluer. Dans notre cas, l'apprentissage pourra se faire à partir de l'ensemble des scénarii de départ, c'est-à-dire ceux que l'on veut confronter à la base de données. Pour chaque nouveau scénario, le filtre ajustera les probabilités de rencontrer ce mot dans un scénario de type S1 ou de type S2 et les stockera dans sa base de données.

Après l'apprentissage, les probabilités des mots (également appelées fonctions de vraisemblance) sont utilisées pour calculer la probabilité qu'un scénario (S) soit un scénario de type S1. Chaque mot du scenario contribue au calcul de cette probabilité. Cette probabilité est calculée en utilisant le théorème de Bayes. Une fois que le calcul de probabilité terminé, on la compare à une valeur arbitraire (95% par exemple) pour le marquer ou non comme un scénario à sélectionner. Entrons maintenant dans le détail du calcul de probabilité.

Etape 1 : Calcul de la probabilité qu'un scénario soit « à sélectionné » (S1) sachant qu'il contient un mot (M).

Le théorème de Bayes donne les formules suivantes :

$$P(S1 \cap M) = P(S1/M) \times P(M)$$

$$P(S1 \cap M) = P(M/S1) \times P(S1)$$

Ici :

$P(S1 \cap M)$ est la probabilité d'avoir un scénario sélectionné S1 contenant le mot M.

$P(M)$ est la probabilité d'apparition du mot M.

$P(S1)$ est la probabilité qu'un scénario soit sélectionné.

$P(M/S1)$ est la probabilité d'apparition du mot M dans un scénarii sélectionné S1.

$P(S1/M)$ est la probabilité qu'un scénario soit sélectionné sachant qu'il contient le mot M.

Cette dernière probabilité est la probabilité qui nous intéresse. Par égalité des deux formules précédentes nous obtenons :

$$P(S1/M) \times P(M) = P(M/S1) \times P(S1)$$

$$P(S1/M) = \frac{P(M/S1) \times P(S1)}{P(M)}$$

Nous pouvons remarquer que, l'ensemble S1 des « scénarii sélectionnés » et l'ensemble S2 des « scénarii non sélectionnés » forment une partition de l'ensemble E :

$$E = S1 \cap S2$$

Cette particularité permet modifier la formule précédente :

$$P(S1/M) = \frac{P(M/S1) \times P(S1)}{P(M \cap S1) + P(M \cap S2)}$$

$$P(S1/M) = \frac{P(M/S1) \times P(S1)}{\{P(M/S1) \times P(S1)\} + \{P(M/S2) \times P(S2)\}}$$

Puisque l'on ne connait pas à priori la probabilité d'avoir un scénario à sélectionner ni celle d'avoir un scénario à ne pas sélectionner, nous partons du postulat qu'il y a autant de chances de sélectionner un scénario que de ne pas le sélectionner. Autrement dit :

$$P(S1) = P(S2) = 0{,}5$$

La formule précédente peut donc être simplifiée de la façon suivante :

$$P(S1/M) = \frac{P(M/S1)}{P(M/S1) + P(M/S2)}$$

Cette formule permet le calcul nécessaire à la première étape. Voyons maintenant la deuxième étape c'est-à-dire la combinaison de toutes ces probabilités pour tous les mots d'un scénario.

Etape 2 : Calcul de la probabilité qu'un scénario soit à sélectionné (S1) sachant qu'il contient les mots (m_1, m_2, … m_n).

La méthode de filtrage bayésien fait l'hypothèse *« naïve »* que les mots présents dans les scénarii sont des évènements indépendants. Cela est faux dans les langages naturels comme le français, où la probabilité de trouver un adjectif, par exemple, est influencée par la probabilité d'avoir un nom. Cela représente une limite à l'utilisation de cette méthode.

L'indépendance de plusieurs évènements implique aux niveaux des probabilités la propriété suivante :

$$P(m1 \cap m2 \cap \ldots \cap mn) = P(m1) \times P(m2) \times \ldots \times P(mn)$$

Le but de cette étape est de calculer :

$$P\,(S1/\{m1 \cap m2 \cap \dots \cap mn\})$$

En appliquant une première fois la formule de Bayes nous obtenons :

$$P\,(S1/\{m1 \cap m2 \cap \dots \cap mn\}) = \frac{P(S1 \cap \{m1 \cap m2 \cap \dots \cap mn\})}{P(m1 \cap m2 \cap \dots \cap mn)}$$

Or les évènements m_1, m_2,…, m_n sont indépendants :

$$P\,(S1/\{m1 \cap m2 \cap \dots \cap mn\}) = \frac{P(S1 \cap m1) \times P(S1 \cap m2) \times \dots \times P(S1 \cap Mn)}{P(m1 \cap m2 \cap \dots \cap mn)}$$

$$P(S1/\{m1 \cap m2 \cap \dots \cap mn\}) = \frac{P(S1)^n \times P(S1/m1) \times P(S1/m2) \times \dots \times P(S1/mn)}{P(m1 \cap m2 \cap \dots \cap mn)}$$

Comme nous l'avons vu plus haut, l'ensemble S1 des scénarii à sélectionner et l'ensemble S2 des scénarii à ne pas sélectionner forment une partition de l'ensemble E des scénarii. Nous pouvons donc écrire :

$$P(S1/\{m1 \cap m2 \cap \dots \cap mn\}) = \frac{P(S1)^n \times P(S1/m1) \times P(S1/m2) \times \dots \times P(S1/mn)}{P(S1 \cap \{m1 \cap m2 \cap \dots \cap mn\}) + P(S2 \cap \{m1 \cap m2 \cap \dots \cap mn\})}$$

Or les évènements m_1, m_2,…, m_n sont indépendants :

$$P(S1/\{m1 \cap m2 \cap \dots \cap mn\}) = \frac{P(S1)^n \times P(S1/m1) \times P(S1/m2) \times \dots \times P(S1/mn)}{P(S1)^n \times P(S1/m1) \times \dots \times P(S1/mn) + P(S2)^n \times P(S2/m1) \times \dots \times P(S2/mn)}$$

Or nous avons vu précédemment que, puisque l'on ne connait pas à priori la probabilité d'avoir un scénario « à sélectionner » ni celle d'avoir un scénario « à ne pas sélectionner », nous partons du postulat qu'il y a autant de chance de sélectionner un scénario que de ne pas le sélectionner.
Autrement dit :

$$P(S1)^n = P(S2)^n = 0{,}5^n$$

De plus, S1 et S2 sont des évènements contraires, c'est-à-dire que :

$$P(S1) = 1 - P(S2)$$

Nous pouvons donc simplifier la formule précédente :

$$P(S1/\{m1 \cap m2 \cap \dots \cap mn\}) = \frac{P(S1/m1) \times P(S1/m2) \times \dots \times P(S1/mn)}{\{P(S1/m1) \times \dots \times P(S1/mn)\} + \{(1 - P(S1/m1)) \times \dots \times (1 - P(S1/mn))\}}$$

Au final si nous appelons :

p la probabilité de sélectionner un scénario sachant un ensemble de mots

p_1 la probabilité de sélectionner un scénario sachant le mot 1

p_2 la probabilité de sélectionner un scénario sachant le mot 2

…

p_n la probabilité de sélectionner un scénario sachant le mot n

Nous obtenons la formule qui servira au filtrage des scénarii sachant à partir un ensemble connu de scénarii :

$$p = \frac{p1 \times p2 \times \ldots \times pn}{p1 \times p2 \times \ldots \times pn + (1 - p1) \times (1 - p2) \times \ldots \times (1 - pn)}$$

Durant l'explication de la méthode, nous avons soulevé le principal inconvénient c'est-à-dire l'hypothèse faite par ces méthodes de filtrage que les mots présents dans les scénarii sont des évènements indépendants. Cependant, ces méthodes ont déjà fait leur preuves notamment pour le filtrage du courrier indésirable c'est pourquoi nous allons les utiliser ici tout en gardant à l'esprit leurs limitations.

L'implémentation de cette méthode sera présentée lors du troisième chapitre de la partie expérimentale.

III.3.3 Le test des outils

Afin de tester la performance des méthodes présentées ci-dessus, nous avions besoin d'un ensemble de scénarii déjà sélectionnés pour servir de témoins. La sélection de ce set de scénarii sera présentée en détails dans le chapitre 1 de la partie expérimentale. Nous allons tout de même, présenter dans ce paragraphe les principes généraux.

Nous savons que pour utiliser et tester les outils précédents, il était nécessaire de posséder un ensemble de scénarii. Avec l'aide des experts, nous avons sélectionné 3 analyses de sécurités qui forment un ensemble d'environ 2200 scénarii. Chaque analyse a été choisie pour être pertinente du point de vue des Facteurs Humains et posséder des particularités systèmes différentes (système complexe, système d'interfaces entre l'Homme et la machine, système source d'énergie). Des pilotes d'essais, des experts en sûreté des systèmes et des experts en Facteurs Humains, ont sélectionné les scénarii à évaluer du point de vue des Facteurs Humains. Cet ensemble de scénarii a permis de tester les différents outils. Notamment en utilisant des ensembles pour l'apprentissage des outils. Une fois les outils calibrés, il a été alors possible de tester le rappel de scénarii sélectionnés mais qui n'étaient pas présents dans la base d'apprentissage. Cette étape permet notamment d'estimer la performance des outils et d'ouvrir des pistes d'amélioration.

Voyons maintenant pour conclure ce chapitre, comment ces outils vont s'intégrer dans le processus industriel

IV Conclusion : outils et processus industriel

Dans les paragraphes précédents, nous avons explicité les besoins de l'industrie et donné les moyens d'y répondre. Ce n'est pas le tout d'avoir un besoin et des outils pour y répondre, il est nécessaire de poser la logique de travail dans laquelle ils seront utilisés. Dans ce paragraphe nous n'allons pas donner tout le détail nécessaire à l'industrialisation potentielle de ces outils, mais seulement les principes de haut niveau. Pour cela, nous allons représenter de manière schématique les processus de conception actuels du point de vu des Facteurs Humains avec en parallèle, les processus de rédaction des analyses de sûreté. Puis, nous verrons où l'utilisation de ces outils serait la plus judicieuse. De plus, nous avons vu, durant le recueil des besoins qu'il était nécessaire de tracer l'extraction et l'évaluation des scénarii issus des FHA. Nous présenterons rapidement comment cela peut être réalisé.

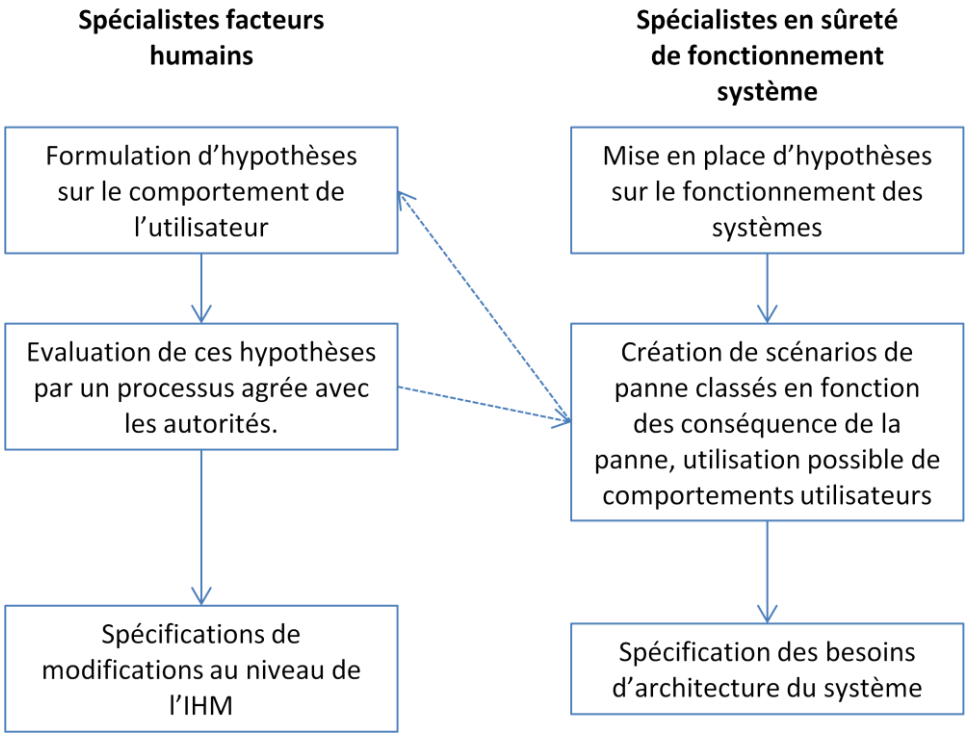


Figure 23 : Processus de conception actuel.

Nous pouvons voir sur la Figure 23, les deux processus de conception concernés en parallèle avec les liens déjà existants entre ces processus (flèches en pointillés). Cependant, il existe des difficultés de

réalisation de ces liens. C'est à ce niveau que les besoins définis au paragraphe II de ce chapitre ce situent. Les outils proposés ici aideraient donc à ce niveau, comme présenté sur la Figure 24.

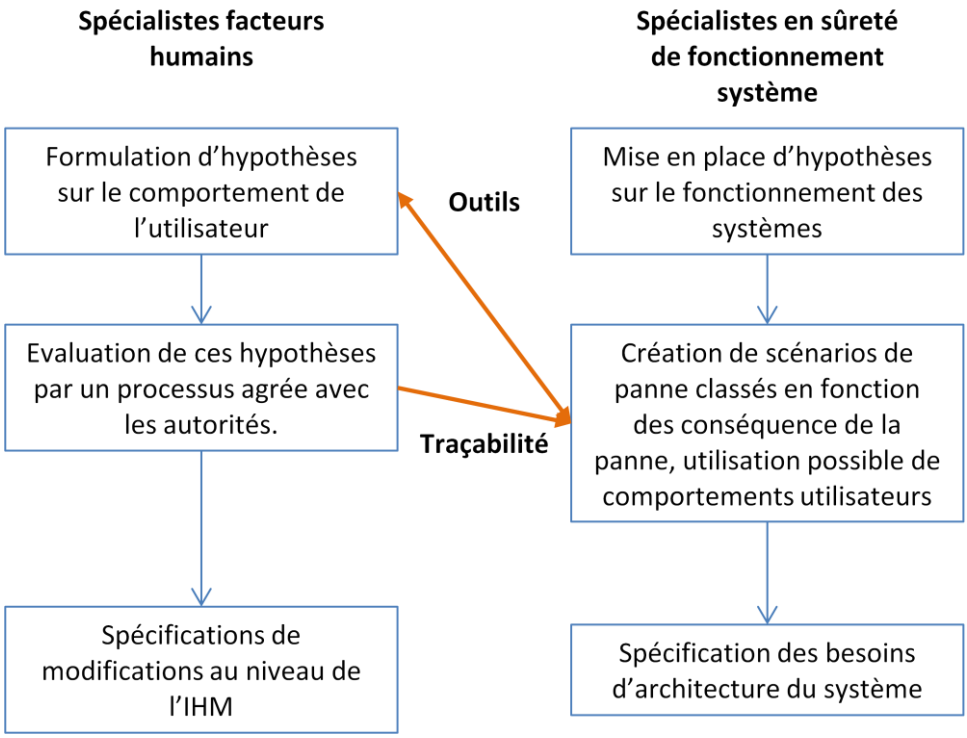


Figure 24 : Processus outillé pour répondre aux besoins des experts.

Les outils présentés dans les paragraphes précédents permettront d'avoir des liens plus aisés entre la création des scénarii par les spécialistes en sûreté des systèmes et la formulation d'hypothèses sur le comportement humain. Ces liens plus forts permettront grâce à une évaluation Facteurs Humains, d'augmenter la confiance que les concepteurs ont dans les hypothèses de comportement de l'équipage prises dans les analyses de sécurité. La traçabilité, assurée à l'aide de méthodes déjà éprouvées par l'industrie (référencement des documents, suivi des recommandations de modification de la conception…), permettra de compléter les scénarii d'analyse de sûreté de fonctionnement avec les hypothèses de comportement consolidées par les études en Facteurs Humains.

Voyons maintenant, grâce à la méthode d'analyse, comment les grandes notions vues dans les chapitres précédents s'articulent pour former le fil conducteur de ce travail de recherche.

CHAPITRE. 4 MODELE D'ANALYSE

L'objectif de ce chapitre est de positionner notre problème dans le cadre théorique. Ce chapitre fera la synthèse des notions abordées précédemment en mettant en lumière les liens entre elles. Tout d'abord, nous allons présenter dans le paragraphe I, la problématique de ce travail de recherche. Puis en nous aidant d'un schéma, nous montrerons les liens entre cette problématique et le socle théorique vu dans les chapitres précédents. Nous présenterons ensuite la démarche mise en place pour répondre à cette problématique. Enfin, le paragraphe III conclura ce chapitre ainsi que la partie théorique.

I Problématique

La problématique de cette thèse s'est précisée tout au long du processus de recherche. Pour expliquer sa construction, nous allons ici donner les questions que nous nous sommes posées avant d'aboutir à la problématique finale.

Tout d'abord, notre première question fut la suivante : *Comment optimiser la conception centrée utilisateur pour concevoir un avion de ligne ?*

A travers une meilleure conception centrée sur l'utilisateur, répondre à cette question permettra de mieux gérer les conséquences des erreurs de pilotage qui sont, dans ce domaine, une source importante d'incidents et d'accidents. Pour optimiser la conception centrée utilisateur, nous souhaitons la doter d'outils adaptés pour les concepteurs. Nous avons vu, dans les chapitres précédents, que la conception des avions de lignes possède des contraintes et des méthodes de travail propres qui doivent être intégrées afin de garantir les outils les plus adaptés. Pour cela, il est donc nécessaire de répondre aux questions suivantes :

- *Où devons nous outiller (c'est-à-dire à quelle étape de la conception) pour optimiser la conception centrée utilisateur ?*

- *Quels outils choisir ?*

Comme nous avons pu le voir précédemment, la conception est conduite par des hypothèses **(hypothèses d'architecture et de fonctionnement des systèmes et hypothèses de comportement des utilisateurs).** Les hypothèses de comportement des utilisateurs sont élaborées, en partie, de la façon suivante : des experts en analyse de sûreté des systèmes établissent la liste de l'ensemble des pannes possible de l'avion. Pour chacune d'entre elles, ils indiquent le comportement attendu de l'équipage pour résoudre la panne. Ces hypothèses de comportement servent alors de guide à la conception.

Des hypothèses de comportement erronées (c'est-à-dire ne correspondant pas à ce que l'équipage ferait réellement face à une panne), pourraient donc entraîner des faiblesses dans l'objet final. Le test de ces hypothèses peut s'effectuer lors d'évaluations en facteurs humains, réalisées sur des simulateurs ou lors d'essais en vol. Cependant pour un système aussi complexe qu'un avion de ligne, le nombre d'hypothèses émises est gigantesque et malgré ce fait, la réalité exposera des situations bien plus variées.

Une base de données pouvant contenir un grand nombre d'hypothèses, et pourtant à l'heure actuelle peu exploitée, est constituée par ce qu'on appelle **les analyses de sûreté des systèmes.**
Ces analyses sont un recueil de scenarii rédigés par les experts en sûreté qui contiennent notamment, les hypothèses de comportement de l'équipage face à la défaillance d'un système quelconque (freinage, conditionnement d'air, Pilote Automatique, etc.).

Les hypothèses élaborées dans ces analyses sont issues de situations anormales (présence d'une panne d'un système). Le nombre de panne recensé est considérable, il est donc difficile pour les concepteurs, de faire le tri entre les hypothèses (ou les scenarii) discutables et ceux qui ne le sont pas. Les hypothèses discutables sont celles pour lesquelles le comportement réel de l'équipage n'aura pas été correctement prédit par les experts.

Face a cette base considérable de données, il n'est pas possible de tester (sur simulateurs ou par essais en vol) l'ensemble des scenarii. Dans ce contexte l'approche que nous avons choisie consiste à identifier et étudier uniquement les hypothèses que des experts en facteurs humains jugent discutables. L'objectif de notre thèse était donc d'élaborer et de valider des outils facilitant l'extraction de ces hypothèses discutables.

Pour atteindre cet objectif, nous avons réalisé les étapes suivantes :

1) Des experts avaient pour tâche de sélectionner les scenarii qu'ils considéraient comme étant discutables dans un ensemble réduit de la base de données initiale (analyse de sûreté). Nous présenterons ces résultats dans le chapitre empirique n°1 ;

2) Ensuite, ces scenarii ont été étudiés afin d'identifier des thématiques récurrentes, des mots clés, des descripteurs, des relations entre descripteurs, etc. L'objectif à terme est de pouvoir utiliser ces informations pour réaliser des requêtes, c'est-à-dire des extractions automatisées des scénarii discutables à partir de mots clés.). Nous présenterons ces résultats dans le chapitre empirique n°2 ;

3) Toujours dans le but de créer des outils d'extractions automatiques, nous avons utilisé les scénarii sélectionnés par les experts pour tester la pertinence de 2 méthodes existantes d'extraction d'information dans une base de données ; la méthode des plus proches voisins et la méthode du filtrage bayésien. L'évaluation de ces méthodes fait l'objet de notre chapitre empirique n°3.

II Schéma du modèle d'analyse : articulation des notions du cadre théorique

Le problème posé par cette étude est à la croisée de nombreux domaines dont nous avons essayé de donner un aperçu dans les chapitres précédents. Ces domaines proviennent à la fois de principes théoriques mais aussi de principes liés au contexte industriel de la thèse. Ces nombreux domaines créent un cadre théorique possédant de multiples aspects représentés sur la figure suivante (Figure 25) qui forme donc le modèle d'analyse de ce travail de recherche.

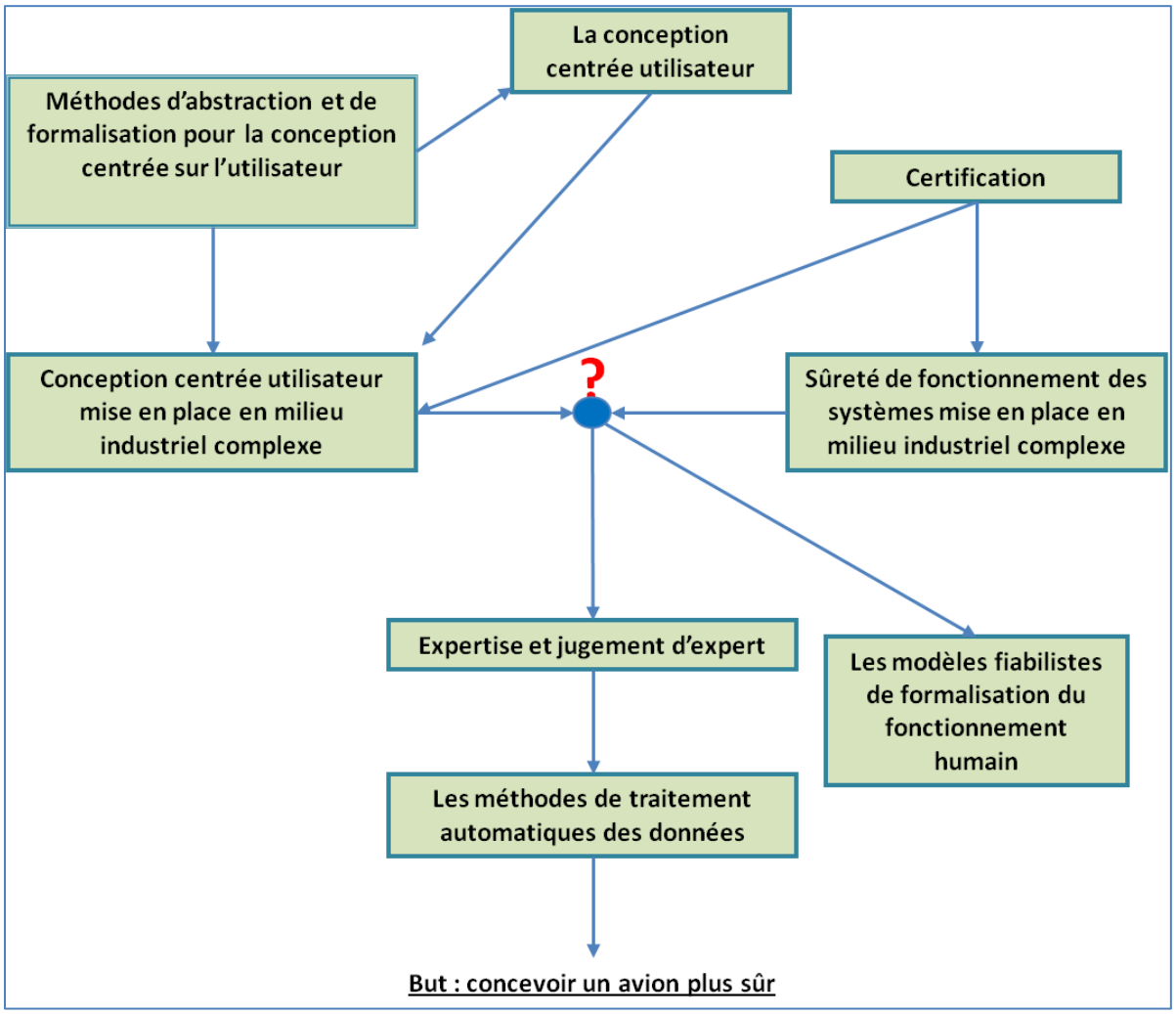


Figure 25 : Modèle d'analyse

Ce schéma reprend donc toutes les notions théoriques présentées dans cette étude :

- **La conception centrée utilisateur.** En effet, notre objectif est d'optimiser la conception des avions afin qu'elle prenne mieux en compte les possibilités d'erreurs de l'équipage.
- **La réglementation de la conception en aéronautique.** Nous avons vu que les processus industriels, mis en œuvre pour concevoir un avions de ligne, répondent à des exigences d'autorités de certification. Ces règlements imposent des exigences et des méthodes de travail.
- **La sûreté de fonctionnement des systèmes.** La sûreté de fonctionnement est un des acteurs du haut niveau de sécurité du transport aérien. La mise en place des analyses de sûreté permet de garantir une architecture des systèmes la plus robuste possible aux pannes les plus graves.

Ces aspects, représentent le cadre théorique dans lequel s'inscrit notre thèse. Cependant comme nous l'avons montré dans le paragraphe donnant la problématique, la réalisation du cycle de conception d'un système aussi complexe qu'un avion de ligne, n'est pas sans poser des difficultés. Afin d'optimiser la mise en place du cycle de conception centrée sur l'utilisateur et de mieux prendre en compte les hypothèses de comportement de l'équipage contenues dans les analyses de sûreté des systèmes, nous avons étudié les champs théoriques suivants :

- **L'expertise et le jugement d'experts.** Cet aspect théorique correspond plus au besoin industriel mais nécessite d'être optimisé.
- **Les méthodes de traitement automatique des données.** Ces méthodes permettront d'optimiser le jugement d'expert.

III Conclusion

Le positionnement du problème a souvent été source de réflexion, en étant à la frontière de domaines scientifiques en perpétuelle évolution et d'enjeux industriels considérables. La façon dont a été défini le problème, c'est-à-dire en concertation permanente avec les experts des domaines concernés, a certes été source d'incertitudes et de remises en question tout au long du travail, mais a permis de consolider un besoin, et d'adapter les méthodes existantes au contexte et aux exigences industrielles.

Pour conclure, bien que la demande émane spécifiquement de l'aéronautique, de nombreuses industries sont soumises à des contraintes similaires. Ces travaux pourraient à termes trouver des échos dans de très nombreux domaines industriels et en particulier les industries à risques tels que l'industrie nucléaire, le transport ferroviaire, l'industrie spatiale, etc.

Nous allons maintenant commencer la deuxième partie : **Partie expérimentale**.

Partie II : DEMARCHE EXPERIMENTALE

Pour rappel, voici le contenu des chapitres que nous allons aborder ici :

Chapitre 1 : Des experts avaient pour tâche de sélectionner les scenarii qu'ils considéraient discutables dans un ensemble réduit de la base de données (analyse de sûretés).

Chapitre 2 : Les scenarii sélectionnés dans le chapitre 1 ont été étudiés afin d'identifier des thématiques récurrentes, des mots clés, des descripteurs, des relations entre descripteurs, etc. L'objectif à terme est de pouvoir utiliser ces informations pour réaliser des requêtes, c'est à des extractions automatisées des scénarii discutables à partir de mots clés.

Chapitre 3 : Toujours dans le but de créer des outils d'extraction automatique, nous avons utilisé les scénarii sélectionnés par les experts pour tester la pertinence de 2 méthodes existantes d'extraction d'information dans une base de données ; la méthode des plus proches voisins et la méthode du filtrage bayésien.

CHAPITRE. 1 SELECTION DES SCENARII PERTINENTS PAR UN GROUPE D'EXPERTS

L'objectif de ce chapitre est de présenter la première étape de notre démarche expérimentale. Cette étape est constituée de deux sous parties :

1) Sélection de scenarii pertinents par des experts ;
2) Etude approfondie de ces scenarii à l'aide du modèle de prise de décision de Reason, (1990).

Ce deuxième point a pour but de caractériser les hypothèses de comportement présentes dans les scenarii sélectionnés.

Les résultats présentés ici ont été publiés dans les actes de la conférence Lambda Mu, 17ième Congrès de Maîtrise des Risques et de Sûreté de Fonctionnement (Papaïx, et al., 2010).

Sélection des scénarii pertinents par un groupe d'experts : Problématique

Dans la partie précédente, nous avons présenté le contexte (scientifique et industriel) de ce travail de recherche. Notamment pour les systèmes sociotechniques complexes, nous avons montré l'intérêt « d'outiller » la conception centrée sur l'utilisateur. Pour l'aéronautique, qui constitue le terrain de cette thèse, nous avons vu que les hypothèses de conception **(sur l'utilisateur, l'architecture des systèmes, le fonctionnement des systèmes)** sont très nombreuses et répertoriées à travers de nombreux documents différents. Les analyses de sûreté des systèmes font partie de ces documents, et sont aussi une source d'hypothèses prises sur le comportement de l'utilisateur : soit ces comportements sont connus de longue date car ils font partie des « basic airmanship » , soit ils sont estimés comme reproductibles d'un avion sur l'autre car non contredits par le retour d'expérience, soit ils sont estimés par les experts pilotes. (rappelons que les analyses de sûreté sont réalisées par des experts, qui répertorient pour tous les systèmes, tous les types de pannes envisagés et, le comportement probable[5] de l'équipage pour y faire face – ce qui s'appelle donc une hypothèse de comportement). Cependant les hypothèses contenues dans ces analyses n'évoluent pas aussi vite que

[5] Basé sur le retour d'expérience en compagnie, les retours des pilotes d'essais et des instructeurs, etc.

les compétences réelles des pilotes. Or, ces analyses contiennent de très nombreux scénarii, et il n'est donc pas possible d'en faire une étude exhaustive pour garantir une mise à jour efficace. Il faut donc prioriser les scénarii à étudier (dans la suite du document nous utiliserons indifféremment les notions de scenarii ou d'hypothèses pour définir les descriptions réalisées par les experts) et se concentrer sur ceux que nous appellerons « pertinents » dans la suite de notre étude.
D'une façon générale, un scenario pertinent sera un scenario pour lequel la réponse de l'équipage en situation réelle, pourrait dévier du comportement attendu par les concepteurs. Ces hypothèses sont considérées comme pertinentes parce que ce sont celles qui, après étude en facteurs humains, pourront confirmer ou infirmer des hypothèses de conception.

Mais, la définition des scénarii pertinents a été précisée avec l'aide des experts en Facteurs Humains afin de permettre une sélection la meilleure possible. Un scenario pertinent a été défini comme suit :

1/ Il contient une hypothèse sur le comportement de l'utilisateur (face à une panne)
Exemple, perte des informations d'altitude sur le coté droit du cockpit : *En cas de perte des informations d'altitude sur un coté du cockpit, le pilote concerné utilisera les informations en provenance de l'autre coté mais aussi de l'instrument de secours.*

2/ Il contient une hypothèse pouvant être remise en cause.
Exemple, *dans certains cas, les experts font l'hypothèse que le pilote détecte que certaines informations présentes dans le cockpit sont erronées du fait de la panne d'un système. En fonction de la situation ce n'est pas toujours évident.*

3/ La classification du scenario pourrait être dégradée si le pilote ne réagit pas comme le prévoit l'hypothèse.
Nous avons vu que chaque scénario de panne est classifié sur une échelle donnée par les autorités de certification (European Aviation Safety Agency, 2011). *Cette échelle classe les scénarii de « Minor » à « Catastrophic » avec une probabilité d'occurrence associée*[6]. Or, dans certains cas, les

[6] Pour rappel, nous donnons ici les définitions de ces catégories (voir plus de détails figure 2) :
Minor (MIN) : Effets sur l'avion : Faible baisse des marges de sécurité ; Effets sur les passagers : Inconfort physique ; Effets sur l'équipage : Augmentation légère de la charge de travail. (Probabilité d'occurrence : inferieure à 10-3 par heures de vol.)
Major (MAJ) : Effets sur l'avion : Baisse significative des capacités opérationnelles ou des marges de sécurité. ; Effets sur les passagers : Douleur physique incluant possiblement de graves blessures. Effets sur l'équipage : Inconfort physique et augmentation significative de la charge de travail. (probabilité d'occurrence < 10-5 par heures de vol.)
Hazardous (HAZ) : Effets sur l'avion : Baisse très importante des capacités opérationnelles ou des marges de sécurité. Effets sur les passagers : Très graves blessures incluant possiblement des morts dans les passagers et dans les membres d'équipage en cabine. Effets sur

conséquences d'une panne seront aggravées si l'équipage ne réagit pas de la façon définie par le constructeur et approuvée par les autorités de certification. Dans ces cas là, le passage d'une classification donnée à une classification supérieure (par exemple de *minor* à *major*) pour une même panne, sera également le signe de la pertinence d'un scénario. En effet, les hypothèses sont considérées comme pertinentes parce que ce sont celles qui, après étude en facteurs humains, pourront donner lieu à des améliorations lors de la conception. Garantir la classification permet de garantir la meilleure conception.

Sélection des scénarii pertinents par un groupe d'experts : Objectifs

Cette première étape, est une étape clef de notre démarche empirique. Elle a deux objectifs principaux.

Le premier objectif est d'identifier un ensemble de scénarii témoins. Ces scénarii sont sélectionnés par jugement d'experts selon qu'ils sont perçus comme pertinents (voir critères ci-dessus) pour être évalués dans une démarche classique en facteurs humains (analyse experte via décomposition de la tâche à réaliser, expérimentations en simulateur ou en vol par exemple). Cette étape a été réalisée sur un sous ensemble de la base de données (c'est-à-dire les analyses de sûretés). Ce sous ensemble contenait les systèmes suivants : *Control and display System (CDS), Pilote Automatique (PA), Fuel System*. Cette étape est importante car elle permettra d'évaluer la performance des outils (développés lors de l'étape 3) et de trouver des pistes pour leur amélioration.

Le second objectif de cette étape 1 est d'analyser les hypothèses de comportement des pilotes contenues dans les scénarii sélectionnés : *Quels sont les caractéristiques des comportements attendus des pilotes, remis en question par les experts ?* Pour ce faire, nous allons utiliser le modèle de Reason qui formalise les étapes de la prise de décision (Reason, 1990).

Par exemple, imaginons que les experts aient sélectionné l'hypothèse de comportement suivante (rappelons qu'une hypothèse sélectionnée est une hypothèse considérée comme discutable) :

« En cas d'affichage d'une information erronée de vitesse, le pilote détectera la panne, comprendra la situation et aura comme action d'utiliser les informations de l'instrument de secours ».

Dans ce cas, en utilisant le modèle de Reason (1990), les experts préciseront l'étape de décision qu'ils considèreront comme discutable dans cette hypothèse. Par exemple, ils pourront considérer que l'étape

---

l'équipage : Douleur physique ou très grande charge de travail avec diminution des capacités de réalisation des tâches. (probabilité d'occurrence : <10-7 par heures de vol.)

Catastrophic (CAT) : Effets sur l'avion : Perte de l'avion. ; Effets sur les passagers : Nombreux morts. ; Effets sur l'équipage : Morts ou blessures très graves. (probabilité d'occurrence : inferieure à 10-9 par heures de vol.)

de détection ne sera pas réalisée par l'équipage. L'hypothèse discutable portera donc sur l'étape de détection de la panne.
L'intérêt d'identifier ces étapes est double : **1)** proposer une terminologie adaptée pour rédiger, à l'avenir, les scenarii contenus dans les bases de données et **2)** optimiser à l'avenir, les recherches automatiques de scenarii à partir de mots clés pertinents (les mots clés pourront être les différentes étapes du modèle de Reason (Reason, 1990)).

I Méthode pour la sélection des scénarii pertinents (Objectif 1) :

Afin de réaliser cette étape, pour chaque système, 2 experts en sûreté des systèmes et 1 spécialiste en facteurs humains (Il y avait 3 systèmes à étudier ce qui correspondant à **9 experts pour l'étude** complète) ont étudié chaque scénario de panne. Ces experts ont une grande connaissance du système technique étudié. Ils avaient un niveau d'étude minimum de 5 ans après le Baccalauréat (Master ou diplôme d'Ingénieur, 3 experts avaient un Doctorat) avec au moins 10 ans de pratique en sûreté ou en facteurs humains dans l'aéronautique.

1/ Les experts en sûreté des systèmes étaient les responsables de l'analyse de sûreté des systèmes étudiés (Control and Display System, Pilote Automatique, système de gestion du fuel). Ce sont eux qui ont rédigé les 2173 scénarii que nous allons étudier. Ils ont donc une grande connaissance de ces analyses et de l'historique de leur création.

2/ Les experts en facteurs humains étaient les responsables des études en facteurs humains des systèmes étudiés *(Control and Display System, Pilote Automatique, système de gestion du fuel)*. Ils possèdent donc une grande connaissance des problèmes de ce domaine (erreurs observées dans la réalité par exemple). Ils avaient aussi une grande connaissance des problèmes pouvant être liés aux interfaces entre ces systèmes et l'équipage.

Avant de débuter chaque session de sélection, un dialogue avec les experts, permettait de définir les objectifs du projet, d'analyser la complexité du problème et d'identifier des questions clés pour l'étude.

Compte tenu du grand nombre de situations décrites pour chaque système, nous avons donné des consignes assez générales. Cela permettait d'engager une réflexion avec un débat entre experts sur chaque scénario étudié. Les consignes données étaient les suivantes :

« Veuillez, s'il vous plait, sélectionner les scénarii pour lesquels :

- Vous pensez que le comportement d'équipage réel (détection, diagnostic, plan d'action, action) peut être différent de celui prévu, et nous donner les raisons de votre choix.
- Vous pensez que si le comportement prévu d'équipage n'est pas réalisé, alors les conséquences de la panne seront dégradées, nous donner les raisons de votre choix. »

II Méthode d'étude des scénarii pertinents : le modèle de Reason (Objectif 2)

Cette étape a pour objectif de caractériser les hypothèses de comportement du pilote remises en question dans les scénarii sélectionnés. Pour ce faire, nous avons demandé aux experts d'indiquer à l'aide du modèle adapté de Reason (1990), la (ou les) étape(s) du processus de prise de décision (voir Figure 26) qu'ils remettaient en cause dans les scénarii sélectionnés

Pour un équipage et selon Reason, la gestion d'une panne d'un système avionique est réalisée en 4 étapes (Reason, 1990) (voir la Figure 26). La première étape est la **détection de panne**. L'équipage doit se rendre compte qu'une défaillance s'est produite. La deuxième étape est un **diagnostic de la panne.** Celle-ci doit permettre la construction d'une représentation de la situation actuelle et de sa possible évolution. La troisième étape est nécessaire pour **construire la séquence spécifique des actions correctives**. Pour finir, la dernière étape est l'exécution de cette séquence, c'est-à-dire **l'étape d'action**. Ces étapes peuvent être suivies d'une **étape de vérification** qui n'est pas représentée sur la figure. Ces quatre étapes peuvent aussi être divisées en deux composantes globales : la compréhension de la défaillance et la correction de la défaillance(Jambon, 1998).La **compréhension de la défaillance** est composée des étapes de détection et de diagnostic du modèle de Reason. **La correction de la défaillance** est composée de la planification des actions et de l'exécution de ce plan.

Les quatre étapes du modèle de Reason ainsi que les deux étapes globales proposées par Jambon, sont représentées sur la figure ci-dessous.

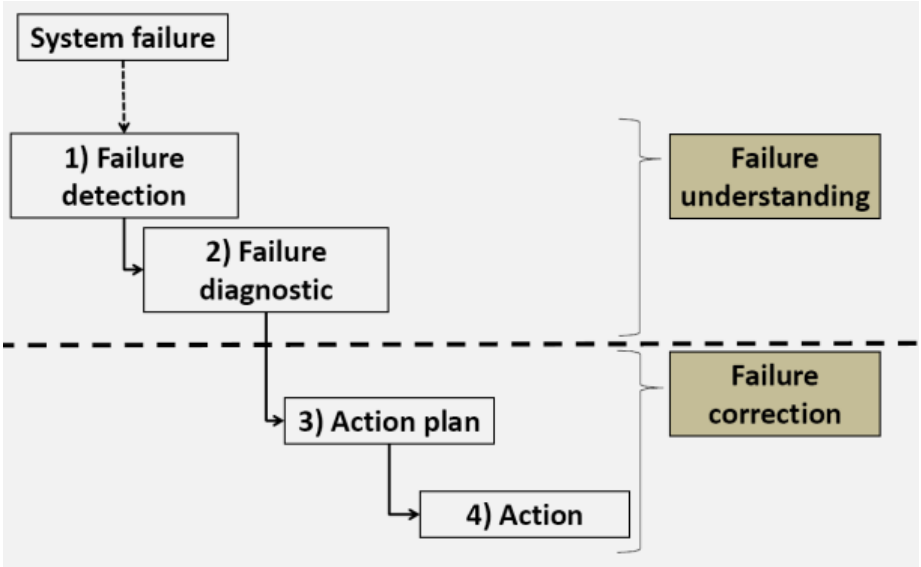


Figure 26 : Processus décisionnel, adapté de Reason (1990)

Voyons maintenant les résultats.

III Résultats :

III.1 La base de données initiale

Les scénarii des analyses de sûreté des systèmes sont le point de départ de cette étape. Afin de limiter le nombre de scénarii, les experts ont choisi les analyses de sûreté de 3 systèmes impliquant différentes sortes d'interactions hommes-machines et pouvant donc être les plus représentatifs de l'intégrabilité des systèmes :

- **Système de commandes et affichages :** CDS (*Control and Display System*) contenant la majorité des commandes et affichages du cockpit.
- **Gestion de l'énergie :** système de gestion du carburant (*Fuel System*).
- **Gestion d'un système complexe :** PA (*Pilote Automatique*)

Le nombre total de scénarii de la base initiale est donc $\mathcal{N} = 2173$. Le nombre de scénarii pour chaque analyse est respectivement $\mathcal{N}_1 = 697$ pour le CDS, $\mathcal{N}_2 = 969$ pour le PA et $\mathcal{N}_3 = 507$ pour le système de gestion du fuel. Comme nous l'avons déjà indiqué, tous ces scénarii sont classés sur une échelle de conséquences croissantes (*Minor, Major, Hazardous, Catastrophic*) donnée dans le règlement de certification. La Figure 27 présente la répartition des scénarii selon leurs classifications (*Minor, Major, Hazardous, Catastrophic*) et le système étudié (*Control and Display System, Fuel System, Pilote Automatique*).

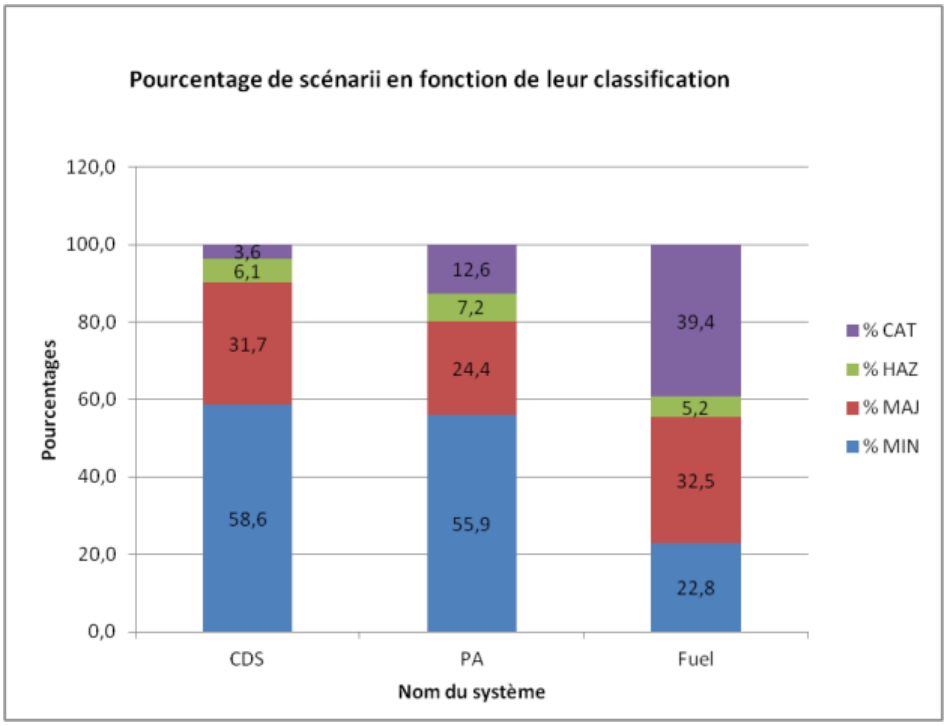


Figure 27 : Répartition des classifications de scenarii (de MIN à CAT) selon le système considéré

Cette figure présente un aperçu global de la répartition des classifications des scenarii selon le système considéré. Ainsi, nous pouvons remarquer que la répartition des classifications n'est pas homogène pour un système donné. Pour le Control and Display System et le Pilote Automatique, les scenarii *minors* sont plus nombreux que les scenarii *catastrophic*s. Ces proportions sont inversées pour le système fuel. Il existe donc aussi des différences inter systèmes. Dans le système fuel les scenarii *catastrophic*s sont beaucoup plus nombreux que dans les 2 autres systèmes.

Voyons maintenant les résultats de la sélection des scénarii pertinents au sein de ces 3 analyses.

III.2 Sélection de scénarii par des experts

Les résultats de la sélection de scénarii sont résumés sur la Figure 28. Parmi les 697 scénarii du système CDS, 84 scénarii (12,1%) ont été sélectionnés ; parmi les 969 scénarii de l'analyse du Pilote Automatique, 174 (17,9%) des scénarii ont été sélectionnés ; parmi les 507 scénarii du système de gestion du carburant, aucun n'a été sélectionné.

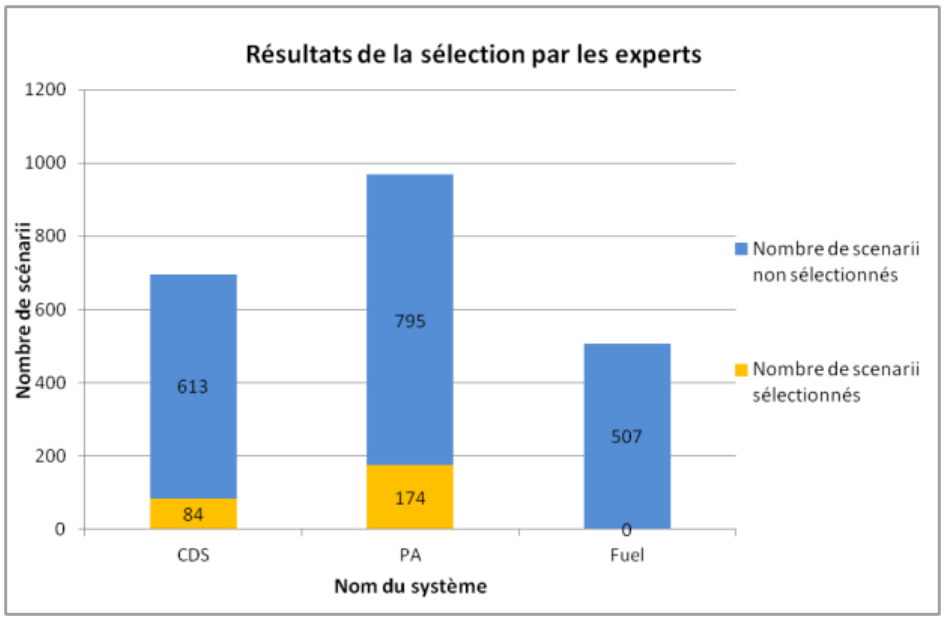


Figure 28 : Résultats de la sélection des scénarii

Il existe une différence significative ($Khi^2 = 31,39$; $p = 2.1206 \times 10^{-8}$) entre les résultats de la sélection des scénarii du *Control and Display System* et du *Pilote Automatique* (voir la Figure 28). Les experts ont donc sélectionné plus de scénarii pour le *Pilote Automatique* que pour le CDS. Ce résultat peut s'interpréter au regard de la nature de ces 2 systèmes. En effet, la pilote automatique doit être en mesure de gérer un très grand nombre de situations dans des contextes opérationnels différents, tels qu'une météo défavorable, la proximité d'un relief, etc. (beaucoup plus que le CDS). De ce fait, les hypothèses peuvent se révéler vraies pour un contexte donné, mais discutables pour un autre contexte. Le nombre d'hypothèses qui peuvent se révéler discutables est donc forcément plus important.

Concernant le fait qu'il n'y ait aucun scénario sélectionné pour le système fuel (gestion du carburant), la Figure 27 peut nous aider à interpréter ce résultat. En effet, comme nous l'avons vu dans la Figure 27, le système fuel contient significativement plus de scénarii *Catastrophic (CAT)* et moins de scénarii *Minor (MIN)* que les 2 autres systèmes. Or comme nous l'avons vu, un des critères de sélection des scenarii pertinents repose sur la possibilité de passage d'une classification donnée à une classification supérieure pour une même panne. Or, les scenarii catastrophiques n'ont aucune possibilité de voir leur classification dégradée (en effet il n'y a pas de classification supérieure à *catastrophic*). Ce fait peut constituer un des éléments d'explication à l'absence de scenarii sélectionnés dans le système fuel.

Pour le reste de l'étude, nous nous concentrerons sur les scénarii sélectionnés du *Control and Display System et du Pilote Automatique*. Etudions maintenant les scénarii sélectionnés à l'aide du modèle de prise de décision proposé par Reason (1990).

III.3 Résultats de l'étude des scénarii grâce au processus de prise de décision :

Dans cette partie, nous allons indiquer, pour chaque scénario sélectionné, quelle étape du processus de prise de décision est remise en cause par les experts. Pour un scénario particulier, plusieurs étapes pouvaient être remises en question. C'est pourquoi, pour 84 scénarii sélectionnés pour le *Control and Display System*, nous avons $\mathcal{N}_1$ = 109 étapes de décision remises en cause (c'est-à-dire que pour au maximum 25 scénarii, plusieurs étapes ont été remises en cause). Avec le même raisonnement, pour le système de Pilotage Automatique, sur les 174 scénarii sélectionnés, $\mathcal{N}_2$ = 327 étapes de décision sont remises en cause.

Les résultats que nous avons obtenus ont été analysés à l'aide des deux étapes globales identifiées par Jambon (Compréhension de la panne, Correction de la panne ; (Jambon, 1998)) et des 4 étapes « de base » du modèle de Reason (Détection, Diagnostic, Planification, Action).
Pour les tests statistiques, $\mathcal{H}_0$ sera l'égalité des distributions entre chaque étape du processus décisionnel.

III.3.1 Etude des scénarii choisis selon les étapes globales de compréhension et de correction de la défaillance

Lorsque nous classons les scénarii sélectionnés en fonction des deux étapes globales du modèle de prise de décision proposé par Jambon (1998), nous obtenons les résultats présentés sur les Figure 29 et Figure 30.

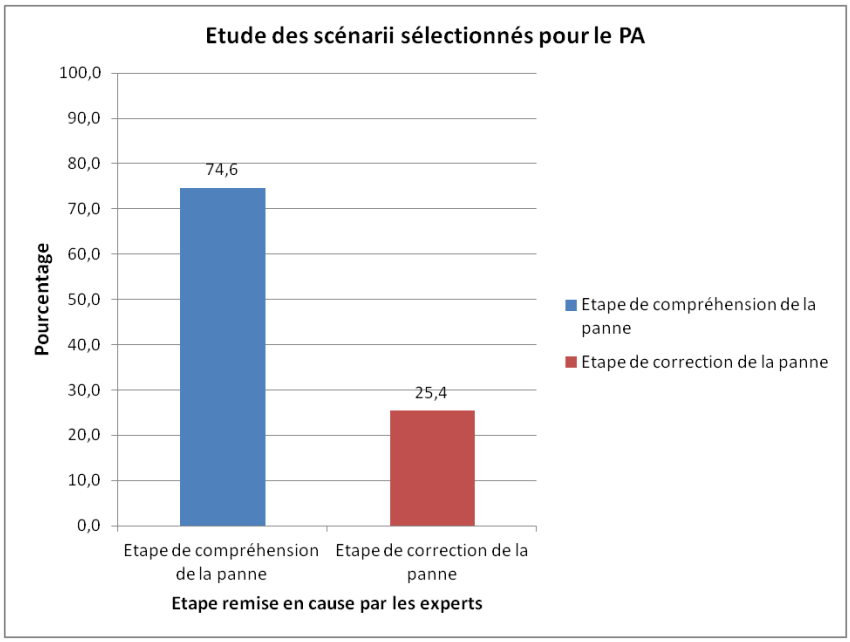


Figure 29 : Résultats obtenus en étudiant les scénarii sélectionnés du PA à l'aide des deux composantes identifiées par Jambon, 1998.

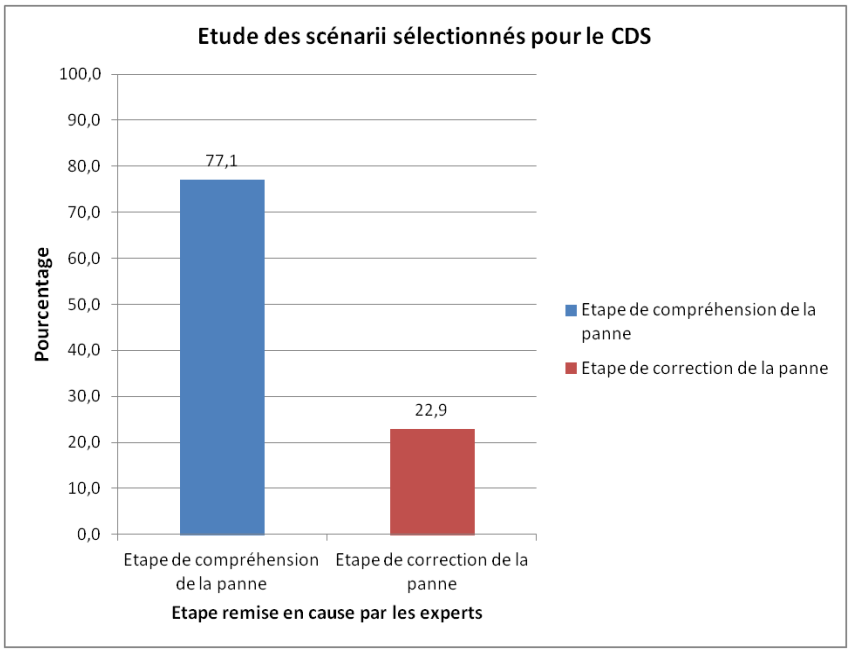


Figure 30 : Résultats obtenus en étudiant les scénarii sélectionnés du CDS à l'aide des deux composantes identifiées par Jambon, 1998.

Ces résultats montrent des différences significatives, entre la phase de compréhension de défaillance et la phase de correction de défaillance. Les valeurs de Khi² sont respectivement Khi² = 31,9 ($p = 1{,}6316\times10^{-8}$) pour le *CDS* et Khi² = 79,2 ($p < 10^{-10}$) pour le système de *Pilotage Automatique*.
Pour ces 2 systèmes, les étapes globales de compréhension de la panne sont plus remises en question par les experts que les étapes de correction de la panne.

III.3.2 Étude des scénarii choisis selon les 4 étapes du processus de prise de décision

III.3.2.1 Étude des scénarii sélectionnés pour le Pilote Automatique

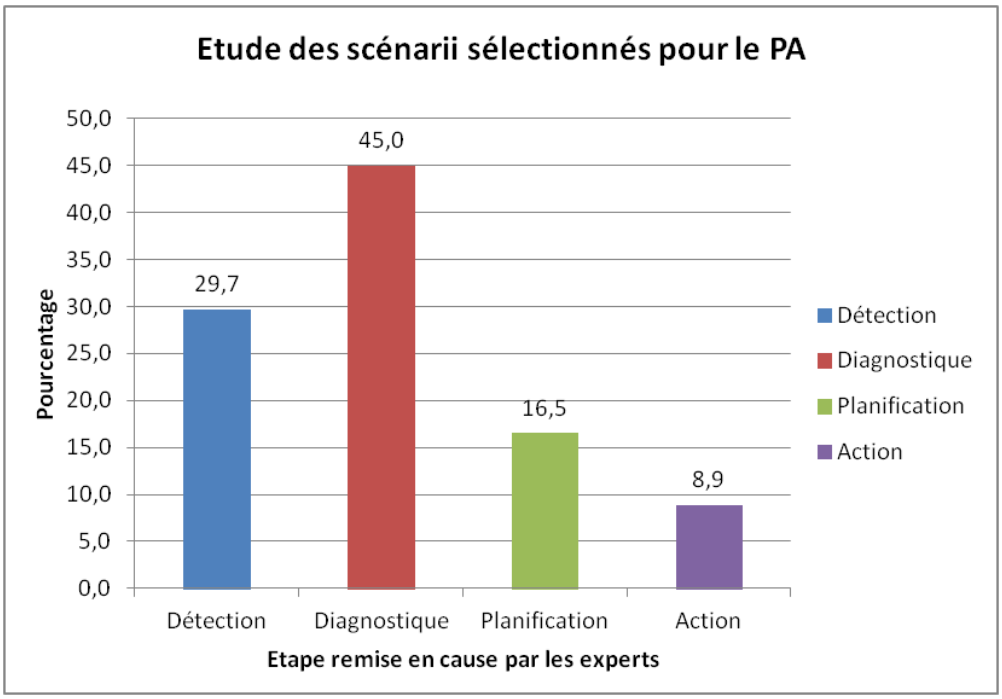


Figure 31 : Résultats obtenus en étudiant les scénarii sélectionnés du PA avec le processus de prise de décision (Etapes locales)

Le tableau ci-dessous (Tableau 7) récapitule les valeurs de Khi² pour la comparaison croisée des résultats présentés sur la Figure 31. Les valeurs surlignées sont des valeurs significatives pour $\alpha = 0{,}05$. Pour information, la valeur du *p* est rappelée dans la table.

Flight management system	Détection	Diagnostic	Planification	Action
Détection				
Diagnostic	10,25 ($p = 0.0013$)			
Planification	12,25 ($p = 0.0004$)	43,03 ($p < 10^{-10}$)		
Action	36,70 ($p < 10^{-10}$)	79,11 ($p < 10^{-10}$)	7,53 ($p = 0.0060$)	

Tableau 7 : Tests du Khi² pour le Pilote Automatique

Pour le pilote automatique, les experts ont plus remis en cause les étapes de Diagnostic, que de Détection, que de Planification et, enfin d'Action.

III.3.2.2 Etude des scénarii sélectionnés pour le *Control and Display System*

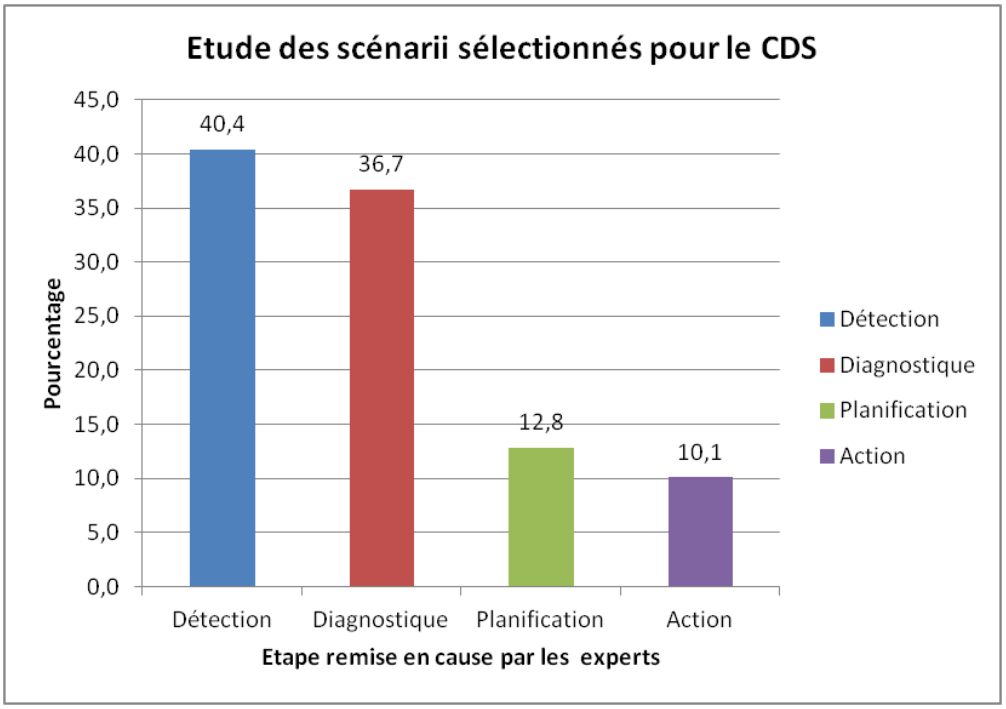


Figure 32 : Résultats obtenus en étudiant les scénarii sélectionnés du CDS avec le processus de prise de décision (Etapes locales)

Le tableau ci-dessous (Tableau 8) récapitule les valeurs de Khi² pour la comparaison croisée des résultats présentés sur la Figure 32. Les valeurs surlignées sont des valeurs significatives pour $\alpha = 0,05$. Pour information, la valeur du *p* est rappelée dans la table.

Control and display system	Détection	Diagnostic	Planification	Action
Détection				
Diagnostic	0,19 (p = 0.6629)			
Planification	27,80 (p = $1,3438\times10^{-7}$)	12,52 (p = 0.0004)		
Action	19,80 (p < 0,00001)	22,95 (p < 0,00001)	0,36 (p = 0.5485)	

Tableau 8 : Tests du Khi² pour le Control and Display System

Pour le *Control and Display System*, les experts ont plus remis en cause les étapes de Détection et de Diagnostic, que de Planification et d'Action. A la différence du Pilote Automatique, il n'existe pas de différences significatives entre les étapes de Détection et de Diagnostic ainsi qu'entre les étapes de Planification et d'Action.

IV Interprétation des résultats

En présentant les résultats de la sélection des scénarii par les experts nous avons remarqué et interprété les différences de sélection inter-systèmes par des différences dans le contenu des analyses de sûreté de ces systèmes. En effet, nous avons montré que le système fuel possède une proportion importante (plus que pour le CDS et le PA) de scénarii catastrophiques. Les scénarii de classification catastrophique sont les plus critiques. Dans un scénario *CAT*, une hypothèse de comportement même erronée de l'équipage ne pourra pas entraîner des conséquences sur sa classification (en effet un scenario *catastrophic* ne pourra être dégradé). Or, nous avons vu que la dégradation d'un scenario constitue un critère d'identification d'un scenario pertinent. Il est donc impossible de considérer un scénario catastrophique comme pouvant être pertinent. De plus, dans ce système fuel, il y a également moins de scénarii *minor*. Or, dans ces scénarii, le pilote peut potentiellement être en charge d'actions correctrices de la panne. Malgré le soin apporté dans la définition des hypothèses de comportement, cet ensemble de scénarii (*minor*) est donc susceptible de contenir des scénarii pertinents. La grande proportion de scenarii *catastrophic* et la faible proportion de scenarii *minor* dans le système fuel explique en partie le fait qu'aucun scénario n'ait été sélectionné dans ce système.

Les résultats de l'étude des scénarii présentés au paragraphe III.3 indiquent que les experts remettent en cause plutôt des scénarii contenant une hypothèse de compréhension de la panne qu'une hypothèse de correction de la panne. Il y a deux façons d'expliquer ce fait. **(1) La première** est plutôt liée au **comportement de l'équipage.** Lorsque le constructeur émet une hypothèse de comportement, il s'appuie sur le retour d'expérience et les compétences établies des pilotes (le *Basic Airmanship*[7]). Cependant, pour établir les hypothèses de détection et de compréhension de la panne du système par l'équipage, il est fréquent que plusieurs interprétations expertes (en lien avec des observations, des retours d'expériences, des apports des pilotes…) se confrontent. Néanmoins, au final, une seule hypothèse doit rester. Ces hypothèses, obtenues par le meilleur compromis possible, sont donc plus souvent sources de discussions et de remises en question. Ce fait est moins présent pour les hypothèses de planification et d'action de l'équipage puisque, généralement, celles-ci découlent de la détection et du diagnostic de la panne. Si les hypothèses de détection et de diagnostic sont établies la mise en œuvre des séquences d'actions correctrices est le plus souvent établie. **(2) La seconde** est en lien avec la première mais cette fois ci plutôt liée aux **systèmes avioniques**. En effet, en cas de défaillance, la complexité générale des systèmes avioniques rend difficile la mise en évidence des paramètres qui seront pris réellement en considération pour la détection et le diagnostic de la défaillance, cela

[7] Compétences de bases des pilotes, voir précisions dans le Chapitre 2 en annexe.

dépendra du contexte opérationnel, de l'état des systèmes avions, de la météo, de l'équipage… La création de ces hypothèses dépend donc beaucoup de la sensibilité et, des connaissances des experts (pilotes, spécialistes des systèmes, spécialistes en sûreté) qui les réalisent. C'est pourquoi elles sont plus faciles à remettre en question.

De plus pour le pilote automatique, et contrairement au CDS, les résultats montrent des différences significatives entre les étapes de Détection et de Diagnostic et les étapes de Planification et d'Action (voir Figure 31).
Pour le PA, au sein de l'étape globale de **compréhension de la défaillanc**e, les experts remettent plus en doute l'étape de diagnostic que l'étape de détection. Au sein de l'étape globale de **correction de la défaillance**, les experts remettent plus en doute l'étape de planification que l'étape de correction.

Au sein de ces étapes globales, ces différences locales peuvent être expliquées par la complexité de ce système. En effet, les étapes de diagnostic et de planification de l'action exigent une bonne conscience de la situation. La construction de la conscience de la situation peut être difficile pour un système extrêmement complexe (pilote automatique) conçu pour couvrir une grande diversité de situations normales et anormales.

Ces résultats pourraient être mis en relation avec ceux de Jones et al. (1996) sur la conscience de la situation. A partir d'une analyse de 143 accidents de l'aviation civile américaine relevant d'une erreur humaine, Jones et al. montrent que les erreurs de niveau 1 de *situation awareness* (défaillance de perception et d'intégration de l'information) peuvent être invoquées dans 76% des accidents, celles de niveau 2 de *situation awareness* (interprétation de la situation) dans 20% des cas, et celle de niveau 3 (anticipation) dans les 4% restants. Pour les erreurs de perception, les auteurs distinguent les cas (13% du total des accidents) pour lesquels l'origine du problème est plutôt à rechercher du côté de la **technologie (ex : information non disponible, mal représentée sur les afficheurs) ou des conditions de vols particulièrement difficiles (11% des accidents), de celles pour lesquelles l'information n'a pas été prise en compte par l'opérateur humain (ex : information non détectée, par omission ou focalisation de l'attention sur autre chose dans 35% des cas ; information négligée car non conforme aux attentes du pilote dans 9% des cas ; information oubliée compte tenu de la charge de travail du pilote dans 8% des** cas). Concernant les erreurs de niveau 2 de la *situation awareness*, elles résultent, d'après les auteurs, d'un modèle mental inapproprié, que celui-ci s'avère d'emblée inadéquat pour représenter la réalité de la situation (ex : erreur de diagnostic et sélection d'un mauvais modèle mental en mémoire), ou qu'il ait conduit à des inférences erronées (ex : en se basant sur les implicites [valeurs par défaut] du modèle). Enfin, pour ce qui est des erreurs

d'anticipation (correspondant au niveau 3 de la SA), elles seraient impliquées dans moins de 4 % des accidents.

Les résultats de Jones et al. (1996) montrent donc que lors de l'étude d'accidents relevant d'une erreur humaine, la perception, l'intégration et l'interprétation de la situation totalisent 96% des accidents. Dans ces cas là, les hypothèses de comportement de l'équipage sur la détection et le diagnostic de la situation sont donc remises en cause, ce qui est cohérent avec notre étude. La remise en cause des comportements attendus de détection et de diagnostic (qui représentent 74,7 % des remises en causes pour le pilote automatique et 77,1% pour le *Control and Display System*) que nous observons dans notre étude est donc retrouvée dans les données de la littérature issues de l'analyse d'accidents. Leur prise en compte en conception pourrait donc jouer un rôle crucial dans l'amélioration de la sûreté du transport aérien.

V Conclusion

En résumé, voici les résultats que nous avons obtenus :

- Pour le fuel, nous avons vu que la répartition des classifications ainsi que la façon dont a été réalisée cette analyse joue certainement un rôle dans le fait qu'aucun scénario n'ai été sélectionné.
- Pour le pilote automatique et le *Control and Display System*, les hypothèses portant sur la compréhension de la panne sont plus représentées que les hypothèses portant sur les étapes de correction dans les scénarii sélectionnés.
- Pour le pilote automatique, lorsque nous étudions au sein de ces étapes globales, les étapes locales les constituant, l'étape de diagnostic est plus représentée que l'étape de détection (pour l'étape globale de compréhension) et l'étape de planification est plus représentée que l'étape d'action (pour l'étape globale de correction).

Bien que des limites soient à prendre en compte dans cette étape de notre étude (seulement 3 analyses de sûreté, utilisation du jugement d'experts…), la concordance de nos résultats avec des études précédentes sur l'analyse d'accident d'avion, conforte certaines de nos observations. L'utilisation du modèle de processus décisionnel adapté à notre étude a permis de préciser les étapes impliquées dans la compréhension et la correction de la panne et a reçu un bon retour des experts. Selon eux, elle permet de disposer d'un bon niveau de formalisme afin de poser la réflexion nécessaire à l'évaluation d'une hypothèse de comportement de l'équipage.

De plus, cette étape a atteint ces deux principaux objectifs :

Tout d'abord, nous possédons maintenant un ensemble de 258 scénarii témoins pour la suite de la méthode (84 pour le CDS et 174 pour le PA).

Enfin, nous avons pu voir que les étapes de détection et de diagnostic de la panne sont souvent remises en cause dans les scénarii sélectionnés. Ces critères, extraits du processus de prise de décision (Reason, 1990), ne sont pas mis en évidence dans les analyses de sûreté aujourd'hui. Notre étude souligne donc la pertinence de les prendre en compte dans les analyses futures. Ces descripteurs de l'hypothèse de comportement de l'équipage pourraient être grandement utiles pour la sélection et la récupération des scénarii.

Passons maintenant à l'étape 2 de cette démarche expérimentale.

CHAPITRE. 2 DES DESCRIPTEURS POUR DES REQUETES PERTINENTES

Dans le chapitre précédent, nous avons identifié et caractérisé à l'aide d'experts des scénarii qui pourraient se révéler pertinents pour une étude en Facteurs Humains. Le modèle de Reason nous a permis d'identifier des descripteurs (étapes de la prise de décision) qui pourront se révéler utile à l'avenir pour rédiger (et donc faciliter leur identification par requêtes) des scénarii pertinents. L'objectif était d'inclure ces critères dans la rédaction des scénarii dans le futur. Dans ce présent chapitre, nous avons focalisé notre attention sur la base de données elle-même et sur les informations (descripteurs) qu'elle contient.

L'objectif de ce chapitre est double :
1/Identifier (dans la base de données existante) les mots clefs (descripteurs) qui pourraient permettre de sélectionner des scénarii pertinents (c'est-à-dire contenant des hypothèses de comportement pouvant dévier du comportement réel de l'équipage)
2/Vérifier si ces mots clefs (descripteurs) sont effectivement plus présent dans les scénarii sélectionnés par les experts que dans l'ensemble de la base de données.
Reprenons l'exemple vu précédemment :
Exemple de titre : Perte non détectée des informations de cap, au décollage ou à l'atterrissage : Minor.
Nous avons les descripteurs suivant : Type de panne : Perte ; *Détection par le système :* Non Détecté ; *Phase de vol :* Décollage ou Atterrissage ; *Classification :* Minor.

En organisant l'information (avec un tableur ou un outil de base de données) nous pouvons assez facilement réaliser des requêtes. Par exemple, sélectionner pour un système, tous les scénarii se passant aux phases de vol : *décollage ou atterrissage* ; pour lesquels le type de panne est : *Perte* ; avec une classification : *Minor.* Pour cette étude, nous avons dans un premier temps utilisé ce formalisme pour décrire l'ensemble des scénarii contenus dans la base de données, soit 2173 rédigés par les experts et que nous avons synthétisé de la façon décrite si dessous.

Sur l'exemple précédent, cela donne :

Type de panne	*Détection par le système*	*Phase de vol*	*Classification*
Perte	Non détecté	Décollage ou atterrissage	Minor

Avec ce formalisme, il est possible de sélectionner toutes les lignes (donc les scénarii) pour lesquelles le type de panne est *Perte* par exemple.

Comme nous l'avons indiqué, notre premier objectif est d'identifier les descripteurs présents dans la base de données et qui auraient pu permettre d'identifier des scénarii pertinents.

Pour cela, les hypothèses que nous avons émises sont les suivantes :
1/ Les descripteurs qui voient leur présence augmenter lorsque la classification s'aggrave sont certainement présents dans les scénarii pertinents.
Nous avons formulé cette hypothèse car, un scénario *minor,* par exemple, qui possède des descripteurs plus souvent associés à des scénarii *catastrophics,* voit, peut-être, sa classification conservée du fait d'une action correctrice de l'équipage. Donc si cette action n'est pas réalisée, il peut se produire une aggravation de la classification.

2/ Les scénarii pertinents ont plutôt une classification faible (Sans perte humaine : *Minor, Major*) car l'augmentation de la classification de ces scénarii est plus probable que pour les scénarii de classifications plus graves (Avec pertes humaines potentielles : Hazardous, Catastrophic).

I Matériel et méthode

Pour cette étape, le matériel que nous avons utilisé est composé des 2 analyses de sûreté des systèmes choisies par les experts et pour lesquelles des scénarii ont été sélectionnés :

- **Commandes et affichages :** CDS (*Control and Display System*) contenant la majorité des commandes et affichages du cockpit.
- **Gestion d'un système complexe :** PA (*Pilote Automatique*)

Le nombre total de scénarii est $\mathcal{N}$ = 1666. Le nombre de scénarii pour chaque analyse est respectivement $\mathcal{N}_1$ = 697 pour les commandes et les systèmes d'affichage, $\mathcal{N}_2$ = 969 pour la gestion de vol (Pilote Automatique).

Tous ces scénarii possèdent des descripteurs permettant d'identifier les caractéristiques de la panne et d'en comprendre le contexte. Ces descripteurs sont contenus dans une phrase de description

synthétique de la panne (phrase accompagnant systématiquement la description libre et détaillée réalisée par l'expert). Il existe **4 grandes familles de descripteurs**, chaque famille pouvant prendre un nombre limité de valeurs.

1. **La phase de vol :** a/toute phase de vol ; b/Décollage et Atterrissage ; c/Autre[8] ;
2. **Le type de panne :** a/Fonction erronée ; b/Fonction perdue ; /c Activation accidentelle de la fonction d/Autre ou Spécifique à l'automatisme[9] ;
3. **Détection par le système :** a/Panne détectée ; b/Panne non détectée ; c/Non précisée
4. **La classification :** a/Minor (MIN) ; b/Major (MAJ) ; c/Hazardous (HAZ) ; d/Catastrophic (CAT)

II Résultats

II.1 Résultats 1 : Les descripteurs qui augmentent avec la classification

Rappelons que pour chaque système, nous avons observé la répartition des descripteurs en fonction de la classification (*Minor, Major, Hazardous, Catastrophic*). Nous évaluerons alors, si ces descripteurs se distribuent de façon homogène en fonction de la classification. Les descripteurs dont la proportion augmente lorsque la gravité de la classification augmente pourront alors être considérés comme nécessaires pour isoler les scénarii pertinents.

Le tableau ci-dessous (Tableau 9) résume les résultats obtenus pour les 2 systèmes (*CDS et PA*) et pour les 3 grandes familles de descripteurs (Phase de vol, type de panne, détection de la panne). Pour chaque descripteur, nous avons indiqué (par un *plus*, un *moins* ou un *égal*) si celui-ci augmentait diminuait ou restait stable lorsque la classification se dégradait.

Les analyses statistiques ont été réalisées à l'aide de tests de Khi² d'ajustement. Pour chaque descripteur, nous avons comparé la répartition des scénarii selon la classification (de *minor* à *catastrophic*) à une répartition aléatoire c'est-à-dire à une équi-répartition. Le Khi² a permis de calculer l'écart entre la répartition observée et la répartition théorique. Un écart significatif indique que la répartition observée n'est pas due au hasard et qu'elle est donc significative.

[8] Autre : Hors décollage & Atterrissage ; Croisière ; Taxi ; Au sol ; En vol (En raison du faible nombre d'occurrences de ces valeurs nous avons choisi de les rassembler pour des raisons de clarté dans les résultats)

[9] Des cas de panne très particuliers ne trouvant pas d'échos dans les catégories présentées ici, ont été rassemblés dans la dernière catégorie : Autre ou Spécifique à l'automatisme (pour des raisons de clarté des résultats).

	Phase de vol			Type de panne			
Système étudié	Toutes phases de vol	Décollage & Atterrissage	Autre	Fonction erronée	Fonction perdue	Activation accidentelle de la fonction	Autre ou Spécifique à l'automatisme
CDS	-	+	=	+	-	=	-
PA	-	=	+	+	-	=	+
	Détection de la panne						
Système étudié	Panne détectée	Panne non détectée	non précisée				
CDS	=	+	-				
PA	=	+	-				

Tableau 9 : Evolution des descripteurs en fonction de la classification

Prenons par exemple, le descripteur type de panne : « fonction erronée ». Le graphique ci-dessous (Figure 33) représente, entre autre, les proportions de scénarii contenant le descripteur « fonction erronée » selon le type de classification. Nous constatons alors que la proportion augmente de la classification *minor* à la classification *catastrophic*. Cette répartition observée est comparée à une équi-répartition à l'aide du test de khi 2 d'ajustement. Dans ce cas , le Khi² est égal à 112,1 et est significatif à un seuil supérieur à 0,00001.

Cette procédure à été appliquée à l'ensemble des descripteurs et pour les 2 systèmes. Les plus et les moins indiqués sont ceux pour lesquelles des différences significatives ont été observées.

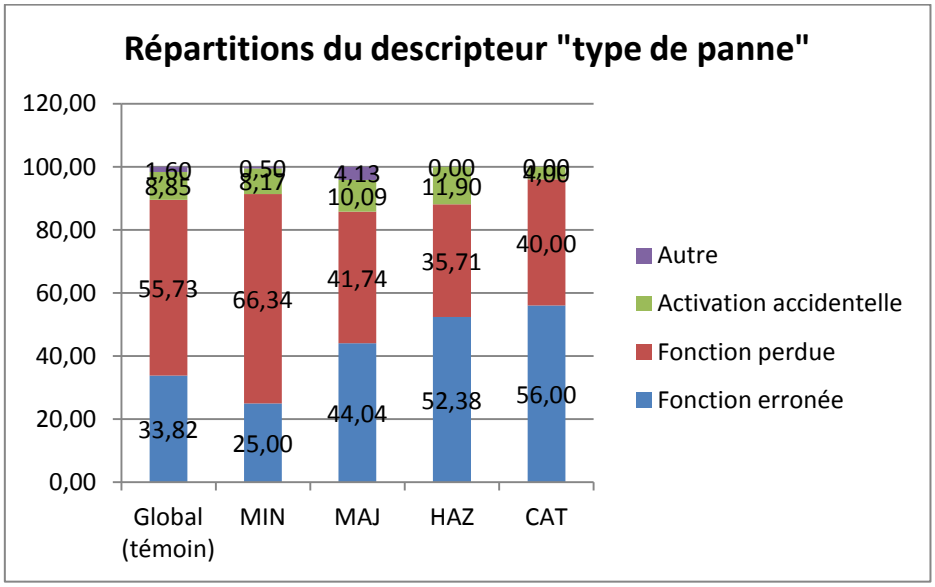


Figure 33 : Répartition du descripteur « Type de panne » (CDS)

Comme nous pouvons le voir dans le Tableau 9 les descripteurs qui augmentent significativement de la classification *minor* à la classification *catastrophic* pour le système CDS sont les suivants :

- **Phase de vol :** Décollage & Atterrissage
- **Type de panne :** Fonction Erronée
- **Détection de la panne :** Panne non détectée

De la même façon, les descripteurs qui augmentent significativement de la classification *minor* à la classification *catastrosphic* pour le PA sont :

- **Phase de vol :** Autre
- **Type de panne :** Fonction Erronée
- **Détection de la panne :** Panne non détectée

Ces résultats seront discutés après la présentation des données ci-dessous, comparant la présence ou non de ces descripteurs dans les scénarii sélectionnés par les experts

II.2 Comparaison des proportions de descripteurs entre les scénarii sélectionnés et de la base de données initiale

Rappelons que dans le chapitre précédent, nous avons identifié, sélectionné et caractérisé à l'aide d'experts des scénarii qui pourraient se révéler pertinents pour une étude en Facteurs Humains (c'est-à-dire qu'ils contenaient des hypothèses de comportement discutables). Dans ce présent chapitre, nous

avons focalisé notre attention sur la base de données elle-même et sur les informations (descripteurs) qu'elle contient.

L'objectif de ce chapitre était double :

1. Identifier (dans la base de données existante) les mots clefs (descripteurs) qui pourraient permettre de sélectionner des scénarii pertinents (c'est-à-dire contenant des hypothèses de comportement discutables)
2. Vérifier si ces mots clefs (descripteurs) sont effectivement plus présents dans les scénarii sélectionnés par les experts que l'ensemble de la base de données.

Le premier point a été réalisé et a permis d'identifier des descripteurs pertinents. Nous allons maintenant présenter les résultats relatifs au second point.

Pour obtenir les résultats relatifs à ce second point, nous avons donc comparé la proportion des descripteurs dans la base de données initiale et dans la sélection des experts. Cette comparaison a porté sur l'ensemble des descripteurs (phase de vol, type de panne, détection de la panne). L'objectif est de vérifier que les descripteurs identifiés comme pertinents dans le paragraphe II.1 de ce chapitre sont plus présents dans les scénarii sélectionnés que dans la base de données initiale sans les scénarii pertinents.

Autrement dit, nous comparons deux méthodes différentes d'indentifications de descripteurs pertinents : 1/la première méthode présentée dans le chapitre 1 a consisté à demander à des experts de sélectionner des scénarii qu'ils jugeaient pertinents. 2/La seconde méthode présentée dans le paragraphe II.1 de ce chapitre a consisté à sélectionner de façon automatique des descripteurs contenu dans la base de données initiale en nous appuyant sur l'hypothèse proposée par les experts : « Les descripteurs pertinents sont ceux dont la proportion augmente lorsque la classification se dégrade ».

Le travail que nous réalisons maintenant consiste à vérifier que les descripteurs identifiés par la seconde méthode sont plus présents dans les scénarii identifiés à l'aide de la première méthode (jugement d'experts) que dans la base de données initiale sans les scénarii sélectionnés.

Le tableau ci-dessous (Tableau 10) résume les résultats obtenus pour les deux systèmes étudiés (*CDS et PA*) et pour les trois grandes familles de descripteurs (Phase de vol, type de panne, détection de la panne). Pour chaque descripteur, nous avons indiqué (par un *plus*, un *moins* ou un *égal*) si la proportion du descripteur considéré était significativement plus importante (ou moins) dans les scénarii sélectionnés que dans la base de données initiale. Le test statistique utilisé est un test de comparaison de deux pourcentages. En effet, pour chaque indicateur nous pourrons comparer le pourcentage de descripteurs issus de la sélection et de la base de données initiale.

	Phase de vol			Type de panne			
Système étudié	Toutes phases de vol	Décollage & Atterrissage	Autre	Fonction erronée	Fonction perdue	Activation accidentelle de la fonction	Autre ou Spécifique à l'automatisme
CDS	-	+	=	+	-	+	=
PA	=	+	=	+	-	=	=
	Détection de la panne						
Système étudié	Panne détectée	Panne non détectée	non précisée				
CDS	=	+	-				
PA	+	=	=				

Tableau 10 : Evolution des descripteurs entre les scénarii sélectionnés par les experts et la base de données initiale.

Comme nous pouvons le voir dans le Tableau 10 les descripteurs qui sont significativement plus présents dans les scénarii sélectionnés sont (pour le CDS),

- **Phase de vol :** Décollage & atterrissage
- **Type de panne :** Fonction erronée et activation accidentelle de la fonction
- **Détection de la panne :** Panne non détectée

De la même façon les descripteurs significativement plus présents dans les scénarii sélectionnés pour le PA sont

- **Phase de vol :** Décollage & atterrissage
- **Type de panne :** Fonction erronée
- **Détection de la panne :** Panne détectée

III Interprétation des résultats

Afin de comparer les résultats obtenus à l'aide des deux méthodes que nous avons utilisées, nous présentons dans les tableaux ci-dessous (Tableau 11 pour le CDS et Tableau 12 pour le PA) les descripteurs pertinents identifiés par chacune d'entre-elles.

1/Pour le CDS nous obtenons les résultats suivants :

		Phase de vol	**Type de panne**		**Détection de la panne**
		Décollage & Atterrissage	Fonction erronée	Activation accidentelle de la fonction	Panne non détectée
CDS	**Méthode 1 (Hypothèse)**	X	X	Ø	X
	Méthode 2 (Jugement d'experts)	X	X	X	X

Tableau 11 : Descripteurs pertinents obtenus en fonction de la méthode utilisée pour le CDS

Pour information, le signe X signifie que le descripteur a été identifié par la méthode considérée. En revanche, le signe Ø signifie que le descripteur n'a pas été identifié par la méthode considérée.

Dans ce tableau, nous pouvons remarquer que les descripteurs sélectionnés à l'aide de la méthode 1 ont également été sélectionnés à l'aide de la méthode 2. Les descripteurs dont la proportion augmente avec la classification (critères de la méthode 1) ont donc été privilégiés lors de la sélection par les experts (méthode 2). Les deux méthodes identifient les descripteurs suivants :

- Décollage et Atterrissage
- Fonction erronée
- Panne non détectée

Une première conclusion est que le critère que nous avons utilisé pour identifier à l'aide de requêtes des hypothèse discutables est pertinent.

Nous pouvons cependant remarquer que les experts ont sélectionné plus de scénarii possédant le descripteur « Activation intempestive de la fonction. Les experts jugent donc ce type de panne pertinent. Mais après analyse, nous avons constaté que les conséquences de cette situation sont souvent modérées (MIN, MAJ) et, c'est la raison pour laquelle nous n'identifions pas ce descripteur à l'aide de la première méthode qui tient compte, pour rappel, des descripteurs plus présents lorsque la classification se dégrade (jusqu'à *Catastrophic*).

2/Pour le PA nous obtenons les résultats suivants :

		Phase de vol		Type de panne		Détection de la panne	
		Décollage & Atterrissage	Autre	Fonction erronée	Autre ou Spécifique à l'automatisme	Panne détectée	Panne non détectée
PA	**Méthode 1 (Hypothèse)**	Ø	X	X	X	Ø	X
	Méthode 2 (Jugement d'experts)	X	Ø	X	Ø	X	Ø

Tableau 12 : Descripteurs pertinents obtenus en fonction de la méthode utilisée pour le PA

Comme précédemment, le signe X signifie que le descripteur a été identifié par la méthode considérée. En revanche, le signe Ø signifie que le descripteur n'a pas été identifié par la méthode considérée.

Dans le cas du PA, à part pour le type de panne : « Fonction erronée », nous n'obtenons pas les mêmes résultats entre les deux méthodes. Ces différences peuvent s'expliquer par la nature complexe du pilote automatique. Rappelons que notre méthode 1 part de l'observation que des pannes très similaires peuvent avoir des conséquences de gravité différentes. Dés lors, lorsqu'une panne donnée est associée à une classification *minor* dans un scénario donné et qu'il existe une panne très similaire associée à une classification catastrophique dans un autre, nous pouvons faire l'hypothèse que la gestion de la panne par l'équipage n'a pas été identique pour les deux scénarii. Mais cette hypothèse nécessite que les pannes en question soient identiques ou extrêmement proches. Or, dans le cas d'un système aussi complexe que le PA, une étude a postériori avec les experts (pour tenter de comprendre les résultats) a montré de grandes disparités possibles entre deux pannes caractérisées par des descripteurs identiques. Cependant, l'analyse des scénarii sélectionnés par les experts a permis d'identifier des descripteurs (« Décollage et atterrissage », « Fonction erronée », « Panne Détectée »). Même si nous n'avons pas été en mesure de prédire les descripteurs pertinents dans le cas de ce système, en revanche les descripteurs identifiés par les experts permettront à l'avenir de réaliser des requêtes pour sélectionner des scénarii pertinents du système PA.

Concernant les deux systèmes, les descripteurs sélectionnés par les experts trouvent des échos dans la littérature sur l'étude des incidents/accidents d'avion. Tout d'abord, concernant les phases de vol, les experts ont sélectionné en priorité des scénarii se passant au décollage et à l'atterrissage. D'après une étude sur les incidents et accidents menée par Boeing (Boeing Company, 2010), ces phases de vols, qui ne totalisent qu'en moyenne 6% du temps de vol effectif, contiennent 53% des incidents/accidents

graves. Durant ces courtes phases de vol, les pilotes doivent gérer la trajectoire, la vitesse, l'altitude de l'avion tout en maintenant une étroite communication avec le contrôle aérien et les membres d'équipage en cabine. La charge de travail est donc élevée. La présence d'une panne à ces moments là est une charge supplémentaire pour l'équipage. D'après une étude de O'Hare, et al.(1994) sur une base de données comportant 596 incidents/accidents attribués à une erreur de l'équipage, 56% des erreurs en causes avaient eu lieu au décollage ou à l'atterrissage. La remise en cause des comportements attendus de l'équipage lors de ces phases de vol observée dans notre étude rejoint donc les données de la littérature issues de l'analyse d'accidents.

Enfin, les descripteurs du type de panne : « Fonction erronée » et « panne non détectée » (dans le cas du CDS) sont privilégiés par les experts. L'étude de Jones et al (1996) portant sur la *situation awareness* ainsi que les études sur les erreurs de prise de décision (Orasanu, et al., 1999) rejoignent les résultats que nous avons obtenus. D'après l'étude d'Orasanu et al., les pilotes peuvent développer une mauvaise interprétation du problème qui entraînera une mauvaise décision. Ce sont des erreurs dites de *situation awareness* dans lesquelles la conscience de la situation et donc du problème à résoudre sont erronées. D'ailleurs l'étude de Jones et al (1996) avait montré que 96% des incidents et accidents provenaient d'erreurs dans l'intégration et l'interprétation de la situation. D'après Orasanu et al. ces erreurs de *situation awareness* proviennent du fait que les indices sur la situation sont **non détectés, mal interprétés, mal diagnostiqués ou encore ignorés** par l'équipage. De ce fait, les types de pannes « erronées et non détectées » préférentiellement sélectionnés par les experts vont jouer un rôle dans l'apparition d'erreurs de *situation awareness,* entraînant chez le pilote, des diagnostics ou interprétations erronés sur la situation dans laquelle il se trouve. La gravité des conséquences explique que les experts aient privilégié ces descripteurs dans leur sélection, ce qui rejoint les données de la littérature.

IV Conclusion :

En résumé, dans le cas du CDS, les résultats de la méthode 1 auraient pu prédire totalement la sélection des scénarii par les experts alors que ce ne fut pas le cas pour le Pilote Automatique.

Les bonnes concordances obtenues avec le CDS laissent penser que l'hypothèse posée (les descripteurs pertinents sont plus présents lorsque la classification s'aggrave) est un bon point de départ pour l'identification rapide de descripteurs pertinents pour réaliser des requêtes automatiques. De plus, dans le cas du pilote automatique, le travail réalisé par les experts a permis d'identifier les descripteurs

pertinents qui permettront à l'avenir de réaliser des requêtes expertes. Ces requêtes pourraient mieux prendre en compte les particularités des systèmes étudiés.

Enfin pour conclure, l'étude approfondie de ces deux grands systèmes (CDS et PA) pourra à l'avenir nourrir les réflexions afin d'identifier les systèmes qui pourront être analysés à l'aide de la méthode 1 et donc par le biais de requêtes automatiques, des systèmes qui nécessiteront une analyse par jugement d'expert.

Voyons maintenant, dans le dernier chapitre, de cette partie expérimentale comment mettre en place des outils d'analyse automatique du langage afin d'aider les experts à l'extraction de ces scénarii.

CHAPITRE. 3 DEVELOPPEMENT ET TEST DES OUTILS DE TRAITEMENT AUTOMATIQUE DE TEXTES

Dans cette dernière partie de notre thèse, l'objectif était de créer des outils basés sur des techniques d'analyse automatique de textes afin de repérer rapidement, au sein de grandes bases de données, les scénarii pertinents pour une étude en facteurs humain.

Dans le chapitre 2, nous avons identifié (dans la base de données existante) les mots clefs (descripteurs) qui pourraient permettre d'isoler des scénarii pertinents (c'est-à-dire contenant des hypothèses de comportement discutables). Ce faisant, nous n'utilisons qu'une fraction de l'information contenue dans les scénarii. En effet, ces descripteurs sont contenus dans une phrase de description synthétique de la panne (phrase accompagnant systématiquement la description libre et détaillée réalisée par l'expert). Dans ce chapitre, nous souhaitons utiliser les scénarii dans leur ensemble c'est-à-dire en incluant aussi les informations détaillées.

Rappelons que l'intérêt et les objectifs liés à la possibilité d'utiliser l'intégralité des scenarii lors de la sélection, ont été précisés par les experts. (cf. Partie Théorique Chapitre 3, paragraphe III.3.1), En particulier 2 points principaux ont été soulignés :

1) La possibilité de récupérer les scenarii les plus proches d'un scenario pertinent : lorsque les experts sélectionnent un scénario pertinent en s'aidant de requêtes construites grâce aux descripteurs sélectionnés (cf. Chapitre 2), ils souhaiteraient alors pouvoir disposer d'un outil permettant de rechercher, dans la base de données, les scénarii qui lui sont proches (c'est-à-dire contenant des mots présents dans le scénarii sélectionné). Pour réaliser cet outil, nous avons utilisé **l'algorithme des « plus proches voisins »** (*Nearestneighbor*). Cet algorithme calcule, grâce à une représentation sous forme de vecteurs, la distance euclidienne entre deux scénarii. Nous indiquerons dans la partie

méthode, comment l'ensemble des scénarii sera formalisé sous forme de vecteurs avant d'expliquer le fonctionnement de l'algorithme des plus proches voisins.

2) L'utilisation de la description complète des scénarii doit aussi permettre **l'évaluation de « la probabilité de pertinence » pour chaque scénario**[10]. Pour être réalisée cette fonction, nécessite : 1/ une phase d'apprentissage qui utilise les scénarii sélectionnés par les experts pour identifier un lexique (mots les plus souvent cités dans ces scénarii) qui constituera le lexique de référence 2/ une phase d'application qui consistera à évaluer la pertinence des autres scénarii de la base de données en en identifiant la présence (et la quantité) de ces mots clefs. Plus ces mots clefs seront présents dans un scénario donné plus il sera considéré comme pertinent (c'est-à-dire contenant des hypothèses de comportements discutables). Comme l'algorithme des « plus proches voisins », cet outil utilisera une formalisation des scénarii sous forme de vecteurs. Pour obtenir la probabilité de sélection d'un scenario, nous utiliserons **l'algorithme de filtrage bayésien.**

I Méthode

I.1 Etape préliminaire : Formalisation sous forme de vecteurs

Comme nous l'avons indiqué dans l'introduction, les deux algorithmes nécessitent d'appliquer un formalisme particulier aux scénarii. Prenons un exemple simplifié. Voici un ensemble de 3 textes (pour nous ce sont les scénarii de la base de données),

Texte 1 : « je je vais bien ».

Texte 2 : « je je je je vais vais ».

Texte 3 : « bien je vais vais ».

Trois mots sont présents dans le lexique : « bien »,« je » et « vais ». « Bien » correspond au vecteur de base [1,0,0], « je » à [0,1,0] et « vais » à [0,0,1]. Les vecteurs correspondant aux trois textes précédents sont donc :

V_1 = [1,2,1] (Il y a 1 fois le mot « bien », 2 fois le mot « je » et 1 fois le mot « vais »).

V_2 = [0,4, 2] (Il y a 0 fois le mot « bien », 4 fois le mot « je » et 2 fois le mot « vais »).

V_3 = [1,1, 2] (Il y a 1 fois le mot « bien », 1 fois le mot « je » et 2 fois le mot « vais »).

[10] Grâce à cette fonction, les experts pourront garantir une analyse complète et rapide de la base de données. En effet les scénarii ayant une forte probabilité de sélection pourront être étudiés en priorité par les experts. Cela permettra de garantir une analyse rapide et exhaustive de la base de données et un donc une sélection plus précise et cohérente des scénarii pertinents.

Cette formalisation est défendue par Losee (1998). Elle permet notamment de décrire, avec un formalisme cohérent, des textes qui peuvent avoir des contenus très divers. Enfin cette formalisation est utilisée dans de nombreux outils de traitement automatique de textes comme le filtrage du courrier indésirable dans les messageries électroniques.

Cette formalisation devra donc être réalisée pour tous les scénarii ($\mathcal{N}$ = 1666). Ce nombre correspond aux scénarii du *Control and Display System* ($\mathcal{N}_1$ = 697) ainsi qu'a ceux du *Pilote Automatique* ($\mathcal{N}_2$ = 969). Nous utilisons ces deux analyses car, pour évaluer la performance des outils développés, nous aurons besoin des scénarii jugés pertinents par les experts (Cf. Chapitre 1).

Pour réaliser cette formalisation nous avons suivi différentes étapes.

1/Tout d'abord, nous devons disposer d'un **« lexique ».** Nous appelons lexique, l'ensemble des mots présents dans les scénarii d'un système. Grâce à un programme informatique que nous avons développé, nous avons extraits tous les mots contenus dans les 697 scénarii du *Control and Display System* et les 969 du *Pilote Automatique*. Au final, il y a 1381 mots différents dans les scénarii du *Control and Display System* et 1237 mots pour le *Pilote Automatique*. Par exemple, pour le Pilote Automatique les 1237 mots du lexique ont été numérotés de « a » =1 et zone = 1237

- a : 1
- ability : 2
- […]
- zero : 1236
- zone : 1237

Dans l'exemple précédent nous n'avions que 3 mots dans le lexique, ce qui correspondait à 3 vecteurs de base ([1,0,0], [0,1,0] et [0,0,1]). Par équivalence, nous allons avoir ici 1381 vecteurs de base pour le CDS et 1237 pour le PA. Chaque vecteur a un nombre de composantes égales au nombre de mots présents dans les scénarii du système considéré. En reprenant le lexique du pilote automatique, voici les vecteurs de base que nous avons obtenu :

- a : [1,0,0,…(1233 fois 0)…,0]
- ability : [0,1,0,…(1233 fois 0)…,0]
- […]
- zero : [0,…(1233 fois 0)…,0,1,0]
- zone : [0,…(1233 fois 0)…,0,0,1]

2/Enfin, à l'aide des vecteurs de base obtenus précédemment, nous avons formalisé **tous les scénarii.** Pour ce faire, à l'aide d'un programme informatique que nous avons développé, nous avons identifié

pour chaque scenario, les mots qu'il contient ainsi que leurs fréquences. Reprenons l'exemple simplifié :

Le texte 1 : « je je vais bien » devient le vecteur [1,2,1] puisqu'il y a 1 fois le mot « bien », 2 fois le mot « je » et 1 fois le mot « vais ».

En appliquant ce principe aux scénarii du *Control and Display System et du Pilote Automatique*, nous avons transformé tous les scénarii à l'aide des 1381 vecteurs de base du CDS et des 1237 vecteurs de base du PA.

I.2 Méthode pour les algorithmes de recherche (Les plus proches voisins, le filtrage bayésien)

Nous avons déjà présenté dans le détail le principe de ces algorithmes (cf. Partie théorique, Chapitre 3, paragraphe III.3.2). Pour rappel nous avons développé 2 algorithmes nous permettant d'atteindre 2 principaux objectifs fixés par les experts:

- Pouvoir sélectionner tous les scénarii proches d'un scénario pertinent
- Pouvoir affecter à chaque scénario une probabilité de sélection permettant ainsi de faire un choix rapide des scénarii les plus pertinents.

Pour atteindre ces objectifs les deux algorithmes que nous avons développés sont les suivants :

L'algorithme des « plus proches voisins », qui va nous permettre de mesurer la distance entre deux scénarii. Par exemple, nous pouvons mesurer la distance entre un scénario sélectionné et les autres scenarii, afin d'identifier ceux qui lui sont le plus proches (et qui pourraient donc être également pertinents).

L'algorithme de filtrage bayésien, qui va nous permettre, à partir d'un ensemble de scénarii utilisés lors d'une phase d'apprentissage, d'estimer la probabilité de sélection des autres scénarii. En classant ces probabilités par ordre croissant, nous pourrons repérer aisément les scénarii ayant la plus forte probabilité de sélection.

Sans rentrer dans le détail de l'implémentation[11], nous avons utilisé Scilab (http://www.scilab.org/) pour réaliser les algorithmes des « plus proches voisins » et celui du filtrage bayésien. Scilab est un logiciel libre de calcul numérique fournissant un environnement de calcul pour des applications scientifiques. Les méthodes qu'il fournit, notamment pour le calcul vectoriel et matriciel, sont

[11] V*oir chapitre 3 en annexe pour plus d'information sur le fonctionnement des algorithmes.*

parfaitement adaptées à nos besoins pour l'implémentation de nos algorithmes. Voyons maintenant les résultats de ces deux algorithmes dans le paragraphe suivant.

II Résultats

II.1 Résultats pour l'algorithme des « plus proches voisins »

Les résultats que nous présentons ici visent à évaluer l**'efficacité de l'algorithme des proches voisins.** Pour ce faire, nous avons utilisé la procédure suivante :

1/A partir d'un scénario sélectionné par les experts, nous avons identifié, à l'aide de l'algorithme ses *n* plus proches voisins (dans la Figure 34, *n* est présenté en abscisse varie de 1 à 697)

2/ Dans ces *n* plus proches voisins (sélectionné par l'algorithme) nous avons alors calculé le nombre de scénarii qui avaient été jugés pertinents par les experts (cf. Chapitre 1).

II.1.1 Résultats pour l'analyse du CDS

Pour ce qui concerne le CDS, les résultats présentés dans le graphique suivant (Figure 34) se lisent de la façon suivante : sur 100 plus proches voisins sélectionnés par l'algorithme (en abscisse) 25 sont des scénarii pertinents (courbe bleue). De la même façon, pour 200 plus proches voisins sélectionnés 37 sont des scénarii pertinents, etc.

Par comparaison, la courbe noire présente une sélection au hasard, sachant qu'environ 11% des scénarii ont été identifiés par les experts comme étant pertinents.

La méthode que nous avons utilisée, permet donc de comparer le nombre de scénarii sélectionnés par l'algorithme (ici 25 pertinents dans les 100 plus proches voisins d'un scénario pertinent donné) avec le nombre de scénarii pertinents qui auraient été sélectionnés aléatoirement (ici 11 pertinents dans les 100 plus proches voisins).

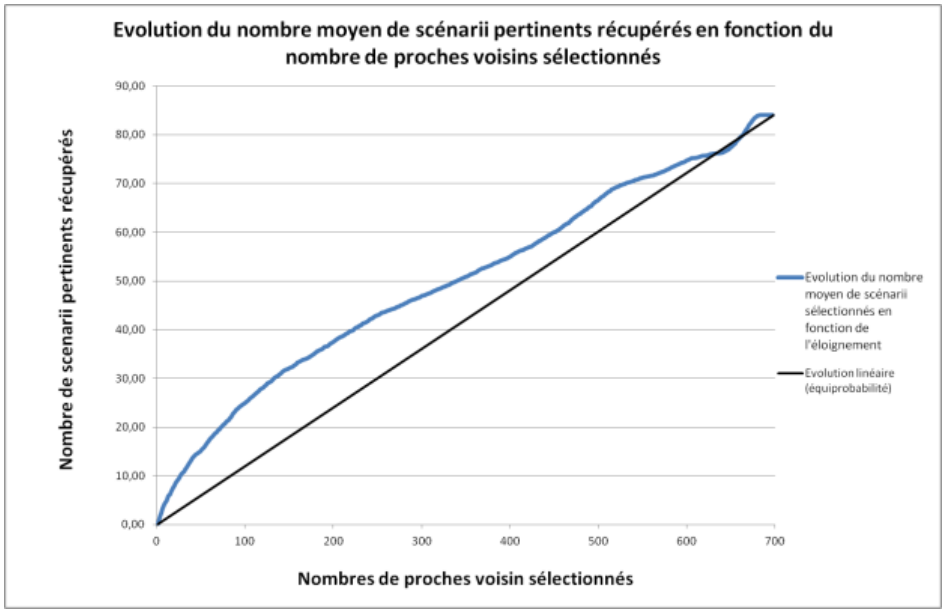


Figure 34 : Nombre de scénarii pertinents dans les plus proches voisins

Pour comparer ces courbes, nous avons utilisé des tests de comparaison de deux pourcentages. Pour chaque *n* (nombre de plus proches voisins), nous avons donc calculé les écarts réduits entre le pourcentage donné par l'algorithme et le pourcentage aléatoire. Dans cette méthode de l'écart réduit, les différences sont significatives quand l'écart calculé entre les deux pourcentages est supérieur à 1,96. Dans la figure ci-dessous (Figure 35), la courbe bleue donne la valeur de l'écart réduit en fonction de *n* en abscisse et la droite horizontale noire représente le seuil 1,96. Nous pouvons donc constater que les différences entre les pourcentages sont significatives à partir de *n* = 12 plus proches voisins jusqu'à *n* = 183 plus proches voisins.

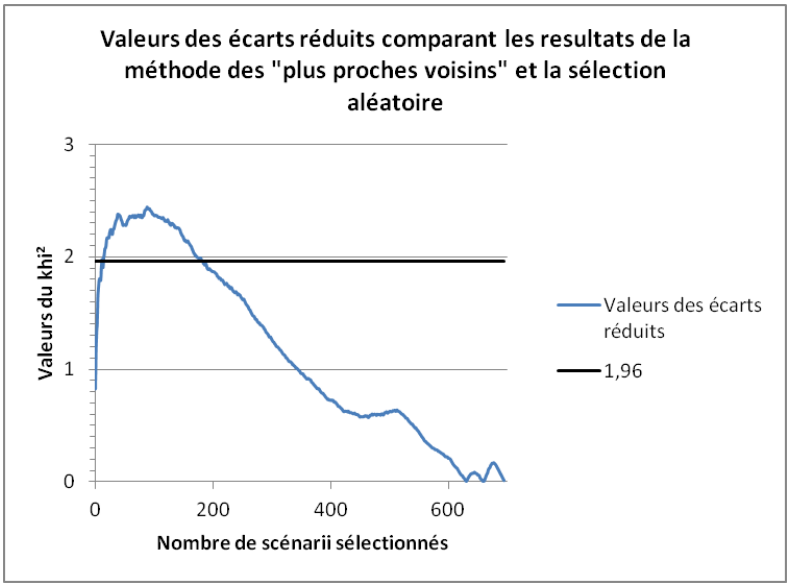


Figure 35 : Valeurs des écarts réduits (CDS)

II.1.2 Résultats pour l'analyse du PA

Pour le PA le graphique ci-dessous (Figure 36) montrent l'évolution du nombre moyen de scénarii pertinents récupérés en fonction de la taille de l'ensemble des « plus proches voisins ».

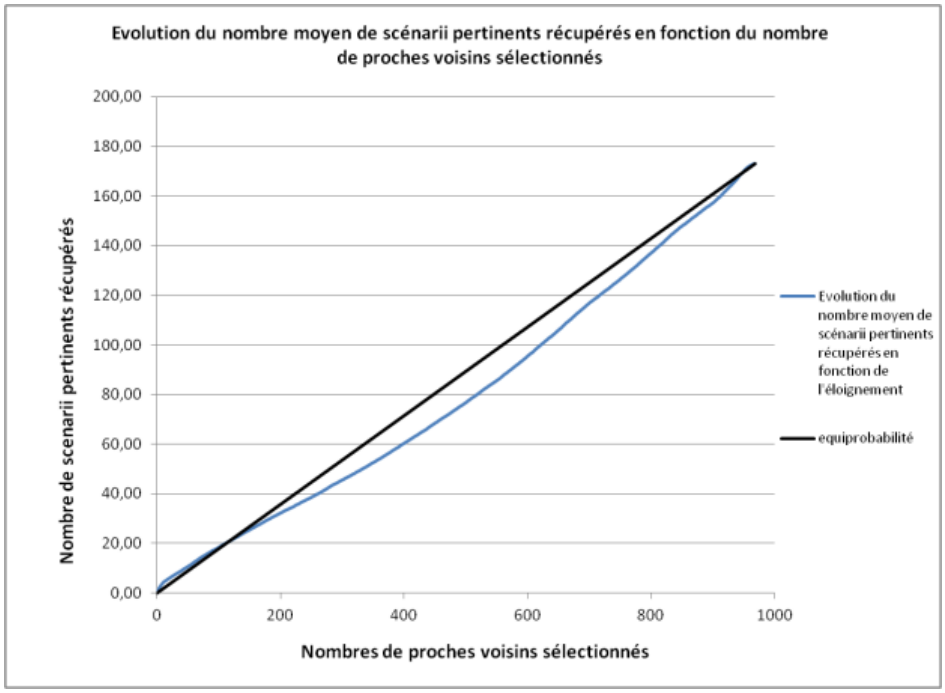


Figure 36 : Nombre de scénarii pertinents dans les plus proches voisins

Cependant, dans le cas du PA, la sélection de scénarii pertinents par la méthode des « plus proches voisins » ne montre pas de différences significatives par rapport une sélection aléatoire.

Après avoir présenter les résultats pour l'algorithme des plus proches voisins, nous allons présenter les résultats obtenus à l'aide de **l'algorithme de filtrage bayésien.**

II.2 Résultats pour l'algorithme du filtrage bayesien

La méthode de filtrage bayesien permet d'attribuer une probabilité de sélection à chaque scénario. Cette méthode s'initialise à l'aide d'une phase d'apprentissage réalisée à partir des scenarii sélectionnés par les experts (ensemble d'apprentissage). Le reste de la base de données constituera l'ensemble de test (ou de rappel). Pour rappel, la phase d'apprentissage utilise les scenarii sélectionnés par les experts pour identifier un lexique (mots les plus souvent cités dans ces scenarii) qui constituera le lexique de référence ; la phase de test consiste à évaluer la pertinence des autres scenarii de la base de données (probabilité de sélection) en identifiant la présence (et la quantité) de ces mots clés. Plus

ces mots clés sont présents dans un scenario donné plus il sera considéré comme pertinent (c'est-à-dire contenant des hypothèses discutables)

Pour le CDS l'ensemble d'apprentissage comprenait 42 scenarii pertinents qui avaient donc été sélectionnés par les experts, et l'ensemble de test comprenait 349 scénarii.

Pour le PA l'ensemble d'apprentissage comprenait 87 scenarii pertinents qui avaient donc été sélectionnés par les experts, et l'ensemble de test comprenait 485 scénarii.

Pour obtenir les résultats que nous présentons ci-dessous, nous avons :

a. réalisé la phase d'apprentissage, puis la phase de test pour les 2 systèmes ;
b. Pour chaque système, classé tous les scenarii par ordre décroissant de probabilité (de sélection)
c. Calculer pour les *n* scénarii premiers scenarii (classés grâce à l'algorithme), le nombre de scenarii qui avaient été sélectionné comme pertinents par les experts.

II.2.1 Résultats pour l'analyse du CDS

Les résultats du CDS, présentés sur la Figure 37, indiquent que sur les 100 premiers scenarii sélectionnés par l'algorithme, nous avons sélectionné 93%des scenarii qui avaient été considérés comme pertinents par les experts. En effet sur les 100 premiers sélectionnés, 40 sont pertinents (sur les 42 au total) sont des scénarii des scénarii pertinents. Par comparaison, la droite noire correspond à une sélection aléatoire, pour 100 scénarii sélectionné nous récupérons 26% des scénarii pertinents.

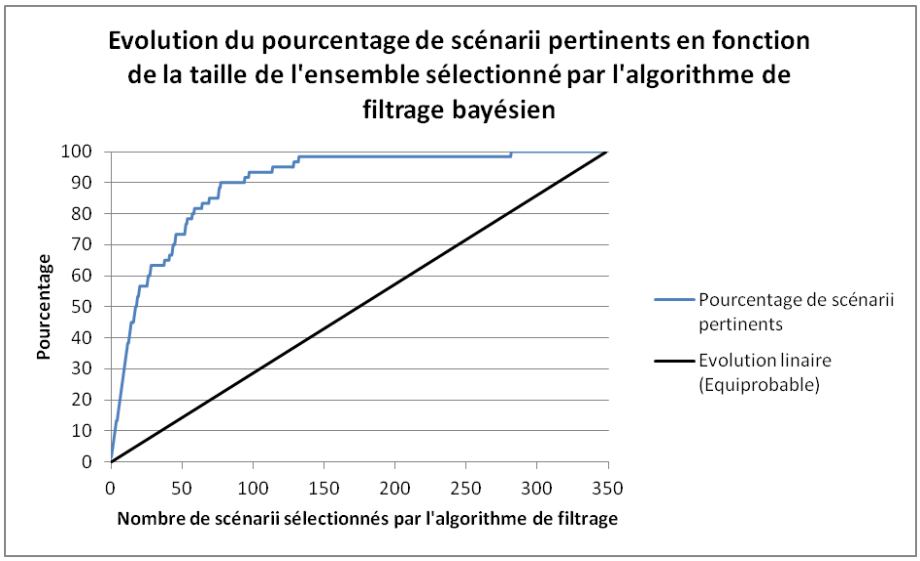


Figure 37 : Evolution du pourcentage moyen de scénarii pertinents récupérés par la méthode du filtrage bayésien (CDS)

Afin d'évaluer la performance de cette méthode de sélection nous avons utilisé une comparaison de deux pourcentages. Pour chaque n (nombre de scenarii sélectionnés) nous avons calculé l'écart entre le nombre de scenarii pertinents sélectionnés par la méthode de filtrage et le nombre de scenarii pertinents qui auraient été sélectionnés de façon aléatoire. Les valeurs des écarts réduits en fonction du nombre de scénarii sélectionnés sont présentées sur la Figure 38.

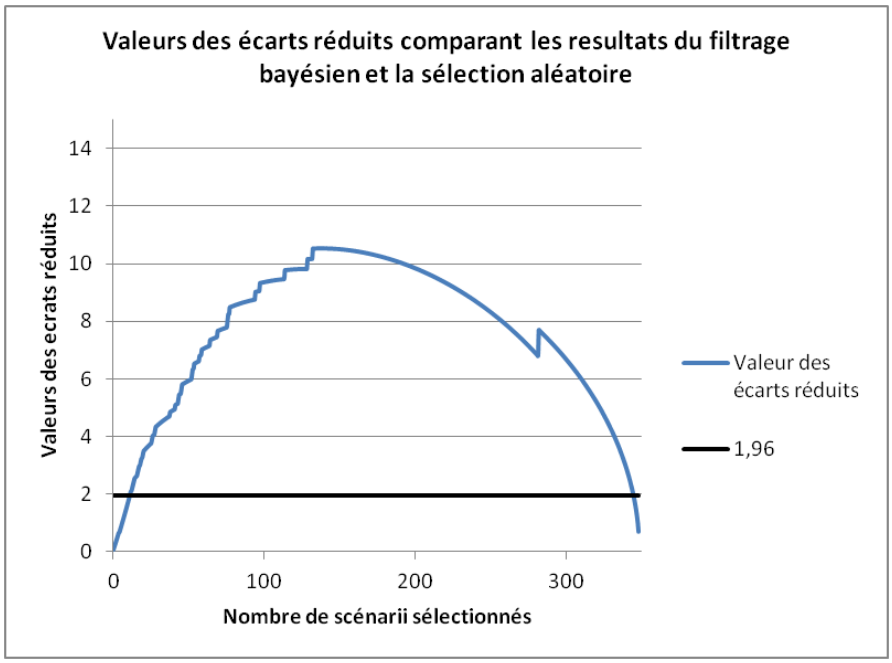


Figure 38 : Valeurs des écarts réduits (CDS)

Les écarts sont significatifs pour *n* (nombre de scenarii sélectionnés) compris entre 12 et 344 scénarii avec un maximum pour 129 scénarii sélectionnés. En sélectionnant 129 scénarii à l'aide de l'algorithme, nous obtenons 96,7% scénarii qui avaient été jugés pertinents par les experts.

Voyons maintenant les résultats du filtrage bayésien pour le pilote automatique.

II.2.2 Résultats pour l'analyse du PA

Les résultats du PA, présentés sur la Figure 39, indiquent que sur les 100 premiers scenarii sélectionnés par l'algorithme, nous avons sélectionné 71% des scenarii qui avaient été considérés comme pertinents par les experts. En effet sur les 100 premiers sélectionnés, 62 sont pertinents (sur les 87 au total) sont des scénarii pertinents. Par comparaison, la droite noire correspond à une sélection aléatoire, pour 100 scénarii sélectionnés nous récupérons 20,6% des scénarii pertinents.

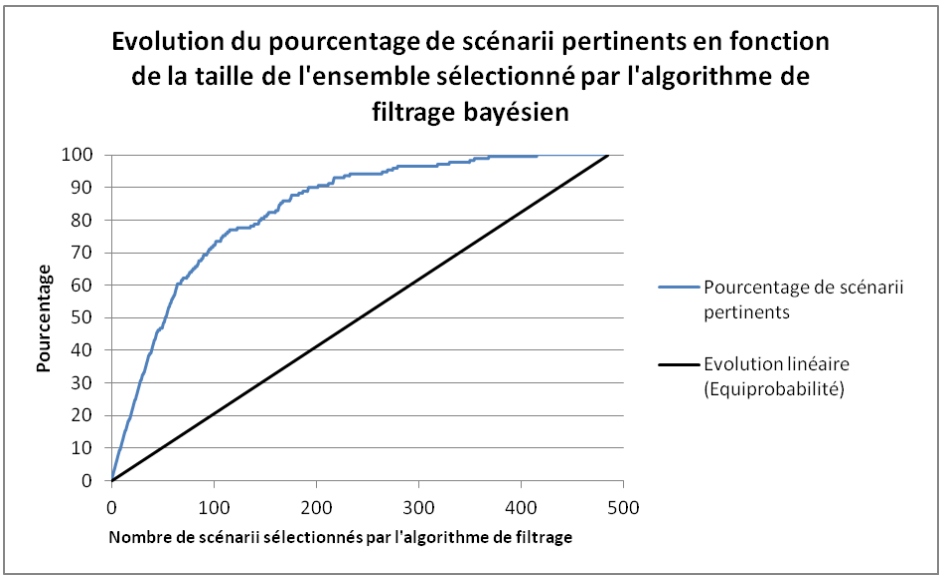


Figure 39 : Evolution du pourcentage moyen de scénarii pertinents récupérés par la méthode du filtrage bayésien (PA)

Afin d'évaluer la performance de cette méthode de sélection nous avons utilisé une comparaison de deux pourcentages. Pour chaque n (nombre de scenarii sélectionnés) nous avons calculé l'écart entre le nombre de scenarii pertinents sélectionnés par la méthode de filtrage et le nombre de scenarii pertinents qui auraient été sélectionnés de façon aléatoire. Les valeurs des écarts réduits en fonction du nombre de scénarii sélectionnés sont présentées sur la Figure 40.

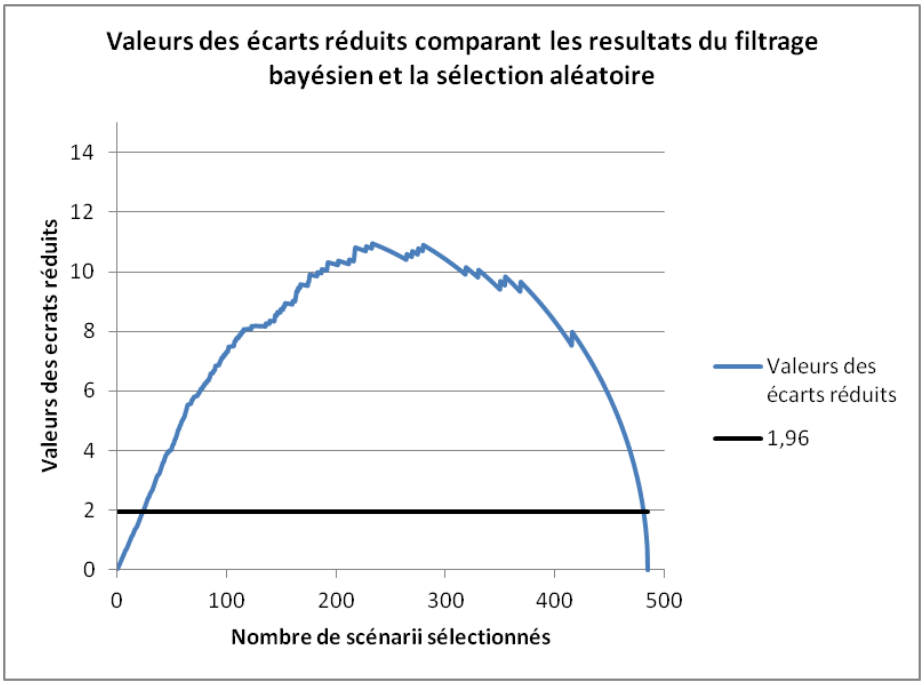


Figure 40 : Valeurs des écarts réduits (PA)

Les écarts sont significatifs pour *n* (nombre de scenarii sélectionnés) compris entre 24 et 480 scénarii avec un maximum pour 233 scénarii sélectionnés. En sélectionnant 233 scénarii à l'aide de l'algorithme, nous obtenons 94 % de scénarii qui avaient été jugés pertinents par les experts.

III Interprétation des résultats

A partir d'un scénario particulier, l'objectif de **l'algorithme des « plus proches voisins »** est d'identifier rapidement au sein d'un grand ensemble de données, les scénarii qui lui sont proches et qui ont donc toutes les chances d'être également pertinents. Cet objectif avait été fixé par les experts afin d'optimiser la recherche d'informations pertinentes au sein des bases de données. Ces scenarii pertinents permettent en effet, comme nous l'avons vu, de tester (sur simulateur ou essais en vol) les hypothèses de comportement des pilotes qui sont considérées comme discutables et ces tests jouent un rôle central dans la conception centrée utilisateur et donc dans la fiabilité et la sécurité du transport aérien.
Mais les résultats que nous avons obtenus à l'aide de cette méthode, dépendent du système considéré :
1/Dans le cas du CDS, l'algorithme permet la sélection d'un nombre important de scenarii pertinents (25 scénarii sélectionnés sur les 100 plus proches voisins) ;.
2/ En revanche, cette méthode c'est révélée inefficace pour le PA.

La méthode de filtrage bayésien permet de calculer la probabilité de sélection des scenarii contenus dans une base de données. Cette méthode s'est révélée extrêmement efficace pour sélectionner de façon rapide et automatique des scenarii qui avaient été jugés pertinents par les experts, et ce aussi bien pour le CDS que pour le PA.

En résumé, **la méthode bayésienne** est efficace pour les deux systèmes contrairement à la méthode **des plus proches voisins** qui c'est révélée inefficace dans le cas du PA.

Au cours des expérimentations précédentes (sélection des scénarii, description selon le modèle de Reason, indentification des descripteurs), nous avons souvent rencontré des différences entre les deux systèmes étudiés que nous avons attribuées à la complexité du système PA. Ce système avait été choisi par les experts car ils le considéraient comme un système complexe. Afin de mieux comprendre les résultats obtenus, nous avons cherché, dans les analyses de sûreté, des indices de cette complexité.
Pour cela, nous allons comparer les fonctions que doivent remplir chacun de ces systèmes. Tout d'abord ils sont en charge d'une fonction principale : 1/Le pilote automatique est en charge de la

fonction de pilotage automatique de l'avion, 2/le *Control and Display System* est en charge de l'affichage d'informations en provenance d'autres systèmes ainsi que des interactions avec ces informations). Ces fonctions principales sont traduites **en fonctions de haut niveau, qui seront-elles même traduites en fonctions de niveau intermédiaire, puis en fonctions de bas niveau**. Cet arbre fonctionnel permet de créer, entre autre, les scénarii des analyses de sûreté. En effet, ces analyses correspondent aux descriptions des conséquences des pannes de ces différentes fonctions[12].

De ce fait, le nombre de fonctions appartenant à ces trois niveaux ainsi que le nombre de scénarii dans l'analyse de sûreté sont le reflet de la complexité du système. Nous avons récapitulé, dans le tableau suivant (Tableau 13) ces différentes informations.

	Pilote automatique	**Control and Display System**
Nombre de fonctions de haut niveau	12	7
Nombre de fonctions de niveau intermédiaires	72	44
Nombre de fonctions de bas niveau	287	82
Nombre de scénarii	969	697

Tableau 13 : Nombres de fonctions (Haut niveau, intermédiaire, bas niveau) et nombre de scénarii pour les deux systèmes étudiés

Nous pouvons vérifier que le nombre de fonctions remplies par le pilote automatique est supérieur à celui du *CDS* (pour les fonctions de bas niveau il y a un facteur 3,5 entre les deux systèmes). La grande diversité de fonctions réalisées par le PA joue donc certainement un rôle dans l'inefficacité de la méthode des plus proches voisins. En effet, cette hétérogénéité fait qu'un scénario pertinent appartenant à une fonction de haut niveau donnée n'aura pas beaucoup de points communs avec un scénario pertinent appartenant à une autre fonction de haut niveau, ce qui rend donc l'analyse des ressemblances (proches voisins) relativement inefficaces. En étudiant comment se répartissent les scénarii pertinents selon ces fonctions nous obtenons le tableau suivant (Tableau 14) :

	Pilote automatique	**Control and Display System**
Nombre de fonctions de haut niveau	11	2
Nombre de fonctions de niveau intermédiaires	38	8
Nombre de fonctions de bas niveau	89	20

Tableau 14 : Nombre de fonctions dans les scénarii sélectionnés

[12] Par exemple, pour le CDS une fonction de haut niveau serait d'afficher l'ensemble des informations issues des données brutes alors qu'une fonction de bas niveau serait d'afficher l'altitude.

Sur ce tableau nous pouvons voir, que les scénarii pertinents du PA comprennent 89 fonctions de bas niveau différentes et seulement 20 pour le CDS. La récupération des scénarii pertinents, à l'aide d'un algorithme calculant la proximité textuelle entre deux scénarii (algorithme des plus proches voisins), se révèle donc inefficace, puisque deux scénarii pertinents qui n'appartiennent pas à la même fonction ne seront pas similaires. En effet, le lexique de description d'une panne donnée sera plus proche de celui d'une autre panne si ces deux pannes appartiennent à la même fonction. Cela donc peut expliquer en partie pourquoi l'algorithme des plus proches voisins c'est révélé inefficace dans le cas du PA.

Néanmoins, puisque la récupération à l'aide du filtrage bayésien fonctionne, on peut faire l'hypothèse que ce problème de différence dans la complexité des deux systèmes peut être contournée lorsque l'algorithme utilisé comporte une phase d'apprentissage (présente pour l'algorithme de filtrage bayésien et non pour l'algorithme des proches voisins).

Cependant, d'après les experts (**retour d'expérience**), l'algorithme des plus proches voisins possède, tout de même, un intérêt dans la construction de l'ensemble des scénarii pertinents et cela quelque soit la complexité du système. En effet, lors de la sélection des scénarii, les experts peuvent vérifier rapidement si après avoir identifié un scénario pertinent, il existe des scénarii proches qui sont eux aussi pertinents. Si la réponse et non (c'est-à-dire que l'algorithme des plus proches voisins n'a pas permis de trouver d'autres scénarii pertinents) ils passeront plus rapidement à la recherche de scénarii pertinents pour d'autres fonctions du système.

Les résultats que nous avons obtenus sont difficiles à mettre en perspective avec les données existantes, néanmoins, une étude d'Hermann et al. (2008) portant sur les analyses d'incidents/accidents est relativement proche de notre problématique. Dans cette étude, réalisée avec le Bureau Enquêtes et Analyses (BEA), les auteurs posent le problème de la manière suivante :
« L'accident aéronautique est traditionnellement analysé comme le résultat de causes et facteurs contributifs multiples. Les principaux modèles d'analyse vont dans ce sens. Or, les autorités sont destinataires d'une information abondante, sous forme notamment de comptes-rendus d'incidents, mais dont la structure ne correspond pas toujours à ce découpage formel. Ainsi, l'analyste rencontre des difficultés pour identifier et extraire les occurrences qui présentent une certaine homogénéité en termes de combinaisons de facteurs. »

Pour répondre à cette problématique, les auteurs ont donc construit un outil d'analyse automatique de textes basé sur des occurrences de mots et groupes de mots (outils semblables aux nôtres) pour extraire des rapports d'incidents et accidents similaires à un rapport donné. Ceci afin « *d'outiller la décision de*

l'expert humain dans la mise en place et la réalisation de la catégorisation d'un corpus de données. » Aujourd'hui, l'exploitation par des experts, de bases de données volumineuses se heurte à l'hétérogénéité des informations enregistrées. Dans le même temps, une harmonisation excessive conduit à appauvrir l'information avant de l'exploiter. Dans ces conditions, il est important d'explorer des outils plus souples, qui prennent en compte la richesse d'information des textes libres.

Quoi qu'il en soit, l'étude de Herman et collaborateurs ainsi que la notre, mettent en lumière que l'utilisation d'outils de traitement automatique de textes pour l'extraction d'information pertinente dans des bases de données d'accidentologies (par exemple), constitue une voie d'avenir dans l'optimisation de la fiabilité des systèmes et la gestion efficace des erreurs humaines en aéronautique.

IV Conclusion

En résumé, l'efficacité de la méthode des « plus proches voisins » pour la récupération des scénarii pertinents dépend du système étudié. Malgré ce fait, les experts restent confiants quand à l'utilité de cette méthode pour gagner en rapidité de sélection. Concernant la méthode de filtrage bayésien, nous avons obtenus des résultats probants qui laissent envisager des applications très prometteuses pour l'analyse des hypothèses de comportement de l'équipage.

Pour conclure, l'étude comparative réalisée pour appréhender les différences de complexité entre les deux grands systèmes (CDS et PA) pourra à l'avenir nourrir les réflexions sur les rapports entre efficacité des algorithmes de traitement et complexité des systèmes. Ces réflexions permettront d'identifier les systèmes pour lesquels la complexité devra être prise en compte afin de garantir l'efficacité des outils d'analyse automatique de textes.

Présentons maintenant la discussion générale de ce travail.

.

DISCUSSION GENERALE ET PERSPECTIVES

Dans un système sociotechnique complexe comme l'industrie aéronautique, la mise en place du processus de conception centrée sur l'utilisateur comprend de nombreuses contraintes. Cette méthode, garante de la prise en compte des caractéristiques des utilisateurs dans le produit conçu, permet de réduire efficacement le risque lié à l'erreur humaine. Cependant, dans le cas de l'aéronautique, la part de l'erreur humaine dans les incidents ou les accidents stagne, autour d'une fourchette de 65% à 80% (Amalberti, 1996) et apparait toujours comme la dernière frontière à franchir pour une sécurité maximum.

L'objectif général de ce travail de recherche était d'outiller la conception centrée utilisateur afin de faciliter sa mise en œuvre (notamment en milieu complexe), améliorant ainsi la sécurité par une meilleure prise en compte des facteurs humains.

En répondant à cet objectif, notre étude a ouvert des axes de discussion qui seront exposés dans ce chapitre. Avant cela, commençons par rappeler les résultats obtenus.

I Synthèses des résultats obtenus

Pour rappel, notre étude était organisée en 3 phases :

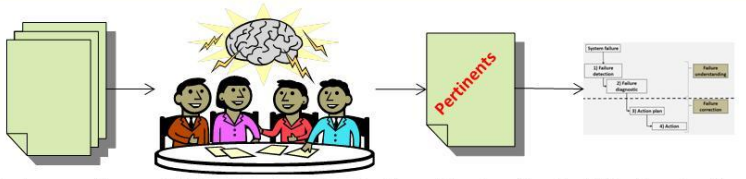
La première, concernait la sélection, par des experts, de scenarii qu'ils considéraient comme étant discutables dans un ensemble réduit de la base de données (analyse de sûreté). Puis, nous avons analysé les hypothèses qu'ils contenaient à l'aide du modèle de prise de décision de Reason (1990).

Ensuite, nous avons étudié les scénarii sélectionnés afin d'identifier des thématiques récurrentes, des mots clés, des descripteurs, des relations entre descripteurs, etc. L'objectif à terme est de pouvoir utiliser ces informations pour réaliser des requêtes, c'est-à-dire, des extractions automatisées des scénarii discutables à partir de mots clés.

Enfin, toujours dans le but de créer des outils d'extraction automatique, nous avons utilisé les scénarii sélectionnés par les experts pour tester la pertinence de 2 méthodes existantes d'extraction d'information dans une base de données : la méthode des « plus proches voisins » et la méthode du filtrage bayésien.

Le paragraphe ci-dessous est organisé en reprenant et synthétisant la méthode et les résultats obtenus dans chacune des trois parties.

I.1 Sélection des scénarii pertinents par un groupe d'experts.

<table>
<tr><td>Méthode</td><td>

Trois ensembles de scénarii : CDS, PA, Fuel
1/Sélection des scénarii par un Groupe d'experts
Ensemble de scénarii pertinents
2/Etude selon le modèle de prise de décision de Reason</td></tr>
<tr><td>Résultats</td><td>1/Sélection des scénarii par un Groupe d'experts
Les experts ont sélectionné : 84 scénarii pour le CDS ; 174 scénarii pour le PA et 0 scénarii du pilote automatique
Interprétation : Le système fuel possède une proportion importante (plus que le CDS et le PA) de scénarii catastrophiques or, il est impossible de considérer un scénario catastrophique comme pertinent ➔ Pas de sélection pour le système Fuel.

2/Etude selon le modèle de prise de décision de Reason
Les résultats de l'étude des scénarii par le modèle de prise de décision de Reason indiquent que les experts remettent en cause plutôt des scénarii contenant une hypothèse de détection ou de diagnostic de la panne qu'une hypothèse de planification ou d'action.
Interprétation : Les hypothèses de détection et de diagnostic sont plus souvent sujettes à des discussions lors de l'établissement de ces hypothèses. Elles sont donc souvent le fruit du meilleur compromis entre les différentes appréciations des experts. De ce fait, elles sont plus sujettes à des remises en cause.</td></tr>
</table>

Tableau 15 : Schéma de synthèse des résultats de l'étape 1

Dans le tableau ci-dessus (Tableau 15), la méthode utilisée pour la sélection de scénarii pertinents reposait sur le jugement d'experts. L'analyse des scénarii sélectionnés par le modèle de Reason, a permis de mettre en évidence l'importance des phases de détection et de diagnostic comme sources possibles d'erreurs humaines.

I.2 Des descripteurs pour des requêtes pertinentes

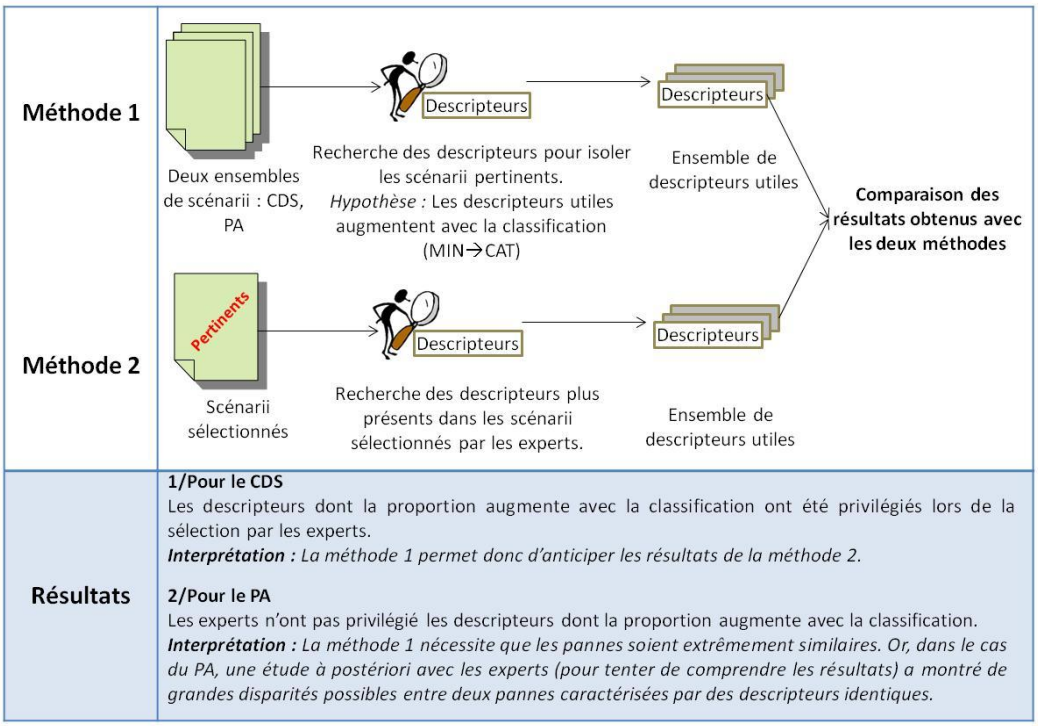


Tableau 16 : Schéma de synthèse des résultats de l'étape 2

L'étape 2, présentée dans le tableau ci-dessus (Tableau 16), a permis la comparaison de deux méthodes. La méthode 1 consiste à rechercher dans la base de données des descripteurs utiles pour la recherche de scénarii pertinents. L'hypothèse utilisée est que la proportion des descripteurs utiles augmente avec la classification. La méthode 2 étudie, a posteriori, les descripteurs qui caractérisent le mieux les scénarii sélectionnés par les experts. Les résultats obtenus dépendent du système étudié. Dans le cas du CDS, la méthode 1 et la méthode 2 donnent des résultats extrêmement encourageants. En effet, les descripteurs identifiés de façon automatique (méthode 1) et ceux identifiés par les experts sont similaires. L'intérêt de la méthode automatique est donc avéré. En revanche pour le PA, les résultats obtenus à l'aide des deux méthodes sont plus divergents.

I.3 Développement et test des outils de traitement automatique de textes

L'étape 3, présentée dans le tableau ci-dessous (Tableau 17), nous a permis d'étudier deux algorithmes de traitement automatique de textes :

- **L'algorithme « des plus proches voisins »** permet de mesurer le degré de similitude entre deux scénarii, l'objectif est alors de pouvoir extraire de façon automatique des scénarii pertinents à partir d'un seul scénario pertinent.

– **L'algorithme de filtrage bayésien** permet d'évaluer la probabilité de sélection d'un scénario, l'objectif ici est de pouvoir identifier rapidement dans la base de données les scénarii à sélectionner.

En résultat, l'algorithme de filtrage bayésien permet de sélectionner très efficacement les scénarii pertinents pour les deux systèmes. Alors que de l'algorithme des plus proches voisins est plus efficace pour le système CDS que pour le PA.

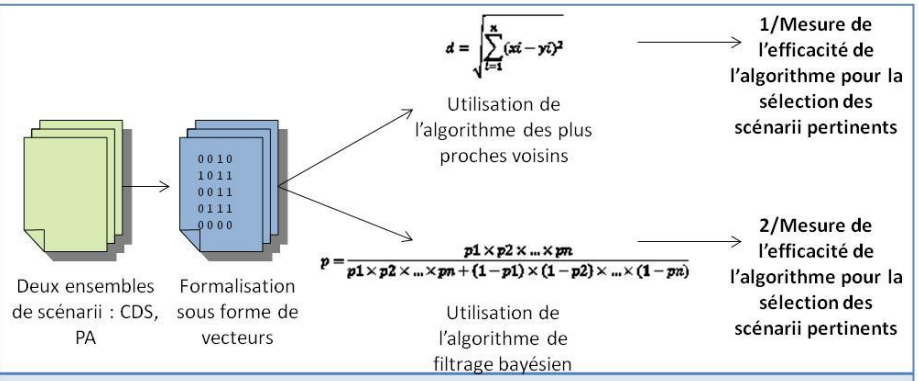
Méthode	
Résultats	**1/Algorithme des « plus proches voisins »** L'algorithme est efficace, pour la récupération des scénarii pertinents dans le cas CDS mais pas pour le PA. ***Interprétation :*** *Le système PA doit remplir plus de fonctions que le système CDS. Or les scénarii pertinents qui n'appartiennent pas à la même fonction ne seront pas similaires et cet algorithme mesure la similarité entre les scénarii.* **2/Algorithme de filtrage Bayésien** L'algorithme est très efficace pour la récupération des scénarii pertinents dans le cas des deux systèmes. ***Interprétation :*** *Puisque la récupération à l'aide du filtrage bayésien fonctionne, la complexité des systèmes peut être surmontée grâce à une phase d'apprentissage de l'algorithme (présente pour l'algorithme de filtrage bayésien et non présente pour l'algorithme des proches voisins).*

Tableau 17 : Schéma de synthèse des résultats de l'étape 3

Après avoir synthétisé les résultats dans le paragraphe précédent, nous allons maintenant voir comment ceux-ci répondent aux attentes industrielles, comment ils s'inscrivent dans la littérature existante et enfin nous conclurons cette discussion sur l'intérêt et les limites des méthodes que nous avons utilisées.

II Inscription de l'étude dans l'environnement industriel et la littérature existante

II.1 Réponses aux attentes industrielles

Le positionnement du problème a souvent été source de réflexion, en étant à la frontière de domaines scientifiques en perpétuelle évolution et d'enjeux industriels considérables. La façon dont a été défini le problème, c'est-à-dire en concertation permanente avec les experts des domaines concernés a certes

était source d'incertitude et de remise en question tout au long du travail, mais a permis de préciser les besoins, et d'adapter les méthodes existantes au contexte et aux exigences industrielles.

Nous avons tout d'abord tenu compte des besoins exprimés par les experts, puis à l'aide d'un processus itératif, nous sommes passés des besoins exprimés, à la conception des outils.

Pour rappel, voici les besoins exprimés :

- 1/ **Donner les moyens aux experts en sûreté de fonctionnement des systèmes, de vérifier l'exhaustivité de leur sélection**

Rappelons que l'exhaustivité de la sélection est un point crucial de la conception car c'est à partir de la sélection des scénarii pertinents que seront réalisées les études approfondies en facteurs humains en vue d'optimiser la sûreté du transport aérien. Nous avons répondu à ce besoin en développant des algorithmes de traitement automatique de textes. En effet, une fois la sélection réalisée l'algorithme « des plus proches voisins » permet de vérifier de façon extrêmement rapide que des scénarii proches et pertinents n'ont pas été oubliés dans la sélection. De plus, grâce à l'algorithme de filtrage bayésien, les experts peuvent obtenir rapidement la probabilité de sélection d'un scénario quelconque. De ce fait, ils pourront vérifier qu'ils n'ont pas oublié de sélectionner des scénarii possédant une forte probabilité de pertinence.
Cependant, nous avons vu que l'efficacité de l'algorithme des plus proches voisins dépendait de la complexité du système. Nous avons alors identifié les indicateurs de complexité suivants : nombre de scénarii de la base de données, nombre fonctions (haut niveau, niveau intermédiaire et bas niveau). Nous avons vu, par exemple, que le système PA contenait 3,5 fois plus de fonctions de bas niveau et 40% de scénarii en plus que le système CDS. Ces résultats nous permettront de prendre en compte, à l'avenir, la variable « complexité des systèmes » dans la conception d'outils adaptés aux besoins des experts.

- **2/ Puisque les experts en Facteurs Humains ne sont pas directement dans le processus d'écriture des scénarii, donner les moyens aux experts en Facteurs Humains d'interroger de la façon la plus efficace possible la base de données.**

Pour donner aux experts en Facteurs Humains la possibilité d'interroger la base de données, nous avons développé, en plus des algorithmes de traitement automatique de textes, une formalisation qui permettra la réalisation de requêtes afin d'extraire des sous ensembles de scénarii. Ces requêtes étaient

basées sur une hypothèse que nous avons élaborée à l'aide des experts. Cette hypothèse consistait à considérer que les descripteurs qui voient leur présence augmenter lorsque la classification s'aggrave sont certainement plus présents dans les scénarii pertinents.
Nous avons formulé cette hypothèse car, un scénario *minor*, par exemple, qui possède des descripteurs plus souvent associés à des scénarii *catastrophics*, voit, peut-être, sa classification conservée du fait d'une action correctrice de l'équipage. Donc, si cette action n'est pas réalisée, il peut se produire une aggravation de la classification (de *minor* à *major* par exemple).
Pour permettre la réalisation de ces requêtes, nous avons formalisé l'ensemble des scénarii (soit 2200 scénarii) en utilisant un mode de présentation standard comprenant des rubriques définies telles que : la phase de vol, le type de panne, la détection de la panne par le système et bien évidemment la classification.
Afin de vérifier l'efficacité de cette hypothèse pour extraire les scénarii pertinents, nous avons comparé les scénarii sélectionnés par les experts avec ceux extraits par requêtes. Les résultats montrent que les descripteurs sont sensiblement les mêmes pour le CDS mais pas pour le PA. Comme pour l'algorithme des plus proches voisins, les deux systèmes CDS et PA ne montrent pas les mêmes comportements. Là encore, une analyse plus approfondie de ces deux systèmes sous l'angle de la complexité, permettra de mieux comprendre ces différences constatées et d'apporter des améliorations en vue d'optimiser la sélection des scénarii pertinents quelque soit le type de système.

- 3/ Garantir le transfert des connaissances acquises d'un avion à l'autre.

L'algorithme de filtrage bayésien permet de garantir le transfert des connaissances d'un avion à l'autre. En effet, comme nous l'avons vu, l'algorithme de filtrage bayésien nécessite une phase d'apprentissage réalisée à partir des scénarii sélectionnés par des experts. Il est, dés lors, tout à fait possible de réaliser l'apprentissage sur une base de données appartenant à un avion particulier et d'utiliser cet apprentissage pour explorer une base de données correspondant à un autre avion. De l'avis des experts, ce transfert d'apprentissages constituerait un gain de temps considérable pour l'extraction des scénarii pertinents dans les premières phases de conception d'un nouvel avion. Cependant, nous n'avons pas encore testé cette possibilité mais cette étape s'inscrit dans la continuité de notre travail de thèse.

Pour conclure, d'une manière générale, nos travaux répondent aux besoins émis par l'industrie. De plus, bien que ces besoins émanent spécifiquement de l'aéronautique, de nombreuses industries sont soumises à des contraintes similaires. Ces travaux pourraient à terme, trouver des échos dans de très

nombreux domaines industriels et en particulier les industries à risques tels que l'industrie nucléaire, l'industrie spatiale, le transport ferroviaire, etc.

Voyons maintenant comment ces résultats s'inscrivent dans la littérature des incidents/accidents d'avion.

II.2 Inscription dans la littérature des incidents/accidents

Tout d'abord, nous avons vu que les résultats de l'étape 1 (Analyse à l'aide du modèle de Reason) pouvaient être mis en relation avec ceux de Jones et al. (1996) sur la conscience de la situation. A partir d'une analyse de 143 accidents de l'aviation civile américaine relevant d'une erreur humaine, Jones et al. montrent que les erreurs de niveau 1 de *situation awareness* (défaillance de perception et d'intégration de l'information) peuvent être invoquées dans 76% des accidents, celles de niveau 2 de *situation awareness* (interprétation de la situation) dans 20% des cas, et celle de niveau 3 (anticipation) dans les 4% restants. Les résultats de Jones et al., (1996) montrent donc que les erreurs de perception, d'intégration et d'interprétation de la situation totalisent 96% des accidents. Dans ces cas là, les hypothèses de comportements de l'équipage sur la détection et le diagnostic de la situation sont remises en cause, ce qui est cohérent avec notre étude.

Ensuite, les descripteurs sélectionnés par les experts (résultats de l'étape 2) trouvent aussi des échos dans la littérature sur l'étude des incidents/accidents d'avion.

- 1/ **Les phases de vol :** les experts ont sélectionnés en priorité des scénarii se passant au décollage et à l'atterrissage. Or, d'après une étude sur les incidents et accidents menée par Boeing (Boeing Company, 2010), ces phases de vol, qui ne totalisent qu'en moyenne 6% du temps de vol effectif, contiennent 53% des incidents/accidents graves. De plus, d'après une étude de O'Hare et al. (1994) sur une base de données comportant 596 incidents/accidents attribués à une erreur de l'équipage, 56% des erreurs en causes avaient eu lieu au décollage ou à l'atterrissage. La remise en cause des comportements attendus de l'équipage lors de ces phases de vol, observée dans notre étude, rejoint donc les données de la littérature issues de l'analyse d'accidents.

- 2/ **Le descripteur du type de panne :** les experts ont sélectionné en priorité des scénarii de type « fonction erronée » et « panne non détectée ». D'après l'étude d'Orasanu et al., les pilotes peuvent développer une mauvaise interprétation du problème qui entraînera une mauvaise

décision. Ce sont des erreurs dites de *situation awareness* dans lesquelles la conscience de la situation et donc du problème à résoudre est erronée (elles représentent 96% des accidents dus à une erreur humaine d'après Jones et al.,1996). D'après Orasanu et al. ces erreurs de *situation awareness* proviennent du fait que les indices sur la situation sont **non détectés, mal interprétés, mal diagnostiqués ou encore ignorés** par l'équipage. De ce fait, les types de pannes « erronées et non détectées » préférentiellement sélectionnés par les experts vont jouer un rôle dans l'apparition d'erreurs de *situation awareness*, entraînant chez le pilote, des diagnostics ou interprétations erronés sur la situation dans laquelle il se trouve. La gravité des conséquences explique que les experts aient privilégié ces descripteurs dans leur sélection, ce qui rejoint les données de la littérature.

Pour finir, rappelons que la problématique adressée par les outils de traitement automatique de textes développés lors de l'étape 3, a déjà été abordée par Hermann et al. (2008) pour les analyses d'incidents/accidents. Nos résultats ainsi que ceux de Herman et al, confirment l'intérêt que présente l'élaboration d'outils de traitement automatique de textes dans l'analyse de bases de données complexes. Ces techniques, encore peu appliquées, permettent d'envisager des méthodologies d'analyse de données respectant le plus possible le contenu de l'information textuelle rédigée par les experts.

III Apports de l'étude à la démarche de conception

Les étapes de la démarche en facteurs humains qui sous-temps notre travail, sont centrées sur :

- Le jugement d'expert ;
- La sûreté de fonctionnement des systèmes, mise en place en milieux industriels complexes ;
- La conception centrée utilisateur.

A l'aide d'un schéma (Figure 41), nous avons représenté à gauche, ces trois étapes et à droite nous avons résumé les principaux apports de notre étude à ces trois étapes.

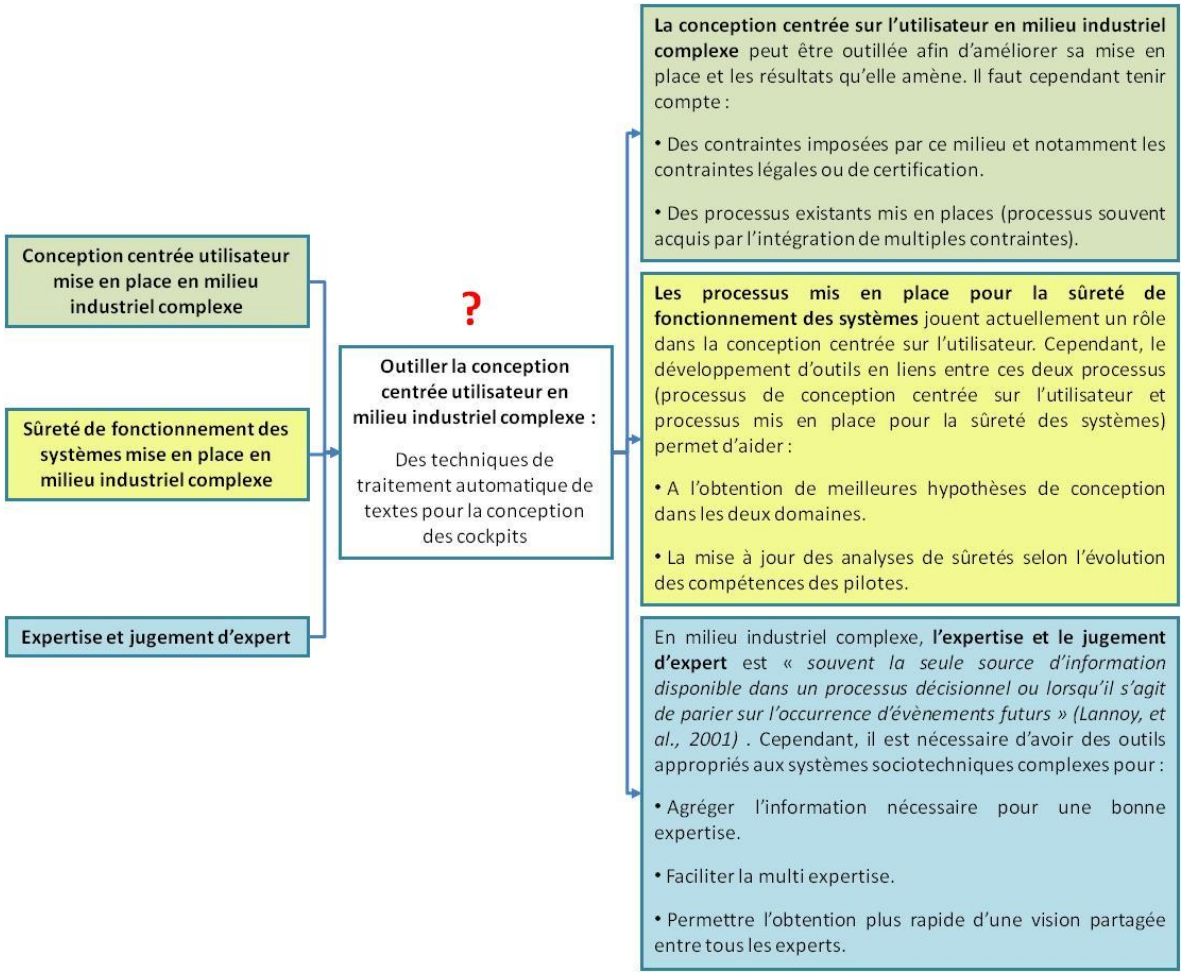


Figure 41 : Apport de l'étude à la démarche industrielle

Les enseignements issus de notre étude serviront de base à une réflexion sur l'évolution des pratiques de conception centrée sur l'utilisateur au sein d'Airbus. En particulier, les outils que nous avons développés, renforceront :

- **1/ L'intégration d'évaluations en facteurs humains dans les processus de création des analyses de sûreté des systèmes.**

En facilitant l'accès aux hypothèses discutables, présentes dans les analyses de sûreté des systèmes, les outils développés rendront possible les évaluations en facteurs humains de ces hypothèses (sur simulateur, essais en vol,…). Ces évaluations permettront de statuer sur la robustesse des hypothèses et le cas échéant de les améliorer.

- **2/ Une meilleure traçabilité des justifications des hypothèses prises dans ces analyses**

En permettant l'évaluation des hypothèses prises dans les analyses de sûreté, les outils développés permettront d'obtenir les arguments « pour ou contre » l'hypothèse discutée (compte rendus

d'évaluation, recueils d'avis de pilote, observation d'erreurs, etc.). Cette traçabilité favorisera la réutilisation de connaissances pour la conception ou la re-conception mais permettra aussi de garantir des justifications solides quant aux choix de conception.

- **3/ Une collaboration plus rapide entre les équipes en facteurs humains et les équipes en sûreté des systèmes.**

En donnant un accès facilité aux informations contenues dans les analyses de sûreté, nos outils garantiront des échanges optimisés entre les équipes en facteurs humains et les équipes en sûreté des systèmes. Puisqu'en milieu sociotechnique complexe, l'utilisation du jugement d'experts est une source importante d'informations utiles pour la conception, assurer une collaboration efficace entre ces équipes aura un impact direct sur l'amélioration du processus de conception.

IV Discussion autour des méthodes utilisées et perspectives

Les méthodes que nous avons utilisées ont bien évidemment fait l'objet d'un choix quant à leur pertinence par rapport au problème posé. Nous proposons ici de discuter ces choix et de proposer d'éventuelles améliorations que nous pourrions prendre en compte en **perspective de ce travail** de thèse.

Le premier choix porte sur la méthode de formalisation des scenarii sous forme de vecteurs (cette formalisation est toujours nécessaire pour les algorithmes d'analyse automatique de textes). Nous avons choisi la formalisation défendue par Losee (1998) qui permet de décrire, avec un formalisme cohérent, des textes qui peuvent avoir des contenus très divers. Cette formalisation est utilisée dans de nombreux outils de traitement automatique de textes comme, par exemple, le filtrage du courrier indésirable dans les messageries électroniques.

Cependant, de nombreuses études sur les mesures de similarité sémantique utilisent des ontologies[13] pour formaliser les textes à analyser. Ces ontologies permettent de faire des liens entre des mots différents qui ont des sens proches. Par exemple « fuel » et « carburant », seront considérés comme proches grâce à une ontologie alors qu'ils seront considérés comme deux mots différents par notre formalisme. Nous pouvons justifier notre choix, car les ontologies sont souvent utilisées pour comparer des textes de domaines différents contenant parfois plusieurs millions de mots. Dans notre

[13] Une ontologie est un recueil de termes organisés en réseaux sémantiques. Grâce à cette organisation, il est alors possible de donner la proximité sémantique entre deux termes présents dans l'ontologie.

cas, les textes étudiés portaient sur le même système, avait été rédigés par un nombre limité de personnes et contenaient au maximum 1381 mots. Cependant, **l'utilisation d'ontologies pourrait constituer un axe d'amélioration et de perspective de nos travaux.**

Nous avons également choisi d'utiliser l'algorithme des plus proches voisins (Nearestneighbor : Cover et Dasarathy, (1967) et Dasarathy, (1991)). qui est principalement utilisé dans les systèmes de raisonnement à partir de cas (Kolodner, 1992 ; Aamodt, et al., 1994 ; Maher, et al., 1995). Même si cet algorithme est particulièrement adapté à notre problème, il existe plusieurs méthodes de mesure de la distance entre deux scénarii :

- La distance euclidienne (que nous avons utilisée) ;
- La distance Manhattan ;
- La distance angulaire ;
- La distance de Minkowski ;
- Etc.

Là encore, en perspective de notre travail, nous pourrons comparer les résultats que nous avons obtenus en utilisant la distance euclidienne, avec ceux issus des autres types de mesures présentées ci-dessus.

Pour finir, il existe beaucoup d'algorithmes permettant, après une phase d'apprentissage, d'inférer des règles de classification de bases de données. Les méthode bayesiennes en font partie, mais aussi les réseaux de neurones (Perceptron Multicouches : Rumelhart, et al., 1986), ou les machines à vecteurs de support (Vapnik, et al., 1963 ; Duda, et al., 1973 ; Duda, et al., 2001).

Nous nous sommes intéressés au filtrage bayésien puisque son application actuelle est particulièrement adaptée à notre problématique. En effet, les réseaux de neurones ont des temps d'apprentissages beaucoup plus long et nécessitent un grand nombre d'itérations durant la phase d'apprentissage. De plus, le choix de l'architecture du réseau et le réglage des paramètres d'apprentissage sont souvent délicats notamment pour des problèmes contenant de multiples informations en entrée.

Cependant, ces méthodes ont montré leur efficacité dans de très nombreux domaines, et pourraient également être comparées à la méthode de filtrage bayésien en perspective de ce travail.

V Epilogue

Les résultats obtenus durant cette thèse sur les méthodes de traitement automatique de textes pour outiller la conception centrée sur l'utilisateur, encouragent le développement d'outils « industriels ». Dès 2012, nous allons définir un cahier des charges guidant la création d'outils d'analyse automatique de textes :

- 1/ Destinés aux experts en facteurs humains et en sûreté des systèmes ;
- 2/ S'insérant dans les processus de conception définis aujourd'hui ;
- 3/ Permettant l'extraction facilitée d'hypothèses discutables prises sur le comportement de l'équipage.

La mise en place effective de ces outils dans le cadre d'une démarche de conception centrée utilisateur participera, nous le souhaitons, à la diminution des risques liés aux erreurs humaines et de ce fait à l'amélioration de la sûreté du transport aérien.

RÉFÉRENCES BIBLIOGRAPHIQUES

European Union. 2008. *REGULATION (EC) No 216/2008 : Common rules in the field of civil aviation and establishing a European Aviation Safety Agency, and repealing Council Directive.* s.l. : Official Journal of the European Union, 2008.

Aamodt, A et Plaza, E. 1994. *Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches.* s.l. : AI Communications. IOS Press, Vol. 7: 1, pp. 39-59., 1994.

AFNOR : Association française de Normalisation. 2003. *NF X50-110 : Qualité en expertise - Prescriptions générales de compétence pour une expertise.* s.l. : AFNOR : Association française de Normalisation, 2003.

Amalberti, René. 2003. *Automatisation, gestion de l'erreur humaine et approche écologique.* s.l. : Ingénierie Cognitive (p 81-98) ; Hermes sciences publications., 2003.

Amalberti, Réné. 1996. *La conduite de systèmes à risques.* s.l. : Le Travail Humain, 1996.

—. 2001. *The paradoxes of almost totally safe transportation systems.* s.l. : Safety Science, 37(2-3), 109-126., 2001.

Annett, J et Duncan, K. D. 1967. *Task analysis and training design.* s.l. : Journal of Occupational Psychology, 41, 211-221., 1967.

Baber, Chris et Stanton, Neville A. 2002. *Task analysis for error identification: theory, method and validation.* s.l. : Theoretical Issues in Ergonomics Science, 3 (2) 212-227. Ed Taylor & Francis, 2002.

Bainbridge, L. 1987. *Ironies of automation.* s.l. : New technologies and human errors. New York : Willey., 1987.

Bennett, H L. 1983. *Remembering drink orders: The memory skill of cocktail waitresses.* s.l. : Human Learning, 2, 157-169., 1983.

Boeing Company. 2010. *Statistical Summary of Commercial Jet Airplane Accidents Worldwide Operations 1959 - 2010.* s.l. : Boeing Company, 2010.

Bowers, C, et al. 1995. *Impact of automation on aircrew communication and decision-making performance.* s.l. : International Journal of Aviation Psychology, 1995.

Boy, Guy. 1998. *Cognitive Function Analysis.* s.l. : Stanford,CT : Abbex publishing Corporation, 1998.

Caelen, J, Jambon, F et Vidal, A. 2006. *Conception participative : des « Moments» à leur instrumentation.* s.l. : Revue d'Interaction Homme-machine, Vol 6, N°2/2005-12-01, 2006.

Carlier, Xavier et Hoc, Jean-Michel. 1999. *Role of a common frame of reference in cognitive cooperation: sharing tasks in air-traffic control.* s.l. : CSAPC'99, 1999.

Cover, T et Hart, P. 1967. *Nearest neighbor pattern classification.* s.l. : IEEE Transactions on Information Theory 13, 21-27., 1967.

Dasarathy, B V. 1991. *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques.* . s.l. : IEEE Computer Society Press., 1991.

De Groot, A. 1978. *Thought and choice and chess.* s.l. : The Hague: Mouton. (Original work published 1946)., 1978.

De Terssac, G et Chabaud, C. 1990. *Référentiel commun et fiabilité.* s.l. : Les Facteurs Humains de la fiabilité, Toulouse Octares, 1990.

Demaria, C. 2006. *Understanding Airmanship.* s.l. : CFI Aviation Channel, 2006.

Deutcher, I. 1974. *What we say/ What we do.* s.l. : The Sociological Quarterly,vol. 15 (3), pp 457-461., 1974.

Dreyfus, Stuart E et Dreyfus, Hubert L. 1980. *A Five-Stage Model of the Mental Activities Involved in Directed Skill Acquisition.* . s.l. : Washington, DC: Storming Media., 1980.

Duda, Richard O., Hart, Peter E. et Stork, David G. 1973. *Pattern classification.* s.l. : Wiley-interscience, 1973.

—. 2001. *Pattern classification.* s.l. : Wiley-interscience, 2001.

Endsley, M R et and Kiris, E O. 1995. *The out-of-the-loop performance problem and level of control in automation.* . s.l. : Human Factors, 37 (2), p 381-394, 1995.

Endsley, M. R. 2004. *Situation awareness: Progress and directions.* s.l. : In S. Banbury & S. Tremblay (Eds.), A cognitive approach to situation awareness: Theory, measurement and application (pp. 317–341). Aldershot, UK: Ashgate, 2004. pp. 317–341.

Ericsson, K A et Polson, P. 1988. *Experimental analysis of the mechanisms of memory skill.* s.l. : Journal of Experimental Psychology: Learning, Memory and Cognition 14, 305–31, 1988.

European Aviation Safety Agency. 2011. *Certification Specifications and Acceptable Means of Compliance for Large Aeroplanes.* s.l. : EASA, 2011.

Fischer, G, et al. 1996. *Making argumentation serve design.* s.l. : In Design Rationale: Concepts, techniques and use (p 267-293). Hillsdale : Laurence Erlbaum., 1996.

Grice, H.P. 1975. *Logic and conversation.* s.l. : P.Cole & J. Morgan (Eds), Syntax and semanctics 3: Speech acts., 1975.

Grudin, J. 1996. *Evaluating opportunities for design capture.* s.l. : In Design rationale: Concepts, techniques and use (p 453-470), Hillsdale : Laurence Erlbaum., 1996.

Heinrich, H W, Petersen, D et Ross, N. 1980. *Industrial accident prevention.* s.l. : McGraw-Hill: New York, 1980.

Hermann, Eric, Leblois, Sylvie et Mazeau, Michel. 2008. *Outils de traitement automatique des langues appliqués au comptes-rendus d'incidents et d'accidents.* s.l. : Actes du congrés Lambda mu 16, 2008.

Hoc, Jean-Michel. 2003. *Coopération humaine et systèmes coopératifs.* s.l. : Ingénierie Cognitive (p139-187), Hermes sciences publications, 2003.

—. 1998. *How can we evaluate the quality of human-machine cooperation?* s.l. : COOP'98, Third International Conference on the Design of Cooperative systems (p121-130), 1998.

Hollnagel, E et Woods, D. D. 2005. *Joint cognitive systems: Foundations of cognitive systems engineering.* s.l. : Boca Raton, FL: CRC Press/Taylor & Francis., 2005.

Hollnagel, E. 1991. *The Phenotype of Erroneous Action: Implication for HCI Design.* s.l. : Human-Computer Interaction in Complex Systems. Academic Press, 1991.

Hollnagel, E, et al. 2008. *A functional resonance accident analysis of Comair flight 5191.* s.l. : Paper presented at the 8th International Symposium of the Australian Aviation Psychology Association. Sydney, Australia., 2008.

Hollnagel, E, Woods, D. D et & Leveson, N. 2006. *Resilience engineering:Concepts and precepts.* s.l. : Aldershot: Ashgate, 2006.

Hollnagel, Eric. 2004 . *Barriers and accident prevention.* s.l. : Aldershot, UK: Ashgate., 2004 .

Hopkin, V.D. 1994. *Human performance implication of air traffic control automation.* s.l. : Human performance in automated systems: Current research and trends (p 314-319). Hillsdale, NJ: Laurence Erlbaum Associates., 1994.

Hughes, David et Dornheim, Michael. 1995. *Accidents Direct Focus on Cockpit Automation.* s.l. : Aviation Week and Space Technology, 52–54, 1995.

International Organization for Standardization. 1999. *ISO 13407 : Human-centred design processes for interactive systems.* s.l. : International Organisation for Standardization, 1999.

—. 2011. *ISO 15504 : Information technology -- Process assessment.* s.l. : International Organization for Standardization, 2011.

—. 1998. *ISO 9241 : Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability.* s.l. : International Organisation for Standardization, 1998.

—. 2002. *ISO/TR 16982 : Ergonomics of human-system interaction -- Usability methods supporting human-centred design.* s.l. : International Organization for Standardization, 2002.

—. 2000. *ISO/TR 18529 : Ergonomics -- Ergonomics of human-system interaction -- Human-centred lifecycle process descriptions.* s.l. : International Organization for Standardization, 2000.

Jambon, Francis. 1998. *Taxonomy for Human Error and System Fault Recovery from the Engineering Perspective.* s.l. : International Conference on Human-Computer Interaction in Aeronautics (HCI-Aero'98), Montréal, Canada, 27-29 may 1998. p. 55-60., 1998.

Jones, D G et Endsley, M R. 1996. *Sources of situation awareness errors in aviation.* s.l. : Aviation, Space, and Environmental Medicine, 67(6), 507-512., 1996.

Jouanneaux, Michel. 1999. *Le pilote est toujours devant : Reconnaissance de l'activité du pilote de ligne.* s.l. : Octares Editions, 1999.

Kantowitz, B.H. 1994. *Pilot workload and fly deck automation.* . s.l. : In R. Parasuraman & M. Mouloua (Eds.), Human performance in automated systems : current research and trends, (pp 212-223). Hillsdale: Laurence Erlbaum Associates., 1994.

Karsenty, Laurent. 2000. *Cooperative work: the role of explanation in creating a share problem representation.* s.l. : Le Travail Humai, 63, 289-309, 2000.

—. 1996. *Une définition psychologique de l'explication.* s.l. : Intellectica, 96, 327-345, 1996.

Kern, A. 1997. *Redefining Airmanship.* s.l. : McGraw-Hill Professional. pp. 21, 1997.

Kolodner, J L. 1992. *An introduction to case-based reasoning.* . s.l. : Artificial Intelligence Review, 6(3), 3-34., 1992.

Kotonya, G et Sommerville, I. 1998. *Requirements engineering processes and techniques.* s.l. : John Wiley & sons, 1998.

Lannoy, André et Procaccia, Henri. 2001. *L'utilisation du Jugement d'Expert en Sûreté de Fonctionnement.* s.l. : Editions : Technique et Documentation, 2001.

Lee, J. 1997. *Design rationale systems: understanding the issues.* s.l. : IEEE expert: Intelligent Systems and their applications, Issue 2, (p78-85), 1997.

Lee, J et Lai, K Y. 1996. *What's in design rationale?* s.l. : In Design Rationale: Concepts, techniques and use (p 21-51). Hillsdale : Laurence Erlbaum., 1996.

Losee, Robert M. 1998. *Text retrieval and Filtering : Analytic Models of Performance.* s.l. : Kluwer Academic Publishers, 1998.

MacLean, A, et al. 1996. *Question, Options and Criteria: Elements of Design Space Analysis.* s.l. : In Design Rationale: Concepts, Techniques and use (p 53-105). Hillsdale : Laurence Erlbaum., 1996.

MacLean, A, Young, R M et Moran, T P. 1989. *Design rationale: the argument behind the artifact.* s.l. : In CHI'99 (p 247-252). New York: ACM., 1989.

Maher, M L, Balachandran, M. B. et M., Zhang. D. 1995. *Case-based reasoning in design.* . s.l. : Mahwah, NJ: Lawrence Erlbaum Associates., 1995.

Marwel, G et Schmitt, D. 1975. *Co-operation - an experimental analysis.* s.l. : Londres, Academic Press, 1975.

Mawhin, B. 2011. *Integrating aircrew resources variability into the design of future cockpits.* s.l. : Proceedings of the 16th international sympsium on aviation psychology, 2011.

Mayhew, D.J. 1992. *Principles ans Guidelines in Software User Interface Design.* s.l. : Englewood Cliffs, NJ : Prentice-Hall, 1992.

Moran, T P et Carroll, J M. 1996. *Design Rationale: Concepts, techniques and use.* s.l. : Hillsdale : Laurence Erlbaum, 1996.

Nielsen, Jakob. 1993. *Usability Engineering.* s.l. : Cambridge, MA : Academic Press, 1993.

Norman, Donald A. 1988. *The Design of Everyday Things.* s.l. : New York, Doubleday, 1988.

Norman, Donald Arthur et Draper, Steve. 1986. *User centered system design : New perspectives on human-computer interaction.* s.l. : Lawrence Erlbaum Associates, 1986.

O'Hare, D., et al. 1994. *Cognitive failure analysis for aircraft accident investigation.* s.l. : Ergonomics, 1994.

Orasanu, J. et Martin, L. 1999. *Errors in Aviation Decision Making: A Factor in Accidents and Incidents.* s.l. : HESSD ’98, 1999.

Ormerod, T C, et al. 1999. *Desperado: Three in one indexing for innovative design.* s.l. : Human Computer Interaction INTERACT'99 (p 336-343). Amsterdam : IOS Press., 1999.

Papaïx, Benoît, et al. 2010. *System Safety and Human Factors: How to forecast crew behaviour confronted to a system failure.* s.l. : Actes du Colloque Lambda Mu 17, 2010.

Parasuraman, R et Riley, V. 1997. *Humans and Automation: Use, Misuse, Disuse, Abuse.* s.l. : Human Factors: The Journal of the Human Factors and Ergonomics Society, 39(2), 230-253., 1997.

Parasuraman, R, Sheridan, T. B et Wickens, C. D. 2000. *A model of types and levels of human interaction with automation.* s.l. : IEEE Transactions on Systems, Man, and Cybernetics — Part A: Systems and Humans, 2000.

Patel, V L et Groen, G J. 1991. *The general and specific nature of medical expertise: A critical look.* s.l. : In K.A. Ericsson & J. Smith (Eds.). Toward a general theory of expertise: Prospects and limits (pp. 93-125). New York: Cambridge University Press, 1991.

Patel, V L et Groen, G L. 1986. *Knowledge based solution strategies in medical reasoning.* s.l. : Cognitive Science, 10, 91-116., 1986.

Piaget, J. 1965. s.l. : Etudes sociologiques. Genève : Droz, 1965.

Plat, M et Amalberti, R. 2000. *Experimental crew training to deal with automation surprises.* s.l. : Cognitive Engineering in the Aviation Domain (p 287-308). Hillsdale, New Jersey : Lawrence Erlbaum., 2000.

Rabardel, Pierre. 1995. *Les hommes et les technologies.* s.l. : Armand Colin, 1995.

Rasmussen, J et Goldstein, L. 1985. *Decision support in supervisory control.* s.l. : Proceeding of the 2nd conference on task analysis., 1985.

Rasmussen, Jens. 1983. *Skills, rules, knowledge; signals, signs, and symbols, and other distinctions in human performance models* . s.l. : IEEE Transactions on Systems, Man and Cybernetics, 13, 257-266., 1983.

Reason, James. 1990. *Human Error.* s.l. : Cambridge University Press, 1990.

—. 1997. *Managing the Human and Organizational Response to Accidents.* s.l. : Brookfield, Vermont, Ashgate Publishing, 1997.

—. 1997. *Managing the risks of organizational accidents.* s.l. : Aldershot, UK: Ashgate, 1997.

Reuzeau, Florence. 2000. *Assister l'évaluation participative des systèmes complexes : rôle des savoirs et savoir-faire des utilisateurs dans la conception d'un poste de pilotage d'avion.* s.l. : Thèse de doctorat, CNAM., 2000.

Reuzeau, Florence et Papaïx, Benoît. 2010. *De l'intérêt d'observer l'erreur humaine en situation de simulation de vol : prédire et concilier la relation humain-machine.* s.l. : Ergo'IA 2010 proceedings. New York : ACM., 2010.

Reuzeau, Florence. 2005. *La conjuration de l'erreur humaine: accepter ou refuser le risque, lequel ?* s.l. : IHM 2005 Proceedings of the 17th international conference on Francophone sur l'Interaction Homme-Machine. New York: ACM., 2005.

Robert, Jean-Marc. 2003. *Que faut-il savoir des utilisateurs pour réaliser des interfaces de qualité?* s.l. : Ingénierie cognitive (p249-283), Hermes sciences publications, 2003.

Roscoe, A. 1989. *Pilot workload and automation.* s.l. : Flight Safety Foundation, 42nd IASS Athens, 1989.

Rosenblatt, Frank. 1958. *The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain.* s.l. : Cornell Aeronautical Laboratory, Psychological Review, v65, No. 6, pp. 386–408, 1958.

Rumelhart, D. E., Hinton, McClelland et Williams, R. J. 1986. *Learning Internal Representations by Error Propagation.* s.l. : Parallel Distributed Processing: Explorations in the Microstructure of Cognition, 1986.

S-18, Aircraft And System Developement And Safety Assessment Committee. 2010. *Guidelines for Development of Civil Aircraft and Systems.* s.l. : SAE International, 2010.

S-18, Aircraft and System Development and Safety Assessment Committee. 1996. *Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment.* s.l. : SAE international, 1996.

Sahami, M., et al. 1998. *A Bayesian Approach to Filtering Junk E-Mail.* s.l. : AAAI'98 Workshop on Learning for Text Categorization, 1998.

Sarter, N B et Alexander, H M. 2000a. *Error Types and Related Error Detection Mechanisms in the Aviation Domain: An Analysis of ASRS Incident Reports.* s.l. : International Journal of Aviation Psychology, 10(2), 189-206., 2000a.

Sarter, N B et Woods, D D. 1995. *Strong, silent and " out-of-the-loop "; properties of advanced automation and their impact on human automation interaction.* s.l. : Colombus: Ohio State University Press., 1995.

Sarter, N et Woods, D D. 1992. *Pilot interaction with cockpit automation. Operational experiences with the flight management system.* s.l. : International Journal of Aviation Psychology, 1992.

—. 2000b. *Team play with a powerful and independent agent: a full-mission simulation study.* s.l. : Human Factors, 39, pp. 553-569, 2000b.

Schneider, Wolfgang, Matthias, Schlagmüller et Mechtild, Visé. 1998. *The impact of metamemory and domain-specific knowledge on memory performance.* s.l. : European Journal of Psychology of Education, 13, 91-103., 1998.

Shanteau, James. 1992. *Competence in experts: The role of task characteristics.* . s.l. : Organizational Behavior and Human Decision Processes, 53, 252-266, 1992.

Shappell, S A et Wiegmann, D A. 2000. *The Human Factors Analysis and Classification System-HFACS.* s.l. : Department of Transportation, Federal Administration Aviation , 2000.

Shappell, S, et al. 2007. *Human error and commercial aviation accidents: an analysis using the human factors analysis and classification system.* s.l. : Human Factors: The Journal of the Human Factors and Ergonomics Society., 2007.

Sommerville, I et Sawyer, P. 1997. *Requirements engineering : A good practice guide.* s.l. : John Wiley & sons, 1997.

Stanton, Neville et Wilson, Jennifer. 2000 . *Human factors: step change improvement in effectiveness and safety.* s.l. : Drilling Contractor, 2000 .

Stanton, Neville. 2003. *Human Error Identification In Human-computer Interaction* . s.l. : In: J. A. Jacko and A. Sears (Eds) The Human-Computer Interaction Handbook, Lawrence Erlbaum Associates, 2003.

Swain, A D et Guttmann, H E. 1983. *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications.* . s.l. : NUREG/CR-1278, USNRC., 1983.

Vapnik, Vladimir et Lerner, A. 1963. *Pattern Recognition using Generalized Portrait Method.* s.l. : Automation and Remote Control, 1963.

Vredenburg, K, Isensee, S et Righi, C. 2001. *User-Centered Design : An Integrated Approach.* s.l. : Prentice Hall, 2001.

WG-63, Complex Aircraft Systems. 2007. *Guidelines and methods for conducting the safety and assessment process on civil airborne systems and equipment.* s.l. : Eurocae, 2007.

—. 2007. *Guidelines for Development of Civil Aircraft and Systems.* s.l. : Eurocae, 2007.

Wiener, E. 1989. *Human factors of advanced technology ("glass-cockpit").* . s.l. : Transport Aircraft, NASA Contractor Report 177528. , 1989.

Wiener, E L et Curry, R E. 1980. *Flight Deck automation: Promises and problems.* s.l. : Ergonomics, 1980.

Williams, J C. 1985. *HEART – A proposed method for achieving high reliability in process operation by means of human factors engineering technology.* s.l. : Proceedings of a Symposium on the Achievement of Reliability in Operating Plant, Safety and Reliability, 1985.

Woods, D D. 1996. *Decomposing automation: Apparent simplicity, real complexity.* s.l. : Automation and human performance: theory and applications, By R. Parasuraman, Mustapha Mouloua, Taylor and Francis Ltd , 1996.

Young, Mark et Stanton, Neville. 2002. *Attention and automation: New perspectives on mental underload and performance.* s.l. : Theoretical Issues in Ergonomics Science (p.178-194) ; Taylor & Francis Ltd, 2002.

TABLE DES FIGURES

FIGURE 1 : EVOLUTION DU NOMBRE D'ACCIDENTS D'AVION SUR LES 50 DERNIERES ANNEES. 12
FIGURE 2 : PROCESSUS DE CONCEPTION CENTRE UTILISATEUR (ISO 13407)..... 24
FIGURE 3 : LE MODELE SLIPS, LAPSES, MISTAKES, VIOLATIONS..... 31
FIGURE 4 : LE MODELE GEMS..... 33
FIGURE 5 : TAXONOMIE DES PHENOTYPES D'HOLLNAGEL. 35
FIGURE 6 : LE MODELE DES DOMINOS 36
FIGURE 7 : LE SWISS CHEESE MODEL 37
FIGURE 8 : HFACS, UNSAFE ACT. 38
FIGURE 9 : HFACS, PRECONDITIONS FOR UNSAFE ACTS..... 38
FIGURE 10 : HFACS, UNSAFE SUPERVISION 39
FIGURE 11 : HFACS, ORGANISATIONAL INFLUENCES 39
FIGURE 12 : ILLUSTRATION DU MODELE FRAM 41
FIGURE 13 : LA PYRAMIDE AUTO..... 42
FIGURE 14 : APPROCHE RECOMMANDEE, EXTRAIT DE L'AMC25-1302 55
FIGURE 15 : PROCESSUS DE CONCEPTION CENTREE UTILISATEUR MIS EN PLACE CHEZ AIRBUS. 59
FIGURE 16 : SIMULATEUR POUR L'A380..... 62
FIGURE 17 : PROCESSUS D'EVALUATION DE LA SECURITE (EXTRAIT DE L'ARP 4754)..... 74
FIGURE 18 : EXEMPLE D'UN SCENARIO EXTRAIT DE LA FHA DU CDS 76
FIGURE 19 : ILLUSTRATION DES SAUTS DE CLASSIFICATIONS 79
FIGURE 20 : THE SYSTEMATIC HUMAN ERROR REDUCTION AND PREDICTION APPROACH (SHERPA)82
FIGURE 21 : PROCESSUS D'APPLICATION DE LA METHODE TAFEI..... 86
FIGURE 22 : REPRESENTATION GRAPHIQUE DES TEXTES PRECEDENTS 99
FIGURE 23 : PROCESSUS DE CONCEPTION ACTUEL. 104
FIGURE 24 : PROCESSUS OUTILLE POUR REPONDRE AUX BESOINS DES EXPERTS. 105

FIGURE 25 : MODELE D'ANALYSE 109
FIGURE 26 : PROCESSUS DECISIONNEL, ADAPTE DE REASON (1990) 117
FIGURE 27 : REPARTITION DES CLASSIFICATIONS DE SCENARII (DE MIN A CAT) SELON LE SYSTEME CONSIDERE 118
FIGURE 28 : RESULTATS DE LA SELECTION DES SCENARII 119
FIGURE 29 : RESULTATS OBTENUS EN ETUDIANT LES SCENARII SELECTIONNES DU PA A L'AIDE DES DEUX COMPOSANTES IDENTIFIEES PAR JAMBON, 1998. 121
FIGURE 30 : RESULTATS OBTENUS EN ETUDIANT LES SCENARII SELECTIONNES DU CDS A L'AIDE DES DEUX COMPOSANTES IDENTIFIEES PAR JAMBON, 1998. 121
FIGURE 31 : RESULTATS OBTENUS EN ETUDIANT LES SCENARII SELECTIONNES DU PA AVEC LE PROCESSUS DE PRISE DE DECISION (ETAPES LOCALES) 122
FIGURE 32 : RESULTATS OBTENUS EN ETUDIANT LES SCENARII SELECTIONNES DU CDS AVEC LE PROCESSUS DE PRISE DE DECISION (ETAPES LOCALES) 123
FIGURE 33 : REPARTITION DU DESCRIPTEUR « TYPE DE PANNE » (CDS) 132
FIGURE 34 : NOMBRE DE SCENARII PERTINENTS DANS LES PLUS PROCHES VOISINS 144
FIGURE 35 : VALEURS DES ECARTS REDUITS (CDS) 144
FIGURE 36 : NOMBRE DE SCENARII PERTINENTS DANS LES PLUS PROCHES VOISINS 145
FIGURE 37 : EVOLUTION DU POURCENTAGE MOYEN DE SCENARII PERTINENTS RECUPERES PAR LA METHODE DU FILTRAGE BAYESIEN (CDS) 146
FIGURE 38 : VALEURS DES ECARTS REDUITS (CDS) 147
FIGURE 39 : EVOLUTION DU POURCENTAGE MOYEN DE SCENARII PERTINENTS RECUPERES PAR LA METHODE DU FILTRAGE BAYESIEN (PA) 148
FIGURE 40 : VALEURS DES ECARTS REDUITS (PA) 148
FIGURE 41 : APPORT DE L'ETUDE A LA DEMARCHE INDUSTRIELLE 161
FIGURE 42 : REPARTITION DU DESCRIPTEUR "TYPE DE PANNE" (CDS) 180
FIGURE 43 : REPARTITION DU DESCRIPTEUR "DETECTION DE LA PANNE" (CDS) 181
FIGURE 44 : REPARTITION DU DESCRIPTEUR "PHASE DE VOL" (CDS) 181
FIGURE 45 : REPARTITION DU DESCRIPTEUR "TYPE DE PANNE" (PA) 182
FIGURE 46 : REPARTITION DU DESCRIPTEUR "DETECTION DE LA PANNE" (PA) 182
FIGURE 47 : REPARTITION DU DESCRIPTEUR "PHASE DE VOL" (PA) 183
FIGURE 48 : REPARTITION DU DESCRIPTEUR "DETECTION DE LA PANNE" (CDS) 183
FIGURE 49 : REPARTITION DU DESCRIPTEUR "PHASE DE VOL" (CDS) 184
FIGURE 50 : REPARTITION DU DESCRIPTEUR "TYPE DE PANNE" (CDS) 184
FIGURE 51 : REPARTITION DU DESCRIPTEUR "CLASSIFICATION" (CDS) 185
FIGURE 52 : REPARTITION DU DESCRIPTEUR "DETECTION DE LA PANNE" (PA) 185

FIGURE 53 : REPARTITION DU DESCRIPTEUR "PHASE DE VOL" (PA) 186
FIGURE 54 : REPARTITION DU DESCRIPTEUR "TYPE DE PANNE" (PA) 186
FIGURE 55 : REPARTITION DU DESCRIPTEUR "CLASSIFICATION" (PA) 187
FIGURE 56 : PREPARATION DES DONNEES POUR LES ALGORITHMES 195
FIGURE 57 : ALGORITHME DES PROCHES VOISINS 196
FIGURE 58 : ALGORITHME DE FILTRAGE BAYESIEN. 197

LISTE DES TABLEAUX

TABLEAU 1 : EXTRAIT DU CS 25 1309 15
TABLEAU 2 : LE MODELE SRK 31
TABLEAU 3 LE MODELE DES PHENOTYPES ET GENOTYPES 34
TABLEAU 4 : CRITICITE ET EXIGENCES QUALITATIVES ASSOCIEES 76
TABLEAU 5 : TABLE DES CLASSES DE LA METHODE HEART 83
TABLEAU 6 : TABLE DES CONDITIONS FAVORISANT L'ERREUR HUMAINE 84
TABLEAU 7 : TESTS DU KHI² POUR LE PILOTE AUTOMATIQUE 122
TABLEAU 8 : TESTS DU KHI² POUR LE CONTROL AND DISPLAY SYSTEM 123
TABLEAU 9 : EVOLUTION DES DESCRIPTEURS EN FONCTION DE LA CLASSIFICATION 131
TABLEAU 10 : EVOLUTION DES DESCRIPTEURS ENTRE LES SCENARII SELECTIONNES PAR LES EXPERTS ET LA BASE DE DONNEES INITIALE. 134
TABLEAU 11 : DESCRIPTEURS PERTINENTS OBTENUS EN FONCTION DE LA METHODE UTILISEE POUR LE CDS 135
TABLEAU 12 : DESCRIPTEURS PERTINENTS OBTENUS EN FONCTION DE LA METHODE UTILISEE POUR LE PA 136
TABLEAU 13 : NOMBRES DE FONCTIONS (HAUT NIVEAU, INTERMEDIAIRE, BAS NIVEAU) ET NOMBRE DE SCENARII POUR LES DEUX SYSTEMES ETUDIES 150
TABLEAU 14 : NOMBRE DE FONCTIONS DANS LES SCENARII SELECTIONNES 150
TABLEAU 15 : SCHEMA DE SYNTHESE DES RESULTATS DE L'ETAPE 1 154
TABLEAU 16 : SCHEMA DE SYNTHESE DES RESULTATS DE L'ETAPE 2 155
TABLEAU 17 : SCHEMA DE SYNTHESE DES RESULTATS DE L'ETAPE 3 156

LISTE DES ABREVIATIONS

AESA :	Agence Européenne de Sureté Aérienne
AFNOR :	Association Française de Normalisation
AMC :	Acceptable Means of Compliance
AP :	Auto-Pilot
ATA :	Air Transportation Association
AUTO :	Artefact User Task Organisation
CAT :	Catastrophic
CBR :	Case Based Reasoning :
CDS :	Control and Display System
CPC :	Common Performance Conditions
CS :	Certification Specifications
DAL :	Design Assurance Level
EASA :	European Aviation Safety Agency
FC :	Failure Condition
FH :	Facteurs Humains
FHA :	Functional Hazard Assessment
FRAM :	Functional Resonance Analysis Method
GEMS :	Generic Error Modelling System
HAZ :	Hazardous
HEART :	Human Error Assessment & Reduction Technique
HF :	Human Factors
HFACS :	Human Factors Analysis and Classification System
HTA :	Hierarchical Task Analysis
MAJ :	Major

MIN :	Minor
ND :	Navigation Display
PA :	Pilote Automatique
PFD :	Primary Flight Display
PSSA :	Preliminary System Safety Analysis
SA :	Situation Awareness
SAS :	Société par Action Simplifiée
SHERPA :	Systematic Human Error Reduction & Prediction Approach
SRK :	Skill Rule Knowledge
SSA :	System Safety Analysis
TAFEI :	Task Analysis for Error Identification
THERP :	Technique for Human Error Rate Prediction

ANNEXES

Les informations présentées ici présentent un intérêt pour la compréhension de cette thèse, mais ne nécessitaient pas une étude approfondie à inclure dans le corps de la thèse à proprement parlé. Les chapitres présents en annexes sont les suivants :

- Chapitre 1 : Des descripteurs pour des requêtes pertinentes : Analyses Statistiques.
- Chapitre 2 : Précisions sur le *« basic Airmanship »*.
- Chapitre 3 : Architecture des programmes informatique.

CHAPITRE. 1 DES DESCRIPTEURS POUR DES REQUETES PERTINENTES : ANALYSES STATISTIQUES

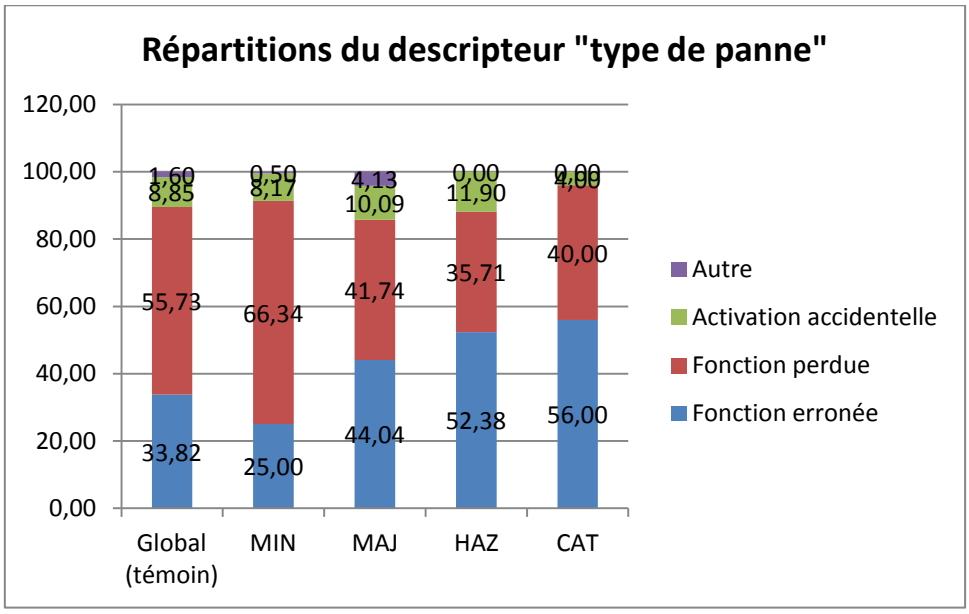


Figure 42 : Répartition du descripteur "Type de panne" (CDS)

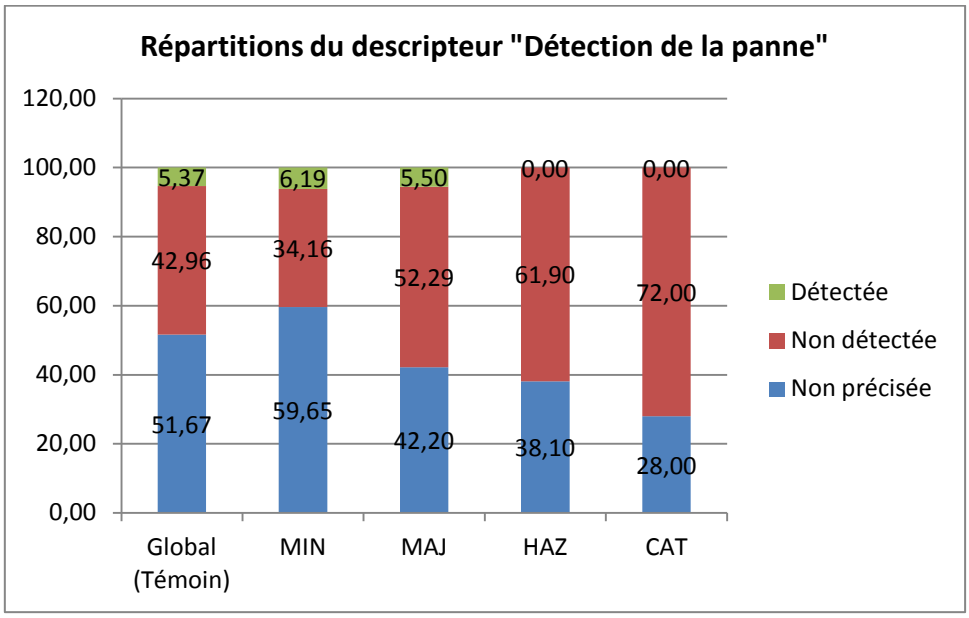


Figure 43 : Répartition du descripteur "Détection de la panne" (CDS)

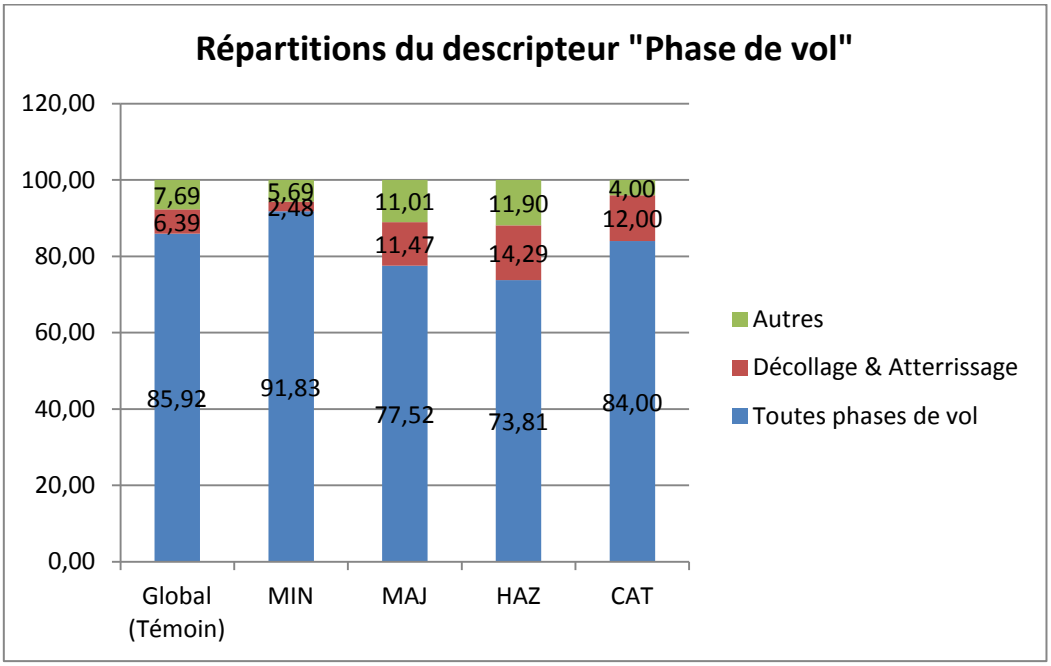


Figure 44 : Répartition du descripteur "Phase de vol" (CDS)

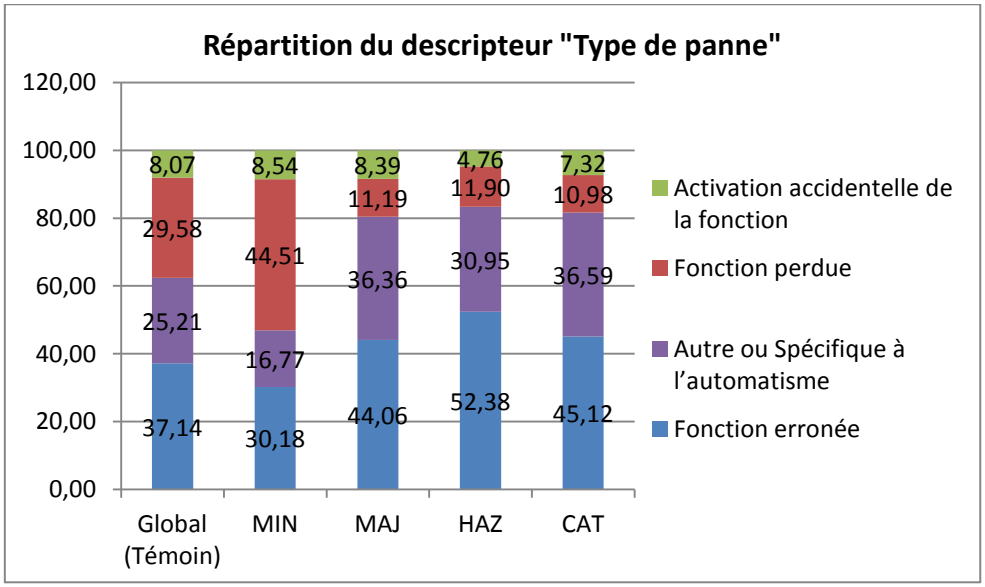


Figure 45 : Répartition du descripteur "Type de panne" (PA)

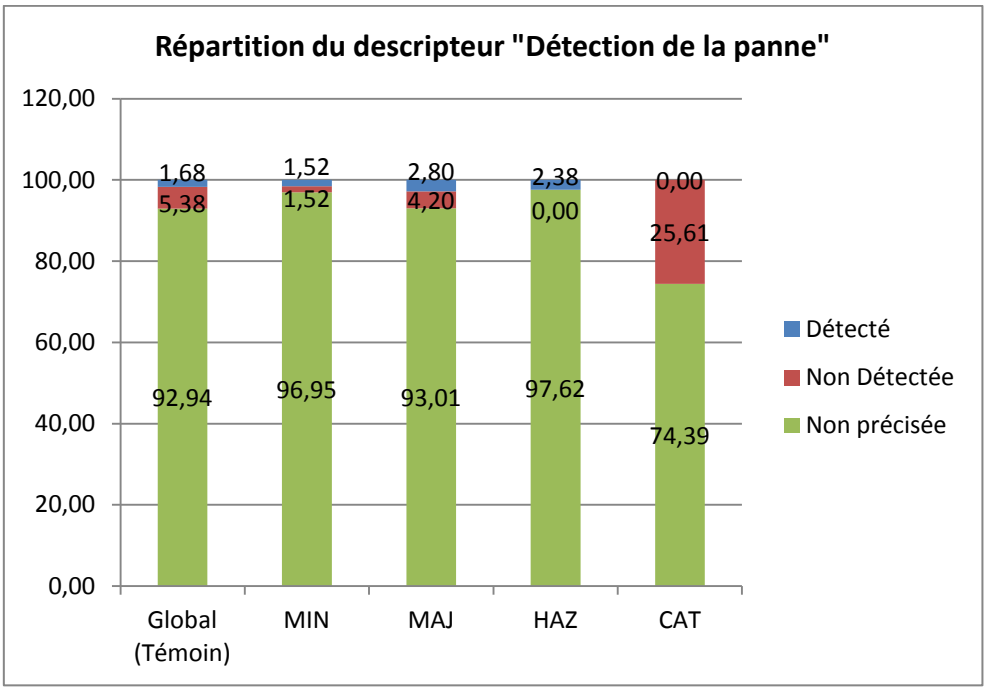


Figure 46 : Répartition du descripteur "Détection de la panne" (PA)

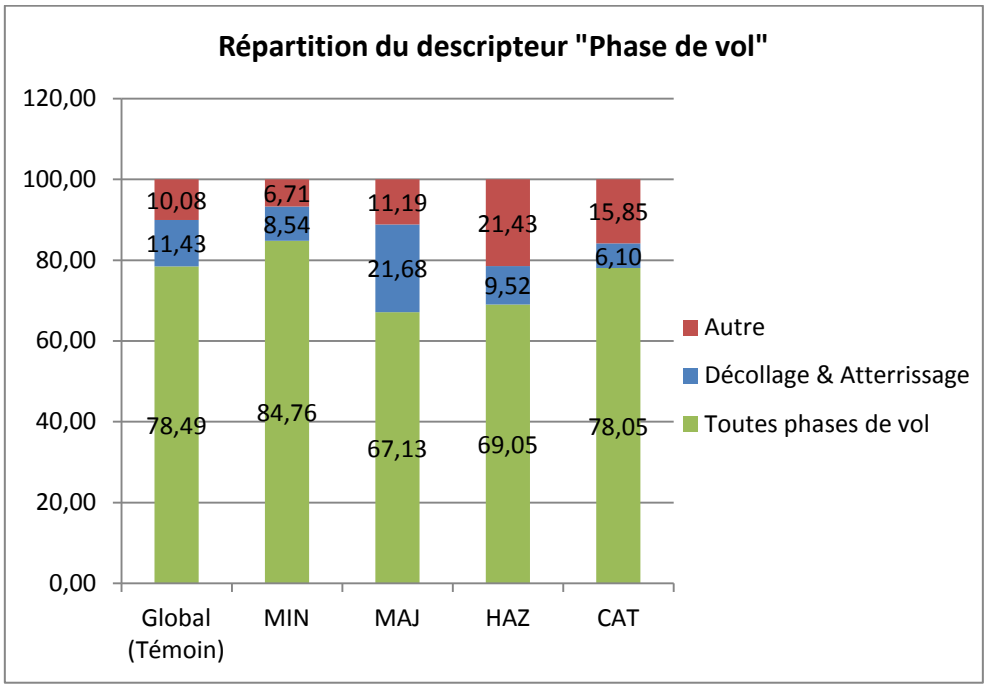


Figure 47 : Répartition du descripteur "Phase de vol" (PA)

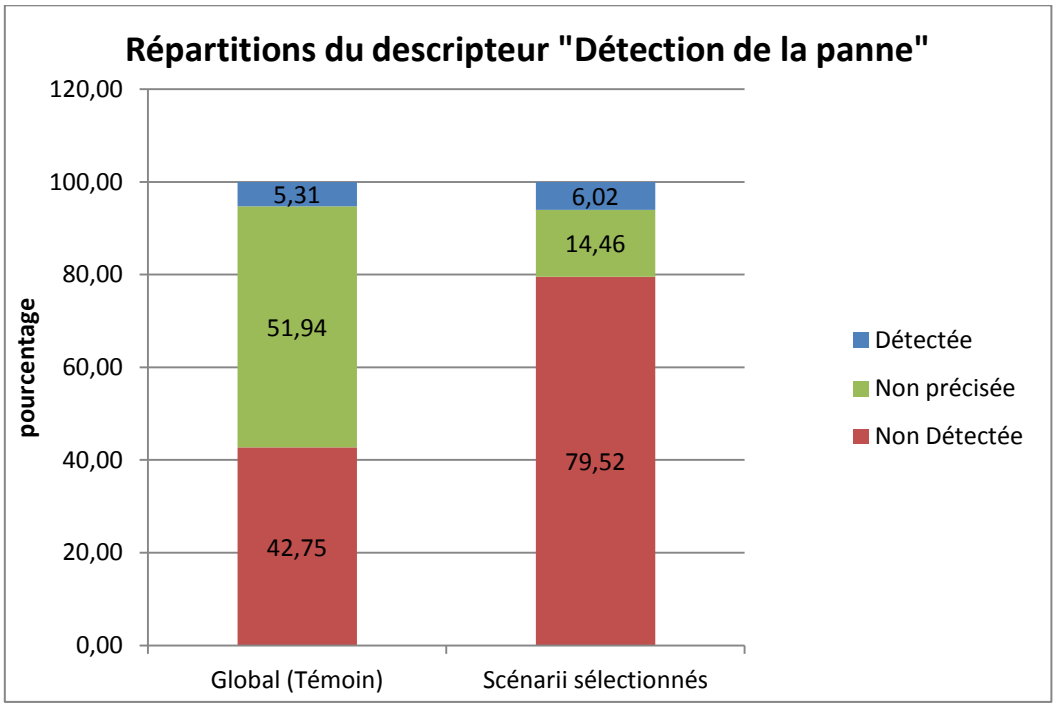


Figure 48 : Répartition du descripteur "Détection de la panne" (CDS)

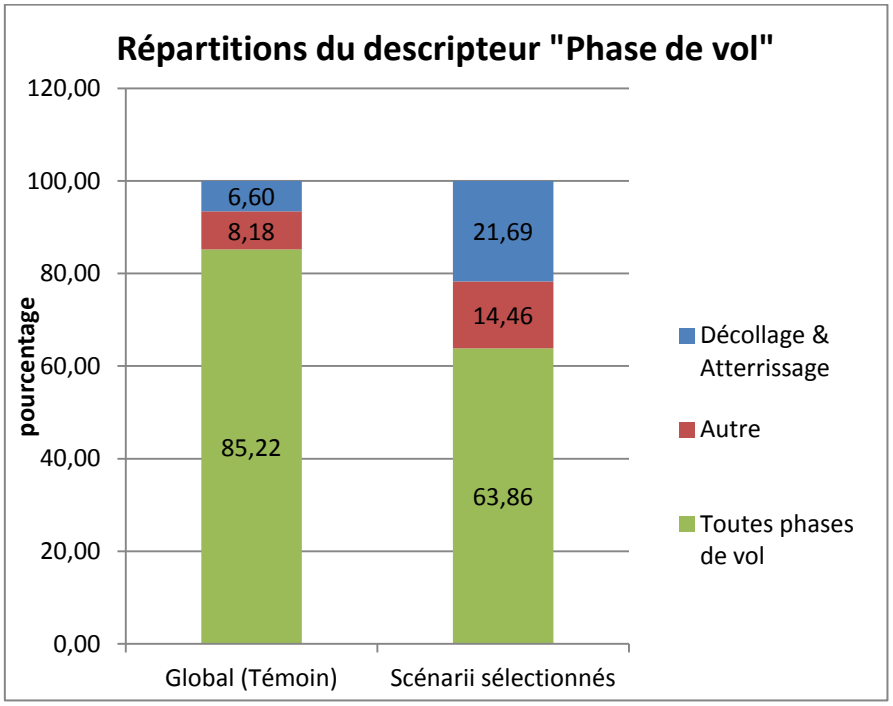


Figure 49 : Répartition du descripteur "Phase de vol" (CDS)

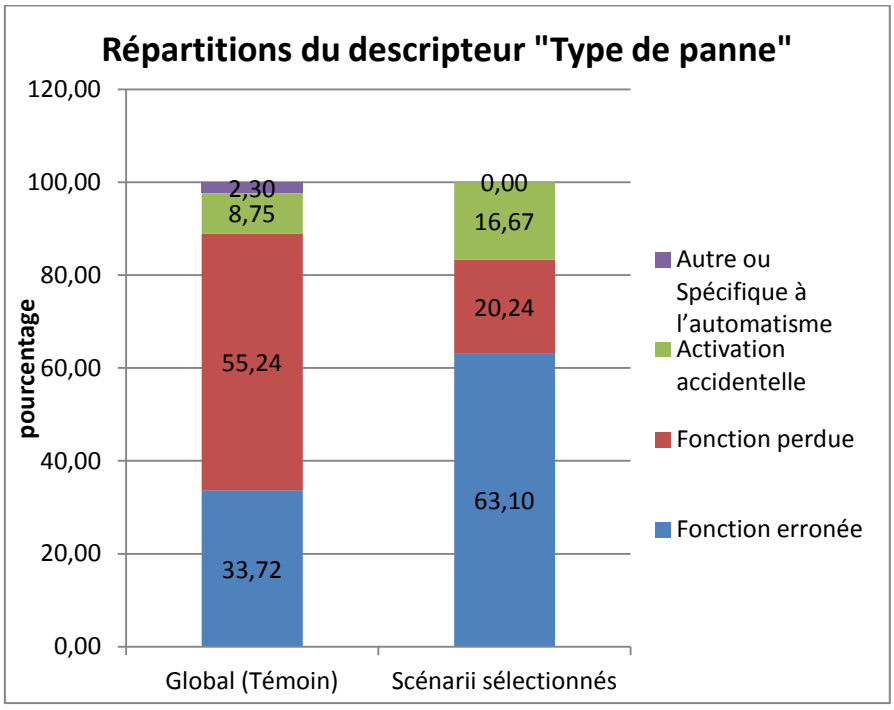


Figure 50 : Répartition du descripteur "Type de panne" (CDS)

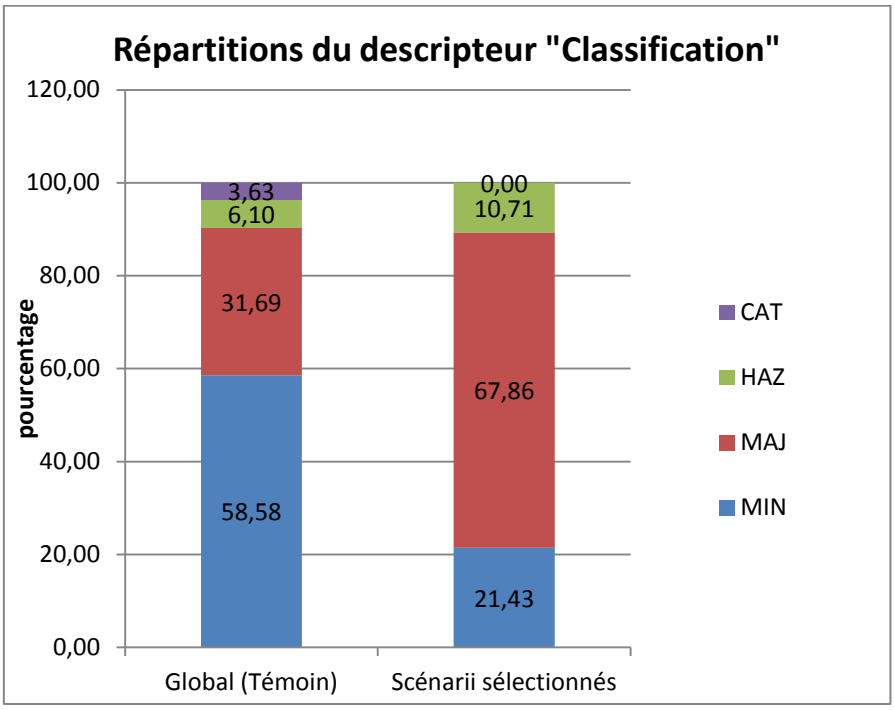


Figure 51 : Répartition du descripteur "Classification" (CDS)

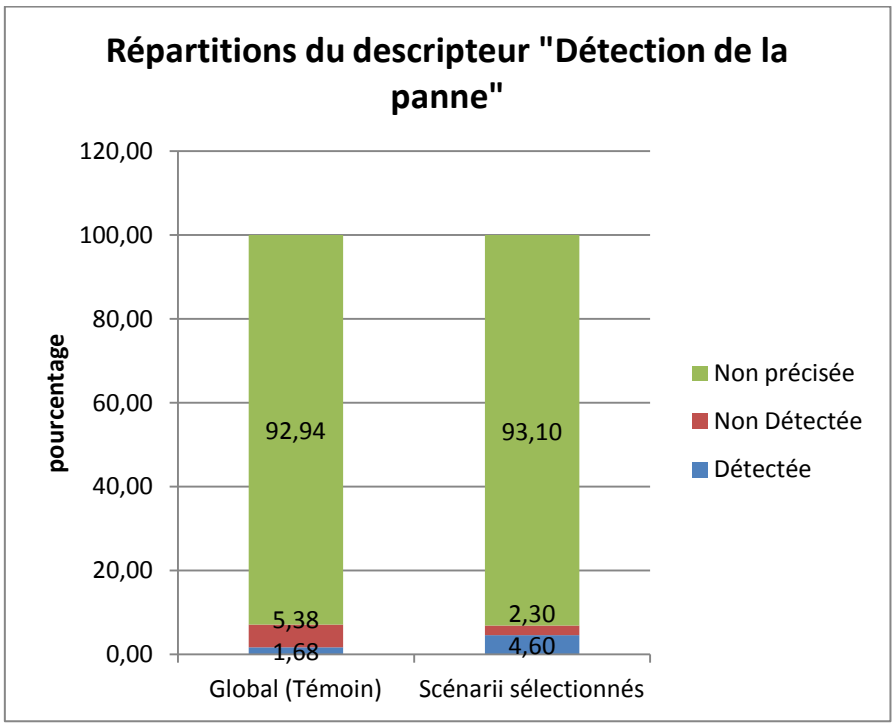


Figure 52 : Répartition du descripteur "Détection de la panne" (PA)

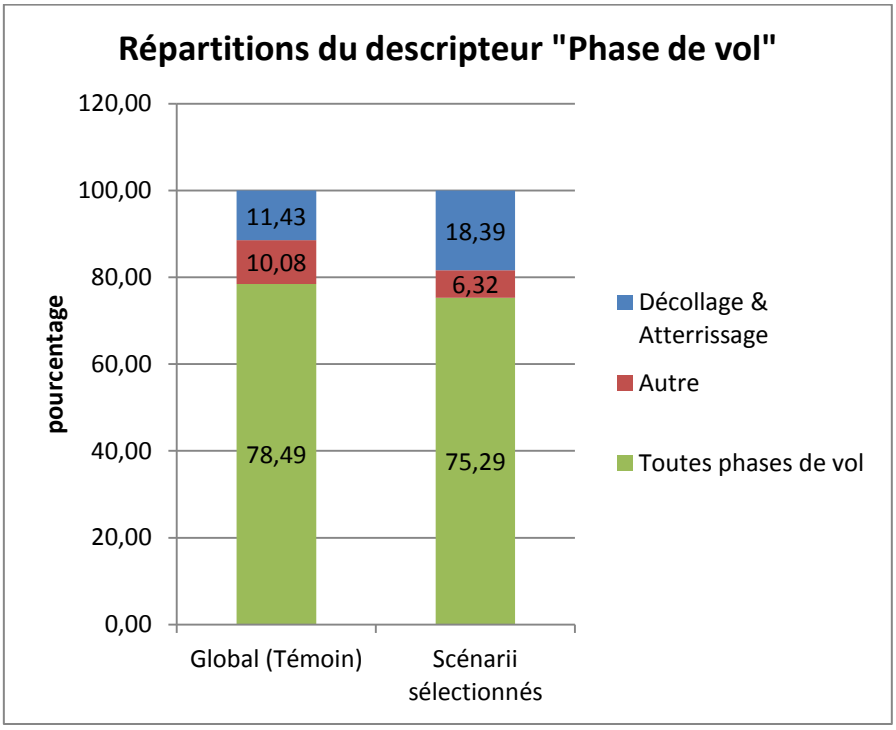


Figure 53 : Répartition du descripteur "Phase de vol" (PA)

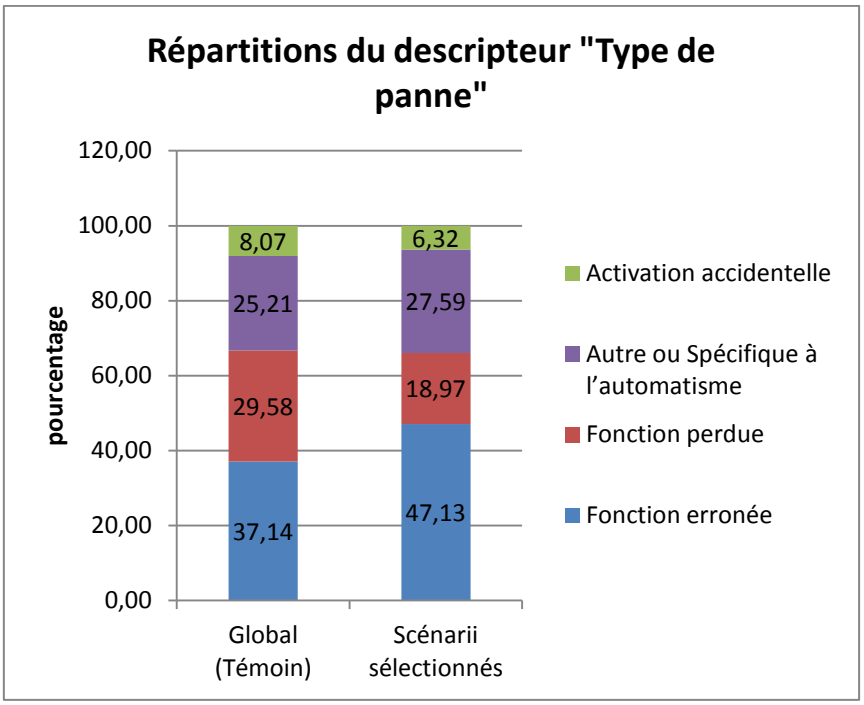


Figure 54 : Répartition du descripteur "Type de panne" (PA)

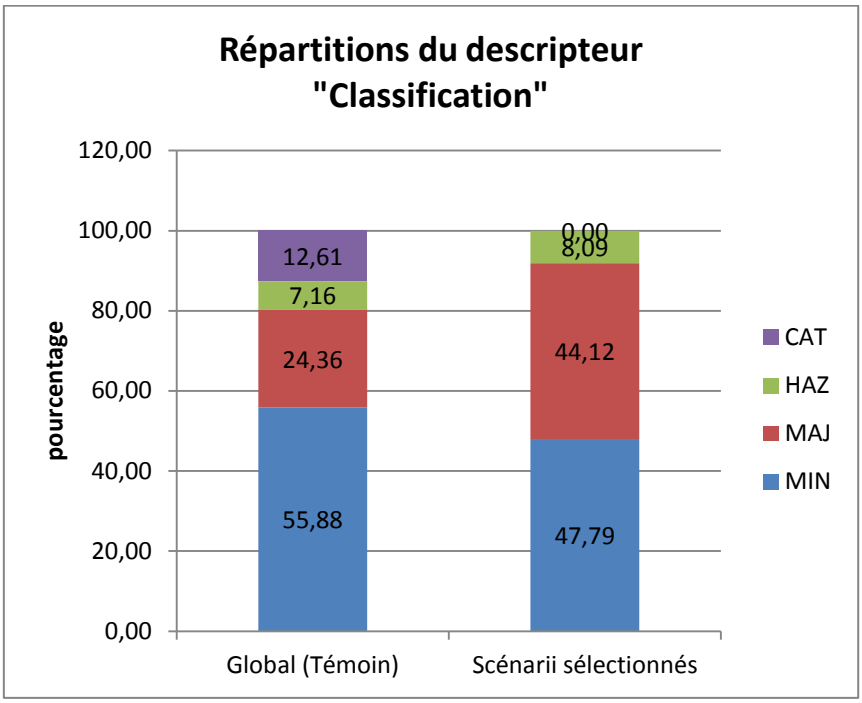


Figure 55 : Répartition du descripteur "Classification" (PA)

CHAPITRE. 2 PRECISIONS SUR LE « BASIC AIRMANSHIP »

Comme son nom l'indique, l'«Airmanship » correspond aux compétences et connaissances mise en œuvre dans la pratique du transport aérien. Le comportement pilote couvre une gamme de comportements et de capacités souhaitables dans l'aviation. Ce n'est pas simplement une mesure de compétence ou de technique, mais également une mesure de la conscience du pilote d'avion, de l'environnement dans lequel il opère et de ses propres possibilités (Jouanneaux, 1999).

Il est important de comprendre qu'il n'y a aucune définition partagée de l'Airmanship mais il y a plusieurs composantes clés que nous pourrions trouver dans plusieurs définitions(Kern, 1997). Afin d'essayer de mieux définir ce concept, nous emploierons la décomposition donnée par DeMaria (Demaria, 2006) parce qu'elle donne, selon nous, une bonne vue d'ensemble de la définition du Comportement pilote : (i) Conscience de la situationnel, (ii) Etat minimum des personnels, (iii) Evaluation et gestion des risques, (iv) Bon sens, (v) Discipline. Nous avons ajouté à ces points précédents (vi) « l'esprit d'équipe ». Chacun de ces points nécessiterait une étude prolongée mais nous donnons juste ici une rapide vue d'ensemble afin de garder à l'esprit les notions nécessaires.

I L'Airmanship

I.1 : Conscience de situation

La conscience situationnelle se rapporte au degré d'exactitude avec laquelle la perception de son environnement courant reflète la réalité. Les composantes clés de la conscience de la situation sont : 1/la capacité de comparer le prévu au réel (perception subjective vs. réalité objective), 2/la compréhension de l'environnement en cours d'évolution (structurant les nouvelles connaissances) et

3/l'usage de la connaissance en résultant pour contrôler la dynamique de son environnement en cours d'évolution (Endsley, 2004).

I.2 : L'état minimum du personnel

Le minimum de personnel se rapporte à l'état spécifique de pilotes tels que la maladie, la prise de médicament, le stress, l'alcool, les drogues, la fatigue, la faim ; mais également peut inclure la formation pilote, l'expérience et les pressions externes.

I.3 : Évaluation et gestion des risques

Afin d'évaluer et de gérer le risque en vol, le pilote doit analyser sa mission (Amalberti, 1996). L'analyse de mission peut se rapporter autant à la capacité de développer des plans à court, à long terme ou d'urgence, qu'à celle de coordonner, allouer et surveiller les ressources de l'équipage et de l'avion. Les phases de l'analyse de mission concernent l'avant-vol (planification du vol, briefing), le vol (planification à court terme, surveillance de la progression du vol, identification et reporting de ce qui a pu poser problème) et l'après- vol.

L'analyse de la gestion des risques mène inévitablement au jugement et au « processus décisionnel » ((Reason, 1990), (Jambon, 1998)), c'est-à-dire : étape de détection, étape de diagnostic, étape de plan d'action et étape de l'action.

I.4 : Bon sens

Le bon sens consiste en plusieurs composants. Ce n'est pas une liste approfondie mais il donne plusieurs étapes clés de bon sens dans l'aviation acquise par la formation et l'expérience : adopter des procédures conformes, pratiques et qu'on peut répéter, connaître les règles, être assuré et conforter sa position d'un endroit de sécurité et d'efficacité, être préparé pour le vol.

I.5 : Discipline

La discipline se rapporte au respect des règles et à la généralisation de bonnes pratiques.

I.6 : « Esprit d'équipe »

La coopération de l'équipage est nécessaire pendant toutes les phases de la mission pour permettre une bonne réalisation de vol.

I.7 : Basic Airmanship

Le « Basic Airmanship » correspond aux comportements élémentaires que les concepteurs pourraient raisonnablement attendre d'un équipage d'aéronef selon la définition du Comportement pilote. Une

description précise « du Comportement pilote de base » n'existe pas en termes concrets, mais cette notion existe chez chacun des acteurs suivants : concepteurs d'avion, pilotes, spécialistes en sécurité, spécialistes facteur humain… Et est employée de façon discrète et indirecte pour construire des avions aujourd'hui. En effet, le partage de connaissance grâce au processus intégré entre plusieurs de ces disciplines, permet l'intégration de la notion « Basic Airmanship » dans la conception d'avion.

II Conception et « Basic Airmanship »:

Dans le cadre d'une conception centrée utilisateur, il y a un besoin réel du Basic Airmanship. La conception centrée utilisateur peut être caractérisée comme un processus à plusieurs étages de résolution de problèmes qui exige non seulement de la part des concepteurs d'analyser et de prévoir comment les utilisateurs pourraient utiliser une interface, mais également d'examiner la validité de leurs suppositions en analysant le comportement d'un utilisateur au cours d'essais en situation réelle avec les utilisateurs réels. De tels essais sont nécessaires car il est souvent très difficile pour les concepteurs d'une interface de comprendre intuitivement ce qu'un utilisateur éprouvera pour la première fois face à leur conception, et ce à quoi chaque courbe d'apprentissage d'un utilisateur peut ressembler (adapté de : (International Organization for Standardization, 1999), (Robert, 2003)). Dans un monde idéal, tous les concepteurs veulent imaginer l'usage du produit et voudraient que toutes les situations réelles d'utilisation soient testées pendant la phase de conception. Néanmoins, la conception de systèmes complexes tels que l'aviation commerciale, peut seulement examiner un petit nombre de scénarii. Même s'ils sont bien pensés, discutés, tracés pendant tous les processus de conception, la réalité explorera une quantité considérable de nouvelles situations : des prévisions simples et approfondies d'utilisation d'avion ne sont pas possibles.

Néanmoins, nous avons à établir des méthodes et outils en vue de mieux prévoir l'utilisation réelle d'avion afin de ramener, autant que possible, un état divergent du système homme-machine à un état stable qui garantira la sûreté du système. Airbus a acquis une grande expertise à prendre en considération les facteurs humains pendant la conception du cockpit et la certification (Reuzeau, et al., 2010). Toutes les études facteur humain entreprises sur des cockpits d'avion sont organisées comme une séquence logique d'expérimentations scientifiques avec la réalisation de tests d'hypothèse, de protocoles expérimentaux, de collecte de données et d'analyses qui ont permis l'intégration des facteurs humains dans le cycle de conception pendant environ quinze ans. La notion de sûreté est une partie fondamentale durant la conception d'avion en particulier la gestion d'erreur d'équipage : comment

prévoir des erreurs humaines ?, comment contrôler leurs conséquences ?, est-ce que des mécanismes de récupération existent ?

Maintenant, la notion de Basic Airmanship entre en scène. Pour répondre à ces questions, plusieurs acteurs interviennent dans un processus commun (pilotes : essai, formation, compagnies ; spécialistes facteur humain ; spécialistes sûreté ; concepteurs d'architecture de système ; …) afin d'intégrer dans le produit des traits élémentaires des utilisateurs d'avion : c'est-à-dire le Basic Airmanship. Cette connaissance est difficile à mesurer, à quantifier, à extraire d'une façon simple. Elle est décrite en partie et éparpillée dans divers documents (procédures opérationnelles, programme de formation, analyses de sûreté…). Généralement, la manière la plus facile de la réutiliser est d'utiliser une partie de la conception précédente pour un nouveau produit.

Afin de faciliter l'innovation et de maintenir le haut niveau de sûreté aérienne, il est nécessaire d'avoir, pendant les phases les plus précoces de la conception, une hypothèse précise, justifiable, traçable sur le comportement d'équipage : que pouvons-nous raisonnablement attendre de l'équipage ?

Dans cette dernière partie nous verrons que ces hypothèses ne peuvent pas être statiques et qu'il est très dur, peut-être impossible, de créer un consensus entre les opinions sur ces hypothèses mais il existe d'autres manières d'atteindre nos objectifs.

III Systèmes et environnement contraignent le « Basic Airmanship » :

Maintenant que nous avons défini ce que sont l'airmanship, le Basic Airmanship et comment ils sont employés pendant la conception, voyons comment ils peuvent être modifiés. Nous avons recueilli ces facteurs de modification : principes d'automatisation et de partage des tâches homme-machine, complexité du système, fiabilité du système, politique de ligne aérienne, modification de la charge de travail. Cette liste n'est pas exhaustive.

III.1 Automatisation et complexité du système

L'automatisation des cockpits au cours des vingt dernières années et l'augmentation de la complexité du système technique ont modifié des compétences de la profession de pilote et le rôle de l'équipage a évolué du pilotage à la gestion. Plusieurs études dans le domaine aéronautique(Parasuraman, et al., 2000),(Plat, et al., 2000) prouvent que l'automatisation de cockpits est une variable significative dans les activités des équipages d'avion. En effet, l'équipage peut déléguer certaines tâches aux automatismes et, de fait, la gestion de l'avion / du vol est plus facile. Mais l'augmentation de la complexité de système a pu potentiellement compliquer la compréhension par l'équipage du

fonctionnement du système en particulier. Ainsi, il y a une interaction entre ces deux effets : aide et complexité. Confrontés à une défaillance du système, les pilotes chercheront à comprendre la panne sans perdre l'aide de l'instrument. Puis Amalberti et Hoc (Amalberti, 2003), (Hoc, 2003), soulignent que le choix entre commande automatique et manuelle dépend de ce que l'équipage pourrait raisonnablement imaginer du comportement du système. En conclusion, Rasmussen et Golstein (Rasmussen, et al., 1985) expliquent qu'il y aura une perte de compétence dans tous les domaines qui sont habituellement pris en charge par l'automatisme. Tous ces faits modèlent le Basic Airmanship.

III.2 Augmentation de fiabilité

L'augmentation de la fiabilité des automatismes mène à un taux très léger de face à face entre le pilote et les défaillances du système. En cas de problème, le pilote aura tendance à mettre à l'épreuve son activité à la place de l'automatisme. Autrement, que ce soit pour un novice ou un expert, le manque de confrontation à des défaillances du système (la formation ne peut pas couvrir la grande diversité de situations réelles) signifie que les pilotes pourraient potentiellement tendre à perdre confiance en leur capacité à gérer l'échec.

III.3 : Politiques de ligne aérienne

Tous les avions d'une famille sont globalement les mêmes quelle que soit la compagnie aérienne (pays, formation des pilotes, politique de coût,…). Des contraintes internes inhérentes aux politiques de chacune des compagnies auront un impact sur le Basic Airmanship.

III.4 : Variation locale du Basic Airmanship

En raison de la variation d'activité en vol et des conditions du pilote, il est impossible de penser que le Basic Airmanship restera le même pendant tout le vol. En effet, le processus de prise de décision de l'équipage évoluera selon des contraintes externes telles que la défaillance du système, les conditions météorologiques, la phase de vol, la fatigue, etc. C'est pourquoi le Basic Airmanship est lié à la charge de travail et à l'état de l'équipage.

L'état de l'équipage vari en fonction de son expérience, du trajet réalisé (décalage horaire), des conditions physiques du moment… La charge de travail d'équipage peut être liée aux systèmes de l'avion et au contexte externe de leur utilisation. En fait, ces systèmes sont construits pour aider à la réalisation des tâches d'équipage ; dans ce contexte l'aide correspond à la réduction de la charge de travail de l'équipage. Globalement, l'automatisation des avions a réduit la charge de travail de l'équipage pour quelques tâches, mais d'une façon paradoxale. En fait, «l'ironie de l'automatisation », comme Bainbridge l'appelle (Amalberti, 2003), est que les systèmes automatiques ont encouragé

l'apparition de nouvelles tâches (comme les tâches de manipulation, de gestion ou de surveillance). Ces nouvelles tâches ont pu être des sources possibles d'erreurs et ont pu être consommatrices de temps. Néanmoins, nous voudrions souligner que l'état de sûreté augmente avec le niveau d'automatisation. En outre, la charge de travail de l'équipage est mal partagée, par exemple, il y a quelques problèmes d'hypovigilance en phase de croisière et à l'opposé, pendant l'approche, la charge de travail est augmentée. En fait, le problème pourrait être formalisé comme suit : les systèmes automatiques tendent à réduire globalement la charge de travail de l'équipage mais ceci pourrait localement créer quelques points de faible charge de travail avec un manque de concentration de l'équipage et une diminution de l'efficience de l'équipage en cas de situations critiques et quelques points élevés de charge de travail avec une augmentation de la consommation de ressources (Hoc, 2003).

L'effet des situations externes, telles que le trafic ou la météo, a pu être une source de charge de travail de l'équipage. La conception des systèmes doit prendre en considération ces faits.

IV Conclusion

Pour conclure, nous voudrions souligner que toutes ces contraintes n'aident pas à la réalisation d'un consensus sur le « Basic Airmanship ». La complexité « Basic Airmanship » est une énigme pour tous les concepteurs d'avion et une source de longues négociations entre tous. Cependant, en favorisant le partage des connaissances entre les acteurs de la conception en aéronautique, il est possible de créer un meilleur compromis, garant d'une plus grande sécurité.

CHAPITRE. 3 ARCHITECTURE DES PROGRAMMES INFORMATIQUES

Dans cette partie, sans donner le code informatique, nous allons présenter schématiquement le fonctionnement des algorithmes développés lors de ce travail de recherche. Les données traitées par ces algorithmes proviennent des scénarii des analyses de sûreté des systèmes normalisés pour être traiter par informatique. Le détail de cette normalisation a été présenté dans la partie II chapitre 3. Cependant cette normalisation, n'est pas suffisante pour entrer directement dans les algorithmes. Elle a permis principalement à rendre les données manipulables par Scilab. C'est pourquoi la premier partie de cette annexe reprend la mise en forme des données. Nous développerons ensuite les algorithmes des proches voisins et du filtrage bayésien. *NB : Tout les développements ont été réalisés en Scilab (http://www.scilab.org/).*

I Préparation des données

La figure ci après montre les préparations qui seront effectuées sur les données brutes.

1/Données brutes

Ce sont les données récupérées lors de la normalisation des scénarii

2/Vectorisation des scénarii

Lors de cette étape, les scénarii vont être transformés en « vecteurs de scénarii » à l'aide d'une base (au sens des ensembles vectoriels) de cet ensemble obtenue grâce au lexique (ensemble des mots présents dans le texte).

Par exemple :
Texte 1 : « je je vais ».
Texte 2 : « je je je je vais vais ».
Texte 3 : « je vais vais ».

Après traitement inital :
Texte 1 : « 1 1 2 ».
Texte 2 : « 1 1 1 1 2 2 ».
Texte 3 : « 1 2 2 ».

Après la vectorisation des scénarios :
Texte 1 : [2, 1]
Texte 2 : [4, 2]
Texte 3 : [1, 2]

Matrice résultante :

|2	1|
|4	2|
|1	2|

3/Ajout des apprentissages

Lors de notre étude nous avons vu que certaines informations étaient pertinentes pour la sélection des scénarios par les experts (Phase de vol, détection de la panne, classification, type de panne), nous avons aussi des scénarii sélectionnés comme pertinents par les experts. Toutes ces informations doivent être rajoutées dans les scénarii.
Par exemple pour l'information de sélection par les experts : [1, 0] signifiera que le scénario a été sélectionné et [0, 1] qu'il na pas été sélectionné. *NB : Afin de ne pas fausser les résultats les information de sélection ne sont pas utilisées pour les algorithme de filtrage mais sont utilisées pour exploiter et conclure sur l'efficacité des méthodes.*

Par exemple
Texte 1 : [2, 1] et sélectionné
Texte 2 : [4, 2] et non sélectionné
Texte 3 : [1, 2] et non sélectionné

Matrice résultante :

|1	0	2	1|
|0	1	4	2|
|0	1	1	2|

4/Matrice prête

La matrice obtenue est alors prête pour être utilisée par les algorithmes des proches voisins et de filtrage bayésien.

Figure 56 : Préparation des données pour les algorithmes

II Algorithme des plus proches voisins

Voyons maintenant l'algorithme de plus proches voisins.

1/Matrice prête

La matrice obtenue est alors prête pour être utilisée par les algorithmes des proches voisins et de filtrage bayésien.

2/Calcul des distances euclidiennes

Nous allons maintenant calculer les distances entre chaque scénarii (sans prendre en compte les information de sélection). Nous allons obtenir une matrice carrée.

Par exemple :

Matrice des scénarii:

V1 =|1	0	2	1|
V2 =|0	1	4	2|
V3 =|1	0	1	2|

Matrice des distances :

	V1	V2	V3
V1 =|	0	d1	d2|
V2 =|	d1	0	d3|
V3 =|	d2	d3	0|

Nous sommes maintenant en mesure de savoir quel est le scénario le plus proche de V1 ou de V2 ou de V3.

3/Efficacité de l'algorithme

A l'aide de la matrice des distances obtenue dans l'étape précédente et des informations de sélection nous allons rechercher à mesurer l'efficacité de l'algorithme pour sélectionner des scénarii pertinents. Nous allons suivre les étapes suivantes :

Matrice des distances :

	V1	V2	V3
V1 :|	0	d1	d2|
V2 :|	d1	0	d3|
V3 :|	d2	d3	0|

Ajout des information de sélection pour les vecteurs sélectionnés (ici V1 et V3 sont sélectionnés)

|0	d1	d2|
|1	0	1|
|d2	d3	0|
|1	0	1|

Classement par ordre croissant des distances (par exemple d2<d1 et d2<d3)

|0	d2	d1|
|1	1	0|
|0	d2	d3|
|1	1	0|

La première colonne correspond à la distance du scénario sélectionné avec lui-même, elle ne nous intéresse donc pas. Pour les autres colonnes, nous sommons chaque valeur de lignes afin de savoir comment évolue le nombre de scénarii sélectionnés au fur et à mesure que la distance augmente. Au final, nous obtenons :

|1	1|
|1	1|

Au final si l'on sélectionne le scénario le plus proche d'un scénario sélectionné en moyenne on obtenons 1 scénario sélectionné. Si on en sélectionne 2, on obtient toujours 1 car le second n'était pas sélectionné.

4/Résultats de la méthode

Figure 57 : Algorithme des proches voisins

III Algorithme de filtrage bayésien

Voyons maintenant l'algorithme de filtrage bayésien

1/Matrice prête

La matrice obtenue est alors prête pour être utilisée par les algorithmes des proches voisins et de filtrage bayésien.

↓

2/Création des ensembles d'apprentissage et calcul de probabilité

Afin de calculer les probabilités de sélection, nous devons posséder un ensemble d'apprentissage. Pour ce faire, nous allons sélectionner de manière aléatoire la moitié des scénarii non sélectionnés par les experts et la moitié des scénarii sélectionnés par les experts.
Ces ensembles vont nous permettre de connaitre la probabilité qu'un scénario soit sélectionné sachant qu'il contient un certain mot cela pour tout les mots de l'ensemble d'apprentissage. A l'aide de ces probabilités unitaires, en fonction des mots des scénarii étudiés, nous allons pouvoir calculer une probabilité de sélection.

Par exemple (imaginons ici que nous ayons un autre ensemble pour l'apprentissage):
Matrice des scénarii:

V1 =|1	0	2	1|
V2 =|0	1	4	2|
V3 =|1	0	1	2|

Matrice des probabilités:

V1 :|P(V1)|
V2 :|P(V2)|
V3 :|P(V3)|

Nous savons maintenant quelle est la probabilité de sélection pour chaque scénario.

↓

3/Efficacité de l'algorithme

A l'aide de la matrice des probabilité obtenue dans l'étape précédente et des informations de sélection, nous allons rechercher à mesurer l'efficacité de l'algorithme pour sélectionner des scénarii pertinents. Nous allons suivre les étapes suivantes :
Matrice des probabilités:

V1 :|P(V1)|
V2 :|P(V2)|
V3 :|P(V3)|

Ajout des information de sélection pour les vecteurs sélectionnés (ici V1 et V3 sont sélectionnés)

V1 :|1	P(V1)|
V2 :|0	P(V2)|
V3 :|1	P(V3)|

Classement par ordre croissant des probabilités(par exemple p(V2) < p(V1)< p(v3))

V3 :|1	P(V3)|
V1 :|1	P(V1)|
V2 :|0	P(V2)|

Au final, les deux scénarii les plus probables sont bien des scénarii pertinents. Parmi les 3 scénarii les plus probables, seuls 2 sont pertinents.

↓

4/Résultats de la méthode

Figure 58 : Algorithme de filtrage bayésien.