

Compte-rendu d'atelier
« Open Data, enrichissement des données et création de valeur »
Forum du GFII 2015

Atelier animé par Alain Chaumet, Chef de pôle, service d'appui aux politiques publiques de l'IGN et co-animateur du groupe de travail "données publiques" du GFII lors du Forum du GFII 2015.

FORUM  **DU GFII**

PROGRAMME DE L'ATELIER

A l'heure où l'administration lance le programme « Etat Plateforme » et où les entreprises commencent à ouvrir et valoriser leurs données, un premier bilan sur la réutilisation économique de l'open data s'impose. Quelles sont les innovations et les offres de services innovantes issues de la réutilisation des données ? Quels sont les modèles économiques existants et quelle est leur pérennité ? De la donnée ouverte à la commercialisation d'offres de services à valeur ajoutée, quels sont les freins rencontrés par les ré-utilisateurs en termes d'accès, de préparation ou de structuration des données ? Quelles sont les perspectives de développement offertes par le croisement et l'enrichissement des données publiques avec d'autres sources (données privées, crowdsourcing ...) ? Cet atelier abordera ces questions à partir du retour d'expérience de start-ups et d'industriels développant des services d'information B2B innovants à partir des données ouvertes.

Intervenants :

- François Bancilhon, CEO, Data Publica
- François-Xavier Bois, Directeur de la R&D, KernixLab
- Victor et Martin Gross, fondateurs, Commoprices
- Stéphane Minard, Directeur, La Centrale des Marchés
- Olivier Tronchon, Directeur marketing, Ellisphere

forum.gfii.fr

DÉROULÉ

Chacun des intervenants a présenté comme prévu l'activité de sa société ou un projet phare choisi en exemple et les problématiques qui y sont liées.

➤ **François Bancilhon, CEO, Data Publica**

François Bancilhon a présenté [Data-Publica et l'offre C-Radar](#). Celle-ci s'appuie sur les données publiques ouvertes de l'INSEE, d'Infogreffe ou de la DILA mais également sur les signaux faibles des réseaux sociaux pour permettre finalement un meilleur ciblage des offres des entreprises. Data-Publica met en œuvre des technologies prédictives innovantes en rupture avec les technologies plus traditionnelles d'analyse de marché. François Bancilhon a rappelé l'inaccessibilité de l'offre tarifaire de la base SIRENE diffusée par l'INSEE qui crée ainsi une situation préjudiciable au développement des applications nécessitant une connaissance des entreprises précise et actualisée.

➤ **François-Xavier Bois, Directeur de la R&D, KernixLab**

François-Xavier Bois a présenté Kernix editrice du site [Accessible.net](#). Ce projet se propose de recenser les établissements recevant du public (ERP) accessibles aux personnes à mobilité réduite pour en constituer une base de données ouverte à l'utilisation mais aussi aux contributions des entreprises ou des collectivités locales intéressées.

➤ **Victor et Martin Gross, fondateurs, Commoprices**

Martin Gross a présenté [Commoprices](#) qui est à la fois la société dont il est le cofondateur, un site et un service présentant l'agrégation la plus importante au monde de prix de matières premières (les « commodités » en français) provenant de diverses sources, notamment des services de douane. L'hétérogénéité des données en entrée impose un travail de standardisation dont le résultat permet la mise en relation avec d'autres données, financières notamment, entraînant à son tour d'autres possibilités d'exploiter et valoriser ces données.

➤ **Stéphan Minard, Directeur, La Centrale des Marchés**

Stéphan Minard a présenté la [Centrale des Marchés](#) et son offre de mise à disposition mise à disposition des marchés publics. Ces marchés sont disponibles sous des formes hétérogènes qu'il s'agit d'intégrer pour en opérer une rediffusion sous une forme standardisée et adaptée à différents

segments de clientèle. L'analyse des sources permet aussi d'opérer un modèle prédictif de sorties de marchés dont une bonne partie est faite de marchés récurrents.

➤ **Olivier Tronchon, Directeur marketing, Ellisphere**

Olivier Tronchon a présenté [Ellisphere](#) pour son offre d'information économique et financière à destination des directions stratégiques, financières et marketing des entreprises. La maturité de cette entreprise centenaire lui permet de disposer d'un réseau relationnel étendu tout en s'appuyant désormais sur les technologies les plus récentes développées par un pôle de data scientists. Olivier Tronchon a souligné l'importance de l'accessibilité de l'information financière d'entreprise sur laquelle se fonde notamment la fluidité du crédit inter-entreprises pour un montant courant de 700 milliards d'euros.

DISCUSSION COLLECTIVE – ÉCHANGES AVEC LA SALLE

Les questions et remarques de l'auditoire ont permis de préciser les approches de chacun sur les points principaux de discussion, disponibilité et accessibilité des données, modèles économiques de valorisation des offres.

➤ **Impact du travail de mise en forme et d'harmonisation des formats de données**

Le catalogue des données publiques accessibles s'enrichit de plus en plus grâce à l'action des pouvoirs publics, notamment par l'action du SGMAP et de la mission Etalab. Toutefois, le seul respect des critères de l'opendata, qui ne prévoit aucun modèle contraignant de diffusion de données métier le plus souvent très complexes, ne permet la réutilisation de ces données qu'au prix d'un travail important de mise en forme. Ce travail de mise en forme est coûteux pour chaque réutilisateur et ces coûts pourraient être mieux valorisés s'ils pouvaient être appliqués à l'intelligence de la donnée ou à la mise en forme des résultats des traitements.

➤ **Le GFII demande l'ouverture des données pivots**

Les administrations comme celle des douanes ou la DILA ont bien compris, du point de vue de l'intérêt général, l'intérêt de faciliter l'accès à leurs gisements de données. Certes, les problèmes techniques sont encore nombreux, comme pour les informations des marchés publics dont la production est très décentralisée et ne fait pas l'objet de norme, mais la démarche d'ouverture des données publiques semble être maintenant assez largement comprise, partagée et mise en oeuvre. En revanche, certains obstacles demeurent dont celui signalé en séance de la condition financière impraticable pour la plupart des entreprises pour accéder au fichier SIRENE diffusé par l'INSEE, administration centrale du ministère de l'économie et des finances. Ce point précis, avec celui des

formats de diffusion des données, a été traité par le GFII dans le cadre de sa demande d'ouverture des données pivots dont il a précisé une définition et fourni une liste¹.

➤ **Modèles de financement des services en ligne**

Le modèle de valorisation principal est très majoritairement celui de l'abonnement auprès de l'utilisateur final avec des barèmes permettant d'adresser différents segments de clientèle. L'adoption très majoritaire de ce modèle provient en fait du modèle technique de services SaaS qui est le seul à être bien compris par les banquiers, ceux-ci restant une partie prenante déterminante dans le développement des entreprises. Seul un service comme accessible.net, dont l'utilisateur final est supposé ne pas avoir de disposition à payer, recourt au financement de tiers payants au titre de l'intérêt collectif privé ou de l'intérêt général public. Les autres modèles de financement des services en ligne, par la publicité, le crowdfunding, le recours aux marchés n'ont pas été évoqués. Ils sont pourtant le support de développement de grands succès de l'internet comme Google, Wikipedia ou Criteo.

CONCLUSION

Les échanges toujours cordiaux, malgré les situations de concurrence ou d'opposition qui pouvaient exister entre certains intervenants de la tribune ou de la salle, et précis, malgré l'ampleur et la complexité du sujet, se sont conclus par des encouragements à progresser encore vers une ouverture des données publiques plus large et plus efficace au bénéfice tant des intérêts privés que de l'intérêt général. Le GFII, avec son groupe de travail « données publiques et opendata », pourra poursuivre ses travaux pour tenter d'identifier plus clairement la part française des 40 milliards d'euros de valeur ajoutée promis par l'opendata lors du lancement de la révision de la directive PSI par la Commission européenne en 2011.

¹ <http://www.gfii.fr/fr/document/reutilisation-des-informations-publiques-le-gfii-demande-l-ouverture-et-la-gratuite-des-donnees-pivots>