



Mesure de la qualité de la documentation et relation au temps de résolution de bogues  
logiciels

par

Noura Khourdaji

Essai présenté au CeFTI

en vue de l'obtention du grade de maître en génie logiciel  
(maîtrise en génie logiciel incluant un cheminement de type cours en génie logiciel)

FACULTÉ DES SCIENCES  
UNIVERSITÉ DE SHERBROOKE

Longueuil, Québec, Canada, juin 2016

## Sommaire

La documentation technique concerne depuis longtemps les experts en génie logiciel [1]. Or, dans le développement de logiciel, les entreprises constatent de plus en plus les coûts considérables qu'entraînent le maintien et l'évolution des applications. À cause de cette réalité en croissance, de plus en plus de professionnels et d'entreprises se questionnent sur des approches pour faciliter et accélérer les correctifs et les modifications des programmes. Une de ces approches interpelle le domaine universitaire et amène les chercheurs à transposer des modèles théoriques à propos de la documentation technique dans les applications commerciales.

Pour comprendre le lien entre la documentation et le maintien, l'effet d'une documentation de qualité est analysé. La qualité est mesurée avec un indice développé durant le projet et corrélé à la durée de résolution des problèmes et défaillances. Les systèmes analysés sont opérationnels et stables. L'essai explore si les applications avec un meilleur indice de documentation présentent des temps de résolution de bogues plus courts.

Le développement d'un indice de documentation repose sur l'élaboration d'une grille de calcul composée d'attributs qui sont des caractéristiques-clés de la qualité en documentation. Les aspects indispensables d'une documentation de qualité (l'exhaustivité, la mise à jour, la précision, la disponibilité, la fiabilité) interviennent dans l'élaboration de la grille. La grille accorde un indice pour chaque attribut, pour chaque document. L'indice total du document est la somme des indices individuels d'attributs. L'indicateur cumulé au niveau du bogue est la somme des indices totaux des documents. Il en résulte pour chaque bogue, un couple (indice cumulatif de tous les documents de l'application, temps de résolution du bogue dans l'application). Une comparaison s'effectue ensuite entre les différents indices et les temps de résolution de bogue.

L'analyse corrélationnelle démontre la relation entre les deux variables. L'hypothèse est que le temps de résolution de bogues diminue quand l'indice de documentation augmente (pente négative). Les faits observés démontrent l'absence d'une linéarité avec un coefficient tellement faible qu'il est à peine perceptible. Avec le traitement des points aberrants, la

tendance observée montre que, contrairement à l'hypothèse, la relation observée a une pente positive. Le temps de résolution de bogues augmente donc quand la documentation des applications considérées présente un indice plus élevé. Une discussion des justifications et des pistes de réflexion pour approfondir la compréhension du phénomène sont proposées. Une analyse plus fine des attributs précis qui forment l'indice de documentation permet de constater que certains parmi eux sont effectivement reliés à une réduction du temps de résolution. Ces corrélations négatives sont alignées avec l'hypothèse et sont scrutées.

Points aberrants exclus, l'indice de documentation cumulatif démontre une tendance linéaire opposée à celle démontrée par les indices individuels de certains attributs. Certains attributs ont nettement plus d'impact sur les temps de résolution que d'autres. Il est possible, et considéré comme une réponse à une des questions secondaires de la recherche, d'évaluer la pertinence du barème proposé. Par exemple, la combinaison d'un attribut à faible incidence sur le délai de résolution avec un autre beaucoup plus influent engendre un résultat qui ne modifie pas la perception du phénomène si les cotes sont égales.

La principale recommandation est un appel à poursuivre les initiatives d'amélioration des processus documentaires en entreprise. Les logiciels pour lesquels le besoin en documentation est grand sont souvent des logiciels dont les défaillances se résolvent le plus lentement. Il est recommandé de mettre en application de prime abord les attributs qui ont une contribution dans la réduction du temps de résolution, telle que la présence d'éléments visuels. La deuxième recommandation est de maintenir les documents en accordant une importance à la cohésion entre l'ensemble des documents, et aussi entre les documents et les logiciels. Les résultats montrent que l'impact d'un seul attribut n'aide pas la résolution de bogue lorsque l'indice de cet attribut est analysé pour chaque document individuellement. Pris dans leur ensemble, les indices donnent un indice moyen qui corrèle de façon remarquable avec le temps de résolution pour plusieurs attributs. L'état de la synchronisation (moyenne des documents) est un exemple de ce phénomène. L'état de chaque document seul impacte le temps de résolution très différemment de la moyenne des états de synchronisation.

Il est recommandé d'effectuer des recherches futures pour trouver d'autres impacts de la documentation sur le cycle de vie du logiciel. Il demeure plusieurs questions sans réponse, mais le présent projet ouvre des pistes nouvelles en abordant un aspect de la question. En

effet, le calcul des caractéristiques par le biais des indices choisis est une interprétation et une application, mais cela n'exclut pas d'autres applications. Un avantage présenté par cet essai est l'application d'un cadre plus théorique au sujet de la documentation dans un environnement de production et la quantification d'éléments peu fréquemment étudiés dans le domaine. Dans cette optique, la recherche peut s'inspirer des pistes d'analyse de cet essai et contrer les obstacles rapportés.

## Remerciements

Je tiens à remercier mon équipe de direction, Messieurs Martin Lapointe et Pierre-Martin Tardif pour leur soutien tout au long du processus de rédaction. En particulier, je remercie Monsieur Lapointe pour les nombreuses discussions et pour m'avoir présenté de nouvelles perspectives de raisonnement. Merci à Monsieur Cardinal pour un appui constant durant mon parcours au diplôme et à la maîtrise.

Un remerciement tout particulier va à Madame Evelina Balut pour son mentorat et son encouragement soutenu. Merci à Madame Victoria Cook pour son sens de l'observation et sa critique constructive.

Je suis reconnaissante pour tout le soutien que m'a offert le personnel du CeFTI et de l'Université Sherbrooke pendant la durée de mes études. Parmi ces individus, je tiens à souligner l'impact positif qu'ont eu les enseignants sur mon développement professionnel, académique et personnel.

Je remercie énormément mon mari Yaser Kabakibo sans qui je n'aurais pu compléter mon essai, ma famille pour l'encouragement continuel et mes fils Mouhammad, Amir et Hashem pour la motivation de réussir.

## Table des matières

Sommaire .....	i
Remerciements.....	iv
Table des matières .....	v
Liste des tableaux.....	ix
Liste des figures.....	x
Glossaire .....	xi
Liste des sigles et des acronymes .....	xiv
Introduction.....	16
Chapitre 1 Mise en contexte .....	18
1.1 La structure du support technologique pour la documentation.....	18
1.2 L'accompagnement de la documentation à la production de logiciel.....	19
1.3 La formation des ressources humaines .....	19
1.4 La population cible .....	20
1.5 Le paradigme de la quantité et de la qualité de la documentation.....	20
1.6 Défis connus.....	21
1.7 Besoin et objectif .....	21
Chapitre 2 Revue de la littérature.....	22
2.1 Méthodologie de recherche des sources .....	22
2.1.1 Mots-clés utilisés .....	22
2.1.2 Outil de recherche .....	23
2.1.3 Langues utilisées et ressources linguistiques .....	23
2.1.4 Thèmes abordés.....	24
2.2 Revue systématique .....	24
2.2.1 Sommaire .....	24
2.2.2 Pertinence .....	25
2.2.3 Discussion .....	25

2.3	Perspective de l'architecture logicielle .....	27
2.3.1	Sommaire .....	27
2.3.2	Pertinence .....	27
2.3.3	Discussion .....	28
2.4	Perspective de la méthodologie de développement.....	29
2.4.1	Sommaire .....	29
2.4.2	Pertinence .....	30
2.4.3	Discussion .....	30
2.5	Perspective de résolution de bogues.....	31
2.5.1	Sommaire .....	31
2.5.2	Pertinence .....	31
2.6	Perspective de l'utilité.....	31
2.6.1	Sommaire .....	31
2.6.2	Pertinence .....	32
2.6.3	Discussion .....	33
2.7	Perspectives additionnelles .....	33
2.8	Aperçu des métriques existantes.....	33
2.9	Conclusion du chapitre .....	34
Chapitre 3 Problématique .....		35
3.1	Description .....	35
3.1.1	Objectifs .....	36
3.1.2	Questions de recherche et hypothèse .....	37
3.1.3	Limites.....	38
3.2	Méthodologie proposée .....	39
3.2.1	Type de recherche.....	39
3.2.2	Méthodes de validation.....	39
3.3	Mise en œuvre .....	40
3.3.1	Déroulement de l'essai .....	41
3.3.2	Environnement de travail .....	42
3.3.3	Population cible .....	42
3.3.4	Mise en œuvre.....	43
Chapitre 4 Approche proposée .....		45
4.1	Description de l'approche .....	45
4.1.1	Outils .....	45
4.1.2	Programmes sélectionnés et échantillon.....	46

4.1.3 Critères, attributs et barèmes.....	47
4.1.4 Description des étapes du développement de la métrique.....	54
4.1.5 Facteurs clés de succès .....	56
4.1.6 Approche de validation des résultats .....	56
4.2 Résultats attendus.....	57
4.2.1 Données de bogues.....	57
4.2.2 Données de bogues non retenus.....	57
4.2.3 Données des attributs de qualité de la documentation.....	58
4.2.4 Données des indices de qualité de la documentation .....	59
4.2.5 Données sommaires.....	60
4.2.6 Autres analyses .....	61
Chapitre 5 Analyse des résultats .....	62
5.1 Résultats obtenus.....	62
5.1.1 Données de bogues.....	62
5.1.2 Données de bogues non retenus.....	62
5.1.3 Données des attributs de qualité de la documentation.....	63
5.1.4 Données des indices de qualité de la documentation .....	63
5.1.5 Données sommaires.....	63
5.1.6 Autres analyses .....	65
5.2 Retour sur les questions et l'hypothèse .....	67
5.3 Démonstration de la validité des résultats .....	70
5.4 Recommandations.....	71
Conclusion.....	73
Liste des références .....	76
Annexe 1 Bibliographie .....	81
Annexe 2 Perspectives additionnelles.....	83
Annexe 3 Métriques existantes .....	88
Annexe 4 Composition de l'échantillon .....	90
Annexe 5 Processus de gestion de projets .....	93
Annexe 6 Captures d'écran pour les recherches de rapports de bogue .....	94
Annexe 7 Données de bogues et analyse.....	98

Annexe 8 Données de bogues non retenus .....	103
Annexe 9 Données brutes des attributs de la documentation pour les bogues .....	105
Annexe 10 Données brutes avec les attributs formatés pour le programme de calcul des indices .....	120
Annexe 11 Ajustement des points extrêmes pour la corrélation sommaire.....	123
Annexe 12 Indices individuels détaillées, cumulés et moyens pour les documents.....	125
Annexe 13 Directives de validation et application pratique dans la recherche.....	140
Annexe 14 Code SQL pour les tables .....	145

## Liste des tableaux

Tableau 4.1 Gabarit des détails des incidents et du temps de résolution de bogues.....	57
Tableau 4.2 Gabarit des données de bogues non retenus pour le projet .....	58
Tableau 4.3 Gabarit de présentation des données brutes des attributs de la documentation	58
Tableau 4.4 Gabarit de présentation d'un attribut de qualité par rapport au temps de résolution.....	59
Tableau 4.5 Gabarit de présentation des données de corrélation .....	60
Tableau 5.1 Données de corrélation .....	63

## Liste des figures

Figure 3.1 Modèle conceptuel du projet, première itération.....	38
Figure 3.2 Modèle conceptuel du projet, deuxième itération .....	38
Figure 3.3 Modèle de séquençement des activités de l'essai.....	41
Figure 4.1 Schéma des étapes de la recherche .....	47
Figure 4.2 Modèle de qualité de la documentation technique.....	49
Figure 4.3 Diagramme de dispersion pour l'analyse des données de l'essai (données fictives) .....	60
Figure 5.1 Graphique du temps de résolution des bogues en fonction de l'indicateur de qualité de la documentation.....	64
Figure 5.2 Graphique du temps de résolution des bogues en fonction du nombre d'éléments visuels (moyenne par document).....	66
Figure 5.3 Graphique du temps de résolution des bogues en fonction de la mise en page des documents (moyenne par document) .....	66
Figure 5.4 Graphique du temps de résolution des bogues en fonction du statut de mise à jour des documents (cumulatif des documents par bogue).....	67

## Glossaire

Accessibilité	Disponibilité d'un document pour la consultation, le partage et la modification de la part d'un usager.
Barème	Règle de calcul pour l'évaluation d'un attribut de qualité de la documentation. Le barème dicte comment incrémenter l'indice de la qualité pour chaque attribut et sert de contribution à la mesure globale.
Bogue	Désigne l'ensemble des erreurs, défaillances, arrêts, alertes ou autres signaux qui causent interruption au traitement normal en production.
Bogue non retenu	Défaillance de système appartenant à l'échantillon, mais dont les données sont exclues de l'analyse.
Code source	Tout artefact logiciel qui sert à produire un traitement informatique ou un résultat électronique quelconque. Cela englobe les logiciels, les applications et les modules peu importe la langue ou l'environnement de développement ou d'opération.
Corrélation négative	Relation entre la variable indépendante (l'indice de qualité de la documentation) et la variable dépendante (le temps de résolution des défaillances). La croissance de la première est liée à la diminution de la deuxième.
Critères	Ensemble des attributs de qualité de la documentation considérés comme des facteurs de calcul de l'indice de documentation.

Défaillance	Le texte emploie ce terme comme synonyme de <i>bogue</i> .
Délai de maintien	Lapsus de temps entre la date et heure de signalement d'un bogue dans un système et la date et heure de fermeture du signalement avec une solution.
Format standard	Types de documents considérés les plus utilisés dans un environnement de travail basé sur l'utilisation d'ordinateurs roulant les logiciels Microsoft, Adobe, et autres similaires.
Gabarit	Une mise en page spécifique et sauvegardée dans un répertoire partagé. Le gabarit est connu et facilement accessible aux usagers qui produisent la documentation.
Grille de calcul	Tableau dans lequel est présenté l'ensemble des barèmes pour calculer un indice de qualité de la documentation
Indice de documentation	Métrique développée par l'auteur pour calculer une valeur qui représente le degré auquel un document est conforme aux critères de qualité. Synonyme de : indice de qualité de la documentation, cote de documentation, cote de qualité de la documentation.
Lisibilité	Aspect de la qualité des ouvrages qui décrit la facilité qu'a un lecteur de parcourir l'ouvrage, de le lire et d'en comprendre le contenu.
Maintien de la documentation	Activité de mise à jour des documents pour en conserver la valeur informationnelle.
Méthodologie de développement	Approche de développement applicatif choisie
Mise à jour	Activité d'ajustement des écrits et des figures pour contenir l'information réelle de l'état actuel des entités, peu importe leur nature.

Métrique de la qualité de la documentation	Valeur résultant d'une formule de calcul à partir d'éléments individuels de la qualité de la documentation
Qualité de la documentation	Évaluation qualitative de la documentation selon un ensemble d'aspects considérés importants et pertinents pour les personnes qui consultent, modifient ou créent des documents. De nombreux chercheurs ont étudié ces aspects et dégagé certains attributs essentiels, sans consensus sur la liste définitive d'attributs de la qualité.
Structure	Ensemble d'attributs qui décrivent l'allure et les caractéristiques visuelles apparentes d'un document
Wiki	Application web qui permet à plusieurs auteurs de collaborer sur la création et la modification de contenu (texte et images) dans un site web.

## Liste des sigles et des acronymes

.doc ou .docx	Document produit par traitement de texte de type <i>Microsoft Word</i> ou similaire
.txt	<i>Text</i> : Fichier de type texte
BABOK	<i>Business Analysis Book of Knowledge</i> : Corpus de connaissances de l'analyse d'affaires
CEI	<i>Cost Effectiveness Index</i> : Indice de retour sur investissement d'un document
CMMI	<i>Capability Maturity Model Integration</i> : Modèle pour l'amélioration des processus
CSV	<i>Comma Separated Values</i> : Fichier de valeurs séparées par des virgules
IEEE	<i>Institute of Electrical and Electronics Engineers</i> : Institut des ingénieurs électriciens et électronicien
ISO	<i>International Organisation for Standardization</i> : Organisation internationale de normalisation
MS Outlook	<i>Microsoft Outlook</i> : Outil de messagerie
PDF	<i>Portable Document Format</i> : format de document multiplateforme, créé à l'aide du langage de description de pages et dont la mise en forme est préservée

PGI	Progiciel de gestion intégré
PMBOK	<i>Project Management Body of Knowledge</i> : Corpus des connaissances en gestion de projets
PMO	<i>Project Management Office</i> : Bureau de gestion des projets
PVCS	<i>Polytron Version Control System</i> : Système de gestion de versions de Serena Software Inc.
RMS	<i>Readability Measuring System</i> : Outil de mesure de la compréhensibilité des documents
SFX	<i>Ex-Libris SFX</i> : un logiciel de recherche bibliothécaire avec résolveur de liens
SQL	<i>Structured Query Language</i> : langage de requête structurée
SQR	<i>Structured Query Report</i> : Rapport en langage structuré
UML	<i>Unified Modeling Language</i> : langage de modélisation unifié
XLS	Extension de fichier de données tabulaires produites dans Microsoft Excel ou similaire

## Introduction

La documentation technique est un sujet qui accompagne la recherche et le développement en science. Son rôle dans les domaines scientifiques est un fait accepté, surtout à la lumière des exigences réglementaires. En TI, le foisonnement d'applications, de langues et de méthodes de développement a permis à cette accélération technologique de prendre le dessus sur la documentation d'une conception réfléchie et vérifiable. La réalité est que le développeur est roi dans un espace technologique nouveau. D'un côté, les défaillances des applications difficiles à maintenir deviennent plus répandues et plus coûteuses. D'un autre côté, les domaines TI où le rôle primordial et incontournable de la documentation technique n'est pas un sujet controversé sont les domaines militaires et paramilitaires, d'infrastructure et d'opérations critiques. La perception installée est que seul le développement par des méthodes formelles lourdes est tenu de contenir une documentation complète.

Avec la maturation des sciences informatiques et du génie logiciel, la documentation technique est parvenue à percer la liste des éléments incontournables de la gestion de projets TI. La motivation pour cette activité doit aussi croître. Cet essai propose d'examiner la documentation en tant que facteur de qualité des logiciels. La préoccupation pour la qualité ne peut que soutenir la motivation pour une documentation pertinente, efficace, et compréhensible.

La recherche à propos de la documentation technique émerge graduellement. De plus en plus d'universitaires soumettent des modèles d'analyse ou développent des prototypes de logiciels d'évaluation de la qualité. Les études empiriques, qui ne forment qu'une minorité des études dans le domaine, croissent considérablement. Les expérimentations en contexte d'entreprise augmentent. La présente étude se sert de ces progrès dans la recherche en retirant les fruits des études les plus récentes et les plus pertinentes.

Plusieurs objectifs sont fixés par cet essai. Tout d'abord, l'objectif principal rejoint la majorité de la recherche dans le domaine : établir l'importance, la validité et l'utilité de la documentation technique dans un contexte de production et de maintien de logiciel en

environnement de production et d'exploitation. Les objectifs spécifiques suivants sont abordés : l'amélioration des modèles de mesure de la qualité de la documentation, l'augmentation de la motivation, la perception de l'importance de la rédaction d'une documentation de haute qualité, la transformation d'aspects qualitatifs observables en facteurs quantitatifs propices à une analyse, la corroboration des faits observés par les études antérieures, la réflexion aux facteurs affectant le maintien de logiciel et le temps de résolution de problèmes.

L'essai présente une analyse du contexte de la pratique documentaire en TI, cerne un besoin et se fixe un objectif de recherche. Ensuite, une revue de la littérature offre une compréhension des travaux existants et des avenues d'exploration. La problématique et le sujet de recherche sont proposés. La question principale et les sous-questions forment la base de l'exploration et la découverte des réponses s'assoit sur la formulation de l'hypothèse de recherche. La méthodologie détaille le type de recherche, les facteurs de validité, le contexte d'entreprise, et les détails et les étapes de la mise en œuvre. Le chapitre 4 contient les informations sur les outils, la population, l'échantillon, les détails du développement de la métrique, les facteurs clés de succès de l'étude, la technique de validation et les gabarits de présentation des résultats. Les données recueillies, l'analyse et les résultats sont ensuite présentés au chapitre 5. La question et les sous-questions sont revisitées à la lumière des résultats pour arriver aux observations finales à propos de l'hypothèse. Enfin, une conclusion et une discussion des pistes de recherches possibles terminent l'étude.

# Chapitre 1

## Mise en contexte

La rédaction technique se définit comme étant un domaine caractérisé par toute communication qui adresse un sujet technique ou spécialisé, emploie un mode de livraison technologique de l'information, ou offre des instructions ou une démarche [2]. Cette définition globale, portée dans les technologies de l'information, demeure un vaste sujet qui englobe tous les artefacts produits ou échangés au cours de la gestion de toutes les phases d'un projet informatique. De plus, la documentation technique dans le domaine du développement logiciel a connu plusieurs changements au cours des décennies. Quelques éléments de ces changements sont abordés dans les sections suivantes : la structure du support technologique pour la documentation, le rapprochement entre la documentation et la production de logiciel, les aspects de la formation des ressources humaines liés à la documentation technique dans le développement du logiciel, la population cible, le paradigme de la quantité par opposition à la qualité de la documentation, les défis connus, ainsi que les besoins.

### 1.1 La structure du support technologique pour la documentation

D'abord composée de fichiers textes accompagnant les logiciels, la documentation a évolué avec les progrès techniques. Au début, l'organisation par répertoires reflète les premières bases de données qui avaient cette structure [3]. Par la suite, la documentation a suivi les nouvelles approches modélisables pour les bases de données (documentation sous forme de modèles). Les modèles ont enrichi la documentation (avec le standard UML). Plus récemment, l'usage des technologies, appliquées à la documentation technique, permet un indexage plus dynamique de la terminologie et des concepts. La popularité des plateformes de collaboration (Sharepoint et Documentum à titre d'exemple) a rendu plus accessible le

travail de production et de consommation des artefacts écrits et illustrés. La prolifération des sites de type wiki a eu, de surcroît, un effet positif sur la documentation en permettant une présentation plus simple, uniforme et hyperliée de l'information, ainsi que des possibilités de création et de modification collaboratrices et sociales.

## **1.2 L'accompagnement de la documentation à la production de logiciel**

Les normes ISO, les publications IEEE et les lignes directrices du PMBoK et du BABoK incluent des listes de documentation suggérée selon le type de projet TI, la phase de conception et les parties prenantes concernées. Celle-ci comprend la documentation fonctionnelle (besoins, analyse de domaine d'affaires, analyse fonctionnelle), la documentation technique (spécifications, modélisation, décisions, cas d'usage), la documentation incorporée, les guides d'accompagnement et tout ouvrage destiné aux usagers, les documents d'architecture et de conception, toute la documentation de gestion de projets, la série de documents pertinents aux tests et à l'assurance qualité. Les ouvrages à produire sont donc variés et exigent des compétences ciblées, compétences parfois difficiles à retrouver, comme discuté dans le sous-chapitre qui suit.

## **1.3 La formation des ressources humaines**

Généralement, le premier contact d'un étudiant en TI avec les règles de la rédaction technique a lieu au cours du premier cycle universitaire. Ceci est souvent un cours requis, mais qui comporte une pondération moins élevée que les cours techniques de programmation, de base de données, et d'architecture [4] [5]. Bien que les études en TI comportent une grande quantité de rédaction académique, l'évaluation porte presque toujours sur l'exactitude d'une formule, l'exécution réussie d'un programme ou la présence de tous les composants essentiels dans un diagramme. Au-delà de la syntaxe et de la qualité de la présentation générale, la capacité de l'étudiant d'expliquer ce qu'il ou elle fait est peu (ou pas) évaluée en comparaison avec la capacité de produire une solution technique. Évidemment, dans la perspective où la solution technique de l'étudiant n'est pas réutilisable,

elle peut tout de même exercer l'étudiant à augmenter la capacité et l'efficacité de la communication. Il n'est donc pas surprenant qu'en industrie, la rédaction technique soit effectuée après le développement, encore seulement par nécessité. On associe souvent la rédaction technique à la production de documentation et de guides aux usagers, bien que la rédaction technique de la conception, des exigences, des spécifications et des tests soit un élément tout aussi important. David Parnas présente deux classes de documents en génie logiciel : les documents descriptifs (*tutorial narratives*) et les ouvrages de référence (*reference works*). La première catégorie regroupe les documents introductoires et constitue le premier contact des usagers et des techniciens avec le logiciel. La seconde catégorie vise les usagers et les professionnels familiers avec le logiciel qui veulent retrouver des informations précises ou parfaire leurs connaissances à propos du système [6].

## **1.4 La population cible**

La problématique de la production d'une documentation utile s'adresse aux développeurs et aux analystes qui rédigent et échangent la documentation, mais aussi aux gestionnaires de projets, aux exécutifs, aux usagers, et le cas échéant, aux parties prenantes législatives ou gouvernementales. Les responsables du budget d'un projet doivent être conscients des coûts de production de documentation et percevoir ces coûts comme étant justifiés. La documentation doit avoir une visibilité pour susciter une plus grande motivation de la mettre à jour.

## **1.5 Le paradigme de la quantité et de la qualité de la documentation**

Les difficultés des architectes avec la documentation des architectures méritent l'attention. Dans ce cas, la documentation est volumineuse et complexe. Toutefois, les études démontrent que la quantité ne garantit pas la compréhensibilité ni l'utilité des artefacts [8] [9]. Les deux aspects de quantité et de qualité sont des attributs d'analyse à considérer.

## **1.6 Défis connus**

La documentation fait face à plusieurs défis dans les projets TI, surtout lors des initiatives de longue haleine, interdépartementales, ou avec des règles d'affaires complexes. Plusieurs projets de recherche ont examiné et tenté de faciliter ces obstacles [8] [9]. Un exposé plus approfondi se trouve au chapitre 2.

## **1.7 Besoin et objectif**

Le besoin de l'essai dans le domaine est d'offrir une justification pratique pour l'amélioration de la création de documentation. Le contenu de la documentation doit être ciblé et la quantité doit être optimale. Cet équilibre produit des documents de qualité qui contribuent au partage des connaissances précieuses.

L'objectif est de motiver les communautés d'affaires et de développement à améliorer les pratiques de documentation. En démontrant une corrélation entre la documentation et la qualité du logiciel, une nouvelle optique sur la réduction des coûts de maintien du logiciel s'introduit. Les études convergent sur une liste d'éléments essentiels à la qualité dans la documentation : la précision, la cohérence, la clarté, la structure, la validité, la lisibilité, la compréhensibilité [9] [11] [12]. Certains de ces attributs sont la base de l'évaluation des documents échantillonnés.

## Chapitre 2

### Revue de la littérature

Le chapitre de la revue littéraire commence par l'exposé de la méthodologie de recherche des sources. Ensuite, une méta-analyse importante est décrite. Les différentes perspectives de l'architecture logicielle, de la méthodologie de développement du logiciel, de la résolution de bogues, et de l'utilité de la documentation sont couvertes. Un sous-chapitre discute des perspectives additionnelles et un autre présente les métriques existantes. Puis, le chapitre clos avec une transition vers le chapitre suivant à propos de la problématique.

#### 2.1 Méthodologie de recherche des sources

Avant de procéder aux thèmes principaux de la revue littéraire, il est justifié d'expliquer la méthode selon laquelle cette revue est entreprise. Le point de départ, la recherche par mots-clés, est suivie d'un exposé des outils de recherche, des aspects linguistiques, et des grandes lignes des thèmes recherchés.

##### 2.1.1 Mots-clés utilisés

La recherche à propos de la documentation aborde la question par des perspectives diverses selon l'usage des textes produits. Cela explique la difficulté d'effectuer une recherche ciblée sur la documentation : ce domaine est un sujet généraliste, employé pour décrire du contenu spécialisé [2]. Pour être efficace, une recherche sur la communication technique doit aborder ce contenu spécialisé. Dans le contexte des TI, les concepts centraux à la conception et au développement priment dans cette élaboration de mots-clés de recherche. Les besoins de production de documentation se trouvent au niveau de la diffusion des connaissances reliées au processus et aux décisions de design. Les mots-clés utilisés sont donc des sous-ensembles de : *documentation, communication technique, architecture logicielle, développement logiciel, méthodologie de développement, et collaboration.*

Cette approche n'est pas exhaustive. Certaines des sources consultées discutaient d'études pertinentes pour cet essai. Les études référencées ne font toutefois pas partie des résultats trouvés par la recherche à partir des mots-clés énumérés précédemment. Un regard plus attentif indique que certaines études considèrent des thèmes apparentés à la documentation sans contenir ce mot dans la liste de mots-clés. Par exemple, une étude évalue les avantages d'une documentation UML sur le maintien de logiciel [12]. Cette étude est indexée avec les termes « *empirical software engineering, UML, modeling, object-oriented programming, software maintainability, quasiexperiment* » sans la notion de documentation. La discussion des résultats de cette recherche par deux équipes différentes a permis de retracer cette étude et de l'ajouter à la base de référence littéraire de cet essai.

### **2.1.2 Outil de recherche**

La recherche s'amorce avec l'outil de gestion des références Mendeley, de la banque de données et de l'outil de recherche du service des bibliothèques et archives de l'Université de Sherbrooke, du service de recherche de la bibliothèque de l'Université McGill, du portail de recherche Google Scholar, du site IEEE, ainsi que du moteur de recherche Google. L'accès au contenu des ouvrages s'opère par les identifiants de l'Université de Sherbrooke (lien SFX pour le contenu électronique) et ceux de l'Université McGill (lien pour le catalogue de références Worldcat).

### **2.1.3 Langues utilisées et ressources linguistiques**

Des sources en langue française ont été activement recherchées. La vaste majorité des études consultées sont toutefois des textes rédigés en anglais, souvent le résultat de conférences internationales ou régionales. Dans d'autres cas, la source est une méta-analyse ou une étude publiée en anglais et basée sur des travaux d'entreprises internationales. Les principales ressources linguistiques employées pour assurer la validité des termes et des propos de cet ouvrage sont les suivants : le grand dictionnaire terminologique de l'Office québécois de la langue française ainsi que sa banque de dépannage linguistique, la banque de données terminologiques et linguistiques Termium Plus du gouvernement du Canada, ainsi que le moteur de recherche multilingue Linguee.

### **2.1.4 Thèmes abordés**

Ce chapitre couvre les perspectives de la documentation selon l'optique de la qualité, du coût et des bénéfices, de la documentation de l'architecture et du cycle de vie, de la méthodologie de développement, du développement collaboratif, ainsi que la résolution des bogues. Pour chaque thème, un exposé de la pertinence de la source suit le sommaire des références consultées. Une discussion est ensuite présentée, au besoin, pour couvrir des considérations supplémentaires et pistes d'intérêt pertinentes.

## **2.2 Revue systématique**

La revue systématique est une méta-analyse des recherches publiées sur le sujet. Elle sert de point de départ pour explorer le sujet à partir d'un survol publié récemment. Les sections couvrent le sommaire du contenu, sa pertinence dans la revue littéraire pour cet essai, et des éléments de discussion qui poussent l'exploration du sujet.

### **2.2.1 Sommaire**

Une équipe de six chercheurs menés par Junji Zhi de l'Université de Toronto, au Canada, a entrepris de synthétiser les connaissances dans une revue systématique [10]. Les chercheurs se sont posé onze questions à propos des coûts, des bénéfices et de la qualité de la documentation en génie logiciel pour générer un survol de l'état de la recherche sur le sujet. Selon une approche itérative, ils ont sélectionné 69 études publiées entre 1971 et 2011 pour leur bilan. Les résultats comprennent les cinq observations principales suivantes :

- La majorité des publications sont des recherches de validation ou des propositions de solution;
- La vaste majorité des études ne précise pas la méthodologie de développement observée, et seules six études font mention de la méthode agile;
- La majorité des systèmes à l'étude sont des prototypes académiques;
- Le sujet le plus adressé est la qualité, suivi des bénéfices alors que peu d'études ont considéré les coûts;

- Les études à propos des bénéfices de la documentation étaient largement qualitatives;
- Les caractéristiques de qualité de la documentation les plus étudiées sont (dans l'ordre décroissant d'importance dans les trouvailles) l'exhaustivité, la cohérence et l'accessibilité.

Un méta-modèle pour la documentation est développé selon chaque dimension (coûts, bénéfices et qualité). Bien que ce soit un nouveau domaine de recherche qui génère un certain intérêt, les chercheurs ont constaté que le domaine est loin d'être mature. Parmi leurs conclusions, l'analyse des coûts et la quantification des bénéfices de la documentation sont des aspects absents de la recherche. D'ailleurs, les auteurs n'ont pas trouvé de modèle ni de méthode documentée de calcul des coûts et des bénéfices de la documentation. De plus, les chercheurs appellent à plus de recherche empirique (études de validation) et à l'étude de projets TI à grande échelle (projets distribués ou ayant un grand nombre de participants issus d'organisations multiples). La collaboration avec l'industrie pour le développement de modèles et de méthodes pour évaluer la qualité, les bénéfices et les coûts de la documentation doit augmenter.

### **2.2.2 Pertinence**

Cette recherche débute par la consultation d'une revue sur le sujet [14]. Cette démarche est appropriée pour un sujet vaste aux thèmes diversifiés. La littérature est dispersée dans des sous-domaines aussi variés que la modélisation, la gestion de projets, l'architecture, le génie logiciel, les aspects humains de la programmation et les tests, pour en nommer quelques-uns. En plus de cette portée élargie, la documentation technique ne reçoit pas assez d'effort de recherche approfondie. Les sources qui se sont véritablement attardées à ce sujet sont doublement pertinentes. Enfin, les conclusions des auteurs torontois reflètent les constatations de l'auteur de cet essai et des professionnels du domaine.

### **2.2.3 Discussion**

Cette revue systématique n'échappe pas à un défi principal du domaine de recherche : l'énormité de la tâche de rentabilisation de la rédaction technique. En effet, la recherche a

abordé 11 questions de recherche, a élaboré trois modèles conceptuels, et a comparé des recherches basées sur des projets de taille inconnue (seulement 12 % des recherches citées dans la revue avaient publié la taille en LOC du logiciel). Sur ce dernier point, les auteurs ont même lancé un appel à la communauté de chercheurs pour les encourager à développer un gabarit normalisé pour la présentation des paramètres de recherche afin de permettre une analyse statistique plus pertinente dans le domaine.

Les auteurs ont seulement regardé les aspects de coût étroitement liés au document. Ils ont créé une catégorie « autre » pour tous les autres facteurs qui auraient pu être plus détaillés (expérience des auteurs de la documentation, la disponibilité des ressources de rédaction et des experts) ([10], p. 185).

Aussi, la recherche sur les coûts d'utilisation de la documentation est manquante selon les auteurs. Ils rapportent ne pas connaître d'article sur ce sujet. L'auteur de cet essai a consulté une publication par Arisholm, Briand, Hove et Labiche. Cette équipe a examiné les coûts d'utilisation de la documentation modélisée (UML) en menant deux expérimentations sur des groupes de développeurs à Ottawa et à Oslo. Les collaborateurs ont observé que l'utilisation de documentation UML a permis la réduction de l'effort de programmation de correctifs de façon générale. Néanmoins, la mise à jour des modèles UML après la modification a engendré des coûts qui annulaient parfois les gains occasionnés par l'utilisation des modèles. Cela est apparu dans deux cas. Lorsque le correctif était très simple, la consultation et la modification du modèle UML n'offraient aucun avantage. Le deuxième cas survient lorsque la formation du sujet était manquante (le programmeur n'avait pas assez d'expérience avec la modélisation de façon générale ou avec l'outil de modélisation employé dans la recherche de façon spécifique). Malgré cela, une comparaison du niveau de validité du correctif (de la perspective du design et du code) indique que l'emploi de modèles UML est avantageux dans le cas de correctifs compliqués [14]. Une précision quant à la notion de complexité des correctifs : les auteurs ont demandé des correctifs selon une échelle croissante de niveau de difficulté de tâche relativement au niveau d'expertise des deux équipes de programmeurs. Ce n'était pas une complexité mesurable statistiquement. La possibilité de reprendre cette étude (mais de la modifier pour d'autres types de documentation), bien que souhaitable et utile, se trouve donc limitée par cet aspect.

## **2.3 Perspective de l'architecture logicielle**

La gestion de la connaissance en architecture logicielle permet de comprendre les enjeux du maintien des documents au sujet des logiciels complexes. La perspective de l'architecture logicielle dans la documentation est explorée sous les trois thèmes : le sommaire des faits de la recherche existante, l'exposé de la pertinence de ces trouvailles à l'essai actuel, et une discussion pour avancer l'exploration.

### **2.3.1 Sommaire**

Dans une étude pour tenter d'améliorer les processus de documentation des architectures logicielles, la réutilisation des composants d'architecture est nommée comme une tendance qui s'accroît étant donné la complexité croissante des systèmes. Cette mise en application requiert une documentation plus complète et plus efficace des connaissances en architecture. Reconnaisant les défis de la documentation, les chercheurs ont proposé une méthode pour enrichir la documentation technique avec une notation formelle des connaissances architecturales. Cette méthode est supportée par un ensemble d'outils développé par l'équipe de chercheurs. Afin d'évaluer leurs résultats, l'équipe a appliqué leur méthode à la documentation d'un système complexe pour un nouveau télescope astronomique pour le Netherlands Institute for Radio Astronomy. Leur recherche consistait à évaluer, de façon empirique, si la documentation avec notation formelle contribue à une compréhension plus élargie et plus approfondie de l'architecture d'un système comparativement à une documentation traditionnelle (non augmentée d'informations en notation formelle). Les résultats de l'étude ont démontré que la compréhension des architectes est plus approfondie lorsque ces architectes ont accès à une documentation enrichie d'une notation formelle [8].

### **2.3.2 Pertinence**

Une contribution importante de l'étude néerlandaise est la présentation détaillée des défis de la documentation des architectures des systèmes, particulièrement des systèmes complexes ou ayant des composants réutilisables. Ces défis sont :

- la compréhensibilité des documents;
- la disponibilité de connaissances pertinentes;
- la visibilité des interconnexions et la traçabilité entre les documents;
- la disponibilité des informations pour une analyse d'impact des changements;
- l'évaluation de la maturité de la solution conçue (architecture);
- la fiabilité de l'information dans la documentation.

Cette énumération concerne la documentation des architectures qui présentent souvent et nécessairement des documents volumineux. Toutefois, en 2015, ces défis sont pertinents à la documentation technique des projets logiciels en général, et non seulement des architectures, puisque la complexité des projets a crû considérablement. De plus, l'approche agile, qui s'établit comme une voie itérative dans les pratiques de conception, engendre des questionnements nouveaux. Les questions sur le rôle et la nécessité de la rédaction technique et l'équilibre de la quantité versus la qualité de la documentation méritent l'attention.

### 2.3.3 Discussion

L'approche développée par les chercheurs permet d'adresser les défis de la documentation d'architecture par l'ajout d'un niveau de formalisme. Une seconde constatation est qu'une meilleure documentation offre des avantages *asymétriques*, c'est-à-dire bénéfiques pour l'individu qui la consomme plutôt que la personne qui la rédige. Ce scénario se manifeste quand l'équipe de maintenance de logiciel (qui a besoin des informations sur le système) diffère de l'équipe de conception (qui a figé les décisions dans une spécification). Les concepteurs risquent de perdre la motivation à livrer une documentation plus complète. Les auteurs proposent que les balises établies par leur outil rendent possible une meilleure vérification de la documentation, et ce, de façon automatisée. Les chercheurs visent à augmenter la motivation des concepteurs en facilitant la vérification de la documentation (selon les trois aspects « *correctness, completeness and consistency* »). Deux de ces aspects sont partagés avec la cartographie : l'exhaustivité et la cohérence [10]. Améliorer l'efficacité de la production de documentation est un facteur clé de motivation des producteurs de documentation, selon Jansen et ses collègues. Le présent essai vise ainsi à

bâtir sur cette observation afin d'accroître la motivation des producteurs de contenu technique.

## **2.4 Perspective de la méthodologie de développement**

Les méthodologies de développement ont une influence sur la documentation parce qu'elles guident la pratique et la perception du rôle de la documentation par une philosophie du logiciel et une vision sous-jacente aux activités de production et de maintien. Cette perspective est résumée par un sommaire de la recherche, une présentation de la pertinence du sujet, et une discussion supplémentaire.

### **2.4.1 Sommaire**

La recherche a exploré les tensions et contradictions entre les besoins en documentation de l'architecture logicielle et les principes des méthodes agiles [8] [16]. Conscients de l'effort considérable déployé par les spécialistes en conception dans la rédaction et la mise à jour de la documentation des architectures, l'équipe de chercheurs a proposé une solution. L'équipe s'est penchée sur un document essentiel, mais souvent incompris en raison de sa longueur et sa complexité : le document de spécification technique. L'objectif de l'étude était de fournir un gabarit pour documenter la spécification technique d'une architecture logicielle conçue selon les principes de développement agile. Cette étude qualitative, menée dans une entreprise mondiale de développement de logiciel, a repris trois des six défis discutés auparavant [15] :

- la compréhensibilité des documents;
- la disponibilité de connaissances pertinentes;
- la fiabilité de l'information dans la documentation.

Par des sondages et des entrevues, les données ont été recueillies pour l'analyse des difficultés de produire une documentation pour un projet agile [7]. L'analyse a révélé que 52 % des concepteurs considèrent la création d'un document d'architecture comme étant une tâche difficile. De plus, 62 % des participants allouent beaucoup de temps et d'effort à la création d'un tel document. Enfin, 68 % de ces experts ont reconnu que l'utilisation d'un

gabarit pour la création de documents aide à générer l'information requise, informations qu'ils auraient ignorées autrement. Un nouveau défi s'ajoute à la liste précédente : la difficulté, au départ, de créer la documentation. Ils ont donc élaboré un gabarit, une liste de contrôle et un outil pour automatiser l'ajout de données requises [7].

### **2.4.2 Pertinence**

La perspective contributive [7] permet au rôle de qualité de la documentation de rester pertinent malgré l'évolution des méthodologies de développement. Dans un énoncé de position, d'autres chercheurs citent que bien que 50 % des développeurs (équipes de développement agile) accordent une grande ou une très grande importance à la documentation, trop peu de documentation leur était disponible dans leurs projets [16].

La production de documentation plus courte et plus précise, avec une insistance sur la disponibilité des informations et la facilité d'accès aux documents sont deux idées centrales au travail de rapprochement entre les besoins des méthodes agiles et les objectifs de la documentation. Ainsi, une automatisation de la documentation du design accorde à la documentation sa place en méthodes agiles [17]. La documentation se doit donc de devenir un support aidant à la collaboration.

### **2.4.3 Discussion**

Un des faits saillants est l'approche d'analyse employée par l'équipe de chercheurs. Le modèle de théorie ancrée utilisé semble a priori contredire la méthode scientifique [18]. Un des principes de cette approche souligne que les chercheurs ne formulent pas d'hypothèse dès le départ. Ils développent leurs théories en analysant les données sur le terrain du champ d'étude. Tout comme cette approche « non traditionnelle » a permis de générer une solution à valeur ajoutée, similairement, une méthodologie « non traditionnelle » de développement, soit la méthode agile, peut contribuer à l'avancement du domaine de la rédaction. À priori, la méthode de recherche semble contradictoire. En réalité, elle met en évidence la possibilité et l'utilité de rapprocher les écarts entre l'approche agile et la documentation.

## **2.5 Perspective de résolution de bogues**

Dans ce sous-chapitre, il est question de la perspective de résolution de bogues et suit deux thèmes : le sommaire de la recherche et la pertinence à l'exercice actuel.

### **2.5.1 Sommaire**

Une équipe canadienne de chercheurs a effectué une étude empirique des facteurs du temps de résolution des bogues. Les chercheurs se sont penchés sur la question du temps entre le signalement d'une défaillance et le moment auquel un développeur commence la correction [19]. Ils ont examiné les rapports de bogues, le code source défectueux et le code du correctif. Plusieurs des facteurs de coût trouvés sont reliés à la structure interne du code source, mais un élément pertinent est discuté dans la section suivante.

### **2.5.2 Pertinence**

Cette étude est pertinente puisque la longueur des commentaires dans le rapport de bogues (un aspect de documentation) constitue un des facteurs importants de délai dans le démarrage de programmation d'un correctif. Les auteurs ont suggéré que les rapports de bogue trop longs soient divisés en rapports plus courts. Les chercheurs ont prouvé empiriquement que l'organisation structurelle (par sa longueur) d'un document de type texte exerce un impact sur la résolution du problème.

## **2.6 Perspective de l'utilité**

L'utilité de la documentation est un sujet substantiel en recherche et dans la pratique, tel que résumé dans le sommaire suivant. La pertinence et la discussion supplémentaire détaillent et nuancent ce thème important.

### **2.6.1 Sommaire**

Une étude de cas examine une entreprise qui produit des logiciels pour systèmes intégrés de navigation. En employant la revue systématique, les chercheurs ont voulu comprendre l'utilité

et l'utilisation des documents techniques produits pendant les phases de production et de maintien du logiciel [11]. L'étude a révélé que la documentation est plus souvent consultée lors de la production de logiciel (documents de conception). Lors du maintien, la source de documentation la plus consultée provient des commentaires dans le code source. Les chercheurs n'ont pas trouvé de types de documents plus particulièrement consultés durant la production et le maintien. De plus, la recherche a confirmé l'hypothèse selon laquelle les facteurs qui engendrent le plus d'impact sur l'utilité de documents sont la précision, l'exactitude et l'actualisation de l'information contenue.

L'équipe a conclu que l'utilisation que faisaient les professionnels de la documentation différait selon le besoin. Ce besoin varie selon la nature des informations requises par les professionnels et des tâches qu'ils ont à effectuer. Les auteurs de l'étude affirment que les recommandations émises dans leur publication ont rapporté des bénéfices à l'entreprise hôte de leur étude de cas.

### **2.6.2 Pertinence**

Tout d'abord, cette étude est pertinente puisqu'elle met de l'avant un cas de partenariat entre les secteurs privé et universitaire. Peu d'études présentent ce type de partenariat [10]. La recherche plus récente tente de répondre à ce besoin [11].

L'étude aide les professionnels à juger de la quantité et de la portée requise pour la documentation technique ([11], p. 665). En étudiant ce qui apparaît utile pour les experts en TI, les chercheurs veulent encourager les professionnels à investir leurs efforts de documentation de façon plus judicieuse. De manière similaire, le présent essai vise à augmenter la motivation pour la production d'une documentation utile. Cet essai retire donc plusieurs réflexions pertinentes du cas NovAtel.

Par exemple, une des recommandations des chercheurs est de fournir une documentation plus complète et plus précise lors de la phase de conception puisque les résultats démontrent que ces documents sont fortement consultés.

Les auteurs évoquent d'autres recommandations : l'amélioration des commentaires dans le code source, la division des documents volumineux et la création de documents succincts. La recherche appuie la formulation d'une métrique d'utilité, le développement d'approches

statistiques pour calculer l'utilité d'un document, et la suggestion d'une métrique du retour sur investissement (*cost-effectiveness index*). Enfin, le concept d'une documentation temporaire (*opportunistic or situational documentation*) et l'identification de la population cible des artefacts sont proposés.

### **2.6.3 Discussion**

L'étude de cas est contrainte par des limites. Ces limites, sans soustraire aux mérites de l'étude comme référence pour cet essai, appellent à la prudence quant à la généralisation des résultats. L'environnement technologique de l'entreprise est « typique » selon les auteurs de l'étude, et l'analyse des données est rigoureuse selon eux. L'étude d'une seule entreprise ne permet pas de formuler des théories concluantes généralisables (ni à l'industrie des TI des logiciels intégrés ni à d'autres entreprises de domaines variés). Le nombre de documents étudiés contraint la portée des conclusions. Des études répétées avec un plus grand échantillon sont requises. Malgré ces limites, la publication a exploré des thèmes originaux et valides à propos du sujet.

## **2.7 Perspectives additionnelles**

La pertinence actuelle du sujet repose sur la réévaluation des perceptions de la documentation de ces débuts dans le développement de logiciel jusqu'aux perspectives nouvelles. Cette réflexion s'inscrit dans un contexte de mise en pratique pragmatique et contemporaine des besoins en documentation. Le rôle de la perspective du développement collaboratif fait partie des nouvelles réalités. Une perspective théorique basée sur les idées de Royce est présentée, suivie d'une exploration des aspects pertinents de la dette technique. Cette discussion, intéressante et valable pour offrir des perspectives supplémentaires d'approfondissement du sujet, est présentée dans l'annexe 2.

## **2.8 Aperçu des métriques existantes**

Les métriques existantes de mesure de la documentation sont présentées à l'annexe 3.

## **2.9 Conclusion du chapitre**

Ce chapitre a présenté quelques pistes d'intérêt qui guident l'élaboration de la question de recherche de cet essai. Le prochain chapitre examine la problématique de recherche et en définit les objectifs et les hypothèses. Le contenu annonce aussi la méthodologie proposée et les limites de l'étude.

## **Chapitre 3**

### **Problématique**

Ce chapitre contient une description de la problématique qui comprend une explication des objectifs, des hypothèses et des limites de l'essai. Par la suite, celui-ci explique la méthodologie et une discussion du type de recherche effectuée et de la validité. Finalement, le chapitre présente la mise en œuvre avec l'explication du déroulement de l'essai.

#### **3.1 Description**

L'absence de documentation et la présence de documents inadéquats sont un facteur de coût de la maintenance de logiciel selon les chercheurs [16], [20]–[22]. Seulement 120 études décrivent les trois aspects coût-bénéfice-qualité de la documentation de logiciel, malgré un processus rigoureux de recherche par combinaison de mots-clés par permutation ([10] p.179). La recherche s'est limitée aux documents en langue anglaise; or, étant donné l'historique du développement de logiciel, cela est une réalité sans impact négatif. Après un processus de vérification itérative, seulement 69 études contiennent un lien de pertinence avec les thèmes (coûts, bénéfices et qualité de la documentation technique en génie logiciel). Les auteurs de la méta-analyse ont déploré le nombre restreint de recherches, ainsi que la quasi-absence de certains thèmes dans la recherche quantitative. De plus, vu le manque de modèles, les auteurs ont proposé un modèle pour chacun des aspects de la documentation technique analysée. Cette équipe n'est pas la seule à revendiquer davantage de rigueur en documentation. En effet, David Parnas a écrit un chapitre important au sujet de la rigueur dans la documentation dans un manuel de référence publié par Springer et employé dans la formation des étudiants universitaires [6]. Il présente l'argumentaire que l'amélioration des logiciels passe par l'adoption d'une documentation précise, exacte et complète (dans la mesure de la portée et du rôle qu'elle doit jouer). L'industrie informatique produit des solutions critiques et ce domaine doit en prendre conscience. Le mandat est d'atteindre un niveau de maturité dans lequel la documentation est l'allié de la conception [6]. La

documentation prend ainsi sa place en informatique comme elle le fait déjà dans les autres domaines scientifiques et d'ingénierie. Rédigée sérieusement dès le début du processus et utilisée comme un moyen d'effectuer le suivi des choix, elle véhicule ainsi la conception d'un système (*documentation as a design médium*) [6].

Pour permettre l'atteinte de cet objectif, une mesure quantifiée du retour sur investissement est une contribution utile. Cette contribution est proposée dans les sections suivantes.

### **3.1.1 Objectifs**

L'objectif principal du projet est de démontrer une corrélation entre la documentation technique d'un logiciel et le temps de maintien du logiciel dans une optique de réduction du temps de résolution de défaillances techniques.

Les objectifs de soutien à cet objectif principal sont les suivants :

- contribuer à combler une des lacunes observées par les experts dans la disponibilité de recherche ciblée;
- offrir une quantification d'un sous-ensemble d'éléments de la qualité de la documentation;
- offrir une grille de calcul pour une nouvelle métrique de qualité de la documentation;
- amorcer une réflexion et un projet pertinents et utiles pour d'éventuelles recherches dans le domaine.

Ces objectifs sont particulièrement valables à cause des lacunes de la recherche actuelle et de ses obstacles. Certains chercheurs informaticiens croient que la documentation du logiciel ne constitue pas un domaine de recherche en informatique parce que la rédaction technique ne se base pas sur des algorithmes ou sur des modèles de calculs ([1] p. 126). Or, un des objectifs de cet essai est justement de concevoir un modèle qui démontre un aspect quantifiable de la documentation technique. Il faut toutefois éviter de tomber dans l'autre extrême, c'est-à-dire la catégorie de chercheurs qui se sont intéressés à la mise au point d'un système métrique pour la documentation sans tenir compte des besoins des praticiens du génie logiciel. Par conséquent, les notations produites s'avèrent peu utiles pour l'industrie du logiciel (et peu utilisées par les professionnels).

### 3.1.2 Questions de recherche et hypothèse

La question principale (QP) abordée par cet essai est :

QP1. Est-ce que la documentation de qualité permet de réduire le temps de résolution de bogues pour un logiciel existant et opérationnel, en environnement de production?

Des questions secondaires (QS) qui complètent les investigations sont :

QS1. Quel est l'ensemble des attributs nécessaires pour faire un bilan complet de la qualité d'un document, en reflétant les aspects de la qualité les plus saillants connus dans le domaine?

QS2. Quels barèmes doivent être accordés à chacun des attributs de la question secondaire QS1?

QS3. Cette métrique est-elle réutilisable pour la recherche subséquente et pour l'industrie?

QS4. Est-ce qu'une telle métrique est considérée comme un facteur d'analyse valable de la qualité de la documentation?

**L'hypothèse** consiste à vérifier qu'une documentation de qualité mesurée permet de réduire le temps de résolution de bogues pour un logiciel existant. Cette mesure se base sur le concept d'un indice de mesure de documentation complète.

Afin de schématiser la problématique, un processus itératif est adopté. Tout d'abord, le schéma de la Figure 3.1 montre le sujet de façon générale. La Figure 3.2 présente la seconde itération du modèle de l'étude.

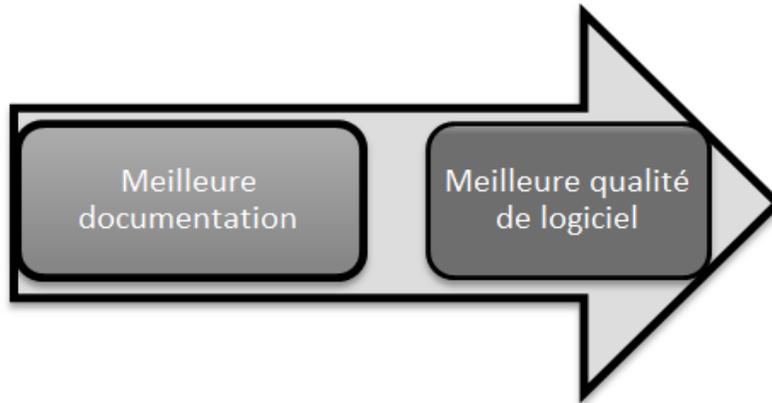


Figure 3.1 Modèle conceptuel du projet, première itération

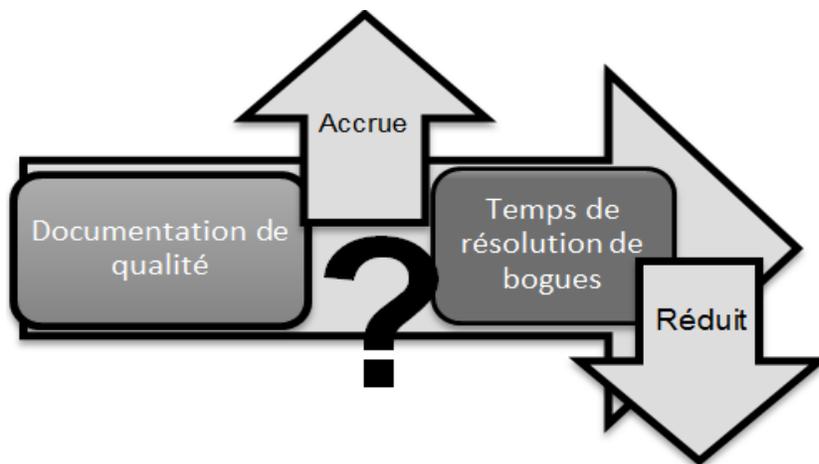


Figure 3.2 Modèle conceptuel du projet, deuxième itération

### 3.1.3 Limites

Les limites de l'essai concernent la quantité de documents à étudier (pour la production de mesure) ainsi que le domaine d'affaires. L'étude est centrée sur un groupe de documents qui pourraient ne pas refléter ce qui est produit dans diverses industries. La quantité idéale de documentation à produire dans chaque type d'organisation n'est pas définie. Les exigences réglementaires, légales et culturelles varient. Toutefois, la variété des formes de documentation, incluant celles internes au programme, permet de mitiger quelque peu cette

contrainte. Cette étude se limite aussi aux logiciels déployés et présents en production au moment de l'étude. Les documents analysés sont seulement ceux de nature technique (conception technique, spécification, documents opérationnels, permissions et sécurité).

## **3.2 Méthodologie proposée**

Les sections et sous-sections suivantes couvrent le type de recherche et les méthodes de validation de manière détaillée.

### **3.2.1 Type de recherche**

Les approches qualitatives caractérisent souvent la recherche en documentation technique à cause de la difficulté de mesure des éléments fondamentaux. Le plus grand besoin étant dans le domaine quantitatif, cela constitue l'approche préconisée pour cette démarche. Un contexte empirique et un autre théorique sont présentés dans le chapitre suivant. La mise en relation de variables est effectuée dans le but d'émettre des recommandations et de guider la pratique. Le type de recherche est donc une recherche empirique, quantitative, corrélacionnelle et prédictive.

### **3.2.2 Méthodes de validation**

La validité est analysée selon quatre axes. Les concepts théoriques qui sous-tendent ce cadre sont repris de la recherche [23].

#### **3.2.2.1 Validité interne**

La qualité de l'analyse effectuée influence la validité interne. Les attributs choisis pour la métrique, les barèmes établis, l'application des calculs et l'analyse corrélacionnelle affectent la fiabilité des conclusions. Des facteurs influencent la preuve de lien entre la métrique et le résultat. Un indice élevé n'est pas forcément lié à une résolution rapide de bogues. En effet, un facteur humain peut en être la cause.

Un des aspects de la validité est le contrôle des variables parasites. En fait, dans la réalité, un ensemble de facteurs, outre la documentation, influence le temps de résolution des

bogues. Une attention est portée au contrôle de ces variables pour en limiter les effets nuisibles : les bogues sont sélectionnés selon une liste fixe de critères (date, type, logiciel, système).

#### **3.2.2.2 Validité externe**

Il faut faire attention à la limitation quant au pouvoir de généralisation des résultats. La validité des résultats de l'analyse de l'échantillon n'est pas toujours applicable à la population. Afin d'obtenir un échantillon plus représentatif de la population, et ainsi réduire les risques à la validité externe, la sélection de l'échantillon est aléatoire. Cela augmente la certitude que les conclusions sont fiables.

#### **3.2.2.3 Validité de conclusion**

Les efforts menés risquent de ne pas indiquer à une relation de causalité avec la variable dépendante basée sur les résultats des statistiques [28]. La nature de la recherche corrélationnelle ne permet pas de valider une relation de cause et d'effet. La structure de la recherche et les choix effectués peuvent causer un biais lors des conclusions.

#### **3.2.2.4 Validité de construit**

Si la mesure de documentation développée ne porte pas une signification réelle ou si le modèle théorique élaboré contient des erreurs fondamentales, cela réduit la validité du construit. Si la métrique développée mesure réellement la qualité, la fiabilité, l'accessibilité, la mise à jour et les autres aspects de la documentation, alors la validité du construit est grande.

### **3.3 Mise en œuvre**

Les sections qui suivent élaborent les étapes de déroulement de l'essai, la description de l'environnement de travail, la population cible, et la logistique de mise en œuvre du travail.

### 3.3.1 Déroulement de l'essai

Les étapes suivies pour la réalisation de cet essai se trouvent dans le diagramme suivant.

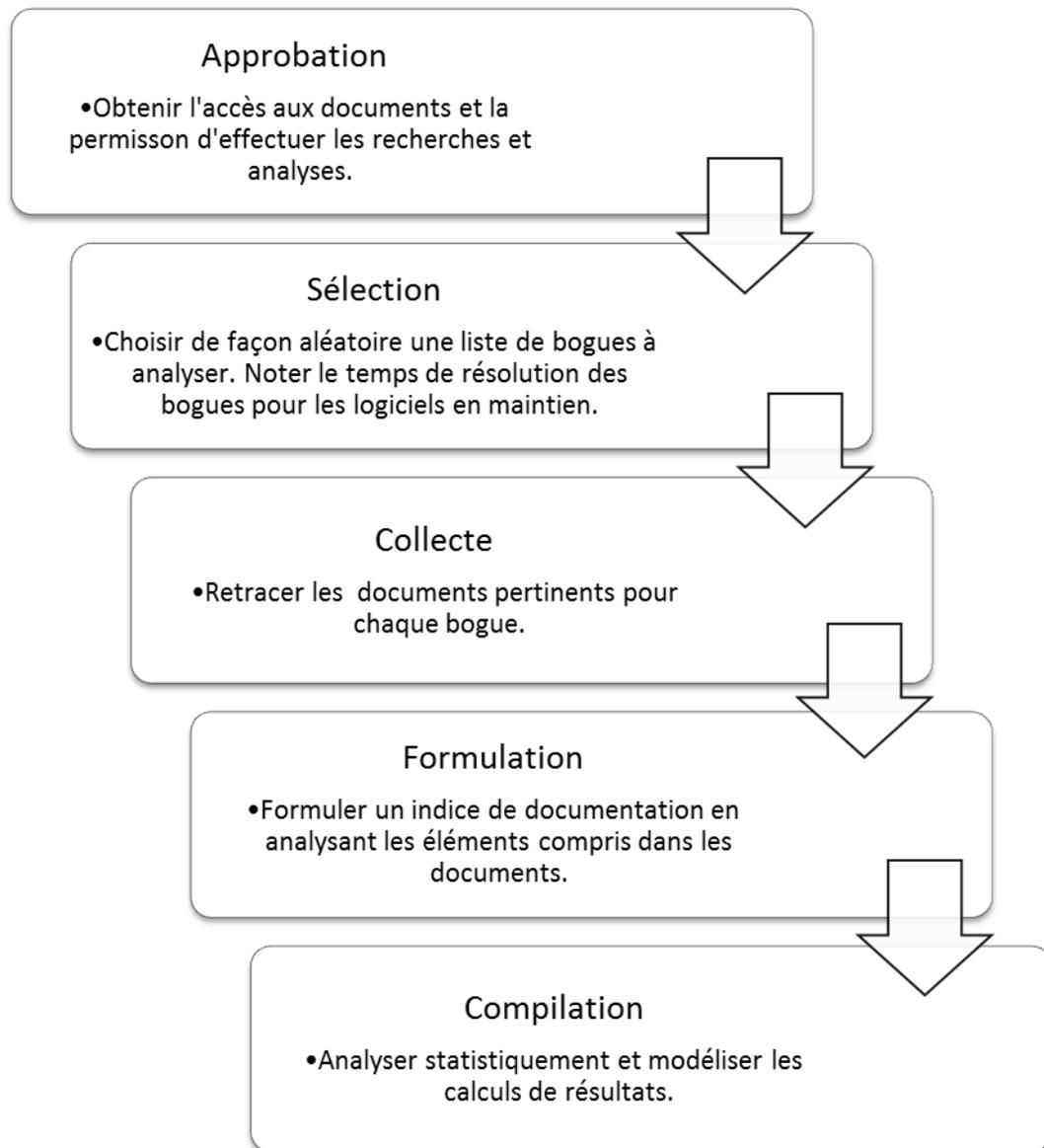


Figure 3.3 Modèle de séquençage des activités de l'essai

Les sections suivantes décrivent l'environnement de travail dans lequel le projet est développé, la population cible, et la logistique de la mise en œuvre.

### **3.3.2 Environnement de travail**

L'environnement de travail est une entreprise dans le domaine parapublic au Québec. Cette entreprise détient un service des technologies de l'information qui contient plusieurs groupes (systèmes informatiques, support applicatif, infrastructure et réseautique, technologie éducative, bureau de gestion des projets). La recherche se fait dans le secteur de l'entreprise qui gère les systèmes informatiques. Cette unité est composée de plusieurs équipes (conception, développement logiciel, assurance qualité, gestion des bases de données, entrepôt de données) qui collaborent étroitement pour la réalisation de leurs mandats. L'équipe dédiée au développement logiciel (toutes les phases) est composée d'une centaine d'employés. Cette équipe développe des systèmes d'envergure, c'est-à-dire utilisés par des milliers d'utilisateurs. Sur un budget opérationnel de plusieurs centaines de millions de dollars, une vingtaine de millions de dollars sont alloués aux TI.

La gestion de projet TI se déroule selon un processus basé sur le guide PMBOK et adapté aux besoins de l'organisme. Le diagramme de l'annexe 5 en résume les principales étapes. Pour un projet, une initiative doit mener à une solution (produit ou service) qui impacte significativement l'offre des TI. Le projet nécessite un effort d'au moins 75 jours personne. L'analyse de l'essai se restreint à la phase d'opérations.

Le budget alloué aux projets TI, la taille et la complexité de la structure organisationnelle et la disponibilité d'une technologie et d'une culture propices au maintien de la documentation sont autant de facteurs favorables à l'étude et à des résultats adéquats pour une analyse.

### **3.3.3 Population cible**

La population cible qui bénéficie de la démarche et des fruits de cet essai est toute entreprise qui effectue du développement logiciel et qui doit aussi assurer l'entretien de ses développements. La documentation est un élément clé, peu importe la taille de l'entreprise. Ce travail est aussi bénéfique aux entreprises qui externalisent le maintien des logiciels afin qu'elles soient au courant des gains qu'elles peuvent cumuler en offrant une documentation

adéquate aux fournisseurs du service de maintien. Dans un contexte de ressources limitées, l'identification des aspects de la documentation qui ont plus le plus grand apport positif est une poursuite avantageuse.

Les acteurs de la population cible sont représentés à la **Erreur ! Source du renvoi introuvable.** L'ensemble comprend quatre groupes particulièrement d'intérêt, soient les analystes (toutes catégories), les développeurs, les rédacteurs, ainsi que les gestionnaires. Cette dernière catégorie contient les gestionnaires de projets, les gestionnaires opérationnels et les gestionnaires techniques.



Figure 3.4 Schéma de rôles occupés par la population ciblée par l'essai

### 3.3.4 Mise en œuvre

L'expérimentation a lieu du 1<sup>er</sup> février au 1<sup>er</sup> mai 2016 aux bureaux du département TI de l'entreprise choisie, à Montréal. L'analyste principale effectue la collecte des documents (étape 1) ainsi que la classification et la catégorisation préliminaires. Les résultats sont présentés à différents moments à une analyste experte de l'entreprise et à un superviseur académique. Leur rétroaction est considérée pour la suite de l'analyse. L'analyste effectue la délimitation de la population de l'échantillon et procède à la sélection de l'échantillon de

bogues à analyser. Cet aspect est validé par un membre du corps enseignant au CeFTI. Les éléments de l'échantillon sont examinés dans l'optique de reconstruire un narratif de la défaillance et des étapes qui mènent à la solution. Étant donné la situation des ressources humaines (nombreux départs à la retraite, roulement de personnel) et la mise en place de processus ITIL dans les années récentes, le repérage des justifications et des raisonnements dans les solutions de bogues n'est pas facilité. Tant que cela est possible, l'objet et la justification de la solution sont retracés dans la mesure qu'ils sont importants pour fixer les aspects temporels de la solution : *quand* le problème a été signalé et *quand* celui-ci est clos.

L'analyste principale complète ensuite le gabarit détaillé de caractéristiques de chaque document pour chaque bogue et présente les résultats sous forme de tableau. Des notes et observations sont conservées sur l'application du gabarit et sur les cas particuliers.

L'analyse des calculs de la mesure est effectuée par l'auteur de l'essai avec validation d'experts. L'application des règles individuelles convertit les caractéristiques et les valeurs observées à propos d'un document en une valeur qui représente la quantification de cet attribut de la qualité. C'est la phase de quantification de la qualité de la documentation, un attribut à la fois, pour chaque document, qui est l'élément distinctif de l'approche.

L'auteur de cet essai entreprend la dernière étape à l'aide du chiffrier Excel pour produire les graphiques des résultats corrélés. Les tendances observées concernent tout d'abord la relation entre l'indice et le temps de résolution. Ensuite, un regard plus attentif et porté sur les attributs individuels et leurs indices pour examiner les tendances plus détaillées. L'interprétation de la signification de la corrélation octroie un sens aux valeurs numériques. Des recommandations émises traduisent les indices à apport significatif en conseils applicables.

Le chapitre suivant adresse l'approche en détail. Les outils, l'échantillonnage, la description des logiciels, la discussion de l'approche et des facettes de la métrique, le sommaire du calcul, la validation et la présentation des résultats attendus constituent le contenu du chapitre.

## **Chapitre 4**

### **Approche proposée**

L'étendue du manque de documentation lors du maintien de logiciels [24] [25]. La correction des bogues est un facteur de coût important de la maintenance [26] [33] [34] [35]. Il est justifié de vouloir déterminer si l'état de la documentation affecte la correction des bogues et de quantifier les bénéfices d'une documentation accrue et pertinente. Le chapitre couvre les aspects saillants de l'approche choisie : population et échantillon, description des logiciels, description détaillée de l'approche, calculs de la métrique et présentation du format des résultats.

#### **4.1 Description de l'approche**

L'approche consiste à cerner les outils utiles sur le terrain de la collecte des données, la sélection de population et d'échantillon, le choix des critères et attributs de qualité, l'établissement des barèmes, leur transformation en indices, la production d'indicateurs et la formulation d'une corrélation, les facteurs clés de succès et la validation des résultats.

##### **4.1.1 Outils**

Les outils à employer comprennent un service web de suivi des problèmes qui relate toutes les informations pertinentes pour l'analyse d'un bogue : temps de signalement, échanges de communication, requêtes, module fautif, temps de résolution et clôture du signalement. Les signaux de défaillance sont émis par l'outil de gestion de l'automatisation Automic. Les outils de développement de base de données sont Ultraedit et Toad. Ce dernier est aussi l'outil de développement du code source de validation des données de l'essai. La gestion des versions est maintenue par l'outil PVCS autant pour la documentation que pour le code source, et rapporte la date et l'heure de la modification, le numéro de la version, l'auteur et la description de la modification.

## **4.1.2 Programmes sélectionnés et échantillon**

Les sous-sections suivantes contiennent un exposé des détails des ressources logicielles, du choix de l'année de l'échantillon et de son contenu,

### **4.1.2.1 Affectation des ressources logicielles**

Les logiciels sélectionnés pour l'étude couvrent les domaines suivants : la gestion des dossiers clients, la production de rapports, la population de données-clés, l'échange électronique de données. Les composants sélectionnés ont l'ampleur suffisante pour créer un grand besoin de communication lors de toutes les phases de la conception, au développement, et au test et déploiement des logiciels. Les projets prennent plus de 75 jours personne, ont une vie utile prévue de plusieurs années et doivent fonctionner selon un horaire établi (quotidiennement, hebdomadairement, mensuellement ou trimestriellement) ou sur demande. Un minimum de six équipes a travaillé sur chaque projet, soit au moins l'équipe du client, l'équipe d'analyse et de conception, l'équipe de développement, l'unité de gestion de la sécurité informatique, l'équipe de gestion des applications et le groupe de gestion des opérations. Dans chaque équipe, un minimum de deux personnes était affecté au projet (un membre de l'équipe et un gestionnaire). La recherche est rétrospective puisque les projets sont tous complétés et que les applications résident en phase d'exploitation. Le modèle de gestion de projets de l'entreprise est placé dans l'annexe 5.

### **4.1.2.2 Choix de l'année de l'échantillon**

La sélection de l'échantillon débute par une liste de bogues survenus durant une période déterminée. La période choisie est l'année 2012. Cette sélection permet de visualiser une année entière de données pour des systèmes stables. Entre 2001 et 2011, les systèmes subissaient des améliorations et plusieurs mises à jour majeures du PGI ont eu lieu. Ces années sont donc exclues de l'échantillon parce que l'intention du projet est d'analyser des éléments opérationnels dans une phase qui ne manifeste pas les particularités et les contraintes des mises à jour majeures. À partir de 2014, le processus pour rapporter les bogues a fait l'objet de remaniement avec de nouvelles procédures, et un nouvel outil de communication entre les équipes a été introduit. Cela a engendré des besoins en formation

des ressources humaines pour développer la connaissance des employés à propos de l'outil et assurer une uniformité dans l'utilisation du nouvel outil pour permettre des analyses métriques cohérentes dans le futur. Les années possibles pour l'échantillon sont donc 2012 et 2013. Parmi celles-ci, la première est choisie au hasard. Une discussion détaillée du choix de l'échantillon se trouve à l'annexe 4.

### 4.1.3 Critères, attributs et barèmes

L'approche utilisée dans cette étude est quantitative et de type corrélationnel puisque l'analyse explore l'existence d'une corrélation entre le niveau de documentation technique et la durée d'existence des défaillances dans une application. L'approche de quantification de la qualité de la documentation et son jumelage à un temps de résolution de bogue se définissent en quatre étapes distinctes. Le schéma 4.2 présente l'approche utilisée.



Figure 4.1 Schéma des étapes de la recherche

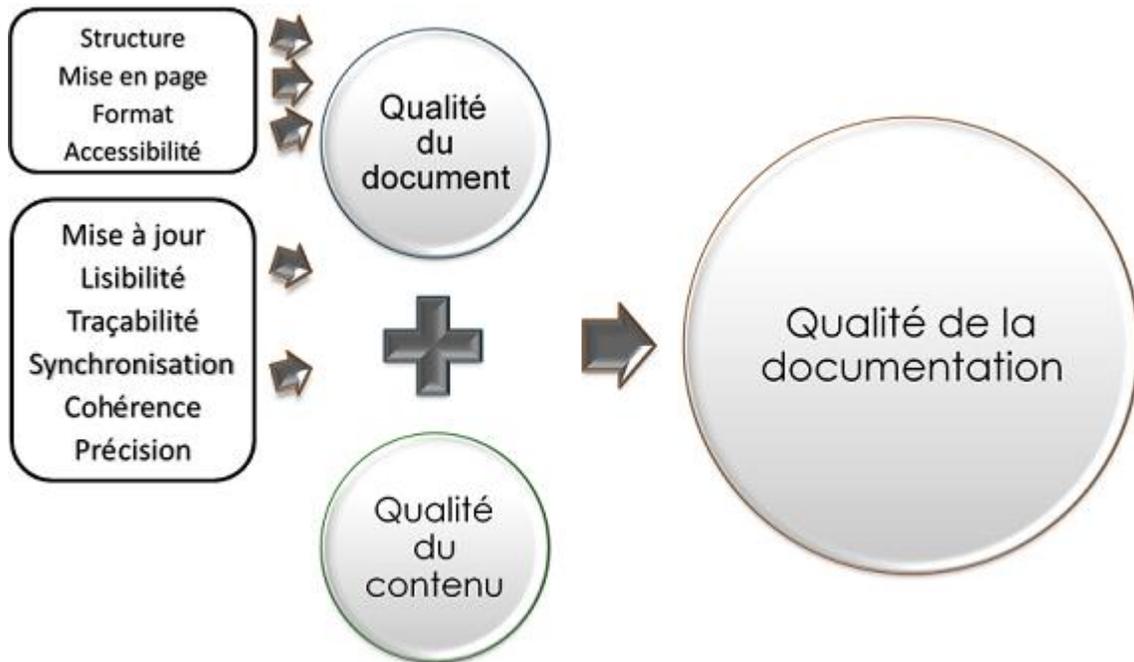
Pour chacun des projets sélectionnés selon la liste aléatoire :

1. Grouper : inventaire des documents existants
2. Filtrer : évaluation des documents en fonction des critères de l'indice
3. Retracer : évaluer le nombre de bogues et leur délai de résolution des projets

#### 4. Corréler : corréler l'indice de documentation et l'effort de support des projets

En premier, il s'agit de regrouper l'ensemble des documents de chacun des projets affectés par un bogue. Par exemple, pour un programme de production de rapport, l'analyste trouve la liste des documents de conception, de spécification et de tests dans le système de gestion de projets et de gestion des révisions. Un répertoire externe est créé pour maintenir ces documents pour éliminer le risque de modification pendant la phase d'analyse. Une liste des caractéristiques de chaque document est créée pour répertorier les détails distinctifs au document. La longueur de chaque document en mots et en pages, de la date de dernière modification et de dernière consultation, du nombre de contributeurs, du nombre de révisions différentes et du nombre de graphiques contenus dans le document, si applicable. Les informations proviennent du document lui-même ou du système de gestion des versions.

Puis, un ensemble de critères est utilisé pour générer la liste finale des documents retenus. Donc, une méthode d'évaluation des documents est développée par l'analyste principale et ensuite validée par deux experts : soit un analyste expert du domaine d'affaires et un expert académique. Quatre éléments sont retracés en vue de définir un indice de documentation. Ces quatre éléments sont basés sur les critères les plus saillants dans les connaissances actuelles. Ils sont la structure, l'exhaustivité, la mise à jour, l'accessibilité, et la traçabilité. La Figure 4.2 montre quelques principaux éléments de la qualité.



**Figure 4.2 Modèle de qualité de la documentation technique**

Traduction libre

Modifié de : Zhi, J., Garousi-Yusifoğlu, V., Sun, B., Garousi, G., Shahnewaz, S., et Ruhe, G. (2016), p. 183

#### 4.3.3.1 La structure

Les quatre éléments pour la pondération sont présentés.

- La longueur totale du document en mots et en pages : ces informations sont fournies par le logiciel de traitement de texte. Dans le cas de documents de formats différents, une conversion vers le format texte est effectuée pour les calculs et la vérification de la qualité. Pour simplification, au besoin, il est considéré qu'une ligne complète se compose de 15 mots et qu'une page complète contient 30 lignes. Étant donnée l'envergure des projets, le barème suivant est établi pour la longueur des documents :
  - Si la longueur excède 20 pages ou n'a qu'une page : +0
  - Si la longueur se trouve entre 10 et 19 pages : +0,5
  - Si la longueur se situe entre 2 et 9 pages : +1

- Si le document contient 1 page seulement, mais que le nombre de mots est plus grand que 300, alors la cote +0,5 est ajoutée au lieu de 0, car la page est bien remplie. Si le document a une longueur de 2 à 3 pages, mais il contient moins de 300 mots, alors la cote +0,5 est accordée au lieu de 1, car plusieurs pages paraissent peu remplies.
- Si le document n'existe pas, la cote est -4.

Il est jugé qu'un document trop court ne recèle pas assez d'information. Un document trop long est aussi peu coté puisqu'il est possible que ce document englobe des sujets pertinents pour deux documents ou plus. La lisibilité en est aussi affectée.

- Le nombre d'erreurs : une vérification du texte est effectuée dans un outil de traitement de texte. Le nombre d'erreurs d'orthographe et de grammaire soulevées par l'outil est retenu. Ensuite, le nombre d'erreurs est évalué selon le nombre de lignes et de pages. Le barème est :
  - S'il y a plus de quatre (4) erreurs par page : +0
  - S'il y a 4-8 erreurs par document : +0,5
  - S'il y a moins de 10 erreurs par document : +1

Les fausses alertes sont omises (les mots techniques, les abréviations et le jargon non reconnu par l'outil) pour produire une évaluation plus juste de la qualité de la rédaction du document.

- La présence d'éléments visuels : des tableaux, des graphiques, des diagrammes UML ou autres. Le barème est :
  - Aucun élément visuel : +0
  - Un à deux items : +0,5
  - Trois ou plus : +1

Les éléments visuels contribuent à la lisibilité des documents et démontrent qu'un effort accru a été consacré à la divulgation des informations et à l'explication.

- La mise en page : la mise en page révèle la structure interne du contenu dans le document et englobe des caractéristiques qui affectent la lisibilité du document. L'emploi d'un gabarit commun et l'organisation spatiale des mots servent d'exemples.

L'organisation du texte, les polices de caractère, les couleurs, et les styles (par exemple, les textes avec extraits biffés forment la base de l'évaluation de la mise en page. Il en résulte le barème suivant :

- Si la mise en page est ordonnée et rédigée selon un **gabarit** : +0,5
- Si la mise en page est **ordonnée**, mais ne respecte pas de gabarit : +0,25
- Si la mise en page est **désordonnée** ou difficile à suivre (ou elle contient une majorité d'éléments de formatage affectant la lisibilité.) : +0

L'ensemble des attributs contribue à une cote minimale de « 0 » (quand un document existe) et une cote maximale de « 3,5 ». Toutefois, un document non trouvable accorde la cote « -4 » par document. Pour qu'un document soit classé comme introuvable, il doit y avoir une insuffisance dans les types de documents trouvés. Pour un module qui exécute une chaîne de logiciel, un document à propos des séquences de logiciels à rouler dans la chaîne joue un rôle essentiel et son absence (ou s'il est introuvable) engendre une cote « -4 ».

La qualité du document n'est pas seulement affectée par la qualité de la langue, mais aussi, grandement, par la qualité de l'information. Un contenu erroné contient des faussetés dans les données fournies, des configurations inexactes, des descriptions trompeuses qui ne correspondent pas à la réalité du logiciel, des erreurs dans les algorithmes ou autres. Cet aspect de validité est décrit par les attributs précision, exactitude, et fiabilité (traduction libre de *correctness*, *accuracy*, *trustworthiness* [10]). Seule une évaluation méticuleuse par un expert peut déceler ce type d'erreur, car son repérage ne peut aucunement être rendu générique. Cette analyse, bien que désirable, est hors la portée de cet essai.

#### **4.3.3.2 L'exhaustivité**

L'exhaustivité est un critère représenté par deux attributs : le cycle de vie du logiciel et la couverture du code source. L'attribut du cycle de vie du logiciel est indiqué par une mention qui décrit le document. Cette mention inclut les opérations, la conception technique, les spécifications, et la sécurité. Le type de document est indiqué, mais ne se traduit pas par un indice. L'attribut de couverture du code source indique si la documentation décrit le code source en entier ou par parties. Cet attribut n'est pas mesuré.

#### 4.3.3.3 La mise à jour

La tenue à jour des documents est importante pour l'indice de documentation vérifiée puisqu'elle représente à la fois un effort fourni dans l'élaboration et le maintien de la documentation et une indication de l'importance et de l'utilité du document. Deux attributs de la mise à jour sont mesurés :

- La fréquence :
  - Aucune mise à jour : -0,5
  - Entre une et trois mises à jour : +0,25
  - Entre 4 et 6 mises à jour : +0,5
  - Plus de 7 mises à jour : +1

Un document peut être utile et de qualité s'il n'a pas subi de révision. Il est toutefois tenu comme vrai qu'un document substantiel (plus de contenu) nécessite une ou plusieurs révisions et mises à jour. Les révisions sont comptées à partir du système de gestion des versions pour les documents qui y résident, ou à partir du système de gestion de la documentation ou des répertoires pour les autres.

Le but principal d'évaluer la mise à jour d'un document est d'en juger l'actualisation ainsi que sa fidélité au logiciel auquel il se rattache. Un document qui est mis à jour lorsque le logiciel l'a été donne une situation idéale et doit se refléter dans le calcul de l'indice de la documentation du projet. Le système de gestion des versions est vérifié pour le valider qu'il y a eu des changements aux deux items dans un intervalle rapproché. Le deuxième attribut mesuré est la synchronisation du document au code source.

- La synchronisation :
  - Lorsque le document est modifié jusqu'à 10 jours avant ou après la modification d'un logiciel (tel qu'indiqué par le dépôt de la version) : +2
  - Si le document n'est pas mis à jour, la déduction est -2
  - Lorsque non applicable : 0

Pour certains documents, une mise à jour n'est pas requise avec chaque modification du code source. Par exemple, un fichier de permissions (classes de sécurité) et un document de création des objets d'un rapport sont des documents essentiels pour le fonctionnement des

logiciels. Une mise à jour du code qui ne conserve pas le statu quo tel que documenté dans ces artefacts produit certainement un problème qui se trouve alors signalé (rapport non visible, erreur d'accès du module aux données d'entrée). Il faut tenir compte des documents non mis à jour mais possiblement non requis. Cela peut être à cause de la nature de la modification du code source. Sans pénaliser outre mesure l'ensemble des indices, le barème est de « 0 » pour les documents sur lesquelles l'évaluation de la synchronisation est jugée non applicable.

#### **4.3.3.4 L'accessibilité**

Plusieurs recherches qualitatives [1] [42] – [45] et quantitatives [43] [44] affirment le rôle de l'accessibilité des documents dans la qualité de la documentation [10]. La facilité de trouver, de consulter, de modifier et de partager un document est évaluée.

- Facilité d'accès :
  - Document non accessible, mal répertorié ou mal nommé (de façon à le rendre inaccessible pour les personnes concernées) : +0
  - Document difficilement accessible seulement sur demande, ou après l'octroi de permissions supplémentaires (et délais) : +0,5
  - Document accessible pour les personnes concernées : +1
- Le type de fichier :

Afin de calculer l'impact du format sur l'accessibilité, ce deuxième attribut est mesuré comme suit :

- Documents de format standard (traitement de texte type .docx, .txt, ou similaire, les PDF, les fichiers Excel) : +0,25.
- Documents de format non standard (.sqr, .ini, etc.) : -0,25.

#### **4.3.3.5 La traçabilité**

Un aspect apparenté à l'accessibilité est la traçabilité des documents. Cela décrit la capacité de retracer les modifications au document à l'aide d'informations qui expliquent quand, où, et par qui un document a été modifié. La traçabilité en rédaction technique requiert une attention particulière parce qu'elle concerne l'évolution d'un document relié à un logiciel, et que ce document doit être mis à jour lorsque le logiciel est modifié. La traçabilité assure que

durant ces évolutions du logiciel, tous les changements à la documentation sont vérifiables et justifiés [10].

Un élément relié à l'accessibilité et la traçabilité est la possibilité d'identifier des ressources humaines vers qui diriger les questions au sujet des documents. Pour cette raison, une évaluation de la mention de ressources-clés dans les documents est incluse. Le barème établi est :

- Si le nom d'un groupe ou un courriel d'envoi de groupe est mentionné : 2
- Si le nom ou le courriel d'une personne est indiqué dans le document : 1
- Aucune mention de personne-ressource : 0

Lorsque le document contient deux types de mentions (un ou plusieurs groupe(s) et une ou plusieurs personne(s)), le barème du groupe d'envoi est appliqué puisqu'une traçabilité minimale est assurée. Lorsque des problèmes surgissent, et qu'un document est consulté, l'indication du courriel d'un groupe à contacter est l'information la plus utile et la plus fiable par sa constance. En deuxième rang, la mention d'une personne offre un indice sur la traçabilité, mais le roulement de personnel est un obstacle à la fiabilité de cette information. Souvent, la traçabilité est vérifiée implicitement et par défaut par les logiciels de gestion des versions. L'inclusion de ce barème permet d'étendre le modèle de mesure aux cas de documents non sauvegardés dans les logiciels de gestion des versions.

#### **4.1.4 Description des étapes du développement de la métrique**

Après les barèmes individuels de chaque attribut, l'étape suivante est la mesure de l'indice de qualité des documents en appliquant la méthode définie. Ainsi, chaque document est coté selon les critères de qualité de la documentation. Les tableaux de gabarit pour la présentation des données se trouvent à la section 4.2.

Puis, les données de maintien sont retracées pour les bogues de l'échantillon.

Par la suite, pour chaque défektivité, la durée de résolution est calculée (nombre d'heures écoulées au calendrier). Les problèmes sont énumérés un par un, mais, par la suite, la somme des délais est la métrique finale pour la comparaison. Il est entendu que le délai de résolution peut induire en erreur si la correction n'est pas entreprise relativement rapidement

ou si des circonstances inattendues ou des priorités différentes ont lieu. Il est aussi compris que le temps écoulé avant la résolution d'un problème n'est pas dédié en entier à la correction sous toutes ses étapes. Néanmoins, il est raisonnable d'assumer que les problèmes prioritaires ou affectant un système essentiel comme ceux de l'échantillon ont été abordés promptement puisqu'ils ont des ressources dédiées. (Pour plus de facteurs, consulter [19].) La durée de temps écoulée est arrondie aux 15 minutes et notée en notation décimale (par exemple, le temps écoulé entre 15:00 et 17:15 est notée 2,25 heures).

À l'étape 4, l'analyse est effectuée avec l'indice de documentation et le délai pour chaque projet pour arriver à une corrélation et une meilleure compréhension du phénomène. La méthode de mesure sélectionnée est l'observation structurée. Par cette méthode, qui est propice à la quantification d'items plus difficiles à compter, les données sont compilées telles qu'observées, directement à partir d'une classification par catégories plutôt que d'être rapportées par des participants.

En observant directement la relation entre l'indice de documentation vérifiée d'un projet, et le délai de résolution de bogues, l'effet de la documentation sur le maintien se définit. Une discussion des résultats est planifiée par l'auteur de l'essai avec les analystes seniors du domaine applicatif ainsi que les directeurs du projet. Il est essentiel de souligner qu'une liste de facteurs variés influence la résolution d'un bogue, dont [19] [45] [46] :

- Le type et l'ampleur du code à modifier;
- Le niveau d'expérience des ressources et leur disponibilité;
- Les ententes de niveau de service;
- L'impact opérationnel;
- Le profil de risque quant à la sécurité;
- Les tests (automatisation, expertise, outils).

La documentation est loin d'être le seul facteur et les spécificités de chaque défaillance ont une plus grande incidence sur le maintien de façon générale à cause de la variabilité des erreurs. Cette étude ne tient pas compte de ces facteurs et se limite à constater l'effet d'une seule variable. De plus, cette étude n'a pas le mandat de vérifier si les documents sont consultés pendant la résolution des défaillances. Cet aspect couvre l'utilité, l'utilisation et le rôle que joue la documentation dans le *comment* du travail technique. Des questionnements

certes pertinents et valables pour d'autres projets, mais la présente étude observe les documents dans leurs caractéristiques immuables et non pas l'*interaction personne-document*.

#### **4.1.5 Facteurs clés de succès**

La mise en œuvre de cet essai se base sur un ensemble de facteurs-clés (facteurs de succès) :

- Obtention des données des bogues et des documents;
- Sélection de projets et de documents représentatifs (éviter le biais d'échantillonnage);
- Maîtrise des outils d'analyse appropriés pour les résultats;
- Possibilité de déceler une relation entre les éléments analysés.

Ultimement, le véritable succès de cet essai repose sur la possibilité de créer un scénario qui peut être reproduit pour valider les étapes et les conclusions, surtout à plus grande échelle. La visibilité que peut obtenir cette recherche est essentielle pour que les résultats exercent un impact dans le monde du développement de logiciel.

#### **4.1.6 Approche de validation des résultats**

David Parnas a étudié le temps perdu à cause de la documentation mal structurée ou qui contient de l'information erronée. D'une autre part, Dzidek, Arisholm et Briand ont mesuré, en 2008, que la documentation sous forme de modèles UML sauve du temps dans le développement orienté objet. La validation de l'essai est donc effectuée à partir d'une comparaison avec les tendances observées dans les études publiées. Les résultats des études qualitatives recensées, dont 27 études sur les bénéfices de la documentation des TI, servent de base pour alimenter une discussion à propos des résultats de cette expérimentation. La validation des résultats auprès d'une analyste experte en systèmes experts contribue à la vérification de la validité des résultats.

Lors de la discussion des résultats, les limites de validité décrites à la section 3.1.3 s'appliquent.

## 4.2 Résultats attendus

Les résultats sont présentés dans des tableaux comme ceux qui suivent (données fictives). Le but est de générer un diagramme dans un chiffrier et de visualiser la relation entre l'indice de documentation et les délais de maintien lors de correctifs. La présence d'une corrélation négative entre les deux variables est vérifiée.

### 4.2.1 Données de bogues

Les informations les plus importantes sur les bogues sont présentées dans le tableau qui suit.

**Tableau 4.1 Gabarit des détails des incidents et du temps de résolution de bogues**

#	Id.	Source	Type	Description	Temps de début	Temps de fin	Temps Écoulé
1	36543	Web	Problème	Arrêt du programme	2012-05-05 à 20:35:55	2012-05-05 à 01:14:48	5
2	78348	Web	Problème	Arrêt de programme. Type de donnée incompatible. Problème de compilation	2012-01-26 à 21:18:25	2012-01-26 à 22:18:13	0,75

### 4.2.2 Données de bogues non retenus

Les défaillances de système sont inspectées et certaines ne peuvent être retenues pour l'étude. Ces données sont inventoriées dans un tableau : identifiant, source, type de rapport, description du problème et justification pour l'exclusion.

**Tableau 4.2 Gabarit des données de bogues non retenus pour le projet**

#	Id.	Source	Type	Description	Justification
1	11526	Web	Rôle de maintien	Demande pour pouvoir supprimer une transaction de client.	Aucun détail du programme fautif n'a été mentionné. Il n'est pas possible de retracer des documents.
2	58965	Web	Rapport de problème	Arrêt de programme	Le composant n'a pas été réactivé.

#### 4.2.3 Données des attributs de qualité de la documentation

Les données avant application du barème pour chaque attribut et du calcul des indices sont présentées dans un tableau comme ci-dessous (cas fictif).

**Tableau 4.3 Gabarit de présentation des données brutes des attributs de la documentation**

Identifiant du bogue		26564	Identifiant du document			
Critères	Attributs	Unités	26564-1	26564-2	26564-3	26564-4
<b>Structure</b>	Longueur	<i>Pages</i>	3	13	1	2
		<i>Mots</i>	926	309	514	64
	Erreurs	<i>Nombre</i>	1	0	10	0
	Éléments visuels	<i>Nombre</i>	1	3	1	2
	Mise en page	<i>Évaluation</i>	ordonnée	ordonnée	gabarit	gabarit
<b>Cycle de vie du logiciel</b>	Type de document		processus	opérations	sécurité	opérations
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	8	1	0	0
	Synchronisée ou non	<i>Notes</i>	Oui	Non	Non	Non

Identifiant du bogue		26564	Identifiant du document			
Critères	Attributs	Unité	26564-1	26564-2	26564-3	26564-4
Accessibilité	Facilité	Oui	Oui	Oui	Oui	Non
		Non				
		Difficile				
	Type de fichier		Standard, .docx	Standard, .doc	Standard, .doc	Standard, .txt
Traçabilité	Qui contacter en cas de problèmes ou de questions	Nombre de personne ou de groupe	0	1 personne	0	1 groupe
Logiciel	Pour identifier le code source pour le jumelage		_rapp2	_rapp2	_rapp2 et 5 autres	_rapp2

#### 4.2.4 Données des indices de qualité de la documentation

Pour considérer l'effet des attributs sur le temps de résolution directement, sans cumuler les indices dans une cote globale, les tableaux des indices sont présentés dans le gabarit du Tableau 4.4. Pour cet attribut de qualité, la valeur cumulative par bogue et la moyenne des éléments par document par bogue sont fournis.

**Tableau 4.4 Gabarit de présentation d'un attribut de qualité par rapport au temps de résolution**

Bogue	Nombre de documents	Éléments	Moyenne des éléments	Cote par document	Temps de résolution
52658	2	2	1	1,5	15
45872	5	3	0,6	3	4,75

#### 4.2.5 Données sommaires

Le graphique de haut niveau : bogue-métrique-temps corrigé comme ci-dessous.

Tableau 4.5 Gabarit de présentation des données de corrélation

Bogue	Indice total	Temps de résolution	Temps de résolution
123	21,5	15	15
432	...	34	13

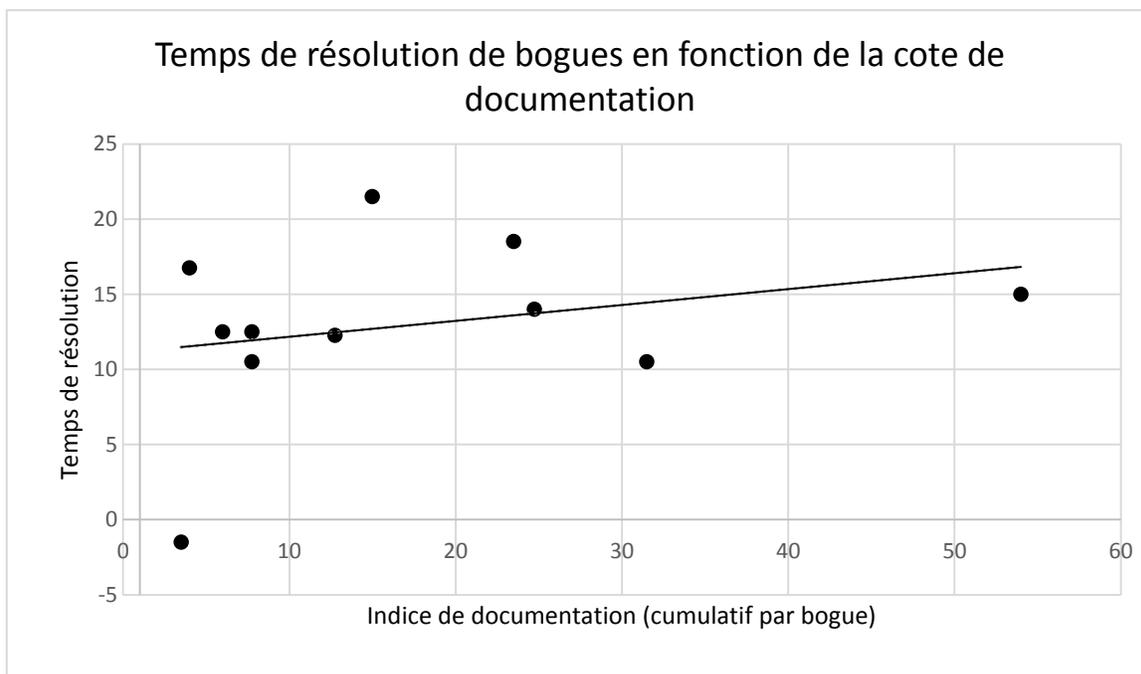


Figure 4.3 Diagramme de dispersion pour l'analyse des données de l'essai (données fictives)

Le prochain chapitre présente les résultats recueillis ainsi que l'analyse produite et les conclusions reliées à l'hypothèse de départ. Une discussion de la formulation de la réponse pour la question principale et des pistes pour les questions secondaires est abordée. Une critique de la validité des résultats est ensuite énoncée.

#### **4.2.6 Autres analyses**

La dernière section contient des éléments d'analyse supplémentaires. Les attributs peuvent interagir différemment avec le temps de résolution de problèmes parce que leurs barèmes sont différents ou parce que leur manifestation a un impact unique sur la variable dépendante. Ces éléments sont présentés dans cette section, accompagnés de graphiques au besoin.

## **Chapitre 5**

### **Analyse des résultats**

Ce chapitre présente les résultats de la collecte de données selon les critères du chapitre précédent, ainsi qu'un retour sur l'hypothèse et une analyse plus détaillée des paramètres de la qualité de la documentation. Le chapitre clos avec une discussion des mesures de vérification de la validité et des recommandations émises.

#### **5.1 Résultats obtenus**

Les résultats sont présentés sur quatre paliers : les données de bogues, les cas de bogues non retenus et la justification de leur exclusion, les données sur les attributs de qualité de la documentation captée par l'analyse, et les données sur les indices de qualité. Les indices de qualité sont le résultat de l'application des barèmes énoncés au chapitre 4.

##### **5.1.1 Données de bogues**

En tout, 14 cas de bogues ont été choisis aléatoirement. Après la sélection, les bogues ont été classés par ordre croissant d'identifiant. L'annexe 7 contient le tableau des données de bogues ainsi que l'analyse détaillée de chaque bogue.

##### **5.1.2 Données de bogues non retenus**

Parmi les cas énumérés dans la sélection aléatoire des bogues, un certain nombre a dû être rejeté lors de l'analyse préliminaire, avant la phase de calcul des indices. Ces demandes ont largement été mises de côté parce que les données étaient non pertinentes. En réalité, l'information essentielle manquante est la mention du programme fautif, dans la vaste majorité des cas. Le tableau des bogues non retenus est présenté dans l'annexe 8.

Parmi les sept cas rejetés, quatre font partie de la catégorie de demande de rôles de maintien (Catégorie B de l'échantillon). Cette catégorie est choisie pour la capture de nouveaux bogues à analyser, mais dans les faits, peu de bogues significatifs sont disponibles. Sur les cas recensés, seulement un rapport de demande de rôle de maintien permet de déceler un nouveau bogue ajouté à l'analyse.

### 5.1.3 Données des attributs de qualité de la documentation

Le gabarit du Tableau 4.3 rempli pour chaque rapport de problème. Chaque colonne contient les données d'un document. Les données brutes sont présentées dans l'annexe 9. Ces données brutes issues de l'observation sont formatées puis soumises à un logiciel de calcul des indices. Ce tableau est fourni à l'annexe 10.

### 5.1.4 Données des indices de qualité de la documentation

Le tableau détaillé de chaque indice se trouve dans l'annexe 12.

### 5.1.5 Données sommaires

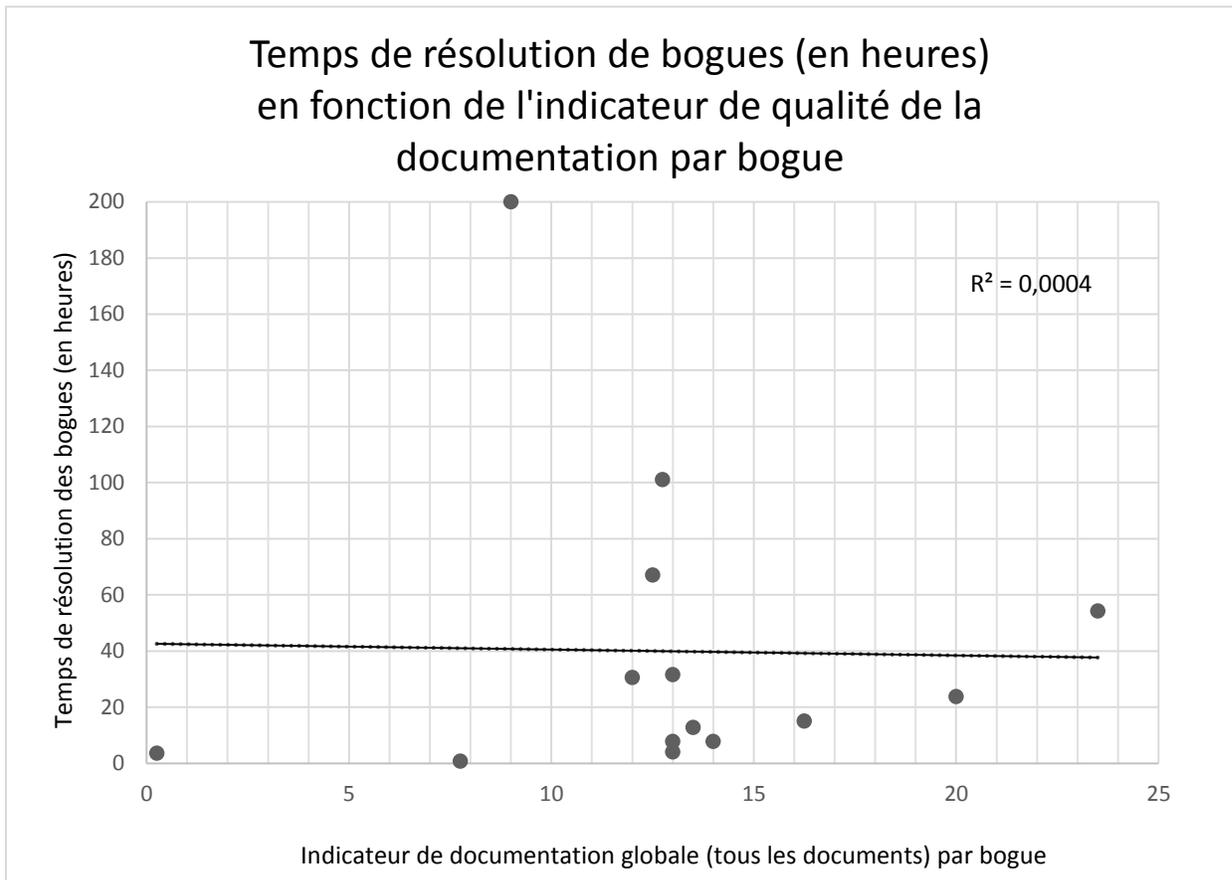
Le graphique de haut niveau du Tableau 5.1 présente les résultats de façon succincte. Chaque cas est lié à son indice total et son temps de résolution.

**Tableau 5.1 Données de corrélation**

<b>Bogue</b>	<b>Indicateur</b>	<b>Temps de résolution</b>
63023	16,25	15
63529	12,5	67
63534	13	31,5
64092	13	4
64150	9	200
64644	13,5	12,75
64650	12,75	101
66594	14	7,75

Bogue	Indicateur	Temps de résolution
69582	13	7,75
69668	0,25	3,5
71536	23,5	54,25
71894	20	23,75
72223	12	30,5
72528	7,75	0,75

La corrélation du temps de résolution en fonction de l'indicateur de la qualité de la documentation est présentée ci-dessous.



**Figure 5.1 Graphique du temps de résolution des bogues en fonction de l'indicateur de qualité de la documentation**

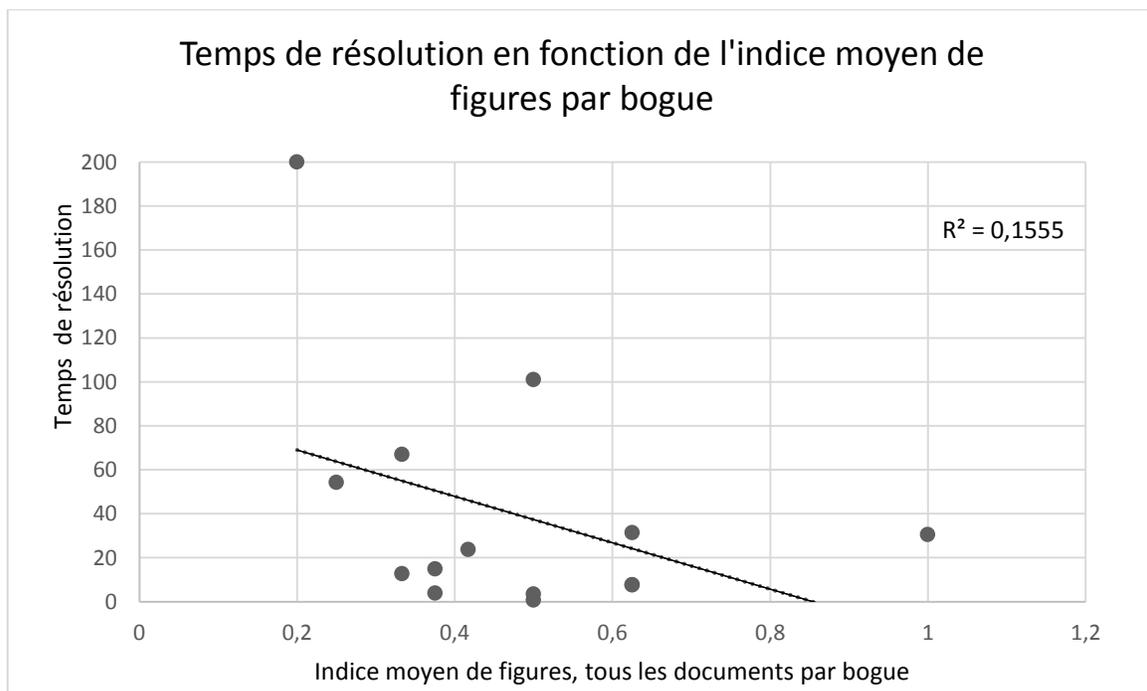
Selon ce graphique, le fait observé est une corrélation positive. L'indice de corrélation  $r$  de Pearson étant de  $-0,02078$ , la corrélation n'est pas assez forte pour émettre des observations. De plus, le coefficient de détermination est de  $0,0004$ . Une valeur à  $0$  indique que la ligne n'explique le positionnement d'aucun des points du graphique par opposition une valeur de  $1$  qui indique que la régression linéaire représente tous les points du nuage.

Une approche d'analyse par modification des points extrêmes est présentée dans l'annexe 11.

Une analyse plus individuelle des attributs donne des éléments supplémentaires de quantification de la qualité selon un critère à la fois. La dernière section de ce sous-chapitre contient les analyses supplémentaires pour les attributs, les indices cumulés, et les indices moyens.

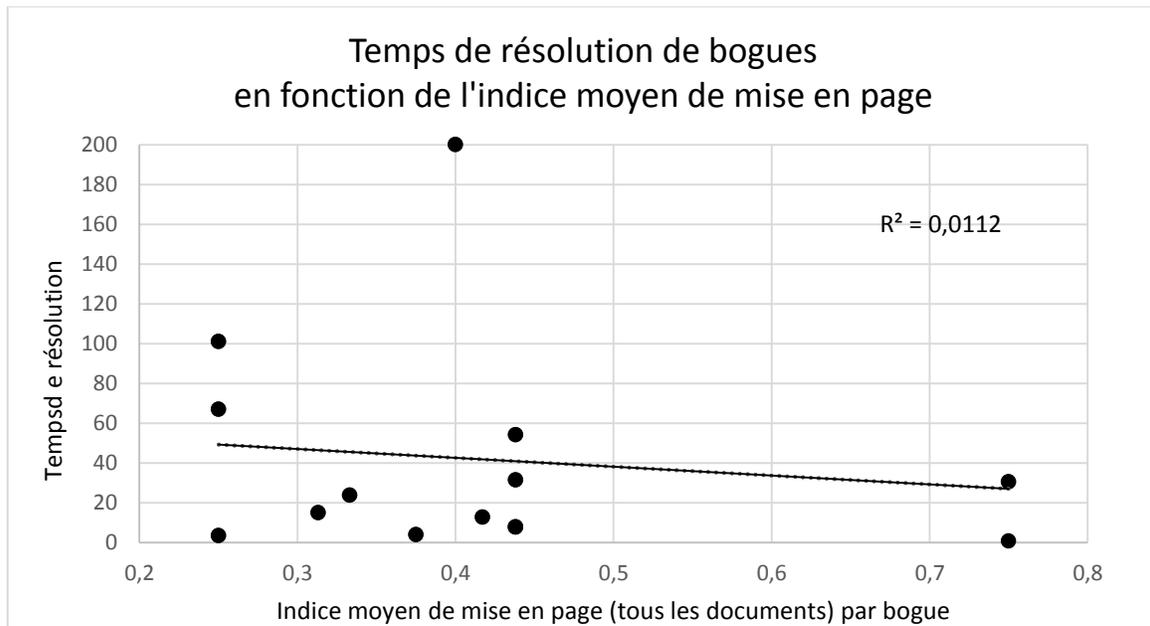
### 5.1.6 Autres analyses

Les analyses supplémentaires sont reprises ici (détails disponibles à l'annexe 12). La figure 5.2 indique une relation linéaire par rapport à la moyenne d'éléments visuels par bogue.



**Figure 5.2 Graphique du temps de résolution des bogues en fonction du nombre d'éléments visuels (moyenne par document)**

La figure 5.3 présente la relation de l'indice moyen de mise en page avec le temps de résolution de bogues. La relation reflète la tendance de l'indicateur global.

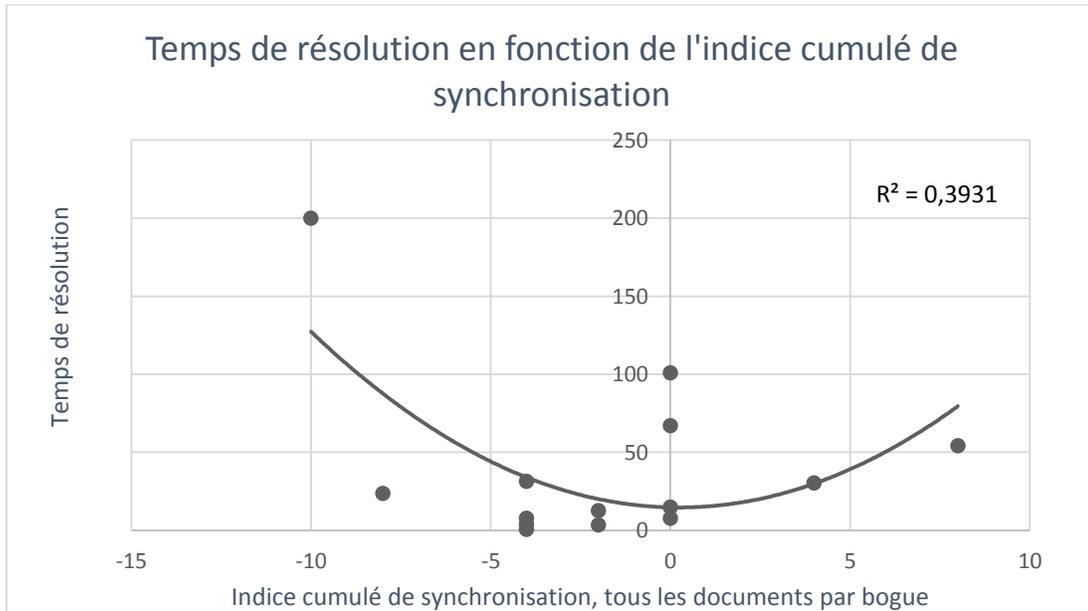


**Figure 5.3 Graphique du temps de résolution des bogues en fonction de la mise en page des documents (moyenne de tous les documents par bogue)**

Le graphique de la figure 5.4 contient une corrélation de l'état de la synchronisation des documents en fonction du temps de résolution. Plutôt qu'une moyenne, cet état est un indice calculé en sommant les indices des états individuels des documents pour chaque bogue.

Il existe un point optimal pour lequel la résolution rapide est favorisée. Cela est visible sur le graphique au point minimal selon l'axe vertical de la courbe de fonction. Lorsque l'indice cumulatif des états de synchronisation des documents est proche de 0, cela veut dire que l'état de la mise à jour est équilibré. Certains documents plus désuets coexistent avec d'autres plus récents, mais l'ensemble tend vers une cohérence avec la mise à jour du

logiciel. Plus les cas ont une documentation cumulative fortement désuète, plus le temps de résolution augmente. Inversement, il y a deux points extrêmes pour lesquels malgré une documentation en excellent état de mise à jour, la résolution a pris beaucoup de temps.



**Figure 5.4 Graphique du temps de résolution des bogues en fonction du statut de mise à jour des documents (cumulatif des documents par bogue)**

Le sous-chapitre qui suit reprend l'hypothèse, la question principale et les sous-questions pour revenir sur les résultats, les observations et leur signification.

## 5.2 Retour sur les questions et l'hypothèse

L'hypothèse posée est que la documentation de qualité permet de réduire le temps de résolution de bogues pour un logiciel existant. Cette qualité se mesure avec un indicateur basé sur les indices individuels des attributs des documents. L'indice a été calculé pour chaque cas trouvé. Les corrélations ont été effectuées à l'aide de Microsoft Excel. Il est attendu qu'une corrélation négative est cohérente avec l'hypothèse formulée.

La corrélation du temps de résolution en fonction de l'indice que la qualité de la documentation est une corrélation très faible. L'indice de corrélation  $r$  de Pearson de -0,02078 prouve que la linéarité est négative mais celle-ci n'est pas statistiquement significative. La correction des points extrêmes (dans l'annexe 11) produit une corrélation avec indice  $r$  de 0.6067 et un coefficient de détermination de 0.3681. Un retour sur la signification de la tendance s'impose. Lorsque les points extrêmes sont ajustés, l'hypothèse est infirmée. En effet, l'énoncé qu'une mesure basée sur la qualité de la documentation augmente le temps de résolution de bogues pour un logiciel existant est la constatation apparente. Toutefois, les experts en méthodologie de recherche avisent contre ce jugement hâtif. Une corrélation n'implique pas une relation de causalité. La direction de la causalité peut aussi être l'inverse de ce qui est présumé ou encore entièrement dû à l'interaction d'une troisième variable inconnue lors de l'expérimentation et donc non contrôlée [37]. En plus de réfléchir à la signification des faits sur l'hypothèse, il est important de réfléchir aux réponses aux questions principales et secondaires ci-dessous.

*QP1. Est-ce que la documentation permet de réduire le temps de résolution de bogues pour un logiciel existant et opérationnel, en environnement de production?*

QP1. En réponse à la question principale de l'essai de la section 3.1.2, il est observé que les temps plus élevés de résolution de bogues sont corrélés avec des indices de documentation très légèrement plus bas mais la corrélation n'est pas jugée assez fiable (selon les indices de corrélation et de détermination) pour suffire à l'affirmation de l'hypothèse et à une réponse claire à QP1. Explorer la contribution des attributs de qualité offre des éléments de réponse plus fiables et fournit une réponse à la QS1 ci-dessous.

*QS1. Quel est l'ensemble des attributs nécessaires pour faire un bilan complet de la qualité d'un document, en reflétant les aspects de la qualité les plus saillants connus dans le domaine?*

Une des questions secondaires de l'essai cherche un ensemble d'attributs nécessaires pour faire un bilan complet de la qualité d'un document, en reflétant les aspects de la qualité les plus saillants. Il est observé que parmi ces attributs saillants sont la présence de figures et d'illustrations, et l'état de synchronisation totale de la documentation pour un logiciel. Dans le barème de l'indice développé, l'indice de fréquence de mise à jour des documents s'est aussi

avéré de peu d'intérêt. De plus, le nombre total de figures dans tous les documents et la moyenne du nombre de figures par document sont apparus comme des éléments valables. L'augmentation de cette moyenne par document et par bogue est liée à une diminution du temps de résolution.

*QS2. Quels barèmes doivent être accordés à chacun des attributs de la question secondaire QS1?*

La deuxième question secondaire de l'essai entend explorer les barèmes qui doivent être accordés à chacun des attributs de la question secondaire QS1. Les barèmes fixés et les indices qu'ils ont générés sont évalués pour la valeur qu'ils contribuent à l'avancement de l'état de la connaissance en matière de quantification de la qualité documentaire. Selon une perspective, des aspects considérés comme critiques reçoivent une pondération élevée dans la grille. Par exemple, il y a le manque de documents, le manque de synchronisation des documents avec le code source et la traçabilité par l'intermédiaire des personnes-clés ou de courriel d'envoi. Dans une autre perspective, les barèmes établis pour ces deux premiers paramètres permettent réellement de traduire cette importance en indice qui démontre une tendance. Le dernier paramètre, soit la traçabilité, ne s'est pas traduit par un barème linéairement significatif. La réponse à la question secondaire se trouve aussi dans les intervalles choisis pour les barèmes. Par exemple, les critères pour la cote pour les erreurs de langue et de typographie sont fixés, comme tous les autres barèmes, avant que la collecte des données ne soit effectuée. Or, il est maintenant connu que les documents présentent une haute qualité. Une façon d'évaluer le barème est d'explorer des barèmes alternatifs pour la cote des erreurs et de décider si une mesure beaucoup plus précise est requise (une mesure qui pénalise la cote pour un nombre d'erreurs moindre). Le questionnement est valable et nécessaire pour une évaluation impartiale de la grille de calcul pour tous les attributs. Dans l'exemple, il est quand même visible qu'une modification de barème n'a pas une conséquence démesurée sur le temps de résolution, mais il se peut que ce soit autrement pour d'autres attributs.

*QS3. Cette métrique est-elle réutilisable pour la recherche subséquente et pour l'industrie?*

La métrique est réutilisable pour des recherches à venir. Étant donné que plusieurs tendances ont été observées après la mise en place de la mesure de qualité, il est

avantageux de pousser ce raisonnement plus loin. Sélectionner les paramètres selon les critères différents, leur accorder des pondérations variées, et les compiler différemment sont des possibilités. La recherche basée sur les données de production et les environnements opérationnels présente des défis qui lui sont particuliers. Le chercheur a de la difficulté à retracer le travail exact des intervenants, à cerner les lacunes véritables, au-delà de l'information projetée par les intervenants eux-mêmes. Le chercheur perd donc une partie du contrôle du cadre et des paramètres pour avoir accès à l'entreprise et cela n'inclut pas les difficultés d'obtention des approbations pour la publication des résultats [37]. La possibilité de personnaliser la métrique et de l'adapter au contexte et au relief de l'entreprise et de l'industrie offre une latitude pour contrer les difficultés exprimées par les chercheurs et vécues par plusieurs praticiens en industrie.

*QS4. Est-ce qu'une telle mesure est considérée comme un facteur d'analyse valable de la qualité de la documentation?*

La mesure développée s'ajoute à une liste de mesures formulée par des chercheurs en génie logiciel. Ces approches transforment les aspects intangibles et divergents de la documentation en une base qui se prête à des comparaisons et des mesures d'ordres de grandeur [1] [11] [48] [49]. La métrique résultante offre une possibilité de plus dans la sélection et la personnalisation d'une approche en analyse de la qualité.

### **5.3 Démonstration de la validité des résultats**

Une équipe de chercheurs statisticiens ont formulé des directives pour la publication des recherches empiriques en génie logiciel. Ces experts ont l'expérience de plusieurs années de recherche, de publication et de révision de textes pour eux-mêmes et autrui. L'approche de validation de la recherche est reprise des recommandations sur la recherche empirique [40]. Les détails se trouvent à l'annexe 13. La validation est assurée au moment de l'établissement du contexte de la recherche, de l'élaboration du cadre de recherche, de la collecte des données, de l'analyse, de l'interprétation et de la présentation des résultats.

Outre les balises méthodologiques, la validation est faite à l'aide de consultations avec des experts. Les calculs des métriques sont reproduits par un code source écrit en PL/SQL (disponible à l'annexe 14).

Le modèle de mesure formulé s'applique aux domaines et aux populations où la création de documents est une pratique établie, et avec une structure générale permettant le repérage et la mesure des attributs des documents. Une certaine maturité au niveau des outils, de la stabilité des systèmes et des capacités de détection, de suivi et de rapport des incidents caractérise cette population.

Enfin, il est important de clarifier la nuance entre une observation statistiquement significative et une observation qui revêt d'une grande importance dans la pratique. Certaines observations ne sont pas statistiquement significatives dans cet essai. Par exemple, l'accessibilité aux documents, leur nomenclature et la qualité de la langue sont des attributs fortement présents dans les données recueillies. La variabilité par rapport au temps de maintien n'est pas statistiquement significative. Toutefois, dans la réalité, ces aspects sont essentiels à l'utilisation des documents. Dans le cas des figures, il est observé que leur présence accélère le maintien. Cet aspect est statistiquement significatif. Par opposition, cela n'exclut pas la grande importance, dans la pratique, de la rédaction de texte explicatif comme accompagnement aux figures ou comme documentation lorsqu'un support visuel n'est pas disponible ou approprié.

## **5.4 Recommandations**

Les recommandations de cet essai se situent dans deux ordres d'idées. En premier lieu, il y a des recommandations spécifiques issues des résultats.

Il est recommandé de considérer l'impact des éléments visuels comme étant une priorité dans la production de documentation. Il y a un lien observé avec une réduction du temps de résolution de problèmes. En plus, l'importance des figures et des illustrations dans la documentation obtient le consensus dans plusieurs domaines cognitifs (apprentissage, gestion de la connaissance, formation, architecture logicielle) et réunit les méthodologies de développement. En effet, une des forces des résultats est de permettre un pont entre les

méthodes dites traditionnelles et les approches agiles. Ce pont représente une activité de documentation qui dépasse les différences méthodologiques.

Il est recommandé de considérer la mise à jour des documents pour un module logiciel en regardant l'ensemble des documents. Il est possible de visualiser les ouvrages comme un groupe cohésif d'informations. Cela permet à l'état de mise à jour des documents de donner des indices sur la rapidité de mise en place des correctifs.

Assurer l'accessibilité aux documents, dont l'état est généralement bon dans les intrants de l'étude, fait partie des recommandations. Bien que la corrélation de l'indice moyen de l'accessibilité présente une corrélation faible avec le temps de résolution, toute tentative future de faire réévaluer ou évoluer la métrique n'est rendu possible que par la disponibilité des documents sources.

Le deuxième type de recommandation rejoint les observations des chercheurs dans le domaine. Il est recommandé d'effectuer plus de recherches puisque les données sont intéressantes, mais nuancées. Une des difficultés de la recherche quantitative est la concordance du sens de la métrique avec la signification réelle des phénomènes.

Les lacunes du barème peuvent être comblées en effectuant des mesures à intervalle régulier pour adapter les catégories du barème et leurs indices selon l'usage voulu et les réalités de l'entreprise. L'auteur recommande donc de faire usage de ce barème pour progresser la qualité dans le temps.

## Conclusion

Dans le contexte informatique actuel, la maintenance des logiciels est devenue une responsabilité considérable. Le vieillissement des systèmes, le roulement de personnel et les attentes accrues des utilisateurs font en sorte que la maintenance des logiciels s'alourdisse. Accompagnant ce fait, les coûts de modification du code source et les conséquences pécuniaires des pannes dictent que les problèmes logiciels soient réglés rapidement. L'implémentation de correctifs repose sur l'acquisition de connaissances, à la fois compréhensives et ciblées, à propos du système fautif. Peu importe l'âge et le type de technologie, la documentation demeure une source d'information importante. La création et la mise à jour des documents sont des tâches peu populaires parmi les praticiens à cause des difficultés d'harmonisation du contenu écrit avec le contenu codé. Les défis dans la rédaction, la longueur et la mise en page n'améliorent pas la facilité de maintien des documents.

La recherche dans le domaine est encore jeune et les recherches quantitatives commencent à apparaître. Dans la méta-analyse, 29 % des recherches recensées sont empiriques. Cette étude entreprend donc de combler le manque noté dans la recherche sur la documentation technique.

Le point de départ de l'essai est de considérer les défaillances techniques qui ont lieu durant une période précise, dans un environnement de production. Les bogues sont répertoriés et analysés du moment de début du problème jusqu'au moment de clôture, tel que cité dans le rapport de bogue. Là où l'analyse des bogues s'arrête marque le début de l'analyse des documents. En effet, chaque logiciel fautif est décrit par une liste de documents. Le but est d'arriver à quantifier les aspects de qualité de ces documents pour procéder à l'analyse corrélationnelle. La contribution principale de cet essai est la mise au point d'un gabarit d'évaluation de la qualité des documents. Cette évaluation segmente la qualité en un ensemble d'attributs issus des études publiées et leur accorde des valeurs numériques. Les caractéristiques des documents deviennent des indices compilés dans des tableaux pour analyse et corrélation. Les données sont analysées horizontalement, c'est-à-dire au niveau d'un indice cumulatif, tous attributs confondus, des documents par bogue, mais aussi

verticalement, par la comparaison d'un seul attribut à la fois pour son impact sur le temps de résolution.

Cet essai démontre que la documentation a un impact sur le maintien des logiciels. L'idée directrice est que la documentation a une incidence positive sur le temps de résolution de bogues. L'hypothèse déclare que la documentation de qualité écourte la durée de résolution des incidents dans les logiciels opérationnels. Les données montrent que cela est valide.

Pour affirmer l'hypothèse, une méthode de quantification de la qualité est appliquée à un ensemble de documents. Ces documents s'apparentent à une liste de bogues qui se sont produits dans des systèmes en production.

Lorsque tous les aspects de la qualité sont analysés dans une seule et même mesure, le résultat n'offre pas de gains ou de pistes d'amélioration. La solution n'est pas dans l'évaluation de tous les attributs pris ensemble. Certains des attributs sont directement liés avec une réduction du temps de résolution des incidents.

L'analyse des données révèle que la présence de figures est un élément clé. Cet élément offre des gains intéressants pour améliorer la qualité de la rédaction tout en maximisant les bénéfices de cet effort. De plus, il est prouvé bénéfique de procéder avec la mise à jour de documents pertinents lorsqu'un logiciel est modifié puisque cela est relié à un gain de temps dans la maintenance.

L'hypothèse est confirmée parce qu'il y a des éléments avec une corrélation négative. Certains éléments confirment une perspective nouvelle sur d'autres études et poussent la connaissance plus loin. La présence de figures corrèle avec un temps de résolution amoindri. Ce résultat fait écho avec l'étude de Dzidek qui rapporte que les diagrammes UML améliorent l'efficacité et la qualité de la programmation. L'avantage du support visuel lorsque les spécifications étaient complexes est prouvé pour le nouveau développement. Le présent essai généralise cette observation en incluant l'ensemble des éléments visuels et étend la portée de leur avantage à la phase de maintenance des logiciels. Plus de recherches sont requises pour confirmer les trouvailles de la présente étude et pour renforcer les meilleures pratiques.

De nouvelles explorations peuvent prendre des avenues non explorées. Sur 69 études rapportées par Zhi, seulement 12 discutent des coûts de la documentation. Maintenant que certains des aspects de la qualité les plus importants pour la maintenance sont détectés, qu'en est-il des coûts que ces attributs permettent d'économiser? Étudier ces gains interpelle les chercheurs à considérer les coûts de production d'ouvrages qui contiennent les attributs à impact significatif. Par contraste, les gains trouvés en évitant les coûts de production de documents à faible impact intéressent les chercheurs et les professionnels en TI. Il en découle que la possibilité d'énumérer et de décrire les facteurs de retour sur investissement de la documentation est une piste d'exploration importante.

## Liste des références

- [1] Parnas, D. L., *Precise documentation: The key to better software*. 2011.
- [2] *Defining Technical Communication*. [En ligne]. Disponible à: <http://www.stc.org/about-stc/the-profession-all-about-technical-communication/defining-tc>. [Consulté le: 15-oct-2015].
- [3] Garcia-Molina, H., Ullman, J. D., et Widom, J., *Database Systems: The Complete Book*, Education, 2008, p. 1248.
- [4] *Department of Computer Science and Software Engineering*. [En ligne]. Disponible à: <http://www.concordia.ca/academics/undergraduate/calendar/current/sec71/71-70.html#b71.70.10>. [Consulté le: 17-oct-2015].
- [5] *Baccalauréat en génie informatique - Bureau de la registraire - Programmes - Université de Sherbrooke*. [En ligne]. Disponible à: <http://www.usherbrooke.ca/programmes/sec/genie/1er-cycle/bac/genie-informatique/>. [Consulté le: 17-oct-2015].
- [6] Parnas, D. L., *The Future of Software Engineering*, *The Future of Software Engineering*, 2011, p. 125-148.
- [7] Hadar, I., Sherman, S., Hadar, E., et Harrison, J. J., *Less is more: Architecture documentation for agile development*, dans *2013 6th International Workshop on Cooperative and Human Aspects of Software Engineering, CHASE 2013 - Proceedings*, 2013, p. 121-124.
- [8] Jansen, A., Avgeriou, P., et van der Ven, J. S., *Enriching software architecture documentation*, *Journal of Systems and Software*, vol. 82, n° 8, août 2009, p. 1232-1248.
- [9] Plosch, R., Dautovic, A., et Saft, M., *The Value of Software Documentation Quality*, 2014 14th International Conference on Quality Software, 2014, p. 333-342.
- [10] Zhi, J., Garousi-Yusifoglu, V., Sun, B., Garousi, G., Shahnewaz, S., Ruhe, G., Garousi-Yusifo, V.,

- Sun, B., Garousi, G., Shahnewaz, S., et Ruhe, G., *Cost, benefits and quality of software development documentation: A systematic mapping*, Journal of Systems and Software, vol. 99, n° 99, 2015, p. 175-198.
- [11] Garousi, G., Garousi-Yusifoglu, V., Ruhe, G., Zhi, J., Moussavi, M., et Smith, B., *Usage and usefulness of technical software documentation: An industrial case study*, Information and Software Technology, vol. 57, n° 2010, 2015, p. 664 - 682.
- [12] Dzidek, W. J., Arisholm, E., et Briand, L. C. L. C., *A Realistic Empirical Evaluation of the Costs and Benefits of UML in Software Maintenance*, IEEE Transactions on Software Engineering, vol. 34, n° 3, mai 2008, p. 407-432.
- [13] Kitchenham, B. et Charters, S., *Guidelines for performing Systematic Literature Reviews in Software Engineering*, Engineering, vol. 2, 2007, p. 1051.
- [14] Arisholm, E., *The Impact of UML Documentation on Software Maintenance: An Experimental Evaluation.*, IEEE Transactions on Software Engineering, vol. 32, n° 6, 2006, p. 365 - 381.
- [15] Jansen, A., Bosch, J., et Avgeriou, P., *Documenting after the fact: Recovering architectural design decisions*, Journal of Systems and Software, vol. 81, n° 4, 2008, p. 536-557.
- [16] Shmerlin, Y., Kliger, D., et Makabee, H., *Reducing Technical Debt: Using Persuasive Technology for Encouraging Software Developers to Document Code*, Advanced Information Systems ..., 2014, p. 207-212.
- [17] Selic, B., *Agile documentation, anyone?*, IEEE Software, vol. 26, n° 6, 2009, p. 11-12.
- [18] Paillé, P., *L'analyse par théorisation ancrée*, Cahiers de recherche sociologique, n° 23, 1994, p. 147-181.
- [19] Zhang, F., Khomh, F., Zou, Y., et Hassan, A. E., *An Empirical Study on Factors Impacting Bug Fixing Time*, 2012 19th Working Conference on Reverse Engineering, 2012, p. 225-234.
- [20] Kajko-Mattsson, M., *A Survey of Documentation Practice within Corrective Maintenance*, Empirical Software Engineering, vol. 10, 2005, p. 31-55.
- [21] Basili, V. R. et Mills, H. D., *Understanding and Documenting Programs*, IEEE Transactions on

Software Engineering, vol. SE-8, n° 3, 1982, p. 270-283.

- [22] Chen, J. C. et Huang, S. J., *An empirical analysis of the impact of software development problem factors on software maintainability*, Journal of Systems and Software, vol. 82, n° 6, 2009, p. 981-992.
- [23] Fortin, M.-F., *Fondements et étapes du processus de recherche*. Cheneliere Education, 2016.
- [24] Bayer, J. et Muthig, D., *A view-based approach for improving software documentation practices*, 13th Annual IEEE International Symposium and Workshop on Engineering of Computer-Based Systems (ECBS'06), 2006.
- [25] Taulavuori, A., Niemelä, E., et Kallio, P., *Component documentation—a key issue in software product lines*, Information and Software Technology, vol. 46, n° 8, 2004, p. 535-546.
- [26] Hunt, B., Turner, B., et McRitchie, K., *Software maintenance implications on cost and schedule*, IEEE Aerospace Conference Proceedings, vol. 20, n° 12, 2008, p. 1-6.
- [27] Bhattacharya, P. et Neamtiu, I., *Assessing programming language impact on development and maintenance: a study on c and c++*, Software Engineering (ICSE), 2011 33rd International Conference on, n° 3, 2011, p. 171-180.
- [28] Bhattacharya, P., *Using software evolution history to facilitate development and maintenance*, 2011 33rd International Conference on Software Engineering (ICSE), 2011, p. 1122-1123.
- [29] Callaghan, D. et O'Sullivan, C., *Who should bear the cost of software bugs?*, Computer Law & Security Review, vol. 21, n° 1, 2005, p. 56-60.
- [30] Sametinger, J., *DOgMA: A Tool for the Documentation & Maintenance of Software Systems*, 1991.
- [31] Arthur, J. D. et Stevens, K. T., *Assessing the adequacy of documentation through document quality indicators*, Proceedings. Conference on Software Maintenance - 1989, 1989, p. 40-49.
- [32] Correia, F. F., *Extending and Integrating Wikis to Improve Software Documentation*, 2008.

- [33] Forward, A. et Lethbridge, T. C., *The relevance of software documentation, tools and technologies: a survey*, DocEng 02 Proceedings of the 2002 ACM symposium on Document engineering, 2002, p. 26-33.
- [34] Lutters, W. G. et Seaman, C. B., *Revealing actual documentation usage in software maintenance through war stories*, Information and Software Technology, vol. 49, n° 6, 2007, p. 576-587.
- [35] Calcagno, C., Distefano, D., Dubreil, J., Gabi, D., Luca, M., Hearn, P. O., et Papakonstantinou, I., *Moving Fast with Software Verification*, NASA Formal Methods, 2015, p. 3-11.
- [36] Garousi, V. et Zhi, J., *A survey of software testing practices in Canada*, Journal of Systems and Software, vol. 86, n° 5, 2013, p. 1354-1376.
- [37] Easterbrook, S., Singer, J., Storey, M.-A., et Damian, D., *Selecting Empirical Methods for Software Engineering Research*, Guide to Advanced Empirical Software Engineering, 2008, p. 285-311.
- [38] Garousi, G., Garousi, V., Moussavi, M., Ruhe, G., et Smith, B., *Evaluating usage and quality of technical software documentation*, dans *Proceedings of the 17th International Conference on Evaluation and Assessment in Software Engineering - EASE '13*, 2013, p. 24.
- [39] Wingkvist, A., Ericsson, M., Lincke, R., et Löwe, W., *A metrics-based approach to technical documentation quality*, Proceedings - 7th International Conference on the Quality of Information and Communications Technology, QUATIC 2010, 2010, p. 476-481.
- [40] Kitchenham, B. a., Pfleeger, S. L., Pickard, L. M., Jones, P. W., Hoaglin, D. C., El Emam, K., et Rosenberg, J., *Preliminary guidelines for empirical research in software engineering*, IEEE Transactions on Software Engineering, vol. 28, n° 8, 2002, p. 721-734.
- [41] Bang, J. Y., Krka, I., Medvidovic, N., Kulkarni, N., et Padmanabhuni, S., *How software architects collaborate: Insights from collaborative software design in practice*. IEEE, 2013.
- [42] Royce, D. W. W., *Managing the Development of large Software Systems*, Ieee Wescon, n° August, 1970, p. 1-9.
- [43] Muir, D. E., Weinstein, E. A., *The Social Debt : An Investigation of Lower-Class and Middle-*

*Class Norms of Social Obligation*, vol. 27, n° 4, 2016, p. 532-539.

- [44] Tamburri, D. A., Kruchten, P., Lago, P., et Van Vliet, H., *What is social debt in software engineering?* IEEE, 2013.
- [45] Mazhelis, A. et Mazhelis, O., *Software Business*, Icsob, vol. 114, 2012, p. 261-266.
- [46] Santos, P. S. M. Dos, Varella, A., Dantas, C. R., et Borges, D. B., *Visualizing and Managing Technical Debt in Agile Development: An Experience Report*, Agile Processes in Software Engineering and Extreme Programming, vol. 149, 2013, p. 121-134.
- [47] Lehner, F., *Franz Lehner*, Degruyter.Com, vol. 29, n° 5, 1993, p. 551-568.
- [48] Johanson, G. A. et Brooks, G. P., *Initial Scale Development: Sample Size for Pilot Studies*, Educational and Psychological Measurement, vol. 70, n° 3, 2010, p. 394-400.
- [49] Legault, L., *No Title*. Laval, 2016.
- [50] *No Title*. [En ligne]. Disponible à : <http://www.statcan.gc.ca/edu/power-pouvoir/ch13/prob/5214899-fra.htm>.

## Annexe 1

### Bibliographie

Albing, B., *Combining human-authored and machine-generated software product documentation*, in *IEEE International Professional Communication Conference, 2003. IPCC 2003. Proceedings.*, 2003, p. 6 pp.

Alshazly, A., Elfatratry, A. M., and Abougabal, M. S., *Detecting defects in software requirements specification*, *Alexandria Eng. J.*, vol. 53, no. 3, 2014, pp. 513–527.

Amare, N. and St. Pierre, S., *Application of Goffman's Social Interaction Theories to the Technical Communicator/Engineer Relationship*, *IEEE Int. Prof. Commun. Conf.*, 2003, pp. 295–300.

Chen, J. C., & Huang, S. J. (2009). An empirical analysis of the impact of software development problem factors on software maintainability. *Journal of Systems and Software*, 82(6), 981–992. <http://doi.org/10.1016/j.jss.2008.12.036>

Garousi, G., Garousi, V., Moussavi, M., Ruhe, G., and Smith, B., *Evaluating usage and quality of technical software documentation*, in *Proceedings of the 17th International Conference on Evaluation and Assessment in Software Engineering - EASE '13*, 2013, p. 24.

Kajko-Mattsson, M. (2005). A Survey of Documentation Practice within Corrective Maintenance. *Empirical Software Engineering*, 10, 31–55. <http://doi.org/10.1023/B:LIDA.0000048322.42751.ca>

Norme IEEE : Standard for Software User Documentation

Rettich, K., *Using the wiki to deliver paperless software documentation*, in *2011 IEEE International Professional Communication Conference*, 2011, pp. 1–8.

Sadri, H. and Flammia, M., *Adapting communication styles and technology use to international environments*, in *IEEE International Professional Communication Conference, 2003. IPCC 2003. Proceedings.*, 2003, p. 5 pp.

Stettina, C. J. and Heijstek, W., *Necessary and neglected?: an empirical study of internal documentation in agile software development teams*, Proc. 29th ACM Int. Conf. Des. Commun., 2011, pp. 159–166.

Zhi, J. and Ruhe, G., *DEVis: A tool for visualizing software document evolution*, 2013 First IEEE Work. Conf. Softw. Vis., 2013, pp. 1–4.

## **Annexe 2**

### **Perspectives additionnelles**

#### **Perspective du développement collaboratif**

Il est important d'explorer le rôle de la documentation dans un contexte de collaboration et de contribution variée et délocalisée.

Des chercheurs ont examiné le sujet de la collaboration entre concepteurs de logiciels [41]. Ils ont voulu rapporter les rôles des concepteurs dans la collaboration lors de l'élaboration d'une architecture. Ils ont voulu aussi examiner les réseaux que ces concepteurs formaient selon leur rôle dans le design, l'impact de la distribution géographique de ces architectes et enfin, les facteurs de coût de la collaboration. L'équipe a mené des entrevues et des sondages échelonnés sur 10 semaines auprès de 18 architectes dans une firme de développement de logiciel mature (CMMI niveau 5) et dont les bureaux sont distribués mondialement. L'étude a mis de l'avant des observations sur la distribution des tâches de conception. Ils ont ciblé les grandes équipes et les réseaux de collaboration selon la nature des projets et selon leur structure organisationnelle (ressources affectées). Les participants ont souligné les dépendances et plusieurs facteurs de coûts.

Les trouvailles de la recherche de l'équipe indo-américaine contiennent des indices qui peuvent guider la présente étude. Certains des commentaires recensés par les entrevues affirment que la collaboration et la communication doivent être fortement présentes au début d'une phase de conception, occupant ainsi un tiers de l'effort. D'autres commentaires mentionnent qu'un effort considérable dans la documentation de l'architecture permettrait d'éviter de tels efforts plus tard dans le processus. Pour atteindre cet objectif, les concepteurs souhaitent conserver le plus grand nombre de détails possible jusqu'aux cas et scénarios d'utilisation à granularité fine (voire au niveau de chaque transaction). Ce commentaire revêt d'une grande valeur puisque ces architectes ont aussi déclaré qu'ils gèrent souvent des

incohérences entre les composants conçus à différents moments du processus de design. Dans la moitié des cas, ces experts reviennent sur leur plan pour reprendre le design d'origine, alors que dans les autres cas, ils émettent des correctifs ciblés ([41], p.47).

Les observations documentées indiquent que les architectes valorisent une documentation plus complète en amont. Cette documentation accompagne les phases initiales de conception et clarifie les choix conceptuels lorsque les designs sont transmis à diverses équipes en aval de la conception [41]. Les auteurs ont exploré le terrain de la collaboration en conception architecturale pour proposer des pistes variées pour de futures recherches. Cet ouvrage entend adopter une approche pour quantifier un des aspects trouvés (soit celui des bénéfices de la documentation), et ainsi ajouter une valeur motivante pour les producteurs et les consommateurs de la documentation. Loin de déplorer la tâche de partager leurs connaissances, ces professionnels contribuent à avancer l'état de la science de production pertinente et utile qui cible le lecteur et ses besoins informationnels.

### **Perspective théorique basée sur W. Royce**

En 1970, Winston Royce a formalisé le concept maintenant connu sous la nomination « modèle *waterfall* ». Son article sur la gestion du développement de logiciels de grande envergure accorde une importance sans équivoque à la mise en place d'une documentation complète. Il explique son approche d'évaluation du progrès d'un projet. Si la documentation présente des lacunes majeures, il recommande de faire remplacer l'équipe de gestion de projet et d'arrêter tous les travaux jusqu'à ce que la documentation parvienne à un niveau acceptable ([42], p.332). En dépit de la volonté d'une bonne gouvernance et d'une communication limpide en vue de réussir le projet selon les échéanciers et le budget fixé, la recommandation de Royce est indésirable de nos jours. La disponibilité des ressources humaines et les cycles de livraison de logiciel rendent l'application de ces recommandations difficile. Pourtant l'argumentaire repose sur plusieurs points encore pertinents de nos jours :

- La communication et la responsabilisation entre les parties prenantes passe par la documentation. En termes actuels, les ententes de niveau de service, les cahiers de charges, les logiciels de suivi de gestion de projet et les logiciels de communication

(courriel, messagerie instantanée, téléconférence) comblent une grande partie du suivi d'un projet.

- Dans les phases initiales du développement logiciel, la documentation est la seule représentation de la conception, des besoins et des spécifications. Pour quantifier ce besoin, Royce cite que le développement d'un logiciel de 5 millions de dollars devrait s'accompagner de 1500 pages de spécifications. Cette quantité de rédaction ne se justifie pas pour la majorité des logiciels. Le besoin peut être maintenu en modifiant les exigences de la documentation en matière de contenu. Cela devient possible par l'inclusion de formes beaucoup plus variées et riches en contenu comparativement aux pages de texte : diagrammes de type UML, cas d'utilisation, cartographie conceptuelle (*mind mapping*), résultats de remue-méninges, ergonomie du logiciel, et prototypes.
- La récupération de l'effort de documentation se réalise lors des tests, du transfert aux opérations et lors du maintien ([42], p.332).

La préparation de la documentation ne s'effectue pas dans un vide. Elle s'effectue dans un contexte de développement qui doit s'aligner avec les besoins de la phase d'exploitation, une phase durant laquelle peuvent jaillir une gamme de défaillances tel que discuté dans la section suivante.

### **Perspective de la dette de documentation**

La notion de dette sociale retrace ses origines aux recherches sociologiques et se définit comme un ensemble de relations et d'interactions tendues qui émergent dans le contexte d'une dynamique débiteurs-créditeurs [43]. Cette notion est valable en génie logiciel non seulement aux dynamiques de groupe, mais aussi aux logiciels produits sous forme de la dette technologique [44]. En 1992, Ward Cunningham a choisi le terme dette technique pour le résultat d'un programme issu de décisions en deçà des meilleures pratiques (*shortcuts*). Cette notion suscite de l'intérêt parce qu'elle transforme le niveau d'abstraction inhérente à la production de logiciel en une valeur pécuniaire. Chaque moment passé sur des correctifs pour un logiciel de basse qualité représente le paiement de l'intérêt (cumulatif) de cette dette[45]. La dette technologique désigne l'écart entre l'état réel d'un système logiciel et son

état idéal. Un type de dette technologique est la dette dans la documentation qui se manifeste quand aucune trace d'information écrite n'accompagne un code. La qualité de la documentation a un effet direct sur la maintenance, la phase la plus dispendieuse dans le cycle de vie d'un logiciel. Une documentation valide et à jour est requise pour accomplir les tâches variées de maintien : corrections d'erreurs, évolution des fonctions, ajustement de la capacité, optimisation, et autres. Les artefacts manquants ou désuets augmentent les coûts de maintien ([16], p. 208). Force est de constater que les nouveaux logiciels ne respectent pas toutes ces recommandations. En effet, faute de documentation appropriée, les efforts de compréhension du code actuel engouffrent de 40 % à 60 % du temps passé sur les tâches de maintien. Ce pourcentage a peut-être fluctué au cours des années. Les professionnels sont possiblement tenus de produire un minimum de documentation. Or, la réalité des programmeurs presque 15 ans plus tard est que la vitesse de modification de la documentation (et des commentaires du code) ne rattrape pas la vitesse de modification des logiciels [16].

Des efforts ont été déployés pour pallier cette tendance. Une équipe de chercheurs brésiliens ont étudié la dette technologique. Leur initiative aidait une équipe de quatre concepteurs (en architecture TI) à réduire la dette technologique dans une compagnie pétrolière brésilienne. Les concepteurs travaillaient avec leurs collègues issus de 25 équipes agiles. L'approche suivait quatre étapes :

1. La conscientisation de la présence de la dette technique et du besoin d'y remédier;
2. La mise en place de techniques de visualisation de la dette technique existante;
3. La mise en place de systèmes de rétroaction (pour voir les fluctuations de la dette);
4. La mise en place de mesures pour corriger les lacunes observées dans les systèmes d'après les indicateurs de la rétroaction [46].

Les résultats ont permis d'apprendre que les petits changements cumulatifs dans le temps occasionnent des gains significatifs pour les équipes et l'organisation[46]. Les principales leçons apprises sont :

1. Augmenter la visibilité au problème sans forcer l'implantation de la solution;
2. Communiquer régulièrement avec plusieurs parties prenantes pour utiliser la visibilité accrue pour engendrer des discussions productives et pour répandre une bonne compréhension des concepts et des outils;

3. Établir un équilibre sain entre la collaboration, la compétition et la ludification pour motiver les équipes à se surpasser dans la réduction de la dette.

L'outil employé pour mesurer la dette technologique est Sonar. Cet outil estime l'effort nécessaire pour clore chaque source de dette dans un projet. L'outil tient compte de facteurs de complexité et de cohésion, des duplicata, de l'absence de commentaires et de certains facteurs spécifiques aux techniques de programmation et de tests [46].

En résumé, bien que les lacunes soient reconnues et persistantes, des expérimentations remarquables inspirent cet essai. Dans ce dernier cas au Brésil, la situation de la dette technologique a reculé de 30 %, un résultat qui avait surpris et ravi les chercheurs. Le présent essai retient des éléments favorables de cette expérimentation. Tout comme des chercheurs ont pu réduire la dette technique dans leurs équipes, la dette de documentation est réductible. Le succès de la recherche inspire une approche structurée et motivante.

## Annexe 3

### Métriques existantes

Le tableau suivant récapitule les métriques recensées dans la littérature. Parmi la liste complète des attributs rapportés, seules les principales métriques d'intérêt sont reprises dans ce tableau. La mesure dans la rédaction est une pratique difficile à cause de la variété des écrits (nature, longueur, complexité, usage). Toutefois, pour mesurer l'impact de la documentation, plusieurs chercheurs ont essayé de fournir des méthodes pour calculer certains aspects de la rédaction technique.

**Tableau récapitulatif des métriques de la documentation**

Nom	Description	Thème(s)	Calcul	Réf.
CEI, <i>Cost effectiveness index</i>	Calcul du retour sur investissement d'un document	Facteurs de coût Mesure de l'utilité de documents	Fréquence d'accès d'un document/Temps écoulé dans la mise à jour de ce document	[11]
Réduction de l'effort	Gain rendu possible par la documentation	Mesure des bénéfiques	Non fourni	[10]
Perception de l'importance	Élément de mesure des bénéfiques	Mesure des bénéfiques	Non fourni	[10]
Longueur du document	Métrique d'utilité des documents	Facteur de qualité et de coût	Non fourni	[12] [19]
Longueur des commentaires dans un rapport de bogue	Métrique qui influence le temps de début d'une résolution de bogue	Facteur de temps	Nombre de lignes de commentaires	[19]

Nom	Description	Thème(s)	Calcul	Réf.
RMS <i>Readability Measuring System</i>	Outil de mesure de la compréhensibilité des documents	Mesure de qualité	Outil programmé	[47]
<i>Clones</i> ou duplication	Comparaison des documents pour détecter la similitude du contenu.	Mesure de l'unicité de la documentation (un aspect du maintien)	Outil d'analyse de la qualité de la documentation VizzAnalyzer	[39]
<i>Code coverage</i> ou couverture du code	Métrique de calcul du pourcentage du code source documenté	Mesure de l'utilité de la documentation	Outil d'analyse de la qualité de la documentation VizzAnalyzer	[39]

Traduction Libre

Sources : Multiples citées dans la colonne de droite du tableau.

## Annexe 4

### Composition de l'échantillon

L'analyse de l'échantillon comprend trois aspects : le choix de catégorie de l'échantillon, le tableau des détails, et l'analyse spécifique.

#### Choix de l'échantillon

Une interface web sécurisée permet de visualiser les signaux de bogue émis par les gestionnaires des opérations à partir de l'outil Automic. L'interface emploie un moteur de recherche pour retracer huit catégories d'évènements : les nouveaux transferts, les rapports de bogue, les demandes de fenêtre de maintien, les demandes pour faire rouler un programme régulier, les questions d'ordre général, l'impression, les demandes pour faire rouler un programme ou un correctif ponctuel, et la restauration de fichier. Les captures d'écran pour le service web sont comprises à l'annexe 6. L'échantillon du projet provient de deux de ces huit catégories. Le suivant contient les caractéristiques de l'échantillon selon les catégories de bogue, la population accessible et l'intervalle visé.

Le tableau ci-dessus contient les détails des catégories d'échantillon sélectionnées.

#### Tableau des détails sur l'échantillon

Catégorie		Catégorie A	Catégorie B
Détails	Année	2012	2012
	Type de rapport	Rapport de problème en production	Demande de permissions de maintien
	Nombre total d'entrées (tous les modules en production)	312	347

Catégorie		Catégorie A	Catégorie B
	Nombre d'entrées du module S1 (accessibles)	96	41
	Taille voulue pour l'échantillon	10-15 % de la population accessible	10-15 % de la population accessible
	Taille minimale de l'échantillon (arrondie au nombre entier)	9	4
	Taille maximale de l'échantillon (arrondie au nombre entier)	14	6

### Analyse spécifique du contenu des catégories de l'échantillon

La catégorie A comprend les entrées de type « Rapport sur un problème » et liste 312 problèmes couvrant tous les modules en production. Parmi cet ensemble, sont retenus seulement ceux qui concernent le module à propos duquel l'auteur de cette analyse peut voir et obtenir des données (module du PGI « S1 » et les rapports et composants reliés). Cela veut dire que sur 312 problèmes, 96 bogues sont accessibles. La population de bogues de 2012 contient donc ces 96 problèmes. Le but est l'analyse d'un échantillon de 10 à 15 % pour amasser des données représentatives, soit entre 9 et 14 cas de bogues. Cette valeur est fixée suite à la consultation d'experts et de recherche de références académiques [48] [37]. Un outil de calcul simple est employé pour préciser la taille de l'échantillon par rapport à la marge d'erreur. Ainsi, avec une population de 96 cas, et un niveau de confiance voulu de 90 % et une marge d'erreur maximale permise de 20 %, l'échantillon nécessaire s'étend à 15 cas.

Le type d'échantillonnage choisi est de type « probabiliste aléatoire simple » qui est une méthode qui permet de produire des estimations fiables, des estimations de l'erreur d'échantillonnage et de faire des inférences au sujet de la population, selon Statistique Canada [50]. Un générateur de liste aléatoire sélectionne la liste des rapports de bogues à analyser. Une liste d'identifiants de problème est fournie et le générateur retourne cette liste dans un ordre aléatoire et ainsi les 15 premiers problèmes sont retenus.

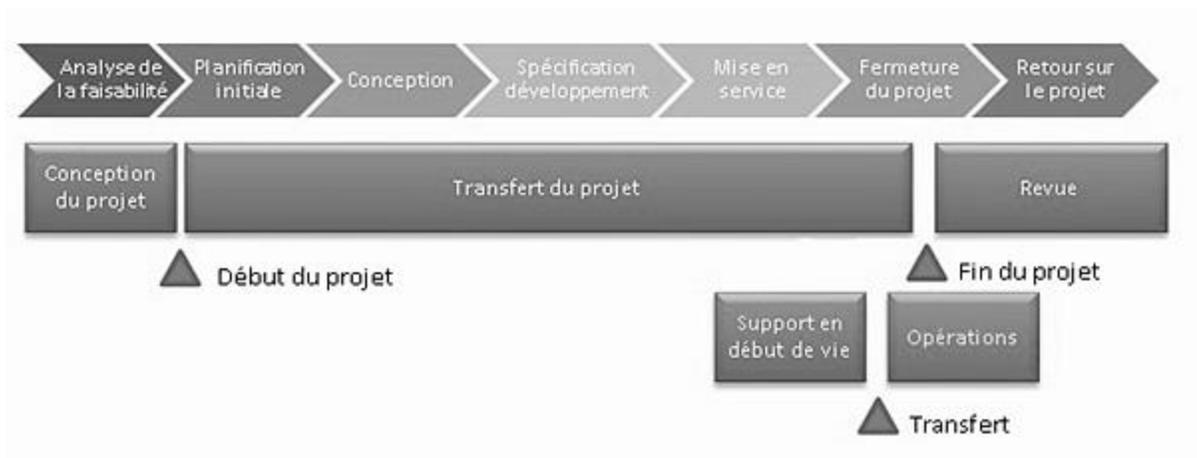
La catégorie B est aussi un groupe de cas de 2012, mais pour lesquels la classification n'est pas un rapport de problème, mais plutôt une demande de fenêtre d'entretien des données (en production). Ces demandes signalent parfois qu'un programme s'est mal comporté ou qu'un autre problème est survenu pour lequel des correctifs sont nécessaires. En 2012, sur tous les modules, la catégorie B liste 347 demandes. Pour le module S1 et ses composants, ce nombre se fixe à 41. Ceci est environ deux à trois fois moins fréquent que les problèmes de la catégorie A, ce qui justifie la sélection d'un échantillon plus restreint, soit entre quatre et six cas.

L'utilité des deux catégories est de permettre de déceler des défaillances de nature différentes. Un problème de la catégorie A marque une interruption d'un programme, par exemple, alors qu'un problème de la catégorie B indique qu'un module a complété son traitement correctement, mais a généré des données erronées.

## Annexe 5

### Processus de gestion de projets

Illustration du processus de gestion de projets de l'entreprise.



#### Processus de gestion de projets TI

Traduction libre

Source : Entreprise hôte de l'étude

## Annexe 6

### Captures d'écran pour les recherches de rapports de bogue

Capture d'écran des paramètres de recherche (rapport de problème).

#### Search Data

**Find Requests** with **all words**

with **exact** phrase

with **at least** word(s)

**without** the words

#### Advanced Search

(Cannot be used in conjunction with top part)

With **all words**

**Request Number(s)**  
(can be a list or a range i.e. 1,2,3 or 1-3)

**Request Type** Return requests of type

**Request Group** Return requests w/ group

**Request Status** Return requests w/ status

**By Requester** Submitted by

**By Provider** Serviced by

**Submitted** Submitted between

**Completed** Completed between

**Required** Required by between

All dates must be in YYYY-MM-DD format

Capture d'écran des résultats de la recherche selon le paramètre de rapports de problèmes.

## Search Data

Number of Records Selected: 312

Request No	Requester Name	Request Type	Request Date	Prov Name
63022	[REDACTED]	2 - Problem Report. According to the BP forecast, D	01/02/2012	SCOTT S.
<a href="#">63023</a>	[REDACTED]	2 - Problem Report. [REDACTED] in ab	01/02/2012	[REDACTED]
63130	B [REDACTED] WASH IM	2 - Problem Report. [REDACTED] run	01/09/2012	WASH D.
63140	B [REDACTED] IM	2 - Problem Report. scp job aborted because it fou	01/09/2012	WASH S.
<a href="#">63169</a>	A [REDACTED] WASH	2 - Problem Report. It appears that SZNEHL and SZ	01/10/2012	WASH S.
63182	B [REDACTED] IM	2 - Problem Report. [REDACTED] alert on	01/10/2012	[REDACTED]

Capture d'écran des paramètres de recherche (rôle de maintien).

## Search Data

<b>Find Requests</b>	with <b>all</b> words	<input type="text"/>	<input type="text"/>
	with <b>exact</b> phrase	<input type="text"/>	<input type="text"/>
	with <b>at least</b> word(s)	<input type="text"/>	<input type="text"/>
	<b>without</b> the words	<input type="text"/>	<input type="text"/>
<input type="button" value="Search"/> <input type="button" value="Reset"/>			

## Advanced Search

(Cannot be used in conjunction with top part)

With <b>all</b> words	<input type="text"/>	<input type="text"/>
<b>Request Number(s)</b>	<input type="text"/>	
(can be a list or a range i.e. 1,2,3 or 1-3)		
<b>Request Type</b>	Return requests of type	<input type="text" value="3: Request Maintenance Roles."/>
<b>Request Group</b>	Return requests w/ group	<input type="text"/>
<b>Request Status</b>	Return requests w/ status	<input type="text"/>
<b>By Requester</b>	Submitted by	<input type="text"/>
<b>By Provider</b>	Serviced by	<input type="text"/>
<b>Submitted</b>	Submitted between	<input type="text" value="2012-01-01"/> ... <input type="text" value="2013-01-01"/> ...
<b>Completed</b>	Completed between	<input type="text"/> ... <input type="text"/> ...
<b>Required</b>	Required by between	<input type="text"/> ... <input type="text"/> ...
All dates must be in YYYY-MM-DD format		
<input type="button" value="Search"/> <input type="button" value="Reset"/>		

Capture d'écran des résultats de la recherche selon le paramètre « rôle de maintien »

## Search Data

Number of Records Selected: 347

Request No	Requester Name	Request Type	Request Date	Prov Name
63039	[REDACTED] N	3 - Request Maintenance Roles. [REDACTED]_maint	01/03/2012	S [REDACTED]
63141	[REDACTED]	3 - Request Maintenance Roles. [REDACTED]_maint	01/09/2012	H [REDACTED]
63144	[REDACTED]	3 - Request Maintenance Roles. [REDACTED]_maint	01/09/2012	P [REDACTED]
63207	[REDACTED] /E	3 - Request Maintenance Roles. [REDACTED]_maint	01/11/2012	P [REDACTED]
63239	[REDACTED]	3 - Request Maintenance Roles. [REDACTED]_maint	01/11/2012	R [REDACTED]
63291	[REDACTED]	3 - Request Maintenance Roles. [REDACTED]_maint	01/13/2012	H [REDACTED]
63301	[REDACTED]	3 - Request Maintenance Roles. [REDACTED]_maint	01/13/2012	P [REDACTED]

## Annexe 7

### Données de bogues et analyse

Tableau des détails des incidents et du temps de résolution de bogues

#	Id.	Source	Type	Description	Temps de début	Temps de fin	Temps Écoulé
1	63 023	Web	Problème	Arrêt du programme	2012-01-02T22:35:55	2012-01-03T16:14:48	15
2	63 529	Web	Problème	Arrêt de programme	2012-01-23T21:18:25	2012-01-26T16:18:13	67
3	63 534	Web	Problème	Arrêt de programme. Type de donnée incompatible. Problème de compilation.	2012-01-24T08:55:56	2012-01-25T16:23:50	31,5
4	64 092	Web	Problème	Arrêt du programme. Données erronées en entrée.	2012-02-13T09:14:08	2012-02-13T13:23:31	4
5	64 150	Web	Problème	Arrêt du programme. Données manquantes. Création de nouveaux profils d'utilisateurs.	2012-02-14T13:03:15	2012-02-22T10:22:55	200
6	64 644	Web	Problème	Arrêt du programme. Problème de traitement d'année bissextile.	2012-02-29T20:28:52	2012-03-01T09:17:51	12,75

#	Id.	Source	Type	Description	Temps de début	Temps de fin	Temps Écoulé
7	64 650	Web	Problème	Erreur dans la production de rapport. Vues créées, mais sans octroi de permissions pour le rapport. Arrêt.	2012-03-01T08:10:41	2012-03-05T13:00:42	101
8	66 594	Web	Problème	Erreur dans la production de rapport. Signal de division par zéro.	2012-05-07T07:23:15	2012-05-07T15:06:10	7,75
9	69 582	Web	Problème	Arrêt de programme. Type de donnée incompatible. Problème de compilation.	2012-08-29T07:16:39	2012-08-29T14:56:00	7,75
10	69 668	Web	Rôle de maintien	Création de doublons et données erronées. Correction manuelle nécessaire.	2012-08-31T08:37:09	2012-08-31T17:00:00	8,5
11	71 536	Web	Problème	Arrêt de programme. Problème de permissions.	2012-11-05T09:00:54	2012-11-07T15:18:57	54,25
12	71 894	Web	Problème	Programme a roulé deux fois de suite, causant des doublons.	2012-11-15T12:43:57	2012-11-16T12:24:03	23,75
13	72 223	Web	Problème	Arrêt de programme. Paramètre trop petit pour les données fournies. Débordement de variable tampon.	2012-11-27T07:19:35	2012-11-28T13:58:25	30,5
14	72 528	Web	Problème	Arrêt de programme. Table non accessible. Ressource bloquée.	2012-12-05T07:56:46	2012-12-05T08:40:03	0,75

## **Explications particulières et nuances sur certains des problèmes recensés**

Pour la durée de résolution, les données brutes ont été recensées, ce qui inclut les soirs et les fins de semaine. Il y a un horaire de surveillance des systèmes durant le soir, mais certains des bogues ne sont pas rapportés exactement au moment même de la défaillance. Malgré cette limitation, les données de temps de résolution brutes sont considérées comme fiables puisque les interruptions déclenchées dans le planificateur Automic sont journalisées.

Pour la défaillance 63529, l'analyste aux opérations s'est renseigné sur le statut du bogue. La réponse de l'analyste de système est que cette personne prévoit adresser la défaillance au cours de la journée. La résolution a été complétée en 5 heures.

Le cas 63534 montre que l'incident a requis 31,5 heures à résoudre en janvier avec l'aide de l'analyste, du gestionnaire des opérations et du programmeur. Le problème a eu lieu une deuxième fois en août, mais cette fois il a nécessité seulement 7,75 heures de travail. L'effet de l'apprentissage par l'accès à la documentation du problème et par l'acquisition d'expérience sur le temps de résolution est manifeste dans ce cas.

Le cas 64092 est intéressant car il démontre une interaction directe entre la documentation et la défaillance. Le programme a fait défaut à cause d'un fichier de données erronées. Le document a été remplacé par un fichier valide et le programme a fonctionné sans défaillance. Toutefois, le gestionnaire des opérations a modifié la chaîne de programmes pour ajouter une étape de plus afin de permettre la visualisation et la validation des intrants et ainsi éviter un arrêt subit du programme. Le gestionnaire a aussi signalé qu'il met à jour la documentation.

Pour le cas 64150, les données de temps de résolution sont considérées comme elles sont rapportées. La correspondance dans le rapport fait état de la solution réelle (et permanente) au problème comme étant le changement d'un fournisseur pour régler la problématique récurrente. Cela a eu lieu environ deux mois plus tard. Toutefois, aux fins du projet, le temps de clôture du rapport est considéré comme étant la fin du problème parce que le système n'était plus en état d'interruption. En tout, il y a plusieurs communications non disponibles dans le rapport du problème, mais après environ 24 heures et 45 minutes une décision a été prise au sujet de la défaillance. L'analyste principale a indiqué aux opérations qu'une solution

temporaire au problème est en voie de développement. Après 200 heures, cette personne assignée a transféré un nouveau code source en production et le rapport de problème est clos.

Dans l'exemple 64644, le cas de la défaillance causée par l'année bissextile non reconnue dans le programme, le logiciel a été arrêté le 29 février 2012. La gestion des opérations n'a pas tenté de faire rouler le programme à nouveau cette journée-là. L'analyste principal a demandé de faire fonctionner le programme au début de la journée le 1<sup>er</sup> mars pour laisser passer la date problématique. La résolution du problème ponctuel a donc pris plus de 12 heures et cette donnée est le temps conservé dans le calcul. En réalité, une demande de changement a été formulée et mise en place en 2016 pour régler le problème de façon permanente avant l'occurrence d'une nouvelle année bissextile cette année-là. Dans ce cas encore, il est possible de constater comment une résolution temporaire est conservée pour la métrique bien que la solution finale ait été planifiée et programmée beaucoup plus tard.

Pour le cas 64650, le problème était résolu, mais il fallait attendre l'ajout des permissions par l'équipe de sécurité. Sur les 101 heures de résolution, 95 heures sont passées dans l'attente de la confirmation des nouvelles permissions.

Le cas 6659 a connu la défaillance deux heures avant le signalement du rapport de problème. Pour maintenir une constance dans la collecte des données, seul le temps de début du rapport est considéré, et non pas le temps réel de la défaillance.

Le cas 69668 est une situation particulière et mérite une analyse attentive. Tout d'abord, c'est un cas d'échantillon de la catégorie B, une demande de rôle de maintien. Contrairement à la majorité des bogues de cette catégorie, ce rapport de bogue mentionne qu'un problème a eu lieu et indique quel programme s'est mal comporté. Cela permet d'inclure ce bogue dans l'analyse parce qu'il est possible de retrouver des documents. La demande par l'analyste sénior a été placée le 31 août 2012 à 8 h 37 et approuvée trois minutes plus tard. La fenêtre de maintien a été permise jusqu'à 17 :00. Le ticket n'a toutefois pas été clos à 17 :00, mais bien le 4 septembre 2012 à 7 :34. Normalement, la méthode de calcul du temps dicte qu'il faut considérer le temps de fermeture du ticket. Or, il est certain que la résolution s'est effectuée avant 17 :00 le 31 août. La preuve en est que le rôle temporaire se termine à 17 :00 et qu'il n'y a pas eu de deuxième demande de rôle de maintien dans les jours subséquents ni

par cet analyste ni par ses collègues. Il peut être compris que la fermeture de cette demande a été retardée pour une raison non citée. Une raison possible est que les demandes de rôles de maintien sont moins prioritaires à fermer parce que les rôles eux-mêmes sont révoqués dès que le temps spécifié par l'opérateur est écoulé. La deuxième explication peut être la nature des demandes, qui sont moins critiques que les signaux de panne (et reçoivent moins d'attention que celles-ci). Il en émerge une observation de la difficulté de traiter les cas de demandes de rôle de maintien au même titre que les rapports de bogues. La nature de la problématique et le degré de précision du contenu du ticket pour demande de rôle de maintien ne sont pas aussi fiables que ceux des rapports de bogue de la catégorie A de l'échantillon.

Le cas 71536 ne résulte pas d'un problème de code, mais plutôt d'une erreur dans la gestion de l'automatisation. La source exacte du problème n'est pas identifiée dans le rapport, mais le gestionnaire des opérations y indique les étapes qu'il a entreprises pour tenter de régler le problème. Le déroulement des heures dans le délai de résolution montre qu'il y a eu une deuxième tentative de fonctionnement du programme et que celui-ci a fait défaut une deuxième fois. Dans ce cas, le rapport sert aussi de documentation pour maintenir des traces des solutions avec lesquelles les professionnels ont expérimenté.

Dans le cas 71894, une chaîne de programmes a été scindée en deux processus distincts. Lors de l'exécution des processus, il y a eu une panne dans l'importation de données. Dans ce cas aussi, la documentation a été mise à jour et signalée.

Dans le cas 72528, l'arrêt a eu lieu à 3 h durant la nuit alors qu'il n'y avait pas de personnel au moment de la panne. L'administrateur a ouvert une demande de suivi dès son arrivée à 7 :56. Comme il est établi en général pour la collecte du temps, le moment de signalement du problème dans le système de gestion est pris en compte, et non pas le temps de la défaillance interne du système. Un fait intéressant : en dehors de la période de l'analyse, le programme a été modifié le 29 mai 2013 pour incorporer la connaissance acquise par la mise en place des correctifs en 2012.

## Annexe 8

### Données de bogues non retenus

Tableau récapitulatif des données de bogues non retenus pour le projet

#	Id.	Source	Type	Description	Justification
1	63 560	Web	Rôle de maintien	Demande pour pouvoir supprimer une transaction de client.	Aucun détail du programme fautif n'a été mentionné. Il n'est pas possible de retracer des documents.
2	65 375	Web	Rôle de maintien	Demande pour pouvoir fusionner deux dossiers client identiques, mais avec des identifiants différents.	Aucun détail du programme fautif n'a été mentionné. Il n'est pas possible de retracer des documents.
3	65 711	Web	Problème	Arrêt de programme. Problème de permissions. Un serveur a été reconfiguré sans l'octroi des permissions nécessaires pour le fonctionnement du logiciel opérationnel.	Aucune intervention des analystes du module S1 requise. Problème réglé par l'équipe de gestion des accès.
4	66 731	Web	Problème	Arrêt du programme. Le programme a été mis en attente puis retiré de la liste. La semaine suivante, il n'y a pas eu de problème.	Détails de résolution non fournis dans le rapport. Temps de résolution non disponible. Le temps de clôture du rapport n'indique pas le temps de fin du problème puisqu'il n'y a pas eu de résolution spécifique dans ce cas. Il est possible que de mauvaises données aient causé la panne, mais il n'y en a pas de trace.

#	Id.	Source	Type	Description	Justification
5	69 822	Web	Rôle de maintien	Demande pour pouvoir ajouter une entrée dans une table en production.	Aucun bogue mentionné.
6	71 721	Web	Problème	Arrêt du programme lorsque lancé en mode vérification. Données erronées. Programme mis en attente.	La résolution du problème a été arrêtée pendant quelques jours pour permettre aux usagers de corriger les intrants. Le temps de résolution n'est pas fiable représentatif (bien qu'il soit repérable).
7	71 839	Web	Rôle de maintien	Demande pour pouvoir modifier des données de production dans un dossier pour aider un usager ayant des permissions restreintes.	Il n'y a pas de bogue.

## Annexe 9

### Données brutes des attributs de la documentation pour les bogues

Les tableaux ci-dessous reprennent les bogues et les données brutes des documents.

Identifiant du bogue		63 023	Identifiant du document			
Attributs (Critères d'analyse)			63023-1	63023-2	63023-3	63023-4
<b>Structure</b>	Longueur	<i>Pages</i>	11	10	1	2
		<i>Mots</i>	1677	2302	28	159
	Erreurs	<i>Nombre</i>	0	5	0	0
	Éléments visuels	<i>Nombre</i>	0	3	0	1
	Mise en page	<i>Évaluation</i>	ordonnée	ordonnée	ordonnée	gabarit
<b>Cycle de vie du logiciel</b>	Type de document		conception technique	opérations	opérations	opérations
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	35	19	80	59
	Synchronisation	<i>Notes</i>	Non	Non	Synchronisé	Synchronisé
<b>Accessibilité</b>	Facile	<i>Oui Non Difficile</i>	Oui	Oui	Oui	Oui
	Type de fichier		Standard, .doc	Standard, .doc	Standard, .txt	Standard, .doc
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	2 personnes, 1 groupe	2 personnes	0	0
Identifiant du bogue		63 023	Identifiant du document			
Attributs			63023-1	63023-2	63023-3	63023-4

(Critères d'analyse)						4
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		Lettres	Lettres	Lettres	Lettres

Identifiant du bogue		63 529	Identifiant du document		
Attributs			63529-1	63529-2	63529-3
(Critères d'analyse)					
<b>Structure</b>	Longueur	<i>Pages</i>	3	5	1
		<i>Mots</i>	393	1453	7
	Erreurs	<i>Nombre</i>	2	5	0
	Éléments visuels	<i>Nombre</i>	2	2	0
	Mise en page	<i>Évaluation</i>	ordonné	ordonné	ordonné
<b>Cycle de vie du logiciel</b>	Type de document		opérations	processus	sécurité
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	0	8	0
	Synchronisation	<i>Notes</i>	non synchronisé mais non requis dans ce document.	Non, car juste en partie	Synchronisé
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	Oui	Oui
		<i>Non</i>			
		<i>Difficile</i>			
	Type de fichier		Standard, .doc	Standard, .doc	Standard, .txt
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	0	2 groupes	0
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		_chkt.sql	_chkt.sql, _load.pc	_load.pc

Pour le bogue 63534, les documents sont les mêmes que pour le bogue 69582.

Identifiant du bogue		63 534	Identifiant du document			
Attributs (Critères d'analyse)			63534-1	63534-2	63534-3	63534-4
<b>Structure</b>	Longueur	<i>Pages</i>	3	3	1	1
		<i>Mots</i>	626	309	514	6
	Erreurs	<i>Nombre</i>	1	0	0	0
	Éléments visuels	<i>Nombre</i>	1	3	1	1
	Mise en page	<i>Évaluation</i>	ordonné	gabarit	gabarit	gabarit
<b>Cycle de vie du logiciel</b>	Type de document		Processus	opérations	sécurité	opérations
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	8	1	0	0
	Synchronisation	<i>Notes</i>	Pas synchronisé	Pas synchronisé	Pas mis à jour, mais pas nécessaire	Pas synchronisé, mais pas nécessaire.
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	Oui	Oui	Oui
		<i>Non</i>				
		<i>Difficile</i>				
	Type de fichier		Standard, .docx	Standard, .doc	Standard, .doc	Standard, .txt
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	0	1 personne	0	0
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		_case	_case	_case	_case

Identifiant du bogue		64 092	Identifiant du document			
Attributs (Critères d'analyse)			64092-1	64092-2	64092-3	64092-4
<b>Structure</b>	Longueur	<i>Pages</i>	2	3	14	1
		<i>Mots</i>	434	554	40	39
	Erreurs	<i>Nombre</i>	0	0	0	0
	Éléments visuels	<i>Nombre</i>	0	3	1	0
	Mise en page	<i>Évaluation</i>	désordonnée	Gabarit	Gabarit	gabarit
<b>Cycle de vie du logiciel</b>	Type de document		conception technique	opérations	opérations	opérations
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	7	5	0	0
	Synchronisation	<i>Notes</i>	Non synchronisé	Document synchronisé avec certains fichiers, mais par les autres, donc non.	Non Synchronisé mais pas nécessaire	Non Synchronisé mais pas nécessaire
<b>Accessibilité</b>	Facile	<i>Oui Non Difficile</i>	Oui	Oui	Oui	Oui
	Type de fichier		Standard, .doc	Standard, .doc	Standard, .doc	Standard, .txt
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	0	2 personnes, 1 groupe	0	0
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		Ldqc et plusieurs autres	Ldqc et plusieurs autres	Ldqc	ldqc

Identifiant du bogue		64 150	Identifiant du document				
Attributs			64150-1	64150-2	64150-3	64150-4	64150-5
(Critères d'analyse)							
<b>Structure</b>	Longueur	<i>Pages</i>	2	14	3	14	13
		<i>Mots</i>	281	4747	321	4767	1759
	Erreurs	<i>Nombre</i>	1	9	0	6	3
	Éléments visuels	<i>Nombre</i>	1	1	0	0	0
	Mise en page	<i>Évaluation</i>	ordonné	ordonné	gabarit	gabarit	gabarit
<b>Cycle de vie du logiciel</b>	Type de document		Conception technique	opérations	opérations	support (routinier)	Support (routinier)
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	3	7	1	3	3
	Synchronisation	<i>Notes</i>	Document en retard d'un an sur le code	Non synchronisé	Non synchronisé	Document en retard d'un an sur le code	Document en retard d'un an sur le code
<b>Accessibilité</b>	Facilité	<i>Oui Non Difficile</i>	Oui	Oui	Oui	Difficile (nom de fichier)	Difficile (nom de fichier)
	Type de fichier		Standard, .doc	Standard, .xls	Standard, .doc	Standard, .doc	Standard, .doc
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	0	0	1 personne	0	0
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		xml.pl wheat.pl 1,3	10 fichiers szr	wheat.pl 1,3	n/d	Ldi, chop et autres

Identifiant du bogue		64 644	Identifiant du document		
Attributs (Critères d'analyse)			64644-1	64644-2	64644-3
<b>Structure</b>	Longueur	<i>Pages</i>	1	2	2
		<i>Mots</i>	324	375	358
	Erreurs	<i>Nombre</i>	0	2	0
	Éléments visuels	<i>Nombre</i>	0	2	2
	Mise en page	<i>Évaluation</i>	ordonnée	gabarit	gabarit
<b>Cycle de vie du logiciel</b>	Type de document		conception technique	opérations	opérations
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	0	1	1
	Synchronisation	<i>Notes</i>	Non. Au moins 5 mises à jour synchronisées, suivies de plus de 5 mises à jour du code non jumelé au document	Pas synchronisé. Document mis à jour en 2005 et 2010 ; code mis à jour en 2005, 2011 et 2016	Synchronisé
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	Oui	Oui
		<i>Non</i>			
		<i>Difficile</i>			
	Type de fichier		Standard, .doc	Standard, .doc	Standard, .doc
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	2 personnes	1 personne, 1 groupe	3 personnes
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		_inst. sql (4 fichiers)	_inst (1)	_inst (2)

Identifiant du bogue		64 650	Identifiant du document	
Attributs (Critères d'analyse)			64650-1	64650-2
<b>Structure</b>	Longueur	<i>Pages</i>	2	7
		<i>Mots</i>	323	1486
	Erreurs	<i>Nombre</i>	2 (inspection visuelle, car vérificateur ne donne pas de statistiques)	4
	Éléments visuels	<i>Nombre</i>	2	1
	Mise en page	<i>Évaluation</i>	Gabarit	Désordonné
<b>Cycle de vie du logiciel</b>	Type de document		Opérations	Conception technique et spécification
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	5	1
	Synchronisation	<i>Notes</i>	Synchronisé	Non. Synchronisé avec certaines mises à jour du code, mais pas toutes.
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	Oui
		<i>Non</i>		
		<i>Difficile</i>		
	Type de fichier		Standard, .doc	Standard, .doc
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	2 personnes, 2 groupes	1 groupe
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		Frez.sql	Frez.sql

Identifiant du bogue		66 594	Identifiant du document			
Attributs (Critères d'analyse)			66594-1	66594-2	66594-3	66594-4
<b>Structure</b>	Longueur	<i>Pages</i>	1	1	3	1
		<i>Mots</i>	142	222	387	8
	Erreurs	<i>Nombre</i>	0	0	0	0
	Éléments visuels	<i>Nombre</i>	3	1	3	0
	Mise en page	<i>Évaluation</i>	ordonné	gabarit	gabarit	gabarit
<b>Cycle de vie du logiciel</b>	Type de document		spécification technique	sécurité	opérations	opérations
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	1	0	0	5
	Synchronisation	<i>Notes</i>	Synchronisé	Non synchronisé mais non nécessaire	Non synchronisé	Non synchronisé, pas nécessaire
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	Oui	Oui	Oui
		<i>Non</i>				
		<i>Difficile</i>				
	Type de fichier		Standard, .doc	Standard, .docx	Standard, .doc	Standard, .txt
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	0	0	0	0
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		_gpaa.sqr	_gpaa.sqr	_gpaa.sqr	_gpaa.sqr

Identifiant du bogue		69 582	Identifiant du document			
Attributs (Critères d'analyse)			69582-1	69582-2	69582-3	69582-4
<b>Structure</b>	Longueur	<i>Pages</i>	3	3	1	1
		<i>Mots</i>	626	309	514	6
	Erreurs	<i>Nombre</i>	1	0	0	0
	Éléments visuels	<i>Nombre</i>	1	3	1	1
	Mise en page	<i>Évaluation</i>	ordonné	gabarit	gabarit	gabarit
<b>Cycle de vie du logiciel</b>	Type de document		Processus	opérations	sécurité	opérations
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	8	1	0	0
	Synchronisation	<i>Notes</i>	Pas synchronisé	Pas synchronisé.	Pas mis à jour, mais pas nécessaire	Pas synchronisé mais pas nécessaire.
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	Oui	Oui	Oui
		<i>Non</i>				
		<i>Difficile</i>				
	Type de fichier		Standard, .docx	Standard, .doc	Standard, .doc	Standard, .txt
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	0	1 personne	0	0
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		_case	_case	_case	_case

Identifiant du bogue		69 668	Identifiant du document	
Attributs (Critères d'analyse)			69668-1	69668-2
<b>Structure</b>	Longueur	<i>Pages</i>	3	Document manquant
		<i>Mots</i>	415	
	Erreurs	<i>Nombre</i>	2	
	Éléments visuels	<i>Nombre</i>	4	
	Mise en page	<i>Évaluation</i>	Gabarit	
<b>Cycle de vie du logiciel</b>	Type de document		Opérations	
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	5	
	Synchronisation	<i>Notes</i>	Non. Document à jour avec un code source (2006, 2007 et 2010), mais la procédure a connu 15 mises à jour, mais les 7 dernières sont absentes du document.	
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	
		<i>Non</i>		
		<i>Difficile</i>		
	Type de fichier		Standard, .doc	
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	2 personnes	
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		Crse et une procédure chg_crse.prc	

Identifiant du bogue		71 536	Identifiant du document			
Attributs (Critères d'analyse)			71536-1	71536-2	71536-3	71536-4
<b>Structure</b>	Longueur	<i>Pages</i>	3	1	1	3
		<i>Mots</i>	482	247	9	731
	Erreurs	<i>Nombre</i>	0	0	0	0
	Éléments visuels	<i>Nombre</i>	2	1	0	0
	Mise en page	<i>Évaluation</i>	gabarit	gabarit	gabarit	ordonné
<b>Cycle de vie du logiciel</b>	Type de document		opérations	sécurité	opérations	processus
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	2	0	0	5
	Synchronisation	<i>Notes</i>	Synchronisé	Synchronisé	Synchronisé	Synchronisé
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	Oui	Oui	Oui
		<i>Non</i>				
		<i>Difficile</i>				
	Type de fichier		Standard, .doc	Standard, .doc	Standard, .txt	Standard, .doc
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	0	0	0	1 personne, 3 groupes
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		_acdi. sql	_acdi. sql	_acdi. sql	_acdi. sql et autres

Tableau 5.3-7a

Le tableau ci-dessous contient les données pour les trois premiers documents du bogue. Les autres documents sont présentés dans un tableau subséquent.

## Partie 1

Identifiant du bogue		71 894	Identifiant du document		
Attributs (Critères d'analyse)			71894-1	71894-2	71894-3
<b>Structure</b>	Longueur	<i>Pages</i>	1	3	15
		<i>Mots</i>	6	1188	3962
	Erreurs	<i>Nombre</i>	0	0	20
	Éléments visuels	<i>Nombre</i>	1	0	2
	Mise en page	<i>Évaluation</i>	ordonnée	ordonné	désordonné
<b>Cycle de vie du logiciel</b>	Type de document		Diagramme de processus	opérations	conception technique et spécifications
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	0	6	16
	Synchronisation	<i>Notes</i>	Pas à jour avec le code	Pas à jour avec le code	Pas synchronisé avec les changements du code en 2014.
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	Oui	Oui
		<i>Non</i>			
		<i>Difficile</i>			
	Type de fichier		Standard, .doc	Standard, .doc	Standard, .doc
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	2 personnes	2 personnes, 1 groupe	0
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		code en C et rapports _etmt, _wapp, _wasu, _waer	Plusieurs	_wapp.sqr

Partie 2

Identifiant du bogue		71 894	Identifiant du document		
Attributs (Critères d'analyse)			71894-4	71894-5	71894-6
<b>Structure</b>	Longueur	<i>Pages</i>	3	2	1
		<i>Mots</i>	1152	215	8
	Erreurs	<i>Nombre</i>	6	0	0
	Éléments visuels	<i>Nombre</i>	3	1	0
	Mise en page	<i>Évaluation</i>	gabarit	gabarit	gabarit
<b>Cycle de vie du logiciel</b>	Type de document		opérations	sécurité	opérations
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	11	5	5
	Synchronisation	<i>Notes</i>	Pas synchronisé. Deux mises à jour manquantes	Pas mis à jour, mais pas nécessaire	Pas mis à jour, mais pas nécessaire
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	Oui	Oui
		<i>Non</i>			
		<i>Difficile</i>			
	Type de fichier		Standard, .doc	Standard, .doc	Standard, .txt
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	1 groupe	0	0
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		_wapp.sqp	_wapp.sqp	_wapp.sqp

Identifiant du bogue		72 223	Identifiant du document	
Attributs (Critères d'analyse)			72223-1	72223-2
<b>Structure</b>	Longueur	<i>Pages</i>	3	2
		<i>Mots</i>	410	443
	Erreurs	<i>Nombre</i>	1	1
	Éléments visuels	<i>Nombre</i>	3	0
	Mise en page	<i>Évaluation</i>	Gabarit	Ordonné
<b>Cycle de vie du logiciel</b>	Type de document		Opérations	Processus
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	0	2
	Synchronisation	<i>Notes</i>	Synchronisé	Synchronisé
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	Oui
		<i>Non</i>		
		<i>Difficile</i>		
	Type de fichier		Standard, .doc	Standard, .doc
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	0	0
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		_cmss	_cmss _cmsi

Identifiant du bogue		72 528	Identifiant du document	
Attributs (Critères d'analyse)			72528-1	72528-2
<b>Structure</b>	Longueur	<i>Pages</i>	1	2
		<i>Mots</i>	239	429
	Erreurs	<i>Nombre</i>	1	0
	Éléments visuels	<i>Nombre</i>	2	0
	Mise en page	<i>Évaluation</i>	Gabarit	Ordonné
<b>Cycle de vie du logiciel</b>	Type de document		Opérations	Processus
<b>Mise à jour</b>	Fréquence	<i>Nombre</i>	12	10
	Synchronisation	<i>Notes</i>	Non synchronisé	Plusieurs mises à jour effectuées, mais le document n'est pas à jour avec le code.
<b>Accessibilité</b>	Facile	<i>Oui</i>	Oui	Oui
		<i>Non</i>		
		<i>Difficile</i>		
	Type de fichier		Standard, .doc	Standard, .doc
<b>Traçabilité</b>	Qui rejoindre en cas de problèmes ou de questions	Nombre de personne ou de groupe	5 personnes	1 groupe
<b>Logiciel</b>	Pour identifier le code source pour le jumelage		_rtruc. sql	_rtruc. sql et 7 autres logiciels

## Annexe 10

### Données brutes avec les attributs formatés pour le programme de calcul des indices

Compteur de bogue	Identifiant du bogue	Identifiant du document	Nombre de pages	Nombre de mots	Nombre d'erreurs	Nombre de figures	Mise en page	Type de document	Nombre de mises à jour	Synchronisation	Accessibilité	Type de fichier	Traçabilité
1	63023	63023-1	11	1677	0	0	O	conception technique	35	NO	Y	.doc	G
1	63023	63023-2	10	2302	5	3	O	opérations	19	NO	Y	.doc	P
1	63023	63023-3	1	28	0	0	O	opérations	80	YES	Y	.txt	N
1	63023	63023-4	2	159	0	1	G	opérations	59	YES	Y	.doc	N
2	63529	63529-1	3	393	2	2	O	opérations	0	NA	Y	.doc	N
2	63529	63529-2	5	1453	5	2	O	processus	8	NO	Y	.doc	G
2	63529	63529-3	1	7	0	0	O	sécurité	0	YES	Y	.txt	N
3	63534	63534-1	3	626	1	1	O	processus	8	NO	Y	.docx	N
3	63534	63534-2	3	309	0	3	G	opérations	1	NO	Y	.doc	P
3	63534	63534-3	1	514	0	1	G	sécurité	0	NA	Y	.doc	N
3	63534	63534-4	1	6	0	1	G	opérations	0	NA	Y	.txt	N
4	64092	64092-1	2	434	0	0	D	conception technique	7	NO	Y	.doc	N
4	64092	64092-2	3	554	0	3	G	opérations	5	NO	Y	.doc	G
4	64092	64092-3	14	40	0	1	G	opérations	0	NA	Y	.doc	N
4	64092	64092-4	1	39	0	0	G	opérations	0	NA	Y	.txt	N

Compteur de bogue	Identifiant du bogue	Identifiant du document	Nombre de pages	Nombre de mots	Nombre d'erreurs	Nombre de figures	Mise en page	Type de document	Nombre de mises à jour	Synchronisation	Accessibilité	Type de fichier	Traçabilité
5	64150	64150-1	2	281	1	1	O	conception technique	3	NO	Y	.doc	N
5	64150	64150-2	14	4747	9	1	O	opérations	7	NO	Y	.xls	N
5	64150	64150-3	3	321	0	0	G	opérations	1	NO	Y	.doc	P
5	64150	64150-4	14	4767	6	0	G	support et maintien	3	NO	D	.doc	N
5	64150	64150-5	13	1759	3	0	G	support et maintien	3	NO	D	.doc	N
6	64644	64644-1	1	324	0	0	O	conception technique	0	NO	Y	.doc	P
6	64644	64644-2	2	375	2	2	G	opérations	1	NO	Y	.doc	M
6	64644	64644-3	2	358	0	2	G	opérations	1	YES	Y	.doc	P
7	64650	64650-1	2	323	2	2	G	opérations	5	YES	Y	.doc	M
7	64650	64650-2	7	1486	4	1	D	conception technique	1	NO	Y	.doc	G
8	66594	66594-1	1	142	0	3	O	spécification technique	1	YES	Y	.doc	N
8	66594	66594-2	1	222	0	1	G	sécurité	0	NA	Y	.docx	N
8	66594	66594-3	3	387	0	3	G	opérations	0	NO	Y	.doc	N
8	66594	66594-4	1	8	0	0	G	opérations	5	NA	Y	.txt	N
9	69582	69582-1	3	626	1	1	O	processus	8	NO	Y	.docx	N
9	69582	69582-2	3	309	0	3	G	opérations	1	NO	Y	.doc	P
9	69582	69582-3	1	514	0	1	G	sécurité	0	NA	Y	.doc	N
9	69582	69582-4	1	6	0	1	G	opérations	0	NA	Y	.txt	N
10	69668	69668-1	3	415	2	4	G	opérations	5	NO	Y	.doc	P
10	69668	69668-2	0	0	0	0	O	0	0	NO	N	0	N

Compteur de bogue	Identifiant du bogue	Identifiant du document	Nombre de pages	Nombre de mots	Nombre d'erreurs	Nombre de figures	Mise en page	Type de document	Nombre de mises à jour	Synchronisation	Accessibilité	Type de fichier	Traçabilité
11	71536	71536-1	3	482	0	2	G	opérations	2	YES	Y	.doc	N
11	71536	71536-2	1	247	0	1	G	sécurité	0	YES	Y	.doc	N
11	71536	71536-3	1	9	0	0	G	opérations	0	YES	Y	.txt	N
11	71536	71536-4	3	731	0	0	O	processus	5	YES	Y	.doc	M
12	71894	71894-1	1	6	0	1	O	processus	0	NO	Y	.doc	P
12	71894	71894-2	3	1188	0	0	O	opérations	6	NO	Y	.doc	M
12	71894	71894-3	15	3962	20	2	D	conception technique	16	NO	Y	.doc	N
12	71894	71894-4	3	1152	6	3	G	opérations	11	NO	Y	.doc	G
12	71894	71894-5	2	215	0	1	G	opérations	5	NA	Y	.doc	N
12	71894	71894-6	1	8	0	0	G	opérations	5	NA	Y	.txt	N
13	72223	72223-1	3	410	1	3	G	opérations	0	YES	Y	.doc	N
13	72223	72223-2	2	443	1	0	O	processus	2	YES	Y	.doc	N
14	72528	72528-1	1	239	1	2	G	opérations	12	NO	Y	.doc	P
14	72528	72528-2	2	429	0	0	O	processus	10	NO	Y	.doc	G

## Annexe 11

### Ajustement des points extrêmes pour la corrélation sommaire

Pour essayer de comprendre les phénomènes qui ont lieu, une analyse de trois cas est effectuée plus attentivement. Il s'agit des trois points extrêmes dans le graphique :

- le point 63529 a un indice total de 12,5 et un temps ajusté de 5 heures;
- le point 64150 a un indice total de 13,5 et un temps de résolution de 200 heures;
- le point 64650 a un indice total de 12,5 et un temps de résolution de 101 heures.

Pour le cas 63529, il y a eu un temps mort au début du rapport de bogue, et le temps ajusté est pris en compte.

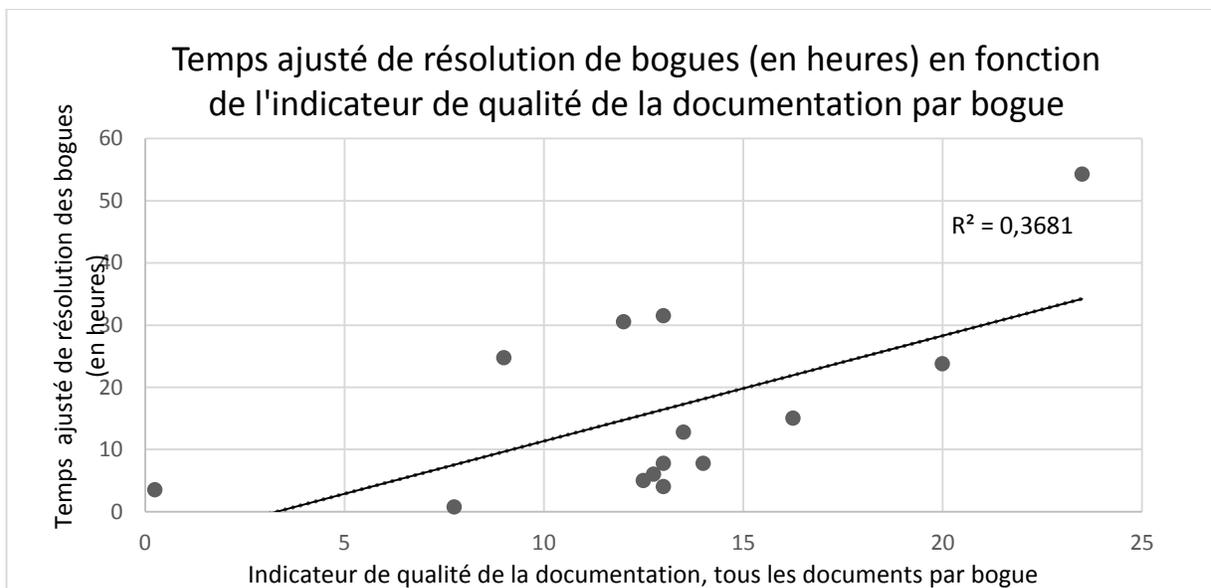
Pour le cas 64150, après environ 24 heures et 45 minutes une décision a été prise au sujet de la défaillance et les responsables ont clos le rapport de problème. La résolution à moyen terme a toutefois requis 200 heures. Un temps de résolution de 6 heures est considéré comme le temps de résolution ajusté.

Pour le cas 64650, le problème était résolu, mais il fallait attendre l'ajout des permissions par l'équipe de sécurité. Sur les 101 heures de résolution, 95 heures se sont écoulées dans l'attente de la confirmation des nouvelles permissions. Un temps de résolution de 6 heures est considéré comme le temps de résolution ajusté.

<b>Bogue</b>	<b>Indicateur</b>	<b>Temps de résolution ajusté</b>
63023	16,25	15
63529	12,5	5
63534	13	31,5
64092	13	4
64150	9	24,75
64644	13,5	12,75

Bogue	Indicateur	Temps de résolution ajusté
64650	12,75	6
66594	14	7,75
69582	13	7,75
69668	0,25	3,5
71536	23,5	54,25
71894	20	23,75
72223	12	30,5
72528	7,75	0,75

Le graphique suivant montre la corrélation pour le temps de résolution ajusté en fonction de la cote totale par cas.



**Graphique du temps de résolution des bogues en fonction de l'indice de qualité des documents des programmes en production**

Les points aberrants sont modifiés pour le graphique de la figure 5.1. Le but est d'observer quelle est la relation lorsque les points extrêmes sont corrigés. En effet, la linéarité s'intensifie avec un indice de corrélation  $r$  de 0,6067, mais la fidélité des points à la pente demeure faible avec un coefficient de détermination de 0,3681.

## Annexe 12

### Indices individuels détaillées, cumulés et moyens pour les documents

Dans cette annexe, les trois tableaux sont le tableau des **indices des attributs** pour chaque document, le tableau des indices **cumulés** des attributs pour chaque document par bogue, et le tableau des indices **moyens** des attributs pour chaque document par bogue. Les deux derniers tableaux (indices cumulés et moyens) sont suivis des graphiques pour les données tabulées.

**Tableau des indices des attributs pour chaque document**

Compteur de bogue	Bogue	Identifiant du document	Longueur	Erreurs	Figures	Mise en page	Nombre de mises à jour	Synchronisation	Accessibilité	Type de fichier	Traçabilité	Indice du document
1	63023	63023-1	0,5	1	0	0,25	1	-2	1	0,25	2	4
1	63023	63023-2	0,5	1	1	0,25	1	-2	1	0,25	1	4
1	63023	63023-3	0	1	0	0,25	1	2	1	0,25	0	5,5
1	63023	63023-4	0,5	1	0,5	0,5	1	2	1	0,25	0	6,75
2	63529	63529-1	1	1	0,5	0,25	-0,5	0	1	0,25	0	3,5
2	63529	63529-2	1	1	0,5	0,25	1	-2	1	0,25	2	5
2	63529	63529-3	0	1	0	0,25	-0,5	2	1	0,25	0	4
3	63534	63534-1	1	1	0,5	0,25	1	-2	1	0,25	0	3
3	63534	63534-2	1	1	1	0,5	0,25	-2	1	0,25	1	4
3	63534	63534-3	0,5	1	0,5	0,5	-0,5	0	1	0,25	0	3,25
3	63534	63534-4	0	1	0,5	0,5	-0,5	0	1	0,25	0	2,75

Compteur de bogue	Bogue	Identifiant du document	Longueur	Erreurs	Figures	Mise en page	Nombre de mises à jour	Synchronisation	Accessibilité	Type de fichier	Traçabilité	Indice du document
4	64092	64092-1	1	1	0	0	1	-2	1	0,25	0	2,25
4	64092	64092-4	0	1	0	0,5	-0,5	0	1	0,25	0	2,25
4	64092	64092-3	0,5	1	0,5	0,5	-0,5	0	1	0,25	0	3,25
4	64092	64092-2	1	1	1	0,5	0,5	-2	1	0,25	2	5,25
5	64150	64150-1	0,5	1	0,5	0,25	0,25	-2	1	0,25	0	1,75
5	64150	64150-2	0,5	1	0,5	0,25	1	-2	1	0	0	2,25
5	64150	64150-3	1	1	0	0,5	0,25	-2	1	0,25	1	3
5	64150	64150-4	0,5	1	0	0,5	0,25	-2	0,5	0,25	0	1
5	64150	64150-5	0,5	1	0	0,5	0,25	-2	0,5	0,25	0	1
6	64644	64644-2	1	1	0,5	0,5	0,25	-2	1	0,25	2	4,5
6	64644	64644-1	0,5	1	0	0,25	-0,5	-2	1	0,25	1	1,5
6	64644	64644-3	1	1	0,5	0,5	0,25	2	1	0,25	1	7,5
7	64650	64650-1	1	1	0,5	0,5	0,5	2	1	0,25	2	8,75
7	64650	64650-2	1	1	0,5	0	0,25	-2	1	0,25	2	4
8	66594	66594-1	0	1	1	0,25	0,25	2	1	0,25	0	5,75
8	66594	66594-2	0	1	0,5	0,5	-0,5	0	1	0,25	0	2,75
8	66594	66594-3	1	1	1	0,5	-0,5	-2	1	0,25	0	2,25
8	66594	66594-4	0	1	0	0,5	0,5	0	1	0,25	0	3,25
9	69582	69582-1	1	1	0,5	0,25	1	-2	1	0,25	0	3
9	69582	69582-4	0	1	0,5	0,5	-0,5	0	1	0,25	0	2,75
9	69582	69582-2	1	1	1	0,5	0,25	-2	1	0,25	1	4
9	69582	69582-3	0,5	1	0,5	0,5	-0,5	0	1	0,25	0	3,25
10	69668	69668-1	1	1	1	0,5	0,5	-2	1	0,25	1	4,25
10	69668	69668-2	-4	0	0	0	0	0	0	0	0	-4

Compteur de bogue	Bogue	Identifiant du document	Longueur	Erreurs	Figures	Mise en page	Nombre de mises à jour	Synchronisation	Accessibilité	Type de fichier	Traçabilité	Indice du document
11	71536	71536-1	1	1	0,5	0,5	0,25	2	1	0,25	0	6,5
11	71536	71536-2	0	1	0,5	0,5	-0,5	2	1	0,25	0	4,75
11	71536	71536-4	1	1	0	0,25	0,5	2	1	0,25	2	8
11	71536	71536-3	0	1	0	0,5	-0,5	2	1	0,25	0	4,25
12	71894	71894-1	0	1	0,5	0,25	-0,5	-2	1	0,25	1	1,5
12	71894	71894-2	1	1	0	0,25	0,5	-2	1	0,25	2	4
12	71894	71894-3	0,5	0	0,5	0	1	-2	1	0,25	0	1,25
12	71894	71894-4	1	1	1	0,5	1	-2	1	0,25	2	5,75
12	71894	71894-5	0,5	1	0,5	0,5	0,5	0	1	0,25	0	4,25
12	71894	71894-6	0	1	0	0,5	0,5	0	1	0,25	0	3,25
13	72223	72223-1	1	1	1	0,5	-0,5	2	1	0,25	0	6,25
13	72223	72223-2	1	1	0	0,25	0,25	2	1	0,25	0	5,75
14	72528	72528-1	0	1	0,5	0,5	1	-2	1	0,25	1	3,25
14	72528	72528-2	1	1	0	0,25	1	-2	1	0,25	2	4,5

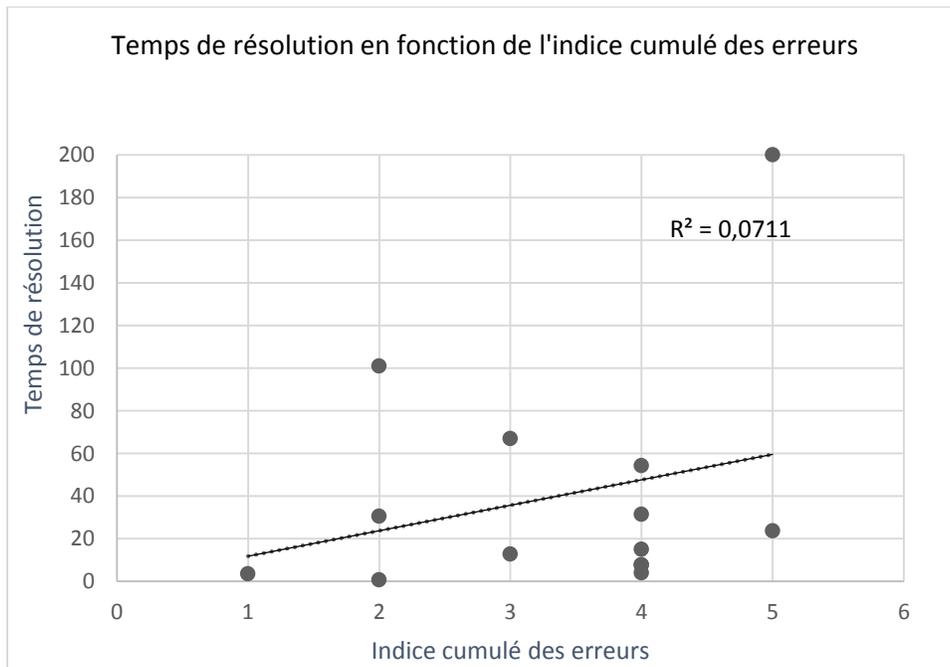
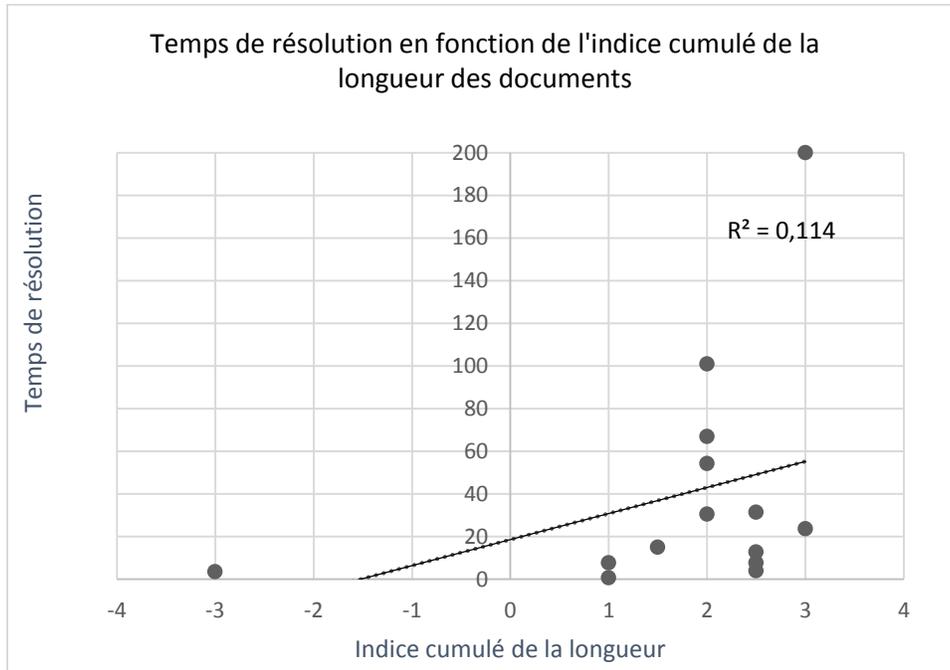
**Tableau des indices cumulés des attributs pour chaque document par bogue**

Bogue	Longueur	Erreurs indice cumule	Figures indice cumule	Mise en page	Fréquence de mises à jour	Synchronisation	Accessibilité	Type de fichier	Traçabilité	Nombre de documents	Temps de résolution	Temps ajusté de résolution
63023	1,5	4	1,5	1,25	4	0	4	1	3	4	15	15
63529	2	3	1	0,75	0	0	3	0,75	2	3	67	5
63534	2,5	4	2,5	1,75	0,25	-4	4	1	1	4	31,5	31,5
64092	2,5	4	1,5	1,5	0,5	-4	4	1	2	4	4	4
64150	3	5	1	2	2	-10	4	1	1	5	200	24,75
64644	2,5	3	1	1,25	0	-2	3	0,75	4	3	12,75	12,75
64650	2	2	1	0,5	0,75	0	2	0,5	4	2	101	6
66594	1	4	2,5	1,75	-0,25	0	4	1	0	4	7,75	7,75
69582	2,5	4	2,5	1,75	0,25	-4	4	1	1	4	7,75	7,75
69668	-3	1	1	0,5	0,5	-2	1	0,25	1	2	3,5	3,5
71536	2	4	1	1,75	-0,25	8	4	1	2	4	54,25	54,25
71894	3	5	2,5	2	3	-8	6	1,5	5	6	23,75	23,75
72223	2	2	1	0,75	-0,25	4	2	0,5	0	2	30,5	30,5
72528	1	2	0,5	0,75	2	-4	2	0,5	3	2	0,75	0,75

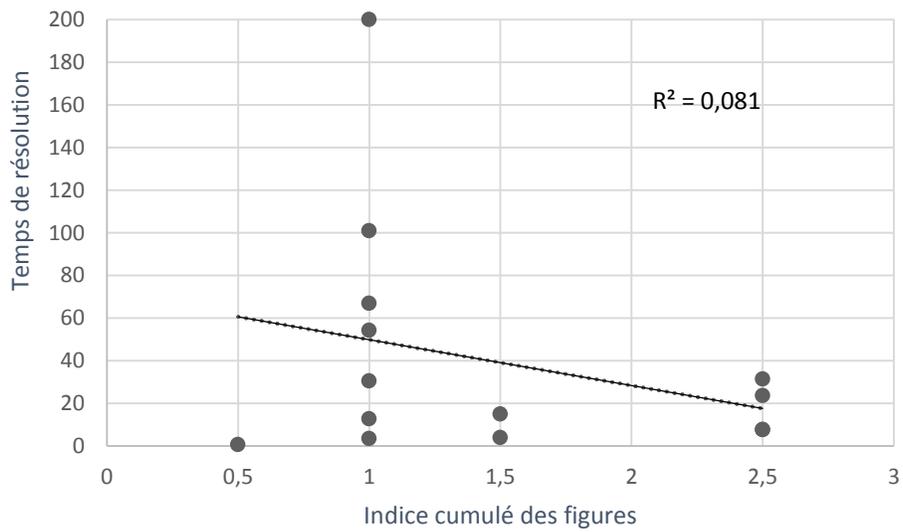
**Tableau des indices moyens des attributs pour chaque document par bogue**

Identifiant du bogue	Longueur	Erreurs	Figures	Mise en page	Fréquence de mises à jour	Synchronisation	Accessibilité	Type de fichier	Traçabilité	Nombre de documents	Temps de résolution	Temps ajusté de résolution
63023	0,375	1	0,375	0,313	1	0	1	0,25	0,75	4	15	15
63529	0,667	1	0,333	0,25	0	0	1	0,25	0,667	3	67	5
63534	0,625	1	0,625	0,438	0,063	-1	1	0,25	0,25	4	31,5	31,5
64092	0,625	1	0,375	0,375	0,125	-1	1	0,25	0,5	4	4	4
64150	0,6	1	0,2	0,4	0,4	-2	0,8	0,2	0,2	5	200	24,75
64644	0,833	1	0,333	0,417	0	-0,667	1	0,25	1,333	3	12,75	12,75
64650	1	1	0,5	0,25	0,375	0	1	0,25	2	2	101	6
66594	0,25	1	0,625	0,438	-0,063	0	1	0,25	0	4	7,75	7,75
69582	0,625	1	0,625	0,438	0,063	-1	1	0,25	0,25	4	7,75	7,75
69668	-1,5	0,5	0,5	0,25	0,25	-1	0,5	0,125	0,5	2	3,5	3,5
71536	0,5	1	0,25	0,438	-0,063	2	1	0,25	0,5	4	54,25	54,25
71894	0,5	0,833	0,417	0,333	0,5	-1,333	1	0,25	0,833	6	23,75	23,75
72223	2	2	1	0,75	-0,25	4	2	0,5	0	2	30,5	30,5
72528	1	2	0,5	0,75	2	-4	2	0,5	3	2	0,75	0,75

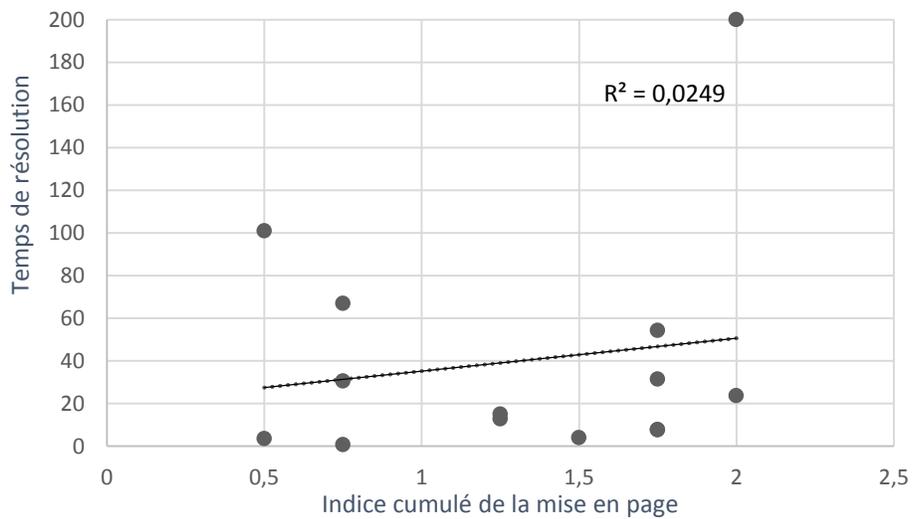
## Graphiques des indices cumulés de chaque attribut

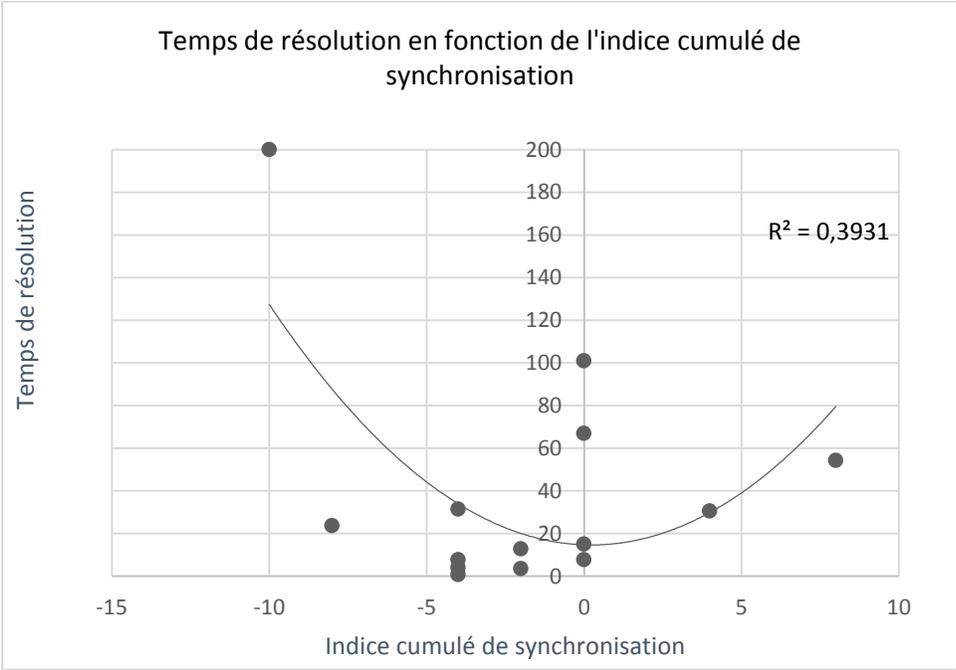
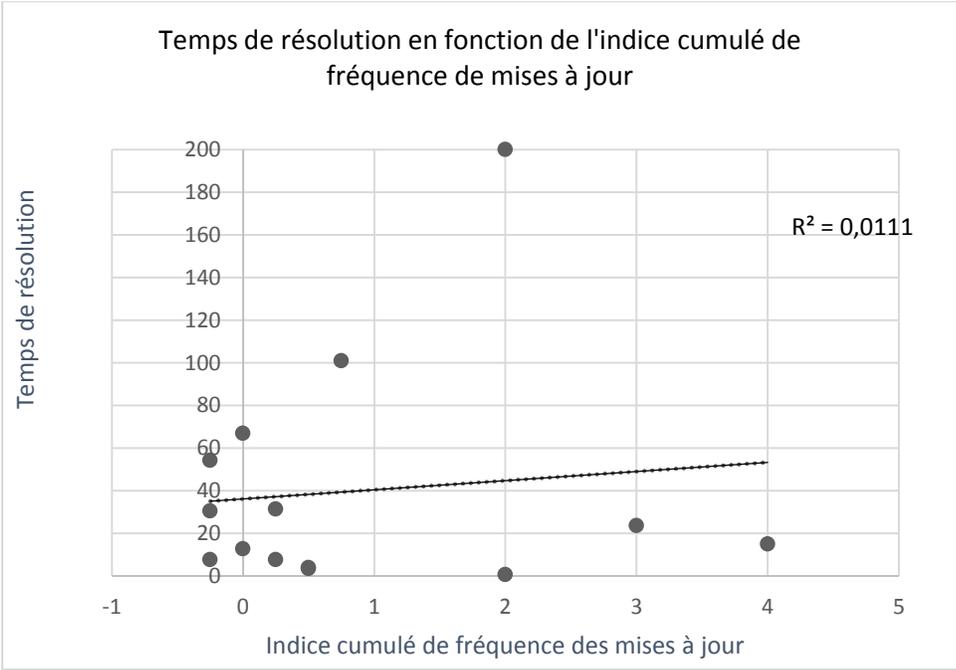


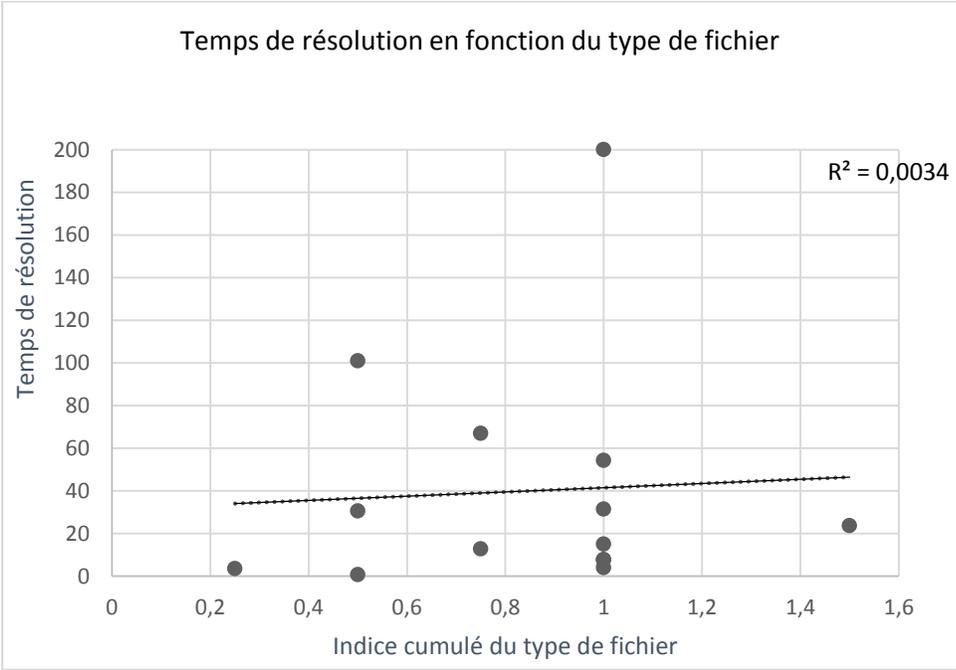
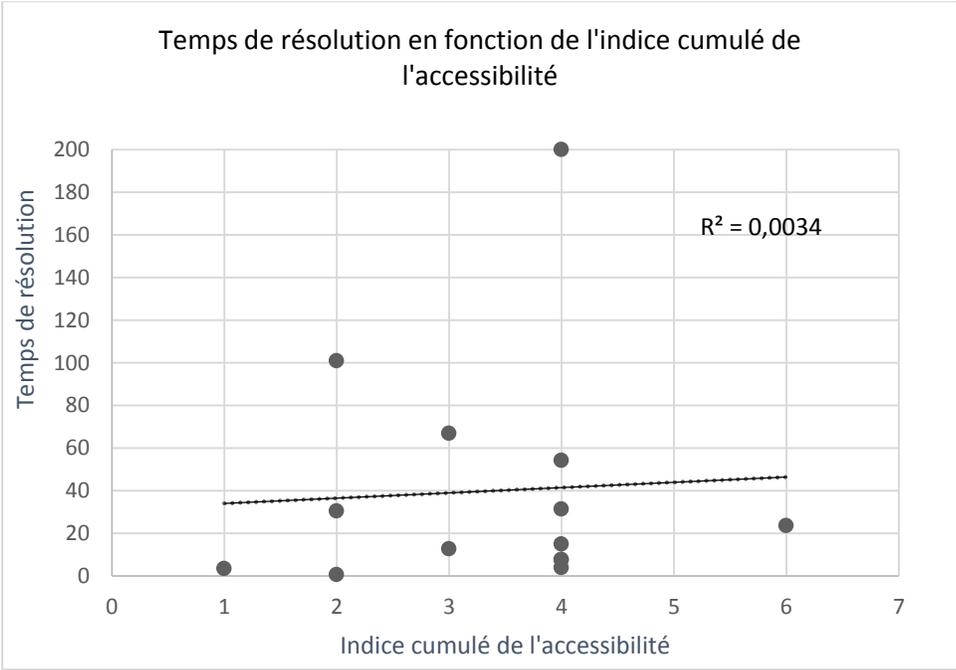
Temps de résolution en fonction de l'indice cumulé des figures

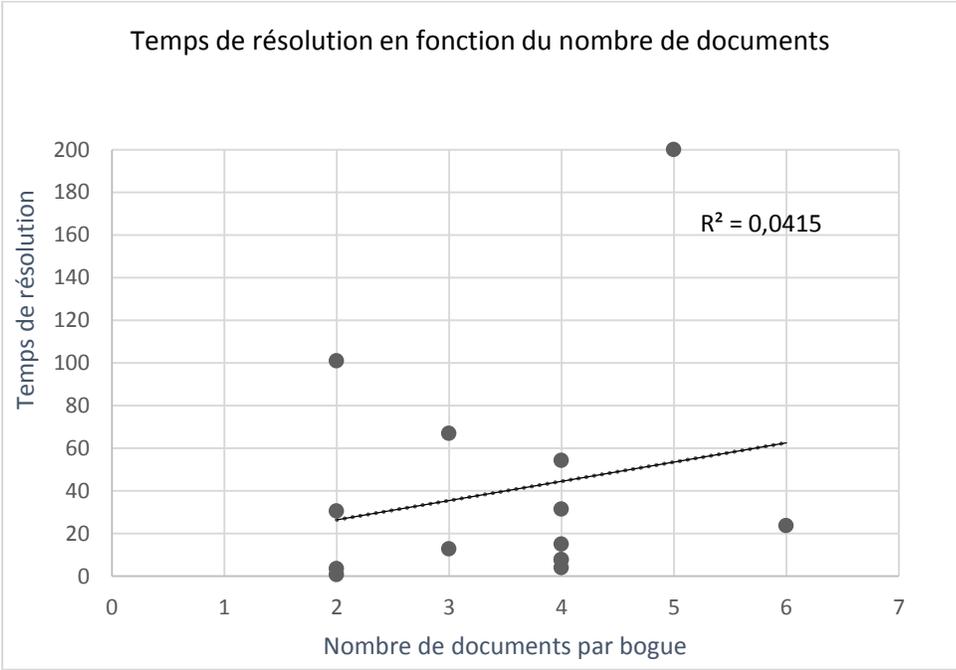
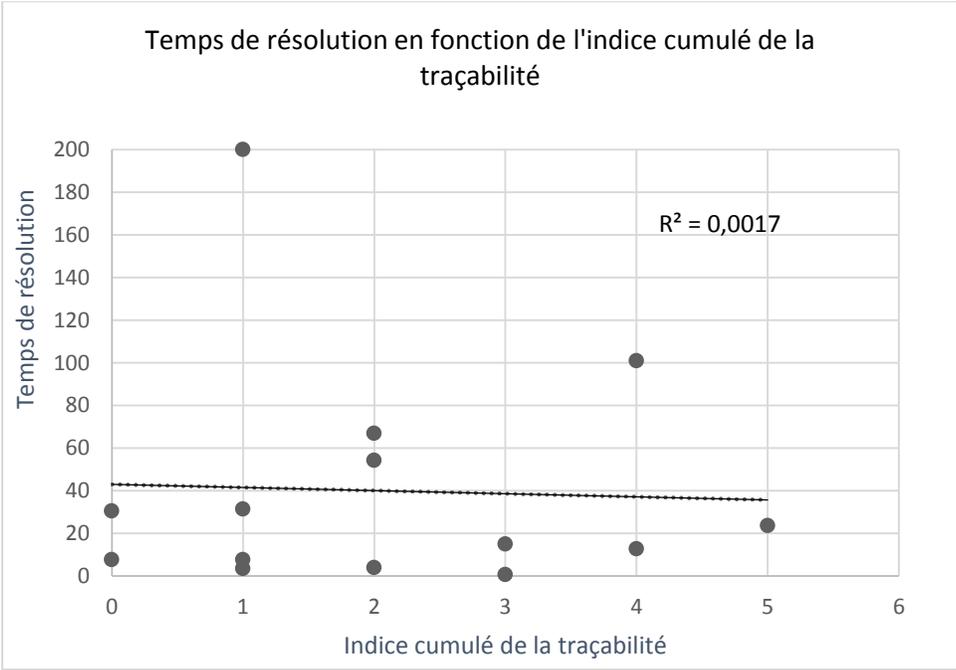


Temps de résolution en fonction de l'indice cumulé de la mise en page

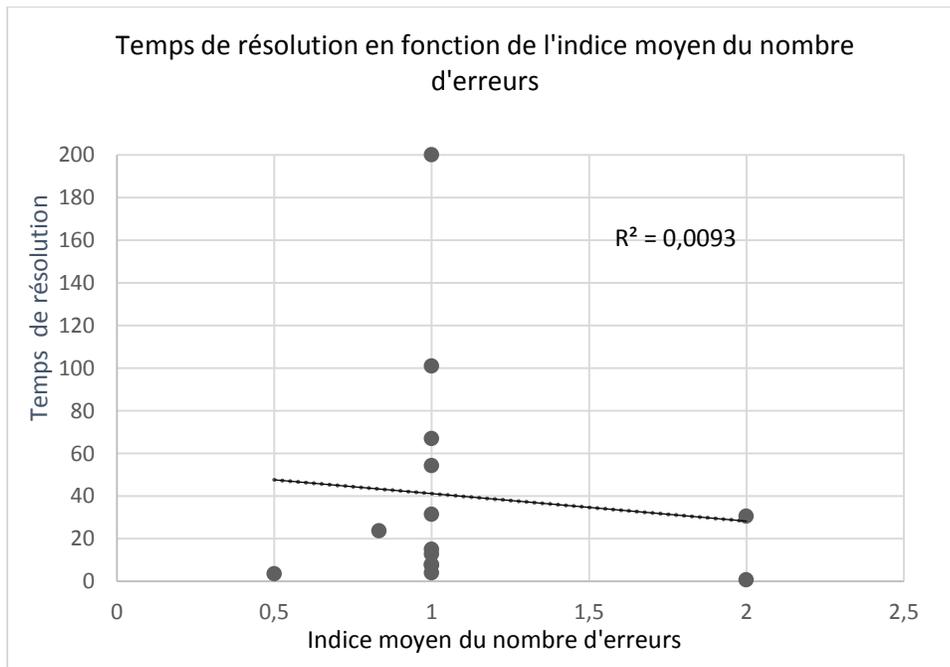
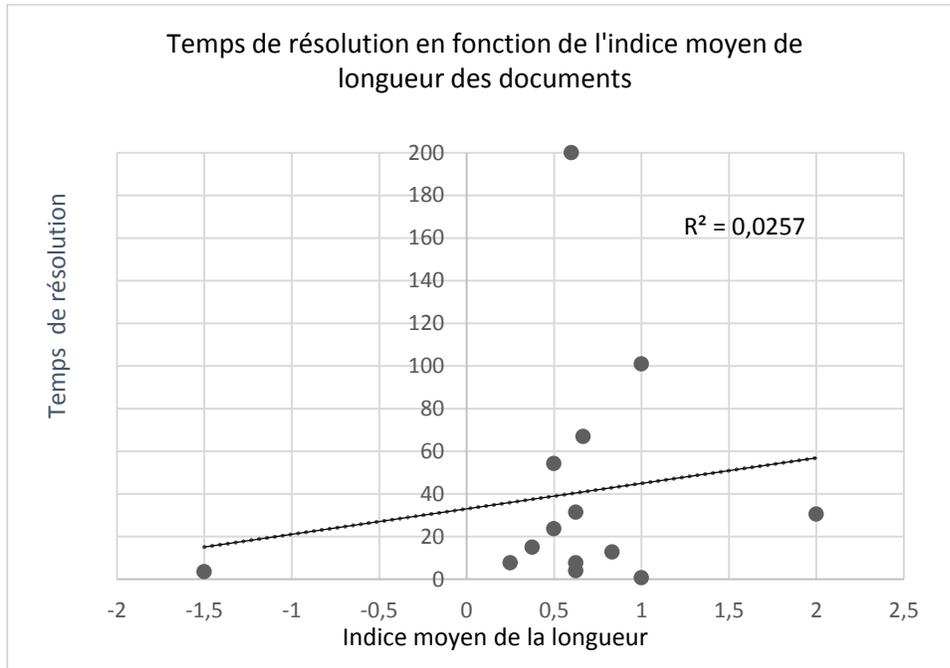


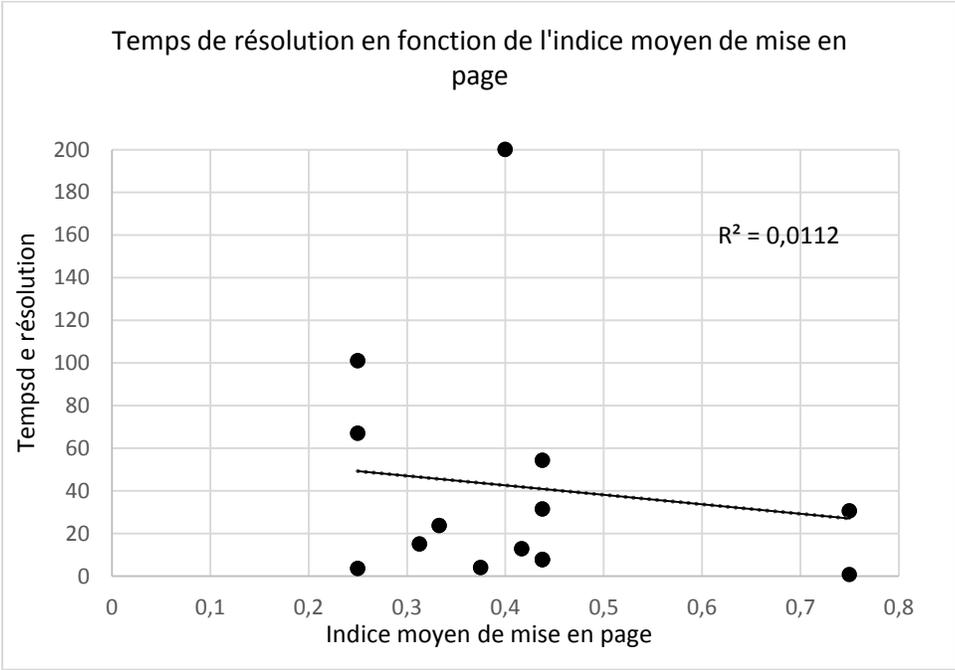
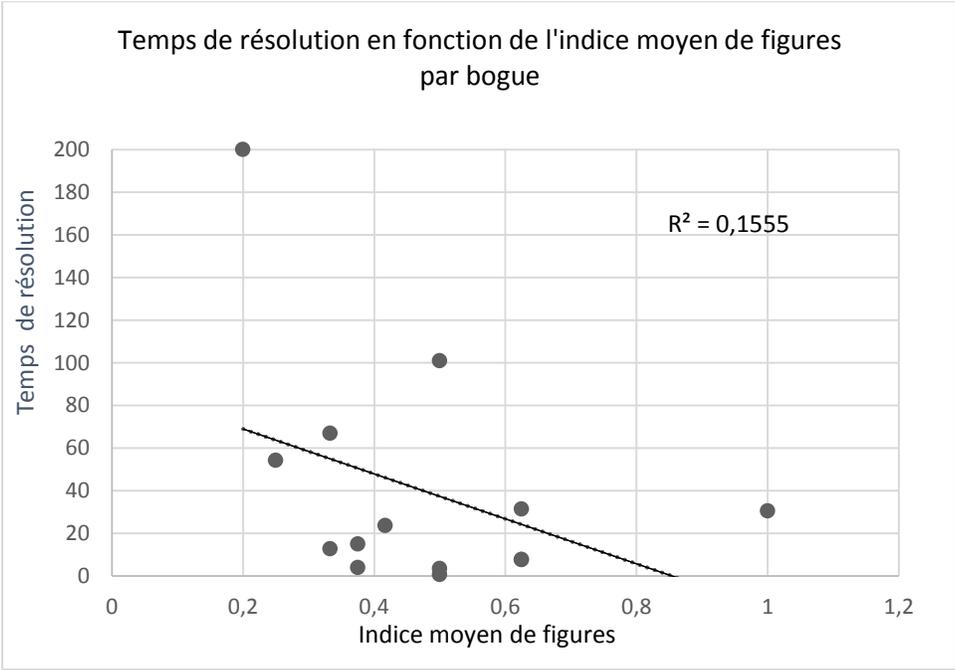




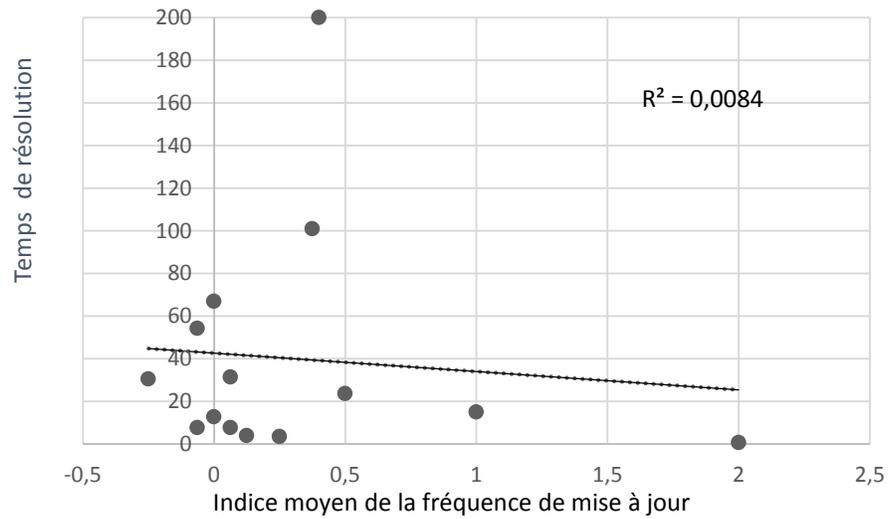


## Graphiques des indices moyens de chaque attribut

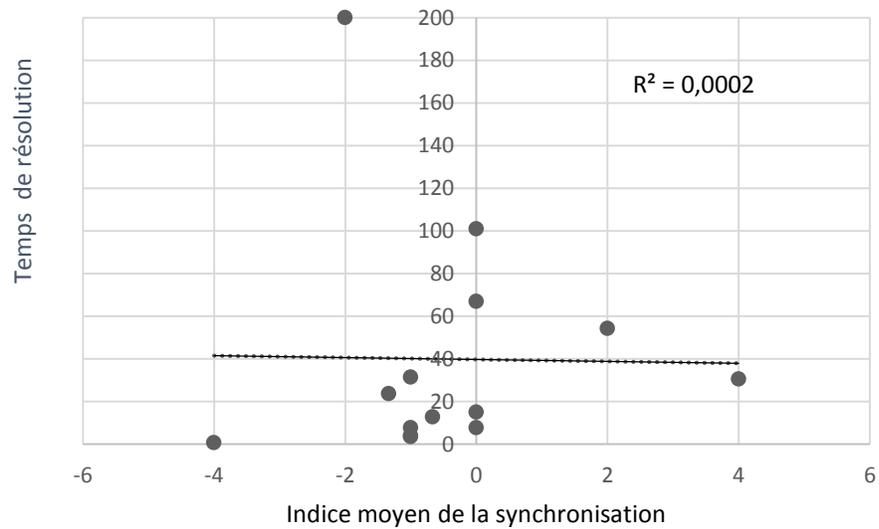


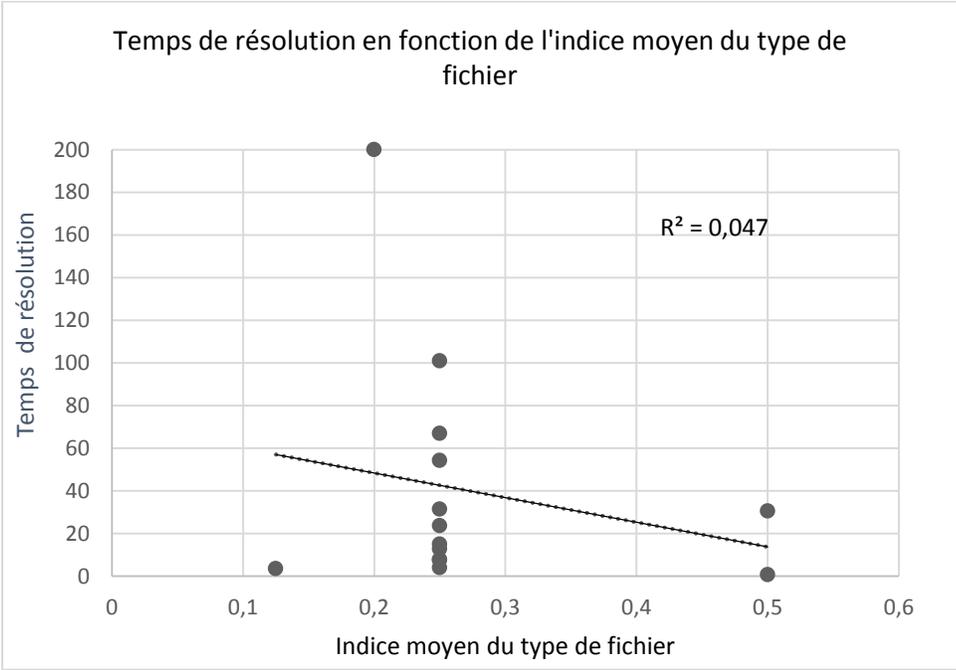
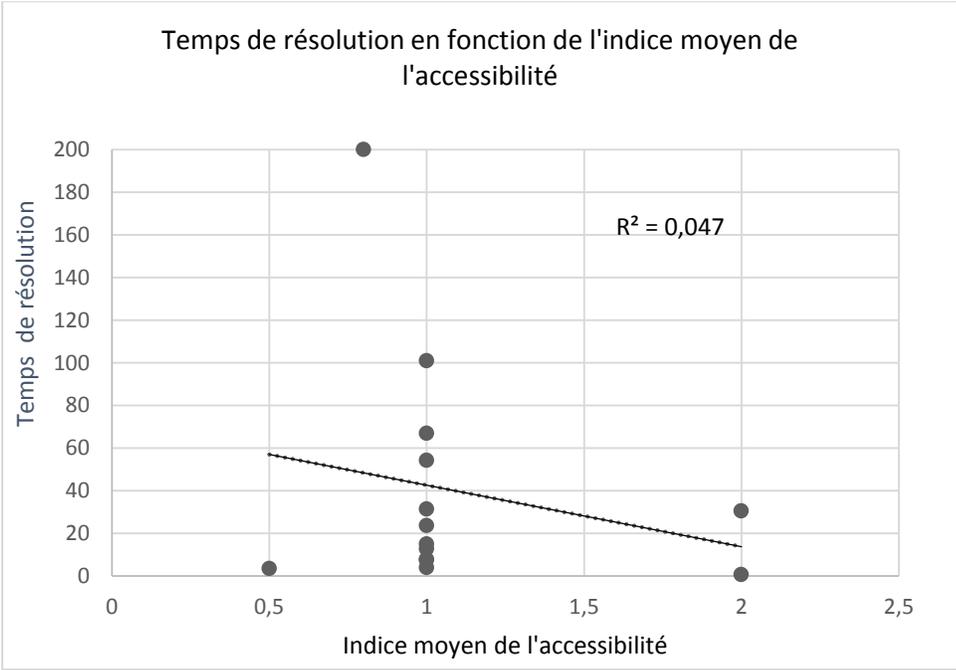


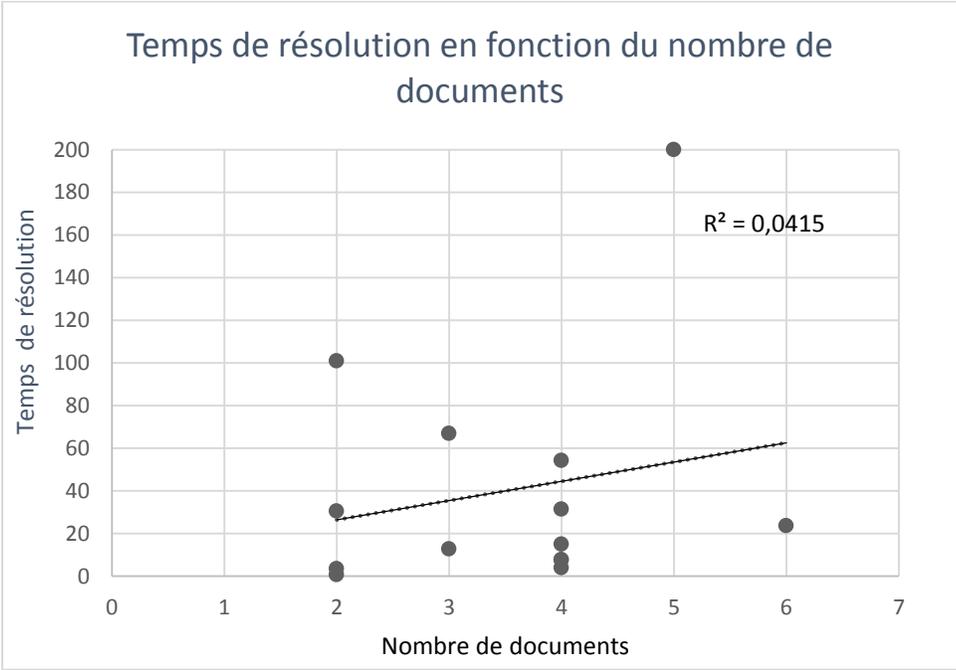
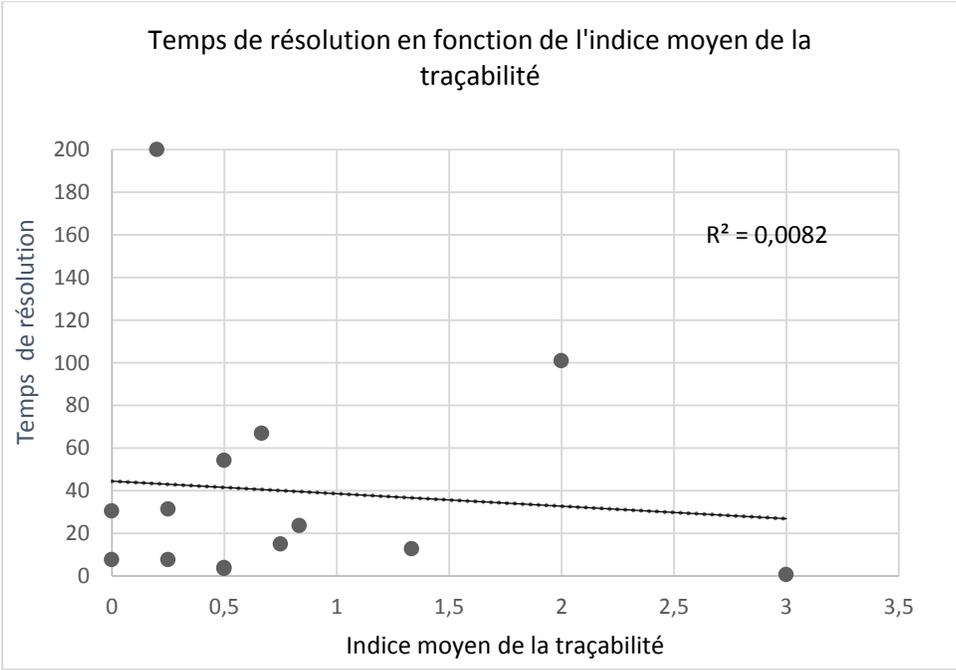
Temps de résolution en fonction de l'indice moyen de la fréquence de mise à jour



Temps de résolution en fonction de l'indice moyen de synchronisation







## Annexe 13

### Directives de validation et application pratique dans la recherche

Tableau des directives de validation à la présente recherche (directives tirées de [40])

Phase	Identifiant	Identifiant et description de la directive	Application dans l'essai
<b>Contexte</b>	C1	Décrire le contexte environnemental le plus précisément possible	Les informations sur le contexte environnemental, l'entreprise hôte de l'étude, les caractéristiques du département TI, l'approche de gestion de projet utilisée sont au sous-chapitre 3.3.
	C2	Énoncer l'hypothèse et ses fondements théoriques avant d'entreprendre l'étude (de façon à ce que l'affirmation de l'hypothèse mène à l'approfondissement de la compréhension au sujet d'une théorie)	Ce critère est appliqué au sous-chapitre 3.1.2.
	C3	Énoncer la question et dire comment elle est abordée	Ce critère est appliqué au sous-chapitre 3.1.2, mais aussi dans le reste du chapitre
	C4	Établir les liens entre l'étude et le corpus actuel des connaissances	Le chapitre 2 couvre amplement ce sujet.
<b>Élaboration</b>	E1	Identifier clairement la population et l'échantillon pour éviter le biais d'échantillonnage	Cet élément est couvert dans 4.1.2.
	E2	Expliquer le processus de	Cet élément est couvert dans

		sélection de l'échantillon	4.1.2
Phase	Identifiant	Identifiant et description de la directive	Application dans l'essai
	E3	Établir comment l'échantillon est divisé pour les expériences différentes	Non applicable car il n'y a pas de division de l'échantillon en groupe pour des expériences différentes.
	E4	Employer des systèmes de modélisations statistiques simples	Emploi de méthodes simples calculées dans Microsoft Excel (régression linéaire simple, coefficients de corrélation $r$ et $r^2$ )
	E5	Bien cerner le groupe expérimental	La complexité du projet ne le requiert pas.
	E6	Pour les recherches plus complexes, employer des méthodes de calcul et d'estimation de la taille de l'échantillon.	Non applicable
	E7	Assurer l'anonymat des groupes ( <i>blinding</i> )	Non applicable car il n'y a pas de groupes
	E8	Identifier et contrer les biais ou conflits d'intérêt apparents surtout lors de l'évaluation de systèmes développés par les auteurs de la recherche	Il n'y a pas de conflit d'intérêt ni de système développé par l'auteur sous une évaluation.
	E9	Éviter d'employer un groupe de contrôle à moins de le définir clairement.	Aucun groupe de contrôle.
	E10	Définir clairement le protocole de recherche	Le chapitre 4 contient tous les détails du protocole ainsi que les tableaux de rapport des données.

<b>Phase</b>	<b>Identifiant</b>	<b>Identifiant et description de la directive</b>	<b>Application dans l'essai</b>
	E11	S'assurer que les métriques choisies satisfont aux usages et aux objectifs de l'étude	Les métriques sont choisies selon les caractéristiques de qualité énumérées par plusieurs chercheurs. La justification des métriques est incluse au chapitre 4.
<b>Collecte de données</b>	CD1	Décrire les métriques explicitement	Les tableaux du sous-chapitre 4.2 expliquent les métriques en détail.
	CD2	Technique spécifique de validation	Non applicable
	CD3	Décrire le modèle d'assurance de la qualité pour valider l'exhaustivité et l'exactitude de la collecte de données.	L'assurance qualité dans la collecte des données comprend la vérification manuelle des documents à deux reprises et de la discussion avec un expert.
	CD4	Validation pour questionnaires	Non applicable
	CD5	Conserver des informations sur les membres de l'échantillon non conservé	Les sections 4.1.2 et 5.1.2 couvrent cet aspect en détail.
	CD6	Prendre note des facteurs autres que ceux ciblés par l'étude et qui peuvent affecter la corrélation	Une discussion des facteurs qui influencent le temps de résolution des bogues est présentée.
<b>Analyse</b>	A1	Clarifier le nombre de tests qui réutilisent les mêmes données sources.	En raison des limites de cette étude et des ressources de l'auteur, les mêmes données sources sont soumises à tous les tests
<b>Phase</b>	<b>Identifiant</b>	<b>Identifiant et description de</b>	<b>Application dans l'essai</b>

		<b>la directive</b>	
	A2	Effectuer une analyse anonyme	Les bogues sont codés selon un identifiant et non selon le logiciel défectueux pour éviter d'identifier le projet auquel c'était jadis associé. De plus, une corrélation préliminaire sur quelques points seulement confirme la taille de l'échantillon requis. Pour le reste de l'analyse, les données (pour les documents et les bogues) sont recueillies indépendamment puis analysées une fois la cueillette complétée.
	A3	Effectuer une analyse des anomalies	Les points aberrants sont considérés dans le chapitre 5 en détail.
	A4	Vérifier que la distribution des données se plie aux exigences des techniques statistiques employées pour les analyser	Seules les techniques statistiques simples sont requises dans cette étude
	A5	Appliquer des mesures de contrôle de la qualité sur les résultats	Un code PL/SQL génère les valeurs de validation des résultats.
<b>Présentation</b>	P1	Décrire ou citer les fonctions statistiques utilisées. Les fonctions statistiques de base sont exemptes de cette directive.	La fonction statistique choisie pour cet essai (l'indice de corrélation linéaire) est exemptée.
	P2	Indiquer le logiciel utilisé pour l'analyse statistique.	Le logiciel décrit est Microsoft Excel.
<b>Phase</b>	<b>Identifiant</b>	<b>Identifiant et description de</b>	<b>Application dans l'essai</b>

		<b>la directive</b>	
	P3	Présenter les données quantitatives ainsi que l'information sur la qualité des données : les intervalles de confiance, les coefficients de corrélation, les ordres de grandeur.	Les données des indices calculés pour chaque attribut ainsi que les résultats sommaires pour chaque bogue sont fournis. Les coefficients de corrélation et les formules sont présentés à titre indicatif.
	P4	Présenter les données brutes	Les données brutes pour les deux variables (les détails des documents et les temps de résolution) sont fournies.
	P5	Fournir des métadonnées sur les données telles que le degré de précision des décimales, la dispersion, les unités d'origine.	Les unités d'origine sont indiquées dans le tableau des données brutes pour les caractéristiques des documents, mais aussi pour les bogues (les étampes de début et de fin des rapports de problème).
<b>Interprétation</b>	I1	Cerner et spécifier la population pour laquelle les modèles s'appliquent.	Expliqué au chapitre 5
	I2	Clarifier la nuance entre une observation statistiquement significative et une observation qui revêt d'une grande importance dans la pratique.	Expliqué au chapitre 5.
	I3	Rappeler le type d'étude	L'étude est de type corrélationnel, et de nature à explorer les relations entre les variables et non pas à identifier une cause et un effet.
	I4	Énumérer les limites de l'étude	Une section est dédiée à cette discussion des limites (section 3.1.3).

## Annexe 14

### Code SQL pour les tables

Fichier dmet\_table\_creation.sql

```
ALTER TABLE nk.dmet DROP PRIMARY KEY CASCADE;
```

```
DROP TABLE nk.dmet CASCADE CONSTRAINTS;
```

```
create table NK.dmet
```

```
(  
  dmet_seqno number(3) NOT NULL,  
  dmet_bpno number(7) "Identifiant du bogue" not null,  
  dmet_dseqno number(3) "Identifiant du document" NOT NULL,  
  dmet_dtitle varchar2(50) "Titre du document", -- doc title varchar  
  dmet_cpages number(3),  
  dmet_cwords number(5),  
  dmet_err number(3),  
  dmet_cvisuals number(3),  
  dmet_layout varchar2(40), -- text layout description varchar  
  dmet_sdlic varchar2(40), -- phase description varchar  
  dmet_updfreq number(3),  
  dmet_updstatus varchar2(4),  
  dmet_access varchar2(15),  
  dmet_filetype varchar2(20),  
  dmet_contact varchar2(10),  
  dmet_activity_date date); --P for people, G for group or M for mixed or 0 for no entry
```

```
alter table nk.dmet add (constraint dmet_pk primary key (dmet_seqno, dmet_bpno, dmet_dseqno));
```

```
COMMENT ON TABLE nk.dmet IS 'Table des donnees brutes issues des documents.';
```

```
COMMENT ON COLUMN nk.dmet.dmet_seqno IS 'Numero de sequence genere automatiquement.';  
COMMENT ON COLUMN nk.dmet.dmet_bpno IS 'Numero identifiant unique du bogue.';  
COMMENT ON COLUMN nk.dmet.dmet_dseqno IS 'Numero identifiant du document appartenant au  
bugue.';  
COMMENT ON COLUMN nk.dmet.dmet_dtitle IS 'Titre du document.';  
COMMENT ON COLUMN nk.dmet.dmet_cpages IS 'Nombre de pages du document.';  
COMMENT ON COLUMN nk.dmet.dmet_cwords IS 'Nombre de mots du document.';  
COMMENT ON COLUMN nk.dmet.dmet_err IS 'Nombre d"erreurs dans le document.';  
COMMENT ON COLUMN nk.dmet.dmet_cvisuals IS 'Nombre d"elements visuels dans le document.';  
COMMENT ON COLUMN nk.dmet.dmet_layout IS 'Description de la mise en page du document.';
```

```

COMMENT ON COLUMN nk.dmet.dmet_sdlic      IS 'Description de la phase du cycle de vie du logiciel
decrete par le document.';
COMMENT ON COLUMN nk.dmet.dmet_updfreq    IS 'Frequence de la mise a jour du document.';
COMMENT ON COLUMN nk.dmet.dmet_updstatus IS 'Etat de la mise a jour du document.';
COMMENT ON COLUMN nk.dmet.dmet_access     IS 'Disponibilite du document et acces.';
COMMENT ON COLUMN nk.dmet.dmet_filetype   IS 'Format du fichier de documentation.';
COMMENT ON COLUMN nk.dmet.dmet_contact    IS 'Personne ou groupe a contacter tel que nomme dans
le document.';

```

```

ALTER TABLE nk.dmet_result DROP PRIMARY KEY CASCADE;
DROP TABLE nk.dmet_result CASCADE CONSTRAINTS;

```

```

create table nk.dmet_result (
dmet_seqno  number(3) not null,
dmet_bpno   number(7) not null,
dmet_dseqno number(3) NOT NULL,
dmet_length_i number(6, 2),
dmet_err_i  number(6,2),
dmet_cvisuals_i number(6,2),
dmet_layout_i number(6,2),
dmet_sdlic_i number(6,2),
dmet_updfreq_i number(6,2),
dmet_updstatus_i number(6, 2),
dmet_access_i number(6,2),
dmet_filetype_i number(6,2),
dmet_contact_i number(6,2),
dmet_metric number(6,2),
dmet_doc_counter number(3),
dmet_activity_date date);

```

```

alter table nk.dmet_result add (constraint dmet_r_pk primary key (dmet_seqno, dmet_bpno,
dmet_dseqno));
alter table nk.dmet_result add ( CONSTRAINT FK1_dmet_key FOREIGN KEY (dmet_seqno, dmet_bpno,
dmet_dseqno) REFERENCES nk.dmet (dmet_seqno,dmet_bpno,dmet_dseqno));

```

```

ALTER TABLE nk.dmet_total DROP PRIMARY KEY CASCADE;
DROP TABLE nk.dmet_total CASCADE CONSTRAINTS;

```

```

create table nk.dmet_total (
dmet_seqno  number(3) not null,
dmet_bpno   varchar2(7) not null primary key,
dmet_doc_count number(3) NOT NULL,
dmet_met_total number(6,2),
dmet_temps  number(6, 2),
dmet_temps_ajuste number(6, 2),
dmet_activity_date date);

```

```

create table nk.dmet_result_index_cumul (
dmet_seqno number(3) not null,
dmet_bpno number(7) not null,
dmet_max_dseqno number(3) NOT NULL,
dmet_length_i_cumul number(6, 2),
dmet_err_i_cumul number(6,2),
dmet_cvisuals_i_cumul number(6,2),
dmet_layout_i_cumul number(6,2),
dmet_updfreq_i_cumul number(6,2),
dmet_updstatus_i_cumul number(6, 2),
dmet_access_i_cumul number(6,2),
dmet_filetype_i_cumul number(6,2),
dmet_contact_i_cumul number(6,2),
dmet_metric_cumul number(6,2),
dmet_activity_date date);

```

Fichier dmet\_load.ctl

Ce fichier prend les données de la table de données séparé par des virgules et les sauvegarde dans la table DMET.

LOAD DATA

```

INFILE 'C:\Users\nk\Desktop\toad\Coppy Finalized scripts into here\Data\dmet_data_final.csv'
BADFILE 'C:\Users\nk\Desktop\toad\Coppy Finalized scripts into here\Data\dmet_data_final.bad'
DISCARDFILE 'C:\Users\nk\Desktop\toad\Coppy Finalized scripts into here\Data\dmet_data_final.dsc'

```

```

INTO TABLE "NK"."DMET"
TRUNCATE
FIELDS TERMINATED BY ','
(DMET_SEQNO,
DMET_BPNO,
DMET_DSEQNO,
DMET_DTITLE,
DMET_CPAGES,
DMET_CWORDS,
DMET_ERR,
DMET_CVISUALS,
DMET_LAYOUT,
DMET_SDLCL,

```

```

DMET_UPDFREQ,
DMET_UPDSTATUS,
DMET_ACCESS,
DMET_FILETYPE,
DMET_CONTACT,
DMET_ACTIVITY_DATE SYSDATE)

```

Fichier metric\_logic\_script\_final.sql

Ce code lit les données de chaque attribut de chaque document de la table DMET, produit les données des indices individuels et les sauvegarde dans la table DMET\_RESULT. Ensuite, le code calcule l'indice de documentation sommaire pour chaque bogue et sauvegarde les données dans la table DMET\_TOTAL.

```

/* Formatted on 5/9/2016 5:43:02 PM (QP5 v5.139.911.3011) */
TRUNCATE TABLE nk.dmet_result;
TRUNCATE TABLE nk.dmet_total;
set serveroutput on size 1000000;

```

```

DECLARE
CURSOR alldata_c
IS
    SELECT *
    FROM nk.dmet
    ORDER BY 1 ASC, 3 ASC;

document_rec      alldata_c%ROWTYPE;

document_counter  INTEGER (4);
dm_doc            NUMBER (6) := 0;
dmet_length_i_var NUMBER (6,2);
dmet_err_i_var    NUMBER (6,2);
dmet_cvisuals_i_var NUMBER (6,2);
dmet_layout_i_var NUMBER (6,2);
dmet_sdlic_i_var  NUMBER (6,2);
dmet_updfreq_i_var NUMBER (6,2);
dmet_updstatus_i_var NUMBER (6,2);
dmet_access_i_var NUMBER (6,2);
dmet_filetype_i_var NUMBER (6,2);
dmet_contact_i_var NUMBER (6,2);
dmet_metric_i_var NUMBER (6,2);
dmet_doc_counter  NUMBER (3);

```

*-- déclarations pour la deuxième phase : les totaux*

```

CURSOR tot_alldata_c
IS

```

```

SELECT *
FROM nk.dmet_result
ORDER BY 1 ASC, 3 ASC;

tot_rec          tot_alldata_c%ROWTYPE;
tot_counter      INTEGER (3) := 0;
t_bpno           NUMBER (5);
t_dseqno         NUMBER (5);
t_doc_count      NUMBER (3);
t_met_total      NUMBER (6,2);
BEGIN
  DBMS_OUTPUT.put_line ('Starting Program...');
  -- small index variables that will hold the computer index for that record.
  -- then the index will be stored directly in the table, then the loop will restart, reset the indexes and
  calculate new ones
  document_counter := 0;

  OPEN alldata_c;
  LOOP
    FETCH alldata_c INTO document_rec;
    EXIT WHEN alldata_c%NOTFOUND;
    document_counter := document_counter + 1;

    dm_doc := 0;
    dmet_length_i_var := 0;
    dmet_err_i_var := 0;
    dmet_cvisuals_i_var := 0;
    dmet_layout_i_var := 0;
    dmet_updfreq_i_var := 0;
    dmet_updstatus_i_var := 0;
    dmet_access_i_var := 0;
    dmet_filetype_i_var := 0;
    dmet_contact_i_var := 0;
    dmet_metric_i_var := 0;

    DBMS_OUTPUT.put_line ('First entry BP=' || document_rec.dmet_bpno
      || ' and Document SEQNO=' || document_rec.dmet_seqno);

    -- Set the primary key (dmet_seqno, dmet_bpno, dmet_dseqno), use the ones from the input table
    -- document_rec.dmet_seqno, document_rec.dmet_bpno, document_rec.dmet_dseqno
    -- Add logic: if length = 0 then document is missing

    -- Doc Length
    IF (document_rec.dmet_cpages >= 20 OR (document_rec.dmet_cpages = 1 AND
document_rec.dmet_cwords < 300))
    THEN
      dmet_length_i_var := 0;
    ELSIF ( (document_rec.dmet_cpages BETWEEN 10 AND 20)
      OR (document_rec.dmet_cpages = 1

```

```

        AND document_rec.dmet_cwords > 300))
    OR ( (document_rec.dmet_cpages BETWEEN 2 AND 3)
        AND (document_rec.dmet_cwords < 300))
THEN
    dmet_length_i_var := 0.5;
ELSIF (document_rec.dmet_cpages BETWEEN 2 AND 9)
THEN
    dmet_length_i_var := 1.0;
ELSIF (document_rec.dmet_cpages = 0) THEN
    dmet_length_i_var := -4;
END IF;

DBMS_OUTPUT.PUT_LINE ('Length Index = ' || dmet_length_i_var);

-- Erreurs
IF (document_rec.dmet_err < 10)
THEN
    dmet_err_i_var := 1;
ELSIF ( (document_rec.dmet_err < 8) AND (document_rec.dmet_err > 4))
THEN
    dmet_err_i_var := 0.5;
ELSE
    dmet_err_i_var := 0;
END IF;

-- Éléments visuels
IF (document_rec.dmet_cvisuals = 0)
THEN
    dmet_cvisuals_i_var := 0;
ELSIF (document_rec.dmet_cvisuals BETWEEN 1 AND 2)
THEN
    dmet_cvisuals_i_var := 0.5;
ELSIF (document_rec.dmet_cvisuals >= 3)
THEN
    dmet_cvisuals_i_var := 1;
END IF;

-- Mise en page, layout
IF (document_rec.dmet_layout IN ('D', 'Z'))
THEN
    dmet_layout_i_var := 0;
ELSIF (document_rec.dmet_layout IN ('O'))
THEN
    dmet_layout_i_var := 0.25;
ELSIF (document_rec.dmet_layout IN ('G', 'T'))
THEN
    dmet_layout_i_var := 0.5;
END IF;

```

```

-- SDLC Rien à mettre ici, ignorer
-- Nombre de mises à jour
IF (document_rec.dmet_updfreq = 0)
THEN
    dmet_updfreq_i_var := -0.5;
ELSIF (document_rec.dmet_updfreq BETWEEN 1 AND 3)
THEN
    dmet_updfreq_i_var := 0.25;
ELSIF (document_rec.dmet_updfreq BETWEEN 4 AND 6)
THEN
    dmet_updfreq_i_var := 0.5;
ELSIF (document_rec.dmet_updfreq >= 7)
THEN
    dmet_updfreq_i_var := 1;
END IF;

-- État de la mise à jour
IF lower(document_rec.dmet_updstatus) IN ('yes', 'y')
THEN
    dmet_updstatus_i_var := 2;
ELSIF lower(document_rec.dmet_updstatus) IN ('no', 'n')
THEN
    dmet_updstatus_i_var := -2;
END IF;

-- Accès
IF LOWER (document_rec.dmet_access) LIKE 'n%'
THEN
    dmet_access_i_var := 0;
ELSIF LOWER (document_rec.dmet_access) LIKE 'd%'
THEN
    dmet_access_i_var := 0.5;
ELSIF LOWER (document_rec.dmet_access) LIKE 'y%'
THEN
    dmet_access_i_var := 1;
END IF;

-- Type de fichier
IF LOWER (document_rec.dmet_filetype) IN ('.ini', '.sqr', '.pl', 'sql')
THEN
    dmet_filetype_i_var := -0.25;
ELSIF LOWER (document_rec.dmet_filetype) IN
    ('.doc', '.docx', '.txt', 'csv', 'xls', '.pdf')
THEN
    dmet_filetype_i_var := 0.25;
END IF;

-- Personne-ressource
IF LOWER(document_rec.dmet_contact) = 'p'

```

```

THEN
  dmet_contact_i_var := 1;
ELSIF LOWER(document_rec.dmet_contact) IN ('m', 'g')
THEN
  dmet_contact_i_var := 2;
ELSIF lower(document_rec.dmet_contact) LIKE ('n')
THEN
  dmet_contact_i_var := 0;
END IF;

-- Cas special: un document est manquant. Reinitialiser les indices. Ceci est une etape de validation
IF (document_rec.dmet_cpages = 0 AND document_rec.dmet_cwords = 0) THEN
  dmet_length_i_var := -4; -- longueur = 0 alors sont indice est 0 aussi
  dmet_err_i_var := 0; -- il ne contient pas d'erreurs. la logique lui donne une cote positive,
  mais il faut éviter cela.
  dmet_cvisuals_i_var := 0; -- doc vide
  dmet_layout_i_var := 0; -- doc vide
  dmet_updfreq_i_var := 0; --aucune mise a jour
  dmet_updstatus_i_var := 0;
  dmet_access_i_var := 0;
  dmet_filetype_i_var := 0;
  dmet_contact_i_var := 0;
END IF;

-- Calculer la metrique totale du document
dmet_metric_i_var :=
  dmet_length_i_var
+ dmet_err_i_var
+ dmet_cvisuals_i_var
+ dmet_layout_i_var
+ dmet_updfreq_i_var
+ dmet_updstatus_i_var
+ dmet_access_i_var
+ dmet_filetype_i_var
+ dmet_contact_i_var;

DBMS_OUTPUT.PUT_LINE ('Inserting this record into DMET_RESULT. ');
INSERT INTO dmet_result (dmet_seqno,
  dmet_bpno,
  dmet_dseqno,
  dmet_length_i,
  dmet_err_i,
  dmet_cvisuals_i,
  dmet_layout_i,
  dmet_sdlc_i,
  dmet_updfreq_i,
  dmet_updstatus_i,
  dmet_access_i,
  dmet_filetype_i,
  dmet_contact_i,

```

```

        dmet_metric,
        dmet_doc_counter,
        dmet_activity_date)
VALUES (document_rec.dmet_seqno,
        document_rec.dmet_bpno,
        document_rec.dmet_dseqno,
        dmet_length_i_var,
        dmet_err_i_var,
        dmet_cvisuals_i_var,
        dmet_layout_i_var,
        NULL,
        TO_CHAR (dmet_updfreq_i_var),
        dmet_updstatus_i_var,
        dmet_access_i_var,
        dmet_filetype_i_var,
        dmet_contact_i_var,
        dmet_metric_i_var,
        document_counter,
        SYSDATE);
END LOOP;
COMMIT;

/* PARTIE 2 DU CODE : CALCULER L'INDICE DE DOCUMENTATION AU NIVEAU DU BOGUE */
-- Commencer à agréger les données
OPEN tot_alldata_c;
LOOP
    FETCH tot_alldata_c INTO tot_rec;
    EXIT WHEN tot_alldata_c%NOTFOUND;
    DBMS_OUTPUT.PUT_LINE ('tot_counter = ' || tot_counter);
    DBMS_OUTPUT.PUT_LINE ('processing this t_bpno = ' || t_bpno);
    DBMS_OUTPUT.PUT_LINE ('document counter is ===== ' || t_doc_count);

    IF (tot_counter = 1)
    THEN
        t_bpno := tot_rec.dmet_bpno;
        t_doc_count := tot_rec.dmet_dseqno;
        t_met_total := tot_rec.dmet_metric;
    ELSE
        IF t_bpno = tot_rec.dmet_bpno
        THEN
            -- on continue avec le même bogue
            DBMS_OUTPUT.PUT_LINE ('same bug. ');
            t_doc_count := tot_rec.dmet_dseqno;
            t_met_total := t_met_total + tot_rec.dmet_metric;
        ELSIF t_bpno <> tot_rec.dmet_bpno
        THEN
            -- on change de bogue
            DBMS_OUTPUT.PUT_LINE ('change bug. ');
            INSERT INTO dmet_total (dmet_seqno,
                                dmet_bpno,
                                dmet_doc_count,

```

```

        dmet_met_total,
        dmet_temps,
        dmet_temps_ajuste,
        dmet_activity_date)
VALUES (tot_counter,
        t_bpno,
        t_doc_count,
        t_met_total,
        (select dmet_temps from temps_res where temps_res.dmet_bpno like t_bpno),
        (select dmet_temps_ajuste from temps_res where temps_res.dmet_bpno like t_bpno),
        SYSDATE);

DBMS_OUTPUT.PUT_LINE ('Inserting this record into DMET_RESULT.' || t_bpno);
t_bpno := tot_rec.dmet_bpno;
t_doc_count := tot_rec.dmet_dseqno;
t_met_total := tot_rec.dmet_metric;
END IF;
END IF;
tot_counter := tot_counter + 1;
END LOOP;

INSERT INTO dmet_total (dmet_seqno,
        dmet_bpno,
        dmet_doc_count,
        dmet_met_total,
        dmet_temps,
        dmet_temps_ajuste,
        dmet_activity_date)
VALUES (tot_counter,
        t_bpno,
        t_doc_count,
        t_met_total,
        (select dmet_temps from temps_res where temps_res.dmet_bpno like t_bpno),
        (select dmet_temps_ajuste from temps_res where temps_res.dmet_bpno like t_bpno),
        SYSDATE);
DBMS_OUTPUT.PUT_LINE ('Inserting this record into DMET_RESULT.' || t_bpno);
COMMIT;

END;
```

Fichier pour indices moyens et cumulés

```

drop table dmet_indice_cumule;

create table nk.dmet_indice_cumule as select * from (
select dmet_bpno,
sum(dmet_length_i) "Longueur indice cumule",
sum(dmet_err_i) "Erreurs indice cumule",
```

```

sum(dmet_cvisuals_i) "Figures indice cumule",
sum(dmet_layout_i) "Mise en page",
sum(dmet_updfreq_i),
sum(dmet_updstatus_i),
sum(dmet_access_i),
sum(dmet_filetype_i),
sum(dmet_contact_i),
max(dmet_dseqno) "DOC_NO"
from dmet_result
group by dmet_bpno
order by 1 asc
);
commit;

```

```
drop table nk.dmet_indice_moyen;
```

```

create table nk.dmet_indice_moyen as select * from (
select dmet_bpno,
sum(dmet_length_i)/max(dmet_dseqno) "Longeur indice moyen" ,
sum(dmet_err_i)/max(dmet_dseqno) "Erreurs",
sum(dmet_cvisuals_i)/max(dmet_dseqno) "Figures" ,
sum(dmet_layout_i)/max(dmet_dseqno) "Mise en page",
sum(dmet_updfreq_i)/max(dmet_dseqno) "Frequence de mise a jour",
sum(dmet_updstatus_i)/max(dmet_dseqno) "Synchronisation",
sum(dmet_access_i)/max(dmet_dseqno) "Accessibilite",
sum(dmet_filetype_i)/max(dmet_dseqno) "Format du document",
sum(dmet_contact_i)/max(dmet_dseqno) "Tracabilite"
from dmet_result
group by dmet_bpno
order by 1 asc
);
commit;

```

Autres requêtes

```

select
dmet_bpno "Identifiant du bogue",
--dmet_doc_count "Nombre de documents",
dmet_met_total "Indicateur total",
--dmet_temps "Temps de résolution"
dmet_temps_ajuste "Temps de résolution ajusté"
from dmet_total
order by 1 asc

```

```

select
dmet_seqno "Compteur de bogue",
dmet_bpno "Identifiant du bogue",
dmet_bpno || '-' || dmet_dseqno "Identifiant du document",

```

```

dmet_cpages "Nombre de pages",
dmet_cwords "Nombre de mots",
dmet_err "Nombre d'erreurs",
dmet_cvisuals "Nombre de figures",
dmet_layout "Mise en page",
dmet_SDLC "Type de document",
dmet_updfreq "Nombre de mises à jour",
dmet_updstatus "Synchronisation",
dmet_access "Accessibilité",
dmet_filetype "Type de fichier",
dmet_contact "Traçabilité"
from dmet order by dmet_seqno asc, dmet_bpno asc

```

```

select
dmet_seqno "Compteur de bogue",
dmet_bpno "Identifiant du bogue",
dmet_bpno || '-' || dmet_dseqno "Identifiant du document",
dmet_length_i "Longueur",
dmet_err_i "Erreurs",
dmet_cvisuals_i "Figures",
dmet_layout_i "Mise en page",
dmet_updfreq_i "Nombre de mises à jour",
dmet_updstatus_i "Synchronisation",
dmet_access_i "Accessibilité",
dmet_filetype_i "Type de fichier",
dmet_contact_i "Traçabilité",
dmet_metric "Indice du document"
from dmet_result order by dmet_seqno asc, dmet_bpno asc

```

```

select * from dmet_result

```

```

select *
from dmet_indice_moyen, temps_res where DMET_INDICE_moyen.DMET_BPNO =
TEMPS_RES.DMET_BPNO

```

/

