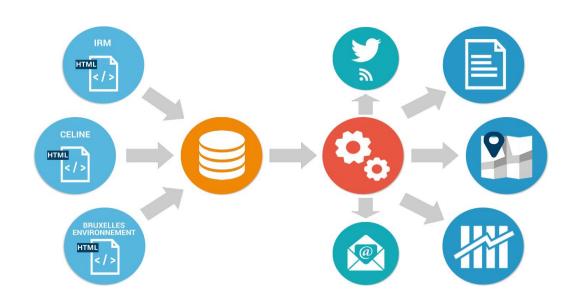




Bxl'air Bot: a journalistic point of view about air quality in Brussels

Laurence Dierickx







1. Context

- PhD research about news automation within the uses and professional practices
- If journalists cannot compete with automated news production, how to make the machines allies rather than adversaries?
- Transdisciplinary research framed by STS theories
- 2 case studies over 1 year: Bxl'air Bot (Alter Echos) and Quotebot (Mediafin)
- Methodology adapted from one case study to another
- Bxl'air Bot = experimental project, developed by the researcher to make the experience of technology (possibilities and limits)
- Final users associated to the design process = first form of use





2. Experimental project

- 2 journalists particularly involved
- Researcher = mediation agent between the newsroom and the bot
- Supposed technical translations of journalistic requirements
- Designed as a tool for journalist by collecting and storing data
 + provide real-time analysis (NLG, maps and graphs)
- Transparency of the project = service to the audience with providing real-time information about the situation
- Observations over one year => to support a wider investigative

Data does not tell anything about the causes and the consequences of air pollution

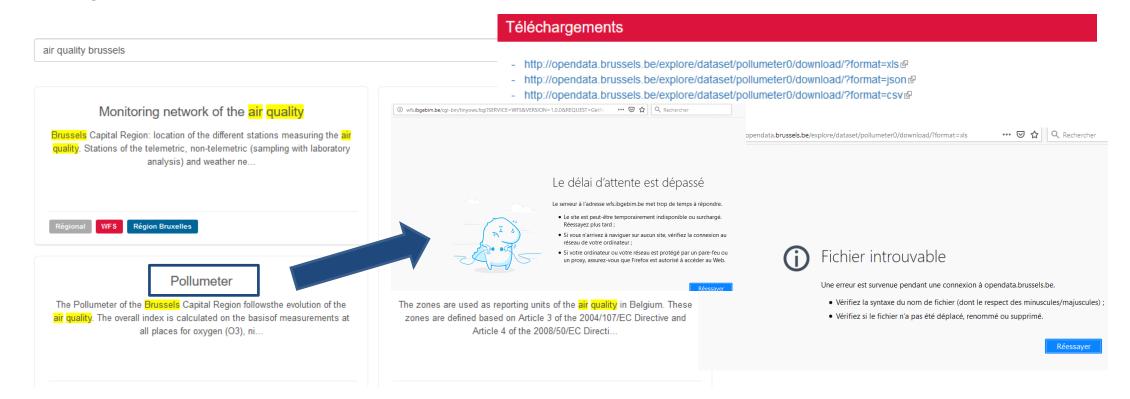




3. The technical challenge

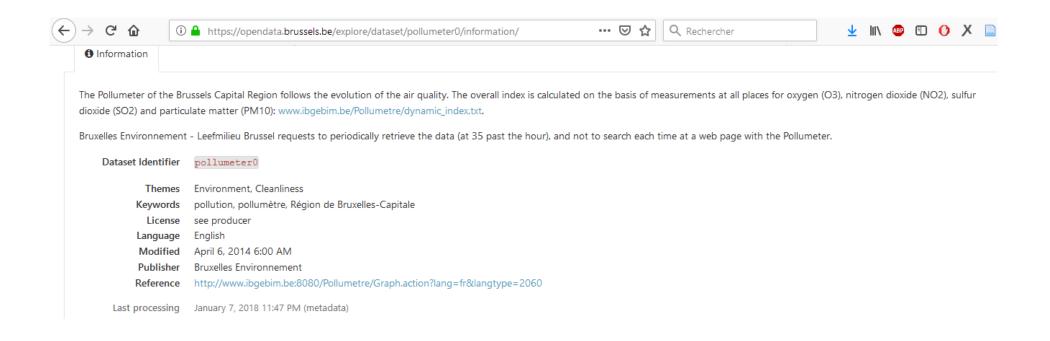
A. Where to find the right data?

Exploration of the possibilities through open data: data.gov.be, opendata.brussels.be, opendatastore.brussels









Not Found

The requested URL "/Pollumetre/dynamic_index.txt" was not found on this server.





Connexion Sinscrire



A / Jeux de données





lire plus

Thème





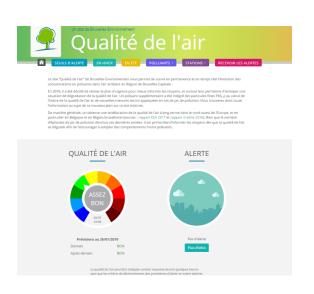
Régions et villes





2 access points: Bruxelles Environnement and IRCELINE

- Availability through HTML pages only, moving format (not made to be reused easily)
- Only unvalidated data (validated data non available)
- Implied scraping + regular monitoring + maintenance of the database to ensure accuracy, correctness







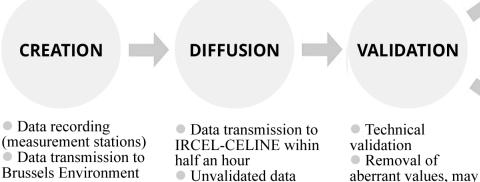


B. When to collect the right data at the right time?

- Moving averages VS fixed averages
- When fixed averages are updated (storage)?
- Requires knowledge of the application domain: 2 interviews with representatives of IRCELINE to understand the data life cycle (data management principle)
- Data are likely to evolve over time (correction of anomalies)

ARCHIVING

Online • Archiving by pollutant type (14 days) • Exceedings' history



published by **IRCEL-CELINE** occur a posteriori

aberrant values, may

Formal validation

Sending data to the European REGULATION authorities (annual)





4. The journalistic challenge

Overview of the situation

- Norms = arbitrary (what is true today can be false tomorrow, the principle of a norm is being violated or respected)
- EU Norms are criticized = compromise between members
 + influenced by industrial lobbying + not enough constraining + outdated
- **Incoherence** of the norms: daily limit rate NO2 = 200 µg/m³ never exceeded at the contrary of the yearly one = 40 µg/m³ systematically exceeded in Brussels
- Overruns in the Brussels' area regularly highlighted by the EU Commission
- WHO recommendations = more severe + outdated too (2005, in a revising process)





- 9 measurement stations at the time of the experience
- Arts Loi, not taken into account within the EU reporting (mandatory) but huge traffic, considered by the authorities as non representative of the ambient air pollution in Brussels for this reason
- Uccle taken into account within the EU reporting but very light traffic.

Choices made regarding public health considerations:

- 1) Comparisons with WHO recommendations
- 2) Comparisons with yearly average rate of NO2
- 3) Averages computed from the data of the 9 stations
- 4) Averages computed including and excluding Uccle
- 5) Critics of the air quality index © often good, rarely bad despite overruns

Technical choices = journalistic choices => not neutral All of the choices explained (online, published articles)

⊕ Bxl'air bot

Bulletin du 13 mai 2017

11h04. L'indice de qualité de l'air est de 2 sur une échelle de 1 à 10, communique Bruxelles Environnement. Le taux moyen de particules fines de type PM10 est de 10.5 µg/m3. Ce taux est endessous de la recommandation de l'Organisation mondiale de la santé (OMS), ainsi que sous la norme européenne. Le taux

moyen de particules fines de type PM2.5 est de 7.3 µg/m3. Ce taux se trouve sous la norme recommandée par l'OMS. Le taux moyen de particules fines de type black carbon (carbone suie) est de 1.4 µg/m3. Le ciel est très nuageux. La température observée est de 14,5 degrés.

Ce texte a été généré de manière automatique à partir de données extraites en







Bxl'air bot, le newsbot d'Alter Echos pour monitorer la qualité de l'air bruxellois



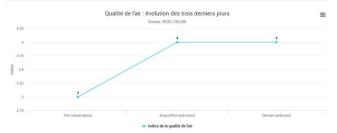
Taux de particules fines

Indice de la qualité de l'air

L'indice de la qualité de l'air est calculé tout au long de la journée, à partir de données récoltées par la Cellule Interrégionale d'environnement (CELINE). Cet indice est basé sur les taux de quatre polluants : ozone, dioxyde d'azote, dioxyde de soufre et particules fines. Il s'agit d'une appréciation qualitative (indice BelATMO), réalisée à partir de mesures techniques opérées en différents endroits tout au long de la journée. Sa valeur scientifique est très limitée, s'agissant avant tout d'un outil de

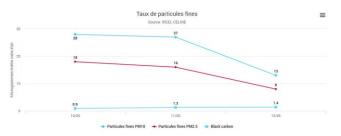


¥ Tweeter ☐ Partager 12



Taux moyen de particules fines

Trois types de particules fines font l'objet de relevés : les PM10, les PM2.5 et les Black carbon (carbone suie). i Voir les recommandations de l'OMS et les normes européennes



Ce graphique a été généré à partir de données extraites et enregistrées quotidiennement de manière automatique



bxlairbot

@bxlairbot Vous suit

Le newsbot qui surveille la qualité de l'air bruxellois. L'expérience est clôturée : compte Twitter fermé, site hors ligne dès le 04/04/19.

Bruxelles, Belgique • bxlairbot.be

5 4 5 5 12 135





Novembre 2017 Reicharger le jeu de données (CSV)						
Date	Indice	PM 10	PM 2.5	Black carbon	Ozone	Dioxyde ďazote
01	4	20	14	1,7	45	7
02	4	34	25	2.7	36	6
03	5	36	28	3.2	-48	8
04	3	18	14	1,5	46	5
05	3	11	9	1,1	50	5
06	4	20	12	3,3	41	9
07	4	33	21	4.6	40	10
08	5	44	35	4.2	5	7
09	5	39	31	2	29	5
10	3	15	10	1,2	39	7
11	3	16	13	1,7	38	7

Statistiques (du 01/04/2017 au 31/03/2018)

En bref pendant 28 jours. Il a été médiocre pendant

L'indice de qualité de l'air a été moven

14 jours. Il a très médiocre pendant 3 jours. Il a été mauvais pendant 3 jours. Il n'a pas atteint le seuil maximal de 10 (exécrable), 14 phases d'information (niveau d'alerte 0) ont été enregistrées en Région bruxelloise depuis le 1er avril 2017. En movenne, pour l'ensemble de la région, les recommandations de l'OMS ont été dépassées 2 fois en ce qui concerne le taux de particules fines de type PM10, et 20 fois en ce qui concerne celui des PM2.5. La recommandation de l'OMS relative au taux d'ozone a été dépassée à 28 reprises. La norme relative au taux de dioxyde d'azote n'a pas été dépassée (moyenne horaire). Le taux moyen de dioxyde d'azote est de 58,54 µg/m³. Il dépasse de 18.4µg/m² la norme relative à la moyenne annuelle. Le seuil de 1 μg/m³ de black carbon (carbone suie), à partir duquel ses effets sont considérés comme plus nocifs pour la santé que les PM10 et PM2.5, a été franchi 157 fois. Le taux de 3 µg/m³ de black carbon a été dépassé 8 fois.

Ce texte a été généré de manière automatique à partir de données publiques (open data) non validées enregistrées de manière quotidienne pendant un an, du 1er avril 2017 au 30 mars 2018. Celles-ci donnent donc une indication sur la qualité de l'air bruxellois sur une période de douze mois. Retrouvez un dossier spécial « qualité de l'air à Bruxelles » dans le numéro d'Alter Echos publié en mai 2018. Source : CELINE, Cellule interrégionale

Indice de la qualité de l'air

Septembre 2017 Septembre 2017 Très bon : 5 jours Bon: 15 jours Assez bon: 8 jours

Moven: 2 lours Moyennes par station

Sélectionner une période Septembre 2017



Maximums par polluant

65 µg/m³ à Haren, le 28/08

57 µg/m² à Berchem-Sainte-Agathe, le 03/03

9.3 μg/m³ à Haren, le 07/11

190 µg/m³ à Berchem-Sainte-Agathe, le 20/06 Dioxyde d'azote 168 µg/m³ à Arts Loi, le 28/08

Septembre 2017 Aucune phase d'information : 30 jours

▼ Tweeter

☐ Partager 41

Polluants atmosphériques

Septembre 2017

Niveau d'alerte

Septembre 2017

Septembre 2017

Moyenne (région bruxelloise)*

Aucun dépassement n'a été constaté.

La recommandation de l'OMS a été dépassée 2

Aucun dépassement n'a été constaté. Dioxyde d'azote

La recommandation de l'OMS et la norme européenne ont été dépassées 25 fois (moyenne annuelle). Aucun dépassement n'a été constaté moyenne horaire). La moyenne mensuelle est de 59µg/m³. Cette moyenne dépasse de 19µg/m³ les norme et recommandation (movenne annuelle).

Back carbon (carbone suie) : si ces très fines particules ne font l'objet d'aucune recommandation ou norme, le seuil de 1 µg/m3 a été dépassé 12 fois.

*Le mode de calcul s'appuie sur une moyenne par 24 heures non-validée. Une marge d'erreur est donc possible.

La "pollucarte" de la Région bruxelloise

Mise à jour une fois par heure. Dernier enregistrement : 13-07-2017 8:36:04

Toutes les stations localisées sur cette carte ne mesurent pas nécessairement les mêmes polluants atmosphériques. Pour l'ensemble des polluants, à l'exception du dioxyde d'azote (moyenne glissante pour 1 heure), il s'agit de moyennes glissantes pour 24 heures, elles évoluent donc pendant ce laps de temps, d'heure en heure, avant de devenir une moyenne fixe. Toutes les données figurant sur cette carte interactive proviennent de la Cellule interrégionale d'environnement (CELINE) et sont exprimées en microgrammes/mètre cube (µg/m³).













Dioxyde d'azote : 27

Ozone:15





Resic CENTRE DE RECHERCHE EN INFORMATION ET COMMUNICATION

5. Journalistic uses

Subject of journalism

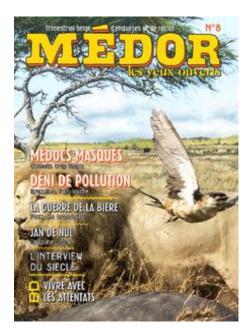
Sciences & Avenir, RTBF

Object of journalism

bxlairbot.be / @bxlairbot.be

Tool of journalism







"

Vous voulez vous renseigner sur la pollution en temps réel autour de chez vous ? C'est ce que propose Bxl'airbot, le robot-journaliste qui surveille la qualité de l'air.

En cas de pic de pollution, il donnera l'alerte. (...) Journaliste parce qu'il écrit des dépêches, qu'il illustre lui-même, à partir des données qu'on lui fournit. Et robot, parce qu'il le fait automatiquement. Il est même journaliste scientifique, puisqu'il fait de la vulgarisation.

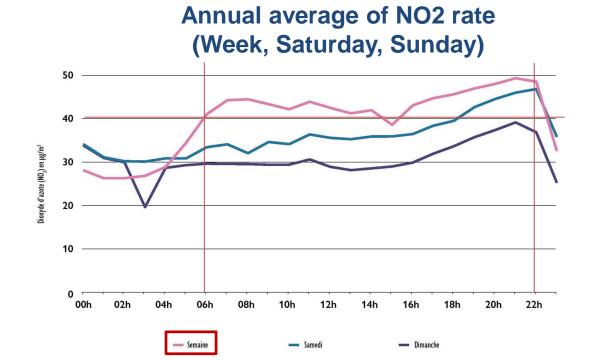
Sciences et Avenir





6. Feedbacks from journalists

- The bot did a job humanly difficult to do: added value of automation = repetitive tasks, gain of time = tool
- Exclusive information (analysis over time of raw facts)
- Make sense when it is integrated within human journalistic practices
- Recognize the subjectivity led by public health preoccupations
- Any data analysis activity cannot be considered as neutral: numbers as subjective as words = interpretation activity (challenge of any ddj project)





AlterEchos @AlterEchos · 11 juin 2018

Pendant douze mois, Alter Échos a accueilli le **Bxl'air bot** à la rédaction. Cette application de datajournalisme nous a aidés à enregistrer, compiler et compter des données sur la qualité de l'air. Le résultat? Pour vous, de... alterechos.be/un-an-avec-un-...





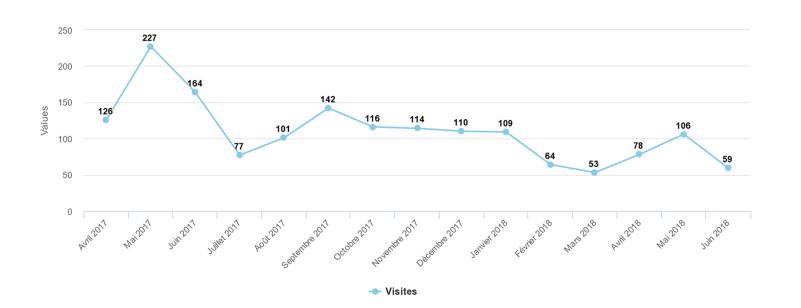
7. Online audience: limited but engaged

Analytics

- 1.646 visitors 2.312 sessions from 04-01-2017 to 06-30-2018
- Monthly average: 110 visitors, bounce rate: 57,40%
- 67,08% Belgium

Automated newsletter

- Launched after a reaction of a reader asking for an alert system based on WHO recommendations
- 50 subscribers







Twitter

- 4.084 tweets automatically pushed over 1 year
- Less than 100 followers
- 30 consultations of the profile/month
- 132 notifications
- 36 engagements/month







AlterEchos @AlterEchos · 27 avr. 2017

BXL'air bot et c'est le premier... fb.me/6a6paGr39

#DATAjournalisme #pollution #Bruxelles #Air Notre nouveau collègue s'appelle

Liévin Chemin @LievinChemin

@bruxselair @nicdas13 @BaptisteHupin @MdM_LeSoir @CleanAirBXL @sethvanhooland RT AIR bot from @ohmyshambles aka L. Diericks for @AlterEchos twitter.com/BRAL_Brussels/...

11 bxlairbot a retweeté

BRAL Brussels @BRAL_Brussels · 15 nov. 2017

En réponse à @jonashelseth @CleanAirBXL et 15 autres

using their stats tool bxlairbot.be/statistiques.p... (@bxlairbot author @ohmyshambles) bxl'AIR bot also delivers monthy #independent analyses of the threshold (both EU & WHO) overruns in Brussels air toxicity



Liévin Chemin @LievinChemin · 1 juin

En réponse à @BRAL_Brussels @AlterEchos et 2 autres





8. Feedbacks from audience

- Journalistic bias never highlighted
- Perceived as an independent information or as a service
- Interpretation of the data / vulgarization viewed as more credible than reinsuring communications (smileys) or more understandable / accessible than scientific communications



Un projet de datajournalisme indépendant qui rejoint les attentes des citoyens! Reprenant en continu les infos de la Cellule Interrégionale de l'Environnement (CELINE) et de Bruxelles Environnement, ce site vous fournit une interprétation digeste et raisonnée des données environnementales brutes. A vous de vous faire une idée!





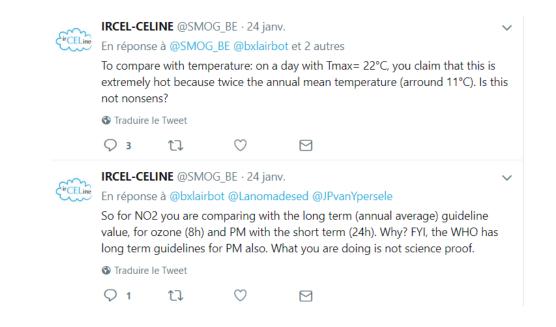
9. The data publisher: science VS journalism

- EU norms are the only referents (institutional isomorphism?)
- Journalists and developers are not scientists (misunderstandings, bad comparisons) = only the authorities able to well communicate
- Communication efforts to deliver their message about comparisons (fixed averages vs moving averages, daily vs yearly averages)
- Did not agree with NO2 comparisons (yearly basis)

"

Il y a des journalistes qui veulent absolument que l'article accroche et ils véhiculent des messages catastrophiques. On a parfois des problèmes avec ceux qui veulent à tout prix faire passer un message dramatique sur parfois un sujet qui ne l'est pas du tout.

Representative of IRCELINE







Conclusion

- No automation without the human factor
- Data must be accurate, reliable, up to date: necessity of a human maintenance of the database over time
- Difficulties because of the HTML (moving format) + no data = no bot
- Journalistic point of view led by public health considerations: as any editorial process, any choice is not neutral + interpretations of data
- Journalism VS communications of public authorities:
 lack of scientific expertise, journalists or developers not specialists
- Institutional isomorphism? Values of the norms are not in question, importance of the respect of the norms despite arbitrary character
- Limited but engaged audiences, taken as an independent and digest information, feedbacks rather positive





Observations du 01/04/2017 au 31/03/2018

L'indice de qualité de l'air a été moyen pendant 28 jours. Il a été médiocre pendant 14 jours. Il a très médiocre pendant 3 jours. Il a été mauvais pendant 3 jours. Il n'a pas atteint le seuil maximal de 10 (exécrable). 14 phases d'information (niveau d'alerte 0) ont été enregistrées en Région bruxelloise depuis le 1er avril 2017. En moyenne, pour l'ensemble de la région, les recommandations de l'OMS ont été dépassées 2 fois en ce qui concerne le taux de particules fines de type PM10, et 20 fois en ce qui concerne celui des PM2.5. La recommandation de l'OMS relative au taux d'ozone a été dépassée à 28 reprises. La norme relative au taux de dioxyde d'azote n'a pas été dépassée (moyenne horaire). Le taux moyen de dioxyde d'azote est de 58,54 µg/m³. Il dépasse de 18.4µg/m³ la norme relative à la moyenne annuelle. Le seuil de 1 µg/m³ de black carbon (carbone suie), à partir duquel ses effets sont considérés



comme plus nocifs pour la santé que les PM10 et PM2.5, a été franchi 157 fois. Le taux de 3 µg/m³ de black carbon a été dépassé 8 fois.

Ce texte a été généré de manière automatique à partir de données publiques (open data) non validées enregistrées de manière quotidienne pendant un an, du 1er avril 2017 au 30 mars 2018. Celles-ci donnent donc une indication sur la qualité de l'air bruxellois sur une période de douze mois. Source : CELINE, Cellule interrégionale d'Environnement.





To know more about the experience

- Dierickx, L. (2017). News bot for the newsroom: how building data quality indicators can support journalistic projects relying on real-time open data. In *Global Investigative Journalism Conference 2017 Academic Track*. Investigative Journalism Education Consortium.
- Dierickx, L. (2018). Du statut d'objet technique à celui d'objet social dans le contexte de la production automatisée d'informations journalistiques. *Terminal. Technologie de l'information, culture & société*, (122).
- Dierickx, L (2019). Why news automation fails?. In proceedings of *Computation* + *Journalism 2019*, February 2019, University of Miami.
- Dierickx, L. (2019). Automated news production within the uses and professional practices. *Ameaças ao Ciberjornalismo*, University of Lisbon (234).

http://ulb.academia.edu/LaurenceDierickx