

**Compte rendu de lecture :**  
**Luigi Lancieri, *Interactions Humaines dans les Réseaux*,**  
**Editions Hermes - Lavoisier, 2005**

Voici un livre pédagogique sur les systèmes d'information. Il est produit de nombreux exemples permettant de comprendre la nécessité dans ce domaine d'une mise en convergence des recherches SHS (sciences humaines et sociales) et STIC (sciences et technologies de l'information et de la communication). En effet, les systèmes d'information de demain seront des espaces informationnels habités. Les personnes qui les utilisent seront informées de l'activité des autres. Ils pourront consulter les informations que ces derniers ont jugé pertinentes, se servir de leur exemple pour orienter leurs propres comportements ou être guidés dans la réalisation de leurs actions individuelles. Dans ces espaces informationnels habités, un travail collectif se fera pour annoter, classer et évaluer des connaissances disponibles, et pour en tirer des conclusions quant aux engagements à prendre. Un tel travail sera assuré simultanément par des hommes et des machines à condition que se mettent en place des organisations socio-cognitives qui sont respectueuses des normes et des valeurs d'un bien-être collectif. Et c'est donc sur ce point précis que les recherches en SHS et en STIC se rejoignent. Le livre de Luigi Lancieri éclaire les dynamiques à l'œuvre dans la co-construction d'espaces informationnels techniquement efficaces et de réseaux de relations socio-cognitivement enrichissantes.

C'est au début des années 1990 que les chercheurs en CSCW (*Computer Supported Cooperative Work*) ont commencé à exploiter les traces laissées par les utilisateurs des réseaux informatiques pour fabriquer des systèmes dits de sélection collaborative (« collaborative filtering »)<sup>1</sup>. Ces systèmes sont couramment exploités en e-commerce où tel ou tel produit est proposé à la considération d'un client potentiel en vertu d'un recoupement observé entre ses centres d'intérêt et ceux des personnes qui se sont déjà intéressées au produit. De nombreuses

---

<sup>1</sup> Munro A.J., Höök K., Benyon D. *Social Navigation of Information Space*, London: Springer, Collection CSCW, 1999.

techniques existent pour construire les profils d'intérêt des usagers et pour calculer leur proximité. Lancieri les illustre avec l'exemple d'un proxy-cache reliant les utilisateurs d'un réseau local à l'Internet.

Un proxy-cache sert à stocker localement une sélection des objets informationnels téléchargés depuis le Web en réponse à des requêtes formulées par les membres d'un Intranet d'entreprise. L'efficacité des proxy-caches dépend de ce qu'un objet recherché y soit stocké avant d'être demandé par un utilisateur. Afin d'améliorer cette efficacité de façon continue, des fichiers de traces peuvent être établis contenant au moins les informations suivantes : l'adresse IP des usagers, l'URL des objets demandés, la date, et le niveau de succès (« hit » si l'objet téléchargé a été précédemment demandé ; « miss » dans le cas contraire). Cette information est traitée statistiquement afin de raffiner dans le temps les téléchargements et ne stocker dans le proxy-cache que des objets informationnels les plus demandés. La modélisation statistique met en évidence la coopération implicite des usagers dans la gestion de leurs ressources informationnelles – une coopération que Lancieri assimile à une forme d'intelligence collective. Cette modélisation statistique de l'intelligence collective constitue l'élément le plus novateur du livre de Lancieri. Compte tenu de son intérêt pour des rapprochements possibles entre SHS et STIC, il convient de l'examiner en détail.

Lancieri constate empiriquement que les usages se concentrent sur une petite minorité d'objets informationnels. C'est-à-dire que l'objet le plus fréquemment demandé dans un proxy-cache (rang 1) a une fréquence d'usage bien supérieure aux objets classés dans les rangs inférieurs. La forme de ces distributions – appelées log-normales – est donc très fortement asymétrique à gauche. Ce type de distributions suscite un intérêt croissant, notamment de la part des économistes, car elles s'expliquent par des mécanismes de sélection et d'avantages cumulatifs<sup>2</sup>. Ces mécanismes interpellent les économistes pour savoir s'ils conduisent ou non à une répartition optimale des ressources collectives. Trois questions sont généralement posées. La première

---

<sup>2</sup> David, P., "Positive feedbacks and research productivity in science: reopening another black box", dans Granstrand, O. (ed.), *The Economics of Technology*, Amsterdam: Elsevier Science, 1994.

concerne l'origine de l'avantage initial. Par exemple, dans le cas des proxy-caches, est-ce que c'est la chance ou le hasard qui font qu'un utilisateur de l'Internet trouve une information qui lui semble pertinente sur le Web ? Une réponse positive à cette question signifie que d'autres sources d'information auraient pu être potentiellement meilleures que celle utilisée, mais qu'elles n'ont pas été détectées. De ce fait, les objets informationnels réunis dans un proxy-cache seraient sous-optimaux du point de vue de l'activité collective en cours.

La deuxième question concerne les mécanismes de renforcement d'un avantage initial, même si celui-ci est produit par hasard. Prenons, par exemple, les travaux fondateurs qu'a mené Derick de Solla Price sur cette question<sup>3</sup>. En considérant le fichier des requêtes individuelles produit par Lancieri comme une source d'information sur les événements retraçant l'histoire des usages d'un proxy-cache, nous pouvons reproduire le raisonnement de Price. Il suffit d'assimiler le contenu de ce fichier à une jarre dans laquelle se trouvent des billes noires et rouges, les rouges signifiant un succès, les noires un échec. Si la composition de la jarre ne variait pas, la probabilité d'un succès ou d'un échec ne se modifierait pas d'un tirage à un autre. Par contre, si la composition se modifiait selon une règle donnée, cette probabilité serait une fonction des tirages déjà effectués. Autrement dit, la question de savoir comment représenter au mieux les mécanismes de renforcement des avantages initiaux est ramenée ici à un problème de calcul : quelle est la probabilité de rencontrer un événement étant donné l'histoire du processus étudié ? Price a constaté que la règle qui convenait le mieux pour ajuster les différentes courbes hyperboliques est de ne pas incorporer dans la définition de la règle un principe de rétroaction négative. Il suffit d'augmenter la probabilité que le succès engendre des succès nouveaux et ignorer l'impact qu'un échec pourrait avoir sur la probabilité d'un succès futur. Il a justifié son raisonnement en expliquant que dans le domaine des processus stochastiques, seuls comptent des événements observables. C'est ainsi, par exemple, que si tel ou tel objet informationnel n'était pas présent dans un proxy-cache,

---

<sup>3</sup> de Solla Price, D., « A general theory of bibliometric and other cumulative advantage processes », *Journal of the American Society for Information Science*, Vol. 27, No 5, 1976, p. 292-306

ceci ne saurait être considéré comme un échec ; il constitue un non-événement.

Enfin, la troisième question est d'ordre normatif et concerne l'optimalité du système qui produit la distribution décrite. Nous venons de voir, en effet, que si la sélection initiale des objets informationnels est biaisée, les processus cumulatifs à l'œuvre dans les proxy-caches renforceraient ce biais dans le temps. Une gestion des proxy-caches qui ne tient pas compte de cette éventualité peut donc conduire à la fabrication des systèmes d'information qui sont collectivement sous-optimaux. Dans son livre, Lancieri n'explore pas la possibilité d'une dérive vers un cercle vicieux destructeur à terme de ce qu'il appelle l'intelligence collective. Il en reste à une définition minimaliste d'un collectif considérant qu'il suffit d'être co-présent dans l'espace informationnel d'un proxy-cache pour être ensemble. Sa modélisation statistique tient les requêtes pour des événements indépendants apportant chacune une information spécifique quant aux besoins collectifs. C'est-à-dire qu'il emploie une démarche typique des sondages : les individus de sa population sont structurellement indépendants les uns des autres ; son étude se concentre sur la ressemblance de leurs profils. Elle établit des relations entre variables mais laisse dans l'ombre les relations entre individus<sup>4</sup>. Or, la compréhension des dynamiques collectives passe, selon moi, par un effort visant à rapporter des distributions hyperboliques aux structures interpersonnelles dans lesquelles les acteurs s'insèrent.

En somme, ce qui retient l'attention dans le livre de Lancieri est l'effort visant à faire converger les recherches en SHS et en STIC pour mieux comprendre les fondements d'une intelligence collective en action dans les réseaux. Les exemples présentés et la précision des définitions conceptuelles proposées permettent de bien comprendre la démarche statistique mise en oeuvre. Ils facilitent une discussion critique de la modélisation adoptée, ce qui est un résultat tout à fait

---

<sup>4</sup> Les limites de ce type de démarche pour comprendre le fonctionnement des structures collectives sont discutées dans Degenne, A. et Forsé, M., *Les réseaux sociaux*, Paris : Armand Colin, Collection Sociologie, 1994. Les méthodes statistiques pour analyser des réseaux sociaux sont présentées dans Newman, M.E.J., « The Structure and Function of Complex Networks », *Society for Industrial and Applied Mathematics*, SIAM Review, Volume 45, No 2, 2003 p. 167-256.

appréciable dans un domaine de recherche en émergence en France où les concepts, les méthodes et les outils d'analyse empirique ont besoin d'être débattus. L'ouvrage de Luigi Lancieri est un bon point de départ pour entamer la construction d'un dispositif théorique et méthodologique adapté à l'objectif fixé et c'est pour cette raison que je le recommande.

William Turner,  
Responsable des recherches sur les pratiques collectives distribuées,  
LIMSI-CNRS