

# Data-journalisme

## @ RTSInfo

### Une approche (-) technique

---

## Agenda

- Présentation
  - Datajournalism(e) / DDJ
  - Contexte @ RTSInfo
  - Contexte IT
  - Etudes de cas
  - Conclusion
- 

## Présentation

- Frédéric Vergez
  - @ikarius
  - fred@ikarius.com
- 

## Présentation

- ilem S.A
  - ~ 80 salariés
  - ex-Migros IT
  - "Rustique", mais honnête !
  - <http://ilemgroup.com>
- 

## Data-Journalism(e)

“

*Le journalisme de données (data journalism en anglais), ou journalisme de bases de données (database journalism), est un mouvement visant à renouveler le journalisme par l'exploitation et la mise à la disposition du public de données statistiques -- Wikipedia*

---

## Contexte

### Radio Télévision Suisse

- Branche langue française de la télévision et radio en Suisse (avec SRF, RSI, RTR, SWI)
- ~ 6000 employés
- 17 chaînes de radios
- 7 chaînes de télévision
- Budget: ~ 1.7mds CHF

---

## Contexte

### Actualité RTSInfo

<http://rts.ch/info>

- Site internet dédié à l'actualité
- un pool de 40 personnes, journalistes
- Synergies avec les émissions télé radio
- 1 "swiss-knife"

---

## Contexte "IT"

- des outils de production (parfois complexes): eScenic, Strada..
- une infra labyrinthique: Akamai, SwissTXT, CDN locaux ou pas
- une démarche forte "d'agilisation" de l'entreprise
- SCRUM essentiellement
- Scrum of Scrum aussi (mais ça marche pas)
- Kanban / Trello, plus que suffisant
- Contexte international

---

## Le "data-journaliste"

- est avant tout un journaliste!
- avec une sensibilité "data"
- avec des sensibilités et spécialisations différentes
- plus que le cursus, les journalistes sont surtout des gens curieux

---

## Dans les médias

- Les décodeurs du Monde
  - Mediapart
  - El Pais
  - Le Guardian
  - Last but not least, Le New-York times
  - ...
- 

## Le job de "data-journaliste"

- investigations et enquêtes journalistiques
  - données ou pas
  - dataviz ou pas
  - storytelling
  - preparation et analyse des données
  - articles, rédaction web
  - Suivi evenementiel live
  - Dossiers et projets de plus grande envergure (enquêtes)
- 

## Les résultats

- de jolis choses interactives:
    - <http://www.nytimes.com/interactive/2012/02/13/us/politics/2013-budget-proposal-graphic.html>
  - ou pas :
    - <http://projects.nytimes.com/madoff/>
    - <http://www.pewresearch.org/fact-tank/2014/12/29/our-favorite-pew-research-center-data-visualizations-from-2014/>
  - Des sujets importants:
    - <http://www.themigrantsfiles.com/>
- 

## Outils utilisés par les journalistes

D'une manière générale, - des outils spécialisés - online ou pas - mais pas de programmation

---

# Outils utilisés par les journalistes

- Tableau
  - Google Spreadsheets + maps
  - Excel(!)
  - Trifacta
  - Réseaux sociaux
- 

## Annonce

**"I'm not a journalist, I'm a IT guy!"**

---

## Le job de "data-analyst"

- difficile de qualifier précisément les intervenants IT dans un contexte de DDJ
  - "data-analyst" semble convenir
  - souvent une longue justification et qualification de poste ("on a toujours fait sans!", "on formera les journalistes")
  - Dev + Ops ? Architecte ? Agiliste ? Data-analyst/scientist ? Front, Back, DBA ?
  - Ben en fait, un peu tout à la fois, bien mixé...
- 

## Coté technique

- Concevoir et développer des WS / API REST
  - Concevoir, structurer et administrer des BD
  - Analyser et préparer les données
  - Extraction de données (scrapping)
  - BD orientées documents (Mongo, Cassandra)
  - BD orientées GIS (PG + PostGIS, Spatialite)
  - Systèmes GIS (QGIS, ArcGIS)
  - Frontend + viz
  - Responsive et/ou mobile natif
- 

## Outils

Généralement: - Python - R - D3 - Tableau - plot.ly - JavaScript - ArcGIS

---

## Outils

Moins généralement (les miens): - Clojure (code is data, data is code) - Incanter - ClojureScript - PostgreSQL + PostGIS - QGIS - Excel (oui, oui)

---

## Méthodes & vision

- Gestion de projet courts
  - Ou plus long et complexes
  - Interactions avec les services techniques (si existant)
  - HHD: High Human Disponibility
  - Plus ou moins agiles
  - Opendata en ligne de mire
- 

## Contraintes

“

*We're not dealing with "big-data", but with "not-so-big-data" and "smart-data"*

---

## Contraintes

“

*Data is beautiful but data needs space*

---

## Etude de cas 1

### Le train de Chlore

Lien vers l'article:

<http://bit.ly/23LD9eZ>

---

## Démarche

- Extrême dangeorisité du chlore, pas qu'en Suisse...

- Densité de population Suisse (BFS)
  - Trajet format KML, retravaillé (homogénéisation)
  - Données étude: Sourcetude
  - Etablir la densité et une idée du risque le long du trajet
- 

## Stack

- Geoding + reverse: Google Maps API
  - Processing / visualisation QGIS + PostGIS + Clojure/Incanter
  - Stockage: PostgreSQL + PostGIS (tracé, densité, communes...)
  - Backend: Clojure / HTTP-Kit + Nginx
  - Cloud: AWS S3 + RDS
- 

## Etude de cas 2

### Projet "EventLapse"

#### L'embrasement du moyen-orient

Version "complète" (article):

<http://bit.ly/1qF5aTq>

Version "simplifiée" (carte seule):

<http://static.ikarius.com/eventlapse>

- présenté au SRG-Hackdays, puis "renforcé"
- 

## Stack

- Clojure (WS)
  - CartoDB
  - Leaflet
  - Postgres + PotsGIS (tracés, densité)
- 

## Pour finir

- sensibilité artistique et curieuse ... tout comme les développeurs au fond
- position unique, complète et gratifiante ... mais très exigeante techniquement

- lente mutation des journalistes en "journaliste augmentés" ... impossible à l'inverse
  - ambiance riche et stimulante ... mais sous pression constante de l'actualité (ex. bruxelles)
- 

**Merci!**

**fred@ikarius.com - @ikarius**