



# **La santé des écosystèmes logiciels ouverts : Une enquête qualitative auprès d'experts**

*Recherche en cours*

*Patrick Marois\**

*Georgia Leïda Mopenza\**

*Josianne Marsan\**

*Mathieu Templier\**

\* Université Laval, Québec (Qc), Canada

## **Résumé :**

Les logiciels libres à code source ouvert (*free/open source*) sont de plus en plus adoptés et deviennent indispensables au fonctionnement de la société numérique. Un écosystème logiciel ouvert, dont la mission est de développer et maintenir un ensemble de logiciels libres à code source ouvert, regroupe non seulement des projets interdépendants techniquement, mais également des échanges sociaux parmi les communautés dans ces projets. Ces écosystèmes peuvent devoir affronter des problèmes de santé qui touchent les projets ou les communautés qui les composent. La place grandissante des écosystèmes ouverts et la complexité accrue au niveau de leur gestion nécessitent de mieux comprendre leur santé et de développer des outils pour sa gestion. En ce sens, la présente recherche en cours s'intéresse aux problèmes de santé de ces écosystèmes, aux causes et conséquences de ces problèmes, ainsi qu'à la propagation de ces problèmes. La métaphore de l'organisation en tant qu'organisme vivant est mobilisée dans une enquête qualitative auprès d'experts des écosystèmes logiciels ouverts. Nous espérons que cette étude contribuera à la fois à la recherche, en développant de nouvelles connaissances sur la santé des écosystèmes logiciels, et à la pratique, en fournissant un catalogue de problèmes de santé, de leurs causes et conséquences, et de leurs routes de propagation, utile à la gestion des écosystèmes.

## **Mots clés :**

Écosystème logiciel, Logiciel libre à code source ouvert, Santé des écosystèmes logiciels, Analyse causale, Enquête qualitative

## 1. Introduction

Les logiciels libres à code source ouvert (LLCSO) ou logiciels free/open source sont des logiciels dont le code source est accessible et modifiable, et dont la licence assure que ce code demeure dans le domaine public (Marsan & Paré, 2013). Ces caractéristiques ont amené la formation de communautés de contributeurs dans les projets de développement de LLCSO. L'adoption accrue des LLCSO amène les organisations utilisatrices à s'insérer dans les communautés open source pour influencer l'évolution des LLCSO (Fitzgerald, 2006; Carillo & Marsan, 2016). Dans le but de conserver leur caractère ouvert et indépendant, des fondations telles que la *Linux Foundation* ont été créées autour des LLCSO les plus populaires. À l'origine, leur mission était de gérer les relations entre communautés et partenaires externes (de Laat, 2007). Cette mission a évolué vers le support à l'émergence et au maintien d'écosystèmes logiciels ouverts (ECLO) pour accélérer le développement technologique et l'adoption des LLCSO (Linux Foundation, 2017). Un ECLO regroupe des projets interdépendants techniquement, mais aussi des échanges sociaux au sein et entre les communautés qui le composent (Manikas & Hansen 2013)<sup>1</sup>. De plus, il existe une interdépendance entre la santé des ECLO et le succès des projets de LLCSO, ces derniers nécessitant que leurs communautés soient actives et impliquées pour la réussite et la pérennité de ces logiciels (Jansen, 2014). La santé d'un organisme vivant se définit comme un état de fonctionnement normal et continu de ses systèmes vitaux (Wang & Lantzy, 2011). Nous définissons ainsi la santé d'un ECLO comme un état de fonctionnement normal et continu des projets constituant l'ECLO. Considérant les interdépendances entre projets au sein des ECLO, la propagation des problèmes de santé est probable et peut s'effectuer à différents niveaux et via différents vecteurs.

Les ECLO se forment, se complexifient et deviennent indispensables au bon fonctionnement de la société numérique (Manikas & Hansen 2013). L'omniprésence des ECLO et la complexité accrue au niveau de leur gestion nécessitent de mieux comprendre la santé des ECLO et de développer des outils pour sa gestion (Kazman & Chen 2010). L'absence de tels outils pose plusieurs problématiques aux gestionnaires des ECLO, qui n'ont alors qu'une vision partielle des problèmes de santé et ne peuvent y apporter les actions correctives appropriées (Kazman & Chen 2010). En ce sens, l'objectif général de cette recherche est de mieux comprendre les problèmes de santé des ECLO. Nous chercherons à répondre aux questions de recherche suivantes : 1) Quels sont les problèmes de santé des ECLO ? 2) Quels sont les causes et les conséquences de ces problèmes ? 3) Quelles sont les routes de propagation des problèmes de santé dans un ECLO ? Nous présentons ici les analyses et résultats préliminaires qui pourraient permettre le développement de politiques et d'outils de gestion.

## 2. Fondements conceptuels

Morgan (2006) décrit l'organisation comme un organisme vivant composé de sous-systèmes interdépendants desquels dépend aussi la santé globale de l'organisme. Carillo et Marsan (2016) reprennent cette métaphore pour étudier les comportements potentiellement toxiques dans les projets et qui peuvent en menacer la santé. Ils proposent de décomposer un projet en cinq sous-systèmes interdépendants (voir l'encadré du bas dans la figure 1). Dans chacun des sous-systèmes, il peut exister des problématiques qui peuvent affecter la santé du projet. En

---

<sup>1</sup> Le projet et la communauté qui l'entoure étant inséparables et fortement entrelacés (Mens, 2016), nous utiliserons simplement le terme projet dans le reste de ce document.

toxicologie, la "distribution" est le processus par lequel une substance toxique voyage à travers un organisme vivant (Gregus, 2013). Carillo et Marsan (2016) ont transposé ce processus au contexte de la santé des projets. Nous proposons de le transposer à la santé des ECLO, suggérant ainsi que les problèmes peuvent se propager entre projets formant un ECLO. Il existe des interdépendances entre les projets au niveau d'un écosystème (Carillo et al., 2017). Un organisme existe dans un environnement externe avec lequel il effectue des échanges. De plus, les organismes dans un même environnement interagissent, peuvent collaborer, et aussi être en compétition entre eux. Le fonctionnement d'un projet dépend donc de l'interaction entre ses composantes et avec les autres projets dans son environnement externe. Des routes de propagation des problèmes de santé dans les ECLO pourraient ainsi exister et être identifiées.

L'analyse causale a pour objectifs de définir les causes et d'évaluer les conséquences des problèmes observés (Rivard, 2013). Un problème est défini comme un écart entre une situation espérée et une situation actuelle (Rivard, 2013), par exemple un nombre de contributeurs actifs plus bas qu'espéré qui empêcherait un projet open source de fonctionner normalement et en continu (Crowston et al., 2006). Une cause est une chose suivie d'une autre, et cette autre chose ne pourrait pas exister si la première chose n'était pas arrivée (Hume, 1975). Dans notre exemple, des comportements déviants ou agressifs, notamment sous la forme de propos inflammatoires (*flaming*), peuvent être la cause du faible nombre de contributeurs (Carillo & Marsan 2016). Une conséquence est la perte encourue lorsqu'une menace, ici un problème, se concrétise (Alberts & Dorofee, 2009). En effet, un problème peut être latent, c.-à-d. que l'écart n'a pas encore été identifié, et même disparaître avant d'avoir des conséquences. Dans notre exemple, la conséquence du faible nombre de contributeurs pourrait être la disparition de projets dépendants (Williams, 2016). La figure 1 présente le cadre conceptuel de notre étude ainsi que les propositions relatives à chacune de nos questions de recherche.

### **3. Méthodologie proposée**

Afin de répondre à nos questions de recherche, nous réaliserons une enquête qualitative auprès d'experts des ECLO (Meyer & Booker, 2001). Les experts seront sélectionnés sur la base de critères d'inclusion. Ils devront (1) être contributeur récent ou actif dans un ECLO, (2) être gestionnaire récent ou actif dans une fondation open source, (3) être impliqué dans le développement de logiciels de gestion de projets open source ou d'ECLO, ou (4) avoir été conférencier invité dans des événements portant sur les ECLO. L'objectif de l'échantillonnage est de couvrir une variété d'expertises à l'égard des concepts qui nous intéressent (Meyer & Booker, 2001). Des entrevues individuelles semi-dirigées seront réalisées (Miles & Huberman, 1994) à l'aide d'un guide d'entrevue élaboré à partir de notre cadre conceptuel et validé par un prétest auprès de deux personnes répondant aux critères d'inclusion. Les entrevues seront enregistrées, transcrites puis codifiées avec le logiciel NVivo. La codification des entrevues sera validée sur un échantillon de 15% des entrevues, sélectionnées au hasard, et l'accord inter-codeurs sera calculé. L'analyse des extraits permettra de construire une réponse de groupe basée sur les réponses individuelles des différents experts (Meyer & Booker, 2001), sous la forme de diagrammes cause-effet et de propagation qui permettront de valider nos propositions.

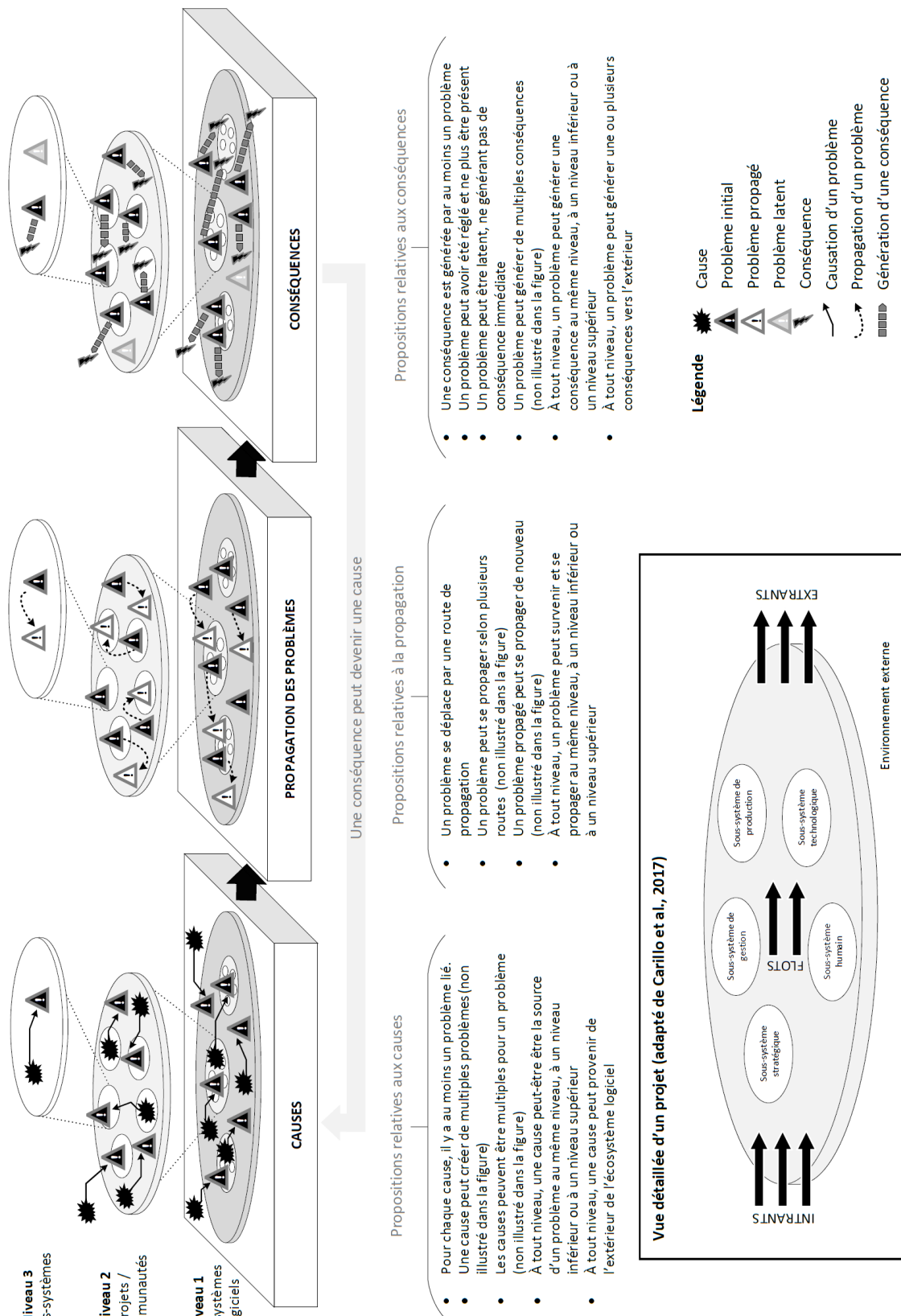


Figure 1 : Cadre conceptuel préliminaire

Une première phase de recrutement d'experts et de collecte de données a été réalisée en octobre 2017 durant l'*Open Source Summit – Europe* par les trois premiers auteurs. Jusqu'ici, 18 entrevues ont été réalisées, dont 17 sur place lors de l'événement et une par téléphone ensuite. Nous prévoyons faire 17 entrevues de plus, à distance, avec d'autres experts recrutés lors de l'événement. Les entrevues ont été retranscrites et révisées par au moins deux membres de l'équipe de recherche. Dans le cadre de cette recherche en cours, 5 des 18 entrevues réalisées sont considérées.

#### 4. Analyse et résultats préliminaires

Les participants aux 18 entrevues réalisées jusqu'ici proviennent de 12 pays différents. La durée moyenne d'une entrevue est de 43 minutes. De ces 18 entrevues, 5 ont été retranscrites et analysées. Basé sur cette analyse préliminaire, le tableau 1 présente des problèmes de santé des ECLO ainsi qu'une cause racine, des causes intermédiaires et une route de propagation pour chacun.

Causes racines	Causes intermédiaires	Problèmes	Propagation
Manque de temps des contributeurs	<ul style="list-style-type: none"> <li>⌘ Tâches demandant du temps</li> <li>⌘ Faible capacité en jour/personne</li> <li>⌘ Diminution de l'allocation de ressources</li> <li>⌘ Absence de refactoring dans un projet</li> <li>⌘ Complexité de maintenir un projet</li> </ul>	<b>Incompatibilité descendante au niveau d'une composante logicielle</b>	Incompatibilité d'une composante logicielle rendant des fonctionnalités inutilisables dans les sous-systèmes des autres projets de l'écosystème logiciel
Formulation inadéquate de questions par un contributeur dirigées vers la communauté	<ul style="list-style-type: none"> <li>⌘ Peu de temps à consacrer par les répondeurs</li> <li>⌘ Réponses désagréables aux questions</li> <li>⌘ Communication devenant malsaines entre les contributeurs (demandeur et répondeur)</li> </ul>	<b>Désengagement des répondeurs à traiter les questions soumises à la communauté</b>	Désengagement à répondre aux questions par un contributeur propagé dans un autre projet de l'écosystème logiciel
Grand attachement affectif au projet par le principal contributeur	<ul style="list-style-type: none"> <li>⌘ Confiance aveugle à la pertinence des orientations</li> <li>⌘ Incapacité à suivre les nouvelles tendances</li> <li>⌘ Manque de pérennité du projet</li> <li>⌘ Perte d'utilisateurs du projet</li> </ul>	<b>Diminution du nombre d'utilisateurs du projet</b>	Diminution du nombre d'utilisateurs du projet propagé dans l'écosystème générant une diminution d'attraction de nouveaux contributeurs au projet

**Tableau 1 : Éléments identifiés en lien avec la santé des ECLO**

Les premiers résultats obtenus nous permettent déjà de faire des liens avec des théories pour mieux comprendre certains problèmes. Le troisième problème identifié, par exemple, peut être expliqué du fait qu'un décideur devient émotionnellement attaché et ainsi trop engagé dans un projet, pouvant alors mener à une escalade sous la forme d'engagement à un plan défaillant (Keil et al., 2000).

## 5. Conclusion et contributions envisagées

Les entrevues avec des experts d'ECLO permettront de mieux comprendre les problèmes auxquels ils font face et de saisir les nuances de leurs expériences dans ce contexte. Cette recherche a plusieurs retombées espérées. D'abord, elle apportera une valeur directe aux divers acteurs clés impliqués dans les ECLO ouverts ou même industriels. Plus précisément, en leur fournissant un catalogue de problèmes de santé et de leurs indicateurs, et de causes et conséquences de ces problèmes de santé, cette recherche permettra aux contributeurs clés et gestionnaires d'ECLO de surveiller et de contrôler leurs ECLO, ce qui est essentiel pour leur survie et leur croissance. Ensuite, elle permettra d'explorer un autre aspect des ECLO, leur santé, jusque-là peu étudié, et d'asseoir de façon plus solide les futures recherches.

## Remerciements

Les auteurs remercient les Fonds de recherche du Québec (FRQ) et Fonds de la recherche scientifique (FNRS) pour le soutien financier obtenu via la subvention 264544 dans le cadre du Programme bilatéral de recherche collaborative Québec – Communauté française de Belgique.

## Références

- Alberts C.J., Dorofee A.J. (2009), *A Framework for Categorizing Key Drivers of Risk*, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.
- Carillo K., Marsan J. (2016), "The Dose Makes the Poison" - Exploring the Toxicity Phenomenon in Online Communities", *International Conference on Information Systems*, Dublin, Ireland.
- Carillo K., Marsan J., Negoita B. (2017), Exploring the biosphere - Towards a conceptualization of peer production communities as living organisms. *AIS SIGOPEN Developmental Workshop for Research on Open Phenomena*, Seoul, South Korea.
- Crowston K., Howison J., Annabi H. (2006), Information systems success in free and open source software development: theory and measures, *Software Process: Improvement and Practice*, vol. 11, n°2, p. 123-148.
- de Laat P.B. (2007), Governance of open source software: state of the art, *Journal of Management & Governance*, vol.11, n°2, p. 165-177.
- Fitzgerald B. (2006), The Transformation of Open Source Software, *MIS Quarterly*, vol. 30, n°3, p. 587-598.
- Gregus Z. (2013), Mechanisms of Toxicity, in Klaassen C.D. (ed.), *Casarett and Doull's Toxicology: The Basic Science of Poisons*, McGraw-Hill Education, China, p. 49-122.
- Hume D.C. (1975), *Enquiries concerning human understanding and concerning the principles of morals*, Oxford University Press, Oxford, England.
- Jansen S. (2014), Measuring the health of open source software ecosystems: Beyond the scope of project health, *Information and Software Technology*, vol.56, n°11, p. 1508-1519.

- Kazman R., Chen H.-M. (2010), The metropolis model and its implications for the engineering of software ecosystems, *Workshop on Future of software engineering research*, p. 187-190.
- Keil M., Mann J., Rai A., (2000), Why Software Projects Escalate: An Empirical Analysis and Test of Four Theoretical Models, *MIS Quarterly*, vol. 24, n°4, p. 631-664.
- Linux Foundation (2017), Home - The Linux Foundation, <https://www.linuxfoundation.org/>.
- Manikas K., Hansen K.M. (2013), Software ecosystems - a systematic literature review, *Journal of Systems and Software*, vol. 86, n°5, p. 1294-1306.
- Marsan J., Paré G. (2013), Antecedents of Open Source Software Adoption in Health Care Organizations: A Qualitative Survey of Experts in Canada, *International Journal of Medical Informatics*, vol. 82, n°8, p. 731-741.
- Mens T. (2016), An ecosystemic and socio-technical view on software maintenance and evolution, *International Conference of Software Maintenance and Evolution*, p. 1-8.
- Meyer M.A., Booker J.M. (2001), *Eliciting and Analyzing Expert Judgment: A Practical Guide*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.
- Miles M. B., Huberman A. M. (1994), *Qualitative Data Analysis: An Expanded Sourcebook*, SAGE Publications, Thousand Oaks, California, USA.
- Morgan G. (2006), *Images de l'organisation*, Les Presses de l'Université Laval, Québec, Québec, Canada.
- Rivard S. (2013), *Le développement de systèmes d'information : Une méthode intégrée à la transformation des processus*, Presses de l'Université du Québec, Montréal, Québec, Canada.
- Wang X., Lantzy S. (2011), "A Systematic Examination of Member Turnover and Online Community Health," *Int'l Conf. Information Systems*, Shanghai, China.
- Williams, C. (2016), How one developer just broke Node, Babel and thousands of projects in 11 lines of JavaScript, The Register, [www.theregister.co.uk/2016/03/23/npm\\_left\\_pad\\_chaos/](http://www.theregister.co.uk/2016/03/23/npm_left_pad_chaos/).