

OPEN DATA: LA SNCF FAIT CIRCULER L'INFO

ENQUÊTE

FRÉQUENTATION DU RÉSEAU TRANSILIEN, STATS SUR LA PONCTUALITÉ... UN AN APRÈS BRÉTIGNY, LA COMPAGNIE PARTAGE LARGEMENT SES DONNÉES. AVEC, EN RETOUR GAGNANT, LE DÉVELOPPEMENT D'APPLIS PAR DES STARTUP.

Par **GABRIEL SIMÉON**
Dessin **ROCCO**

C'était il y a près d'un an, le 12 juillet. Ce n'était pas la grève contre la réforme du rail mais l'accident de l'Intercités Paris-Limoges à Brétigny-sur-Orge (Essonne), qui mobilisait la com de crise de la SNCF : 7 morts, 32 blessés, la découverte de la fameuse éclisse défilante et de nombreuses questions sur l'entretien du réseau ferré... Dans la foulée, 10 000 personnes demandaient à la SNCF dans une pétition « d'aller jusqu'au bout sur la transparence » en publiant « un rapport complet en "open data" sur l'état du réseau et du matériel ». On aurait eu vite fait d'en conclure à l'époque

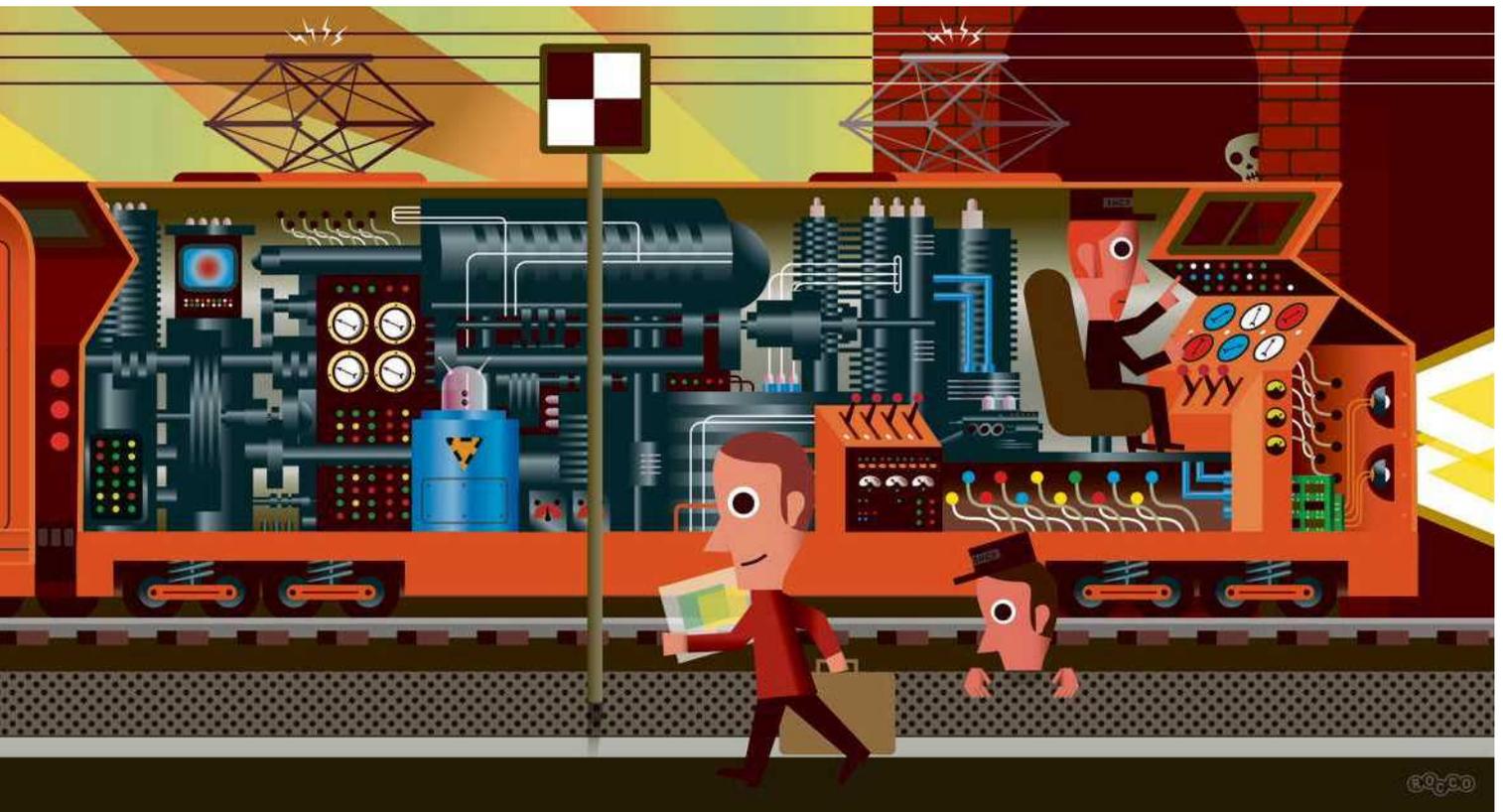
que le dinosaure du rail découvrait le terme d'open data, synonyme de mise à disposition du grand public de données internes. Et on aurait eu tort. Car la SNCF, quelle que soit sa responsabilité dans l'accident, est en pointe en ce domaine.

Alors que la Poste, Veolia ou la RATP se lancent prudemment (*lire encadré*), on n'hésite pas aujourd'hui dans la communauté open data à citer l'opérateur ferroviaire comme un modèle. Pour les non-initiés, il y aurait pourtant de quoi s'étonner à voir le logo du groupe public s'afficher en bonne place aux côtés de ceux de Google, Microsoft et du pôle de compétitivité Cap Digital sur le site du concours Dataconnexions – qui encourage la création de services réutilisant des données publiques, et dont la SNCF est partenaire depuis la première édition. « On est assez légitime dans ce domaine, estime son directeur de la communication, Patrick Ropert, qui conduit depuis plusieurs années la mue numérique de l'entreprise. On a ouvert notre plateforme open data [www.data.sncf.com] en décem-

bre 2011 et on l'a progressivement remplie avec des jeux de données. » Ces derniers, au nombre de 21, ce sont par exemple la localisation des gares, les horaires des lignes de transport express régional (TER) ou le nombre de voyageurs montant dans chaque gare du réseau Transilien d'Ile-de-France. En consultant le tableau sur la ponctualité des TGV, on apprend ainsi que sur les 123 ayant circulé entre Marseille et Lille en décembre, 36 ont eu plus d'un quart d'heure de retard.

Partager ces informations n'allait pourtant pas de soi. « Notre réflexe initial était de ne publier que les "bonnes" données, celles qui ne nous étaient pas défavorables. Maintenant, je demande aux équipes de ne jamais être dans le déni : les faits sont les faits », raconte Patrick Ropert. Le dirigeant reconnaît que parler d'open data au sein d'un groupe qui emploie 250 000 collaborateurs « a conduit à des débats hémériques » : « C'est une dynamique difficile à mettre en œuvre, qui appelle à une révolution culturelle. » Pour quels résultats ? « Difficile à dire. Il y a une part de pari... » Au-delà de la com, la SNCF y trouve plusieurs intérêts. Ouvrir certaines de ses données au grand public accélérerait la « transformation interne ». Cela stimulerait également la R & D et obligerait les décideurs à s'inspirer de l'esprit des start-up avec lesquels ils sont en contact.

PROTOTYPES. Ces données brutes – des milliers de lignes de tableau, de code, des graphiques abscons – restant difficilement lisibles par l'internaute lambda, la SNCF encourage leur intégration dans des applications facilitant la vie de ses clients. Et organise régulièrement des « hackathons », des événements où des équipes de développeurs ont un temps limité pour créer des prototypes d'application. Le « Hack Days » de juin 2012 a ainsi conduit au lancement, un an plus tard, de l'application Tranquillien, affichant à l'avance le taux de remplissage des trains de la banlieue parisienne. « La SNCF nous a énormément supportés et nous a permis de monter une boîte directement profitable », se souvient Rand Hindi, patron de Snips et cocréateur de Tranquillien (dont EcoFutur faisait le portrait



le 2 juin). Le hackathon «Hackcess» de novembre, en partenariat avec l'association OpenStreetMap, a quant à lui débouché sur la création de No Way, une application indiquant aux personnes handicapées les itinéraires adaptés en gare. «On travaille avec eux pour cartographier les informations disponibles sur les 400 gares d'Ile-de-France», raconte Gaël Musquet, porte-parole de OpenStreetMap France (lire EcoFutur du 10 mars). C'est le top du top en matière de réalisation pour le moment.»

A Versailles, la SNCF teste en ce moment un système permettant d'adapter automatiquement l'éclairage d'une gare au flux des voyageurs. La démarche open data consiste ici à permettre à la start-up Lookies de récupérer les informations wi-fi pour déterminer le nombre de smartphones – et donc de personnes – qui transitent. Avec Rail-

L'application No Way, née lors d'un «hackathon» organisé par la SNCF, indique aux personnes handicapées les itinéraires adaptés en gare.

dar, l'opérateur ferroviaire proposera bientôt à ses clients une carte numérique affichant le déplacement des trains en temps réel (un avant-goût est visible sur www.raildar.fr, qui affichait 68% de trains à l'heure et 30% en retard début juillet). «Les développeurs de Raildar ont d'abord «scrapé» [de l'anglais «to scrap», synonyme de hacker, ndr] nos données, se souvient Patrick Ropert. Il y a quelques années, on les aurait considérés comme des pirates, mais on a préféré les inviter au «Datashaker».» Le «Datashaker» est un dispositif de soutien aux start-up mis en place à l'espace de coworking parisien Numa et parrainé par la SNCF. Cette dernière fournit des données, délivre un coaching et autorise les tests grandeur nature sur son réseau. La démarche open data de la SNCF s'inscrit en réalité dans un mouvement plus vaste d'«open innovation»: travailler avec des start-up serait le meilleur moyen pour faire émerger un maximum

de services innovants. Partenaire depuis 2011 du «Camping», un accélérateur de start-up parisien, le groupe public a également lancé le sien, «Voyageurs connectés». «A la SNCF, ils ont mieux compris que d'autres les enjeux de l'innovation ouverte. Ils ont un vrai projet», confie un spécialiste des questions d'open data. «C'est un futur acteur incontournable de l'écosystème start-up français», prophétise Rand Hindi, patron de Snips.

«DÉFENSE». Mais ces éloges ne vont pas sans quelques reproches. Un acteur du secteur déplore les «freins» actionnés par la SNCF sur la libération de certaines données, notamment de sa branche Voyages. Un autre regrette que l'entreprise «la joue en défense», soulignant que l'industrie touristique bénéficierait grandement d'une partie de ses données. De l'autre côté du quai, la SNCF assume: pas question de partager ses données commerciales et celles portant sur les habitudes des clients, qui «pourraient être utiles à la concurrence», défend Patrick

Ropert. Est-ce bien compatible avec une stratégie d'open data? «C'est un sujet un peu différent des critiques qui nous sont faites sur notre gamme tarifaire. Que nous allons d'ailleurs rendre plus lisible grâce aux outils digitaux.» Quant aux données en open data sur l'état du réseau ferroviaire et du matériel... elles se font toujours attendre. Un an après Brétigny, on n'en trouve pas de trace en ligne. «Il est difficile de partager ces informations avec le grand public, les données sont complexes et denses, cela prend du temps, clame-t-on à la SNCF. On ne les cache pas, on a d'ailleurs l'intention de les publier un jour.» Quand exactement? On n'en saura pas plus. Le sujet embarrasse visiblement. La com de l'entreprise a en revanche été plus prompte à publier des données sur les reprises de travail dans les grandes gares pendant les perturbations causées par la grève de juin. Open data... quand ça arrange. ◆

LA POSTE RECHIGNE À PRENDRE LE PLI

LE GROUPE PUBLIC DEMEURE FRILEUX SUR LA PUBLICATION DE SES DONNÉES, NOTAMMENT SA «BASE ADRESSES».

Partenaire de la première heure du concours Dataconnexions – qui récompense depuis 2012 les créateurs de services réutilisant des données publiques – et organisateur des ateliers DataPoste, la Poste semble avoir bien intégré la démarche open data. A première vue. «Je ne dirais pas qu'on est à la pointe dans ce domaine», reconnaît Delphine Desgurse, directrice de la branche innovation du groupe public. Le sujet y est en réalité simplement suivi et pas encore recommandé ou prioritaire. Seuls trois jeux de données – localisation et horaires des enseignes – sont partagés sur le portail open data du gouvernement (1), et ceux mis à disposition durant le Data-Poste n'ont pas eu le droit d'être réutilisés après l'événement.

«La Poste n'est pas un bon exemple à suivre», tacle un pilier de la communauté open data française. Selon un autre spécialiste, la direction de l'innovation manquerait de pouvoir au sein du groupe. «Commençons par partager nos données en interne et après on verra», répond Delphine Desgurse. La Poste refuse d'ouvrir

sa très convoitée «base de données adresses» (codes postaux, noms de rues), dont la vente aux entreprises représente «entre 5 millions et 10 millions d'euros, un chiffre d'affaires significatif» pour le groupe. «C'est pourtant dans leur intérêt de libérer ces données. Même si les retombées ne sont pas immédiates, elles sont énormes. Il faut qu'ils y réfléchissent vite car la communauté peut très vite pallier un défaut de données», prévient-on chez l'association de cartographie OpenStreetMap France.

Un projet collaboratif, «Bano», est justement sur le point de combler le manque, avec déjà plus de 14 millions d'adresses recensées en France. D'ici un an, toutes les voies devraient être renseignées. «Notre donnée est beaucoup plus précise», glisse-t-on à la Poste. Le groupe a d'autres projets, comme permettre à chacun de suivre le parcours du facteur en temps réel. Un service envisageable, selon Delphine Desgurse, «à partir de 2015, lorsqu'ils seront tous équipés du téléphone portable Facteo».

G.Sj

(1) www.data.gouv.fr