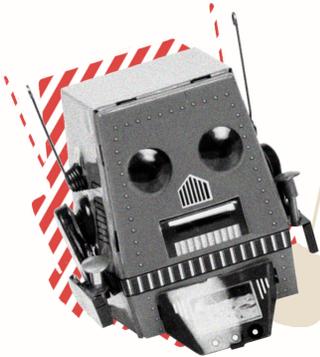


COMMENT MIEUX FAIRE PARLER LES DONNÉES ?

INTRODUCTION

Aujourd'hui, les données sont partout dans l'environnement du journaliste. Il les traite de plus en plus pour trouver des angles d'articles, mener des enquêtes, ou faire émerger des sujets.

Un premier enjeu consiste à établir des corrélations entre ces données. Ensuite, si les bases de données à l'échelle nationale sont plutôt accessibles, lorsqu'on descend à l'échelle locale, voire hyperlocale, ça se complique : les bases de données sont éparpillées, parfois incompatibles entre elles. Des outils sont en cours de développement pour répondre à ces enjeux. Pour la suite, il faudra généraliser les algorithmes et disposer d'une interface claire et utilisable. L'autre inconnue : la question du traitement des données en temps réel, qui deviendra incontournable. Cela permettra aussi de redéfinir la relation entre le journaliste et son lecteur.



PERSPECTIVE

COMMENT MIEUX TRAITER LES DONNÉES POUR FACILITER LE TRAVAIL DES JOURNALISTES DANS LES RÉDACTIONS ?

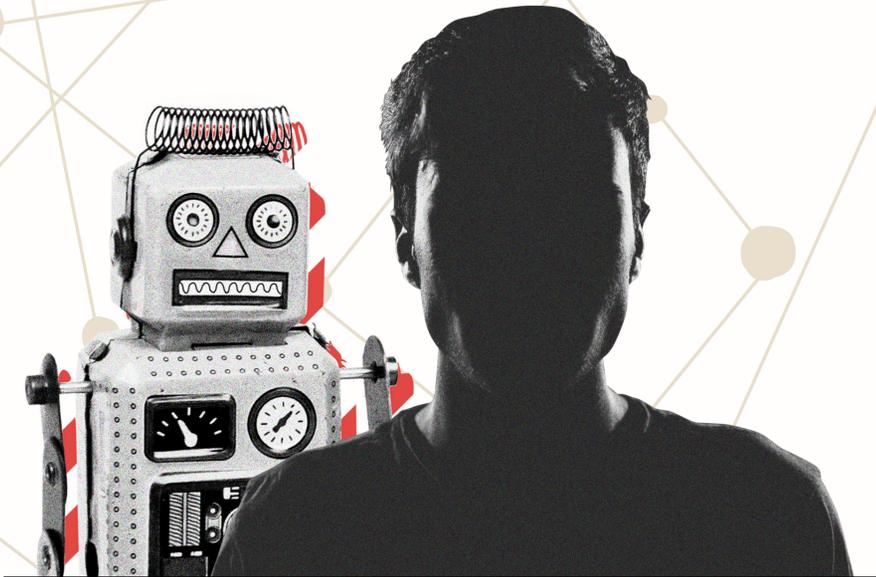
INITIATIVES

DES INITIATIVES AUX USA
PROJET "BIG LOCAL NEWS" STANFORD UNIVERSITY

LE PROJET PERMET DE :

- Traiter des données difficiles à obtenir et les analyser avec plus de finesse.
- Traiter des données au niveau hyper-local. « Le projet a permis de mettre en évidence la discrimination raciale, pas seulement à l'échelle des états mais aussi à l'échelle plus locale. »
- « Big Local News va lancer une plateforme qui permettra aux journalistes de collaborer sur des données communes. Elle aidera les journalistes à ranger les données en désordre, à les standardiser, et pouvoir les analyser et les visualiser. »

DATA



HISTOIRE

Le data journalisme permet de traiter des enquêtes et d'agrémenter des articles depuis des années.

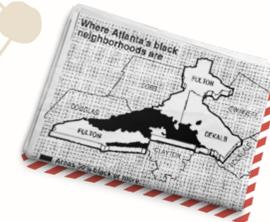
1821 : LE MANCHESTER GUARDIAN
Le Manchester Guardian publie un tableau de données sur des écoles de la ville de Manchester et de Salford : nombre d'élèves dans chaque école / dépenses annuelles de chaque école.



1854 : THE GUARDIAN
The guardian publie une carte du choléra à Londres. Elle montre la concentration du choléra autour des pompes à eau et a permis de comprendre l'origine de l'épidémie.



1854 : THE GUARDIAN
The guardian publie une carte du choléra à Londres. Elle montre la concentration du choléra autour des pompes à eau et a permis de comprendre l'origine de l'épidémie.



INITIATIVES

UNE INITIATIVE RENNAISE
PROJET ICODA, LABORATOIRE INRIA

QUELS SONT LES OBJECTIFS D'ICODA ?

Relier les données des bases de données entre elles pour assister les rédactions locales dans le croisement de leurs infos.

QUELLES LIMITES ?

La première limite, c'est que iCoda fonctionne au cas par cas, sur un cas d'usage : des élections municipales. Leurs outils fonctionnent dans des situations spécifiques. Pour les redéployer sur d'autres sujets, il faudrait redéfinir le langage, car ce ne sont pas les mêmes infos. Le problème, c'est donc la généralisation. Finalement iCoda s'est surtout concentré sur le seul cas d'usage de Ouest France. Le problème aussi est l'accès à des données claires via une interface utilisable par tous.



Mettre en commun le travail de recherche de chaque journaliste collecté isolément : dimension collective. « L'idée c'est aussi de formaliser la connaissance accumulée de manière isolée et informelle par chacun, pour qu'elle puisse être partagée. »

Exemple : pour le projet iCoda, les informaticiens de l'Inria ont beaucoup travaillé avec Ouest-France, notamment sur le cas des municipales. Au lendemain des élections de 2015, les journalistes disposaient de nombreuses données. Mais entre 2014 et 2020, aucune base de données gouvernementale ne détaillait les évolutions de la vie municipale : quels maires ont démissionné, quels maires sont décédés...etc. Pourtant, ces informations existent dans les archives du journal. Le but d'iCoda, c'est d'extraire ces informations via des outils, pour arriver en 2020 avec une vue précise de l'état des conseils municipaux, tenant compte des évolutions depuis 2015.

Erwan Alix : « Chez Ouest France, iCoda nous aide à faire le travail de croisement d'infos, et nous a permis de faire émerger certaines informations extraites de notre propre base de données : les archives de OF. Par exemple : parmi tous les articles, le programme est capable de faire ressortir une liste des maires décédés pdt leur mandat. Il organise les informations, les met en forme. »

QUELS RÉSULTATS POUR L'INSTANT ?

« Pour le moment, nous avons beaucoup travaillé sur les fondements scientifiques, avec divers outils. Ces travaux n'ont pas été traduits par des produits opérationnels au sein des rédactions. Cela demande une transformation, une sensibilisation de la future génération de journalistes, assez longues. »



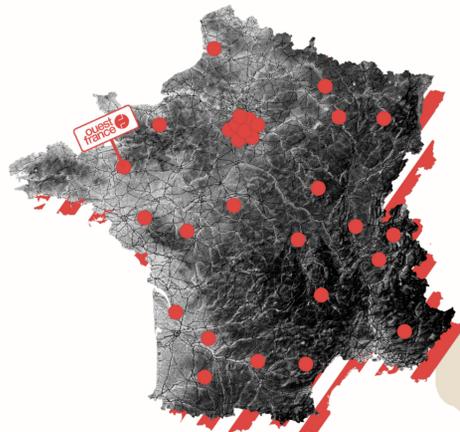
ICODA

UNE INITIATIVE RENNAISE
PROJET ICODA, LABORATOIRE INRIA

Une initiative rennaise pour aider les journalistes à traiter des données : le projet iCoda, du laboratoire Inria, par Guillaume Gravier

Guillaume Gravier : « Je suis parti du constat que le journaliste doit aller chercher des infos dans des bases de données plus ou moins bien organisées, éclatées, éparpillées. Aussi, d'un point de vue informatique, elles ont des formats différents. Sur wikipédia, l'INSEE, "nos députés.fr", on trouve des tas de statistiques, qui traînent, isolés, sous des formats différents... »

Donc Guillaume Gravier et ses collègues planchent sur un agrégat d'outils divers (reconnaissance de noms dans un texte, outils qui établissent des corrélations entre deux entités). Ils permettent aux journalistes d'établir des corrélations entre des bases de données, de repérer des similitudes entre des situations...



CHIFFRES

Il existe de plus en plus de cellules data journalisme dans les rédactions locales. On compte aujourd'hui à peu près 30 journalistes dans les titres de presse quotidienne régionale en France. Il en existe une à Ouest France depuis 2017.

AUJOURD'HUI

COMMENT LE JOURNALISTE AUJOURD'HUI TRAITE-T-IL LES DONNÉES DE SON ENVIRONNEMENT ?

On utilise actuellement plutôt des données en open data (données publiques, notamment mises à disposition par les institutions) ou compilées manuellement.

QUELLES UTILISATIONS DES DONNÉES ?

- Trouver des idées d'angle ou faire émerger des sujets d'articles. Faire émerger des sujets, des questions.
- Faire parler les chiffres à travers des visualisations de données (datavisualisation) : rendre limpides des bases de données complexes via l'infographie, la visualisation de données (Exemple : Datagueule/Compilation des violences policières par Dufresne avec Allo Place Beauvau).

LES ENJEUX DU LOCAL :

Le problème aujourd'hui c'est que plus on descend à l'échelle locale, moins il existe de bases de données précises et mises à jour. Par exemple : Il n'existe pas de base de données qui compile les maires démissionnaires de certaines municipalités, ni la raison de leur démission, alors que cela permettrait d'expliquer des faits. En plus, si les données existent, elles ne sont pas lisibles pour le grand public.

LIMITES :

Les journalistes, faute d'accès à certaines données (qui sont éparpillées, non cohérentes entre elles..) sont contraints de créer leurs propres bases de données, souvent manuellement. Erwan Alix : « Mais le problème, c'est que peu d'entre nous sont capables de les traiter. Il faudrait commencer par des formations pour tous les journalistes. »



CONCLUSION

Aujourd'hui, des start up naissent pour vulgariser l'accès aux données pour des clients divers et variés.

Cette entreprise rennaise fait partie de l'incubateur de start up de Ouest France et sera sûrement amenée à traiter avec les médias pour les aider à développer des bases de données.