

## Au travers de leurs offres RIMS, les grands éditeurs scientifiques tentent de pénétrer le workflow des organismes de recherche

*Analyse I/IST-n°8*



### L'information :

Elsevier a organisé les 2 et 3 novembre dernier à Amsterdam sa première « Pure Conference », entièrement consacrée à son outil Pure, un système de gestion de l'information sur la recherche (en anglais Research Information Management System, RIMS<sup>1</sup>). Les 10 et 11 novembre le consortium euroCRIS ([www.eurocris.org](http://www.eurocris.org)), qui porte l'effort européen de normalisation dans ce domaine, tenait sa conférence stratégique à Barcelone. Ces deux événements signent l'arrivée à une certaine maturité des systèmes d'information intégrés traitant les données « administratives » de la recherche. Ces derniers ont vocation à rassembler et intégrer l'ensemble des données issues du workflow de la recherche et qui ne sont pas des données scientifiques (données sur les sources de financement, les projets de recherche, les chercheurs et leur expertise, les publications etc.). Sont concernés par ces outils les universités, les organismes de recherche, mais aussi les financeurs ou toutes entités de gouvernance de la recherche qui ont des besoins de gestion, de reporting et d'analyse de la production scientifique. Parcequ'ils sont au cœur des fonction d'« intelligence » des processus de recherche et ont vocation à terme à pénétrer tout le système d'information de l'ESR, ces services logiciels intégrés sont des outils stratégiques. Or les trois principales offres d'outils arrivés à maturité, Pure, Converis et Symplectic Elements sont portées respectivement par Elsevier, Thomson Reuters et Digital Science, une société soeur de Springer Nature. La pénétration généralisée de ces offres « propriétaires » (on n'est nullement ici dans le domaine du logiciel libre) pose de redoutables problèmes de dépendance à venir de la recherche publique à ces services logiciels.

### L'analyse d'Intelligence IST :

Dans un billet publié le 11 novembre sur le SparcOAforum ([sparcOAforum@arl.org](mailto:sparcOAforum@arl.org)) Stevan Harnad, l'une des grandes figures de l'Open Access, dénonce l'adoption par certaines universités britanniques du système Pure d'Elsevier vu comme le cheval de Troie du N°1 de l'édition scientifique mondiale pour s'introduire dans les systèmes d'information de ces universités. Si l'on n'est pas obligé d'adopter le ton de Savonarole de ce billet, qui nuit à sa portée, il est évident que la question soulevée par Harnad est essentielle.

Les *Research Information Management Systems (RIMS)* appartiennent à la grande famille des logiciels intégrés de gestion fonctionnelle : ce n'est pas un hasard si SAP, le leader mondial de ce type de service logiciel est (avec Elsevier et Thomson Reuters) l'un des membres actifs de l'association EuroCRIS. Sur le plan conceptuel les RIMS n'ont rien d'innovant : il s'agit de décrire le système de la recherche comme un ensemble d'objets (les agences de financement, les organismes de recherche, les labos, les projets, les chercheurs...) dotés chacun de ses attributs (par exemple la liste de publications d'un chercheur, les collaborations scientifiques d'un labo...). L'intérêt de ces RIMS est de normaliser ces données, d'automatiser autant que faire se peut leur collecte et leur mise à jour, de reconnaître et de décrire l'interdépendance de ces variables, enfin d'automatiser au maximum les tâches de reporting et d'analyse de ces données à des fins administratives ou de planification stratégique. Quand bien même ils ne feraient que cela les RIMS seraient déjà des outils-pivots puisqu'en eux, idéalement, se concentre toute l'intelligence de la recherche. On

---

<sup>1</sup>Ces systèmes de gestion intégrée de l'information sont aussi parfois connus sous l'acronyme CRIS : « Current Research Information System ». La Commission européenne a investi de façon précoce dans les problématiques de normalisation des données relatives à la recherche donnant naissance au Current European Research Information Format (CERIF), qui est aujourd'hui portée par l'association EuroCRIS.

voit l'intérêt stratégique des grands éditeurs, Elsevier, Thomson Reuters ou demain Springer Nature<sup>2</sup> à développer des offres qui ont vocation à devenir l'épine dorsale des systèmes d'information de l'ESR. Si le conflit d'intérêt est évident, ces offres n'en connaissent pas moins une pénétration rapide. Pour des raisons très pragmatiques : ils augmentent de façon drastique la productivité de l'ensemble des tâches administratives liées à la recherche (par exemple le montage de dossiers de financement), ils simplifient les tâches de reporting interne et externe, enfin permettent le développement de multiples axes d'analyse de l'information qu'ils collectent et intègrent. Dans le contexte britannique ces systèmes ont aussi bénéficié des impératifs de « compliance » (conformité) aux mandats de publication en Gold Open Access définis par les pouvoirs publics.

La France à ce stade semble en retard sur ses voisins : dans la liste de références clients de Pure on ne trouve aucun organisme français, alors que les références dans les autres pays européens sont nombreuses. Outre la pénétration du « workflow » des entités clientes, les RIMS ont un autre intérêt stratégique : ils présentent pour les éditeurs des réserves de croissance et de rentabilité intéressantes. Alors que les enjeux financiers et les logiques de gouvernance de la recherche évoluent dans un sens qui exige des gains de productivité, ces outils apportent une vraie réponse. D'autre part les grands éditeurs scientifiques savent bien qu'ils ne pénétreront pas très avant dans le workflow scientifique proprement dit : le processus de recherche, même s'il s'appuie désormais de manière générale sur les technologies de l'information et de la communication, présente de très grandes spécificités disciplinaires et de nombreux développements ad hoc. Les éditeurs auront ici un rôle périphérique, éventuellement lié à l'édition de jeux de données expérimentales (« Datasets »). Par contre les problématiques générales de la recherche (accès aux financements, planification, gestion des input et des output (en premier lieu les publications), reporting, évaluation, sont des problématiques transversales à tout l'écosystème de la recherche et donc représentent un marché qui n'est pas un marché de niche. Ce marché présente de plus deux caractéristiques positives : d'une part les gains de productivité dégagés par ces outils seront tels que les clients ne seront pas trop regardants sur la formation des prix de ces offres de services ; d'autre part une fois qu'une institution a fait le choix d'un RIMS du marché, il lui est très difficile d'en changer tant ces outils deviennent la colonne vertébrale des systèmes d'information des entités clientes.

Les RIMS présentent aussi pour les grands éditeurs scientifiques qui les contrôlent deux avantages de moindre importance. D'une part ces systèmes ont certes recours à des informations internes aux organisations clientes, mais ils interfacent aussi des sources d'information externes, par exemple pour enrichir les profils de chercheurs de références bibliographiques normalisées et des données de citations. Logiquement Pure est donc interfacé chez Elsevier avec Scopus, et Converis avec Web of Science. Or on le sait chacun de ces outils a son profil et ses biais propres, des biais que l'adoption de l'un ou l'autre système risque d'« importer » dans les systèmes d'information des clients. Deuxième avantage : les RIMS incluent généralement une fonctionnalité de dépôt automatique en archive ouverte des articles scientifiques émanant d'une institution cliente. Par le biais des RIMS les grands éditeurs peuvent donc acquérir un rôle clé dans le « Green Open Access ». C'est d'ailleurs ce point qui fait réagir Stevan Harnad, soupçonnant sans doute à tort Elsevier de vouloir jouer de cette caractéristique pour aligner les délais de dépôt en archive ouverte sur ses propres délais d'embargos.

Un autre axe de réflexion suscité par le développement des RIMS est la captation précoce de l'innovation dans le champ de l'e-science par les grands éditeurs scientifiques. Aucun des développements ici mentionnés n'est né au sein des services de R&D de ces grands éditeurs. Mais ces derniers ont su détecter de façon précoce ces développements innovants et les ont cueillis au moment où pour se développer les start-up qui portaient ces développements ont eu besoin de financement. Digital Science, qui au sein du groupe Macmillan/Nature Publishing Group, joue le rôle d'un fonds de capital venture dédié aux applications de la science numérique a dès 2010 participé au financement de Symplectic (développeur du RIMS Symplectic Elements). Fin 2012, Elsevier rachetait la start-up danoise Atira, qui avait développé les premières releases de Pure. Fin 2013 Thomson Reuters rachetait la start-up allemande Avedas qui a conçu le service Converis. Dans les trois cas, les grands éditeurs sont intervenus au moment où ces start-up avaient besoin de fonds et d'appui pour bâtir leur développement commercial et leurs présences internationales. Si les éditeurs savent si bien capter l'innovation au moment où elle a démontré sa viabilité ce n'est nullement en raison d'une quelconque volonté maléfique comme semble le croire Stevan Harnad, mais simplement parce qu'ils exercent une veille acérée sur leur environnement et disposent d'une vision stratégique de ces questions. La Commission européenne avait investi de façon précoce (à la fin des années 90) dans cette réflexion, intervenant en particulier au

---

<sup>2</sup>Si Digital Science qui propose le système Symplectic Elements n'est pas une filiale de Springer Nature – elle est filiale du groupe Macmillan qui est l'actionnaire de référence de Springer Nature

niveau du processus de normalisation matérialisé par le Current European Research Information Format (CERIF), effort de normalisation qui est aujourd'hui porté par l'association EuroCRIS. Mais aucun industriel du logiciel ne s'est saisi de ces développements. Malgré les prises de position en ce domaine d'Elsevier, de Thomson Reuters et de Springer Nature, les RIMS sont à un stade précoce de leur développement. La structuration de ce marché est loin d'être figée. SAP, leader mondial des « Enterprise Information Management » systems, qui sont les précurseurs de toute cette famille d'outils, a avancé quelques pions sur cet échiquier et a l'envergure pour jouer un rôle important. Mais il semble urgent que les opérateurs de la recherche se saisissent de ces questions, reconnaissent l'importance stratégique de ces outils, et soient conscients que leurs choix de clients déterminera la structuration de ce marché. Alors que les grands éditeurs scientifiques sont déjà en situation d'oligopole et contrôlent une part croissante de la communication scientifique, il faudra y réfléchir à deux fois avant de leur confier les clefs des systèmes d'information de la recherche.