

Open Source & Open Data

November 2020 – AFC Brussels

Bart Hanssens

BOSA DG Digital Transformation

BOSA.be

Proposed agenda

- Introduction
- Open Source
 - Licenses, business models ...
- Break
- Open Data
 - Challenges, community, ...
- Q & A

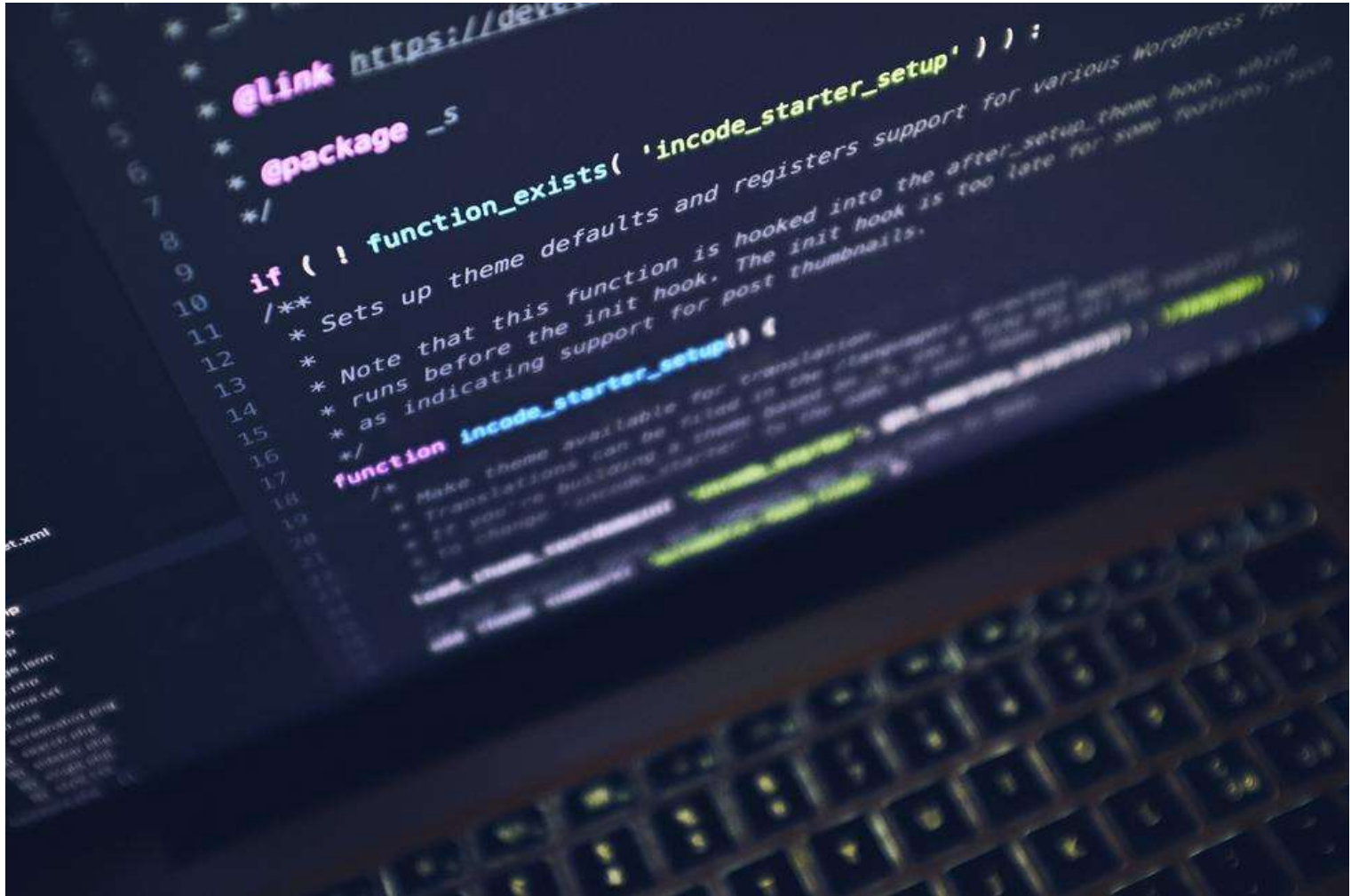
FPS BOSA DG Digital Transformation

- Formerly: Fedict
- Providing services to other federal public services
 - Secure exchange of data between administrations
 - Federal Authentication Service (FAS)
 - Web hosting platform
 - <https://dt.bosa.be/en>

Setting the scene

- Workshop expectations ?
- Who is already familiar with open source ?
 - As a user / as a developer
- Who is already familiar with open data ?
 - As a user / as a developer

Open Source



What is Open Source ?

- Open source / FLOSS
 - Versus proprietary
- Open projects: more than just free code
 - Questions / support ?
 - Dealing with 3rd party contributions
 - “Open” development: issue tracker, roadmap...

Picking “healthy” projects

- Is the code itself actively maintained ?
 - E.g. releases, changes, code commits
- How active is the community (if there is one) ?
 - Activity in mailing list / issue tracker
 - How many different contributors ?
- (References to users / other projects) ?

Open source projects are not that different

- Which programming languages / frameworks ?
 - Almost any (Java, .NET, Swift, Python, NodeJS, ...)
- Open source still needs
 - Documentation
 - Tests
 - Release notes

Open source licenses



Differences and similarities

- Not all licenses can be combined !
- What they have in common
 - No guarantees, non-exclusive, (mostly) irrevocable
 - Right to distribute
 - Can be used commercially
- Important differences between licenses
 - Patent usage protection or not
 - Distribution of source code or not
 - License(s) of software combinations

Types of licenses

- “Strong” copy-left licenses
 - GPL, EUPL, Affero GPL, ...
- “Weak” copy-left
 - LGPL, MPL, ...
- Permissive licenses
 - BSD, Apache, MIT, EPL ...
- <https://choosealicense.com/appendix/>
- <https://opensource.org/licenses>

Strong copy-left example: GNU/Linux

- GPL kernel + GNU “userland” applications
- Lots of additional kernel features developed
 - Different filesystems, firewall, network stack, ...
 - Proprietary drivers possible, but distributed separately

Weak copy-left example: eID middleware

- Belgian eID card middleware
 - <https://github.com/fedict/eid-mw>
- Lesser GPL (LGPL)
 - Can be used by proprietary applications

Permissive example: MacOS / iOS

- Mix of proprietary and open source parts
 - <https://developer.apple.com/opensource/>
- Darwin “core OS”
 - BSD licensed parts
 - Apple open source license ASPL
- Other parts and applications
 - Some Apache licensed parts

Code sharing platforms and communities



Code sharing / versioning

- Focus on sharing code, packaging
- Often include issue tracking, build services
- Free and commercial plans
- Github, Codeberg, SourceForge, ...

Foundations and communities

- Focus on process, consistency, license compatibility
- Often NPO with elected board members
- Use or provide own code sharing platform
 - Apache Foundation, Eclipse Foundation, Drupal.org...

Communication is key

- Open source projects are international
 - Users may not be native (English) speakers
 - Different cultures: direct vs indirect communication
- Often driven by volunteers
 - They are not (free) employees

Example: RDF4J project board on github

The screenshot displays a GitHub Project board for the 'eclipse / rdf4j' repository, specifically for the 'Project Progress' board. The board is organized into three columns: 'Inbox', 'Backlog', and 'In progress'. Each column contains a list of issues, each with a title, issue number, creator, and associated labels.

Inbox

- Add model-api and model-vocabulary to client pom** (#2648) opened by jeebroekstra. Labels: enhancement. Version: 3.5.0.
- implement LINDT UCUM units of measure** (#2643) opened by VladimirAlexiev. Label: enhancement.
- Document use of ValidationException as generic way to get (remote or local) validation report** (#2641) opened by jeebroekstra. Labels: website / documentati..., SHACL.

Backlog (272 items)

- Streamline SPARQL parser** (#2580) opened by jeebroekstra. Labels: sparql, enhancement. Progress: 0 of 3.
- SPARQLConnection.size should use aggregate query** (#2463) opened by jeebroekstra. Labels: performance, enhancement.
- Console: command / filename completion** (#2527) opened by barthanssens. Labels: console, enhancement.
- Upgrade to java.time API for date and time handling** (#2563) opened by jeebroekstra. Labels: epic, model, enhancement. Progress: 0 of 3.
- Support complex wildcard language** (no number visible).

In progress (9 items)

- SHACL - create new AST that supports the full spectrum of SHACL** (#2083) opened by hmottestad. Labels: SHACL, enhancement.
- NativeStore datastore benchmarks** (#2151) opened by hmottestad. Label: internal task. 1 linked pull request.
- Allow HTTP authentication to be configured in repository configuration templates** (#528) opened by jeebroekstra. Labels: server, enhancement. Version: 3.5.0.
- Docker build should act like the maven build and use the local code instead of downloading a different version of RDF4J** (#2374) opened by hmottestad. Label: internal task.

Example: Eclipse Foundation



The screenshot shows the Eclipse Foundation website homepage. The background is a dark, low-angle photograph of several skyscrapers reaching towards a bright sky. The Eclipse Foundation logo is in the top left corner. The navigation menu includes 'Projects', 'Working Groups', 'Members', 'More', and a search icon. A 'Download' button is in the top right. The main headline reads 'The Home of Entrepreneurial Open Source' in orange, followed by 'Community driven.' and 'Code first.' in white, and 'Commercially-friendly.' in orange. A breadcrumb trail below the headline reads 'Home / About Us / The Open Innovation Community'. The main content area contains a paragraph about the foundation's governance framework and a 'Join Us' button. A decorative network graphic is in the bottom right corner.

ECLIPSE
FOUNDATION

Projects Working Groups Members More - Q - [Download](#)

The Home of Entrepreneurial Open Source
Community driven.
Code first.
Commercially-friendly.

Home / About Us / The Open Innovation Community

The Eclipse Foundation provides a proven governance framework and processes for entrepreneurial collaboration on sustainable, commercializable open source software that benefits developers, companies, and users alike.

[Join Us](#)

Open source business models



Dual-licensing

- Multiple licenses for the same code
 - Only the copyright holder can do this
 - E.g. GPL for use in other open source projects...
 - ... and commercial license for “closed” source
- Examples:
 - MySQL Database: GPL and Oracle license
 - iText PDF library: AGPL and proprietary license

Additional services or modules

- Training, extra components, integrations ...
 - Odoo, Alfresco, Pentaho, ...
- Providing “stable” combinations, long-term support
 - RedHat Enterprise Linux, Oracle JDK, ...

Cost savings / synergies

- **Sharing cost of development and maintenance**
 - Companies contributing to tools they use themselves
 - Android, Python, Chromium, Eclipse IDE ...
- **When software is not the core business**
 - Selling hardware or (more profitable) services
 - Network drivers, Kubernetes ...

For marketing, fun and/or for the greater good

- Software developed by/for universities, governments
 - Paid with taxpayers' money
 - E.g. BSD, Decidim, Accumulo ...
- Community projects
 - For fun and glory
 - E.g. GIMP, PostgreSQL, VLC, ...
- Some “fun” projects became quite big and profitable
 - E.g. Linux, PHP, Drupal, ...

Miscellaneous links on open source



Links

- EC Open Source Software Strategy
 - https://ec.europa.eu/info/departments/informatics/open-source-software-strategy_en
- “Blue hats” community (French gov community)
 - <https://www.modernisation.gouv.fr/le-hub-des-communautes/blue-hats>
- Annual FOSDEM conference
 - <https://fosdem.org>

Open Data



What is open data ?

- Free, easy to reuse, non-sensitive data
 - No personal data (GDPR !)
 - Download files or webservice / API
- Who publishes open data ?
 - Governments (local / regional / federal)
 - Companies (voluntary basis)
 - Citizens / crowd-sourcing projects (voluntary basis)

Legal framework

- EU Public Sector Information / Open Data Directive
 - <https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information>
- Transposed into federal law and regional decrees

What kind of data ? Almost any !

- A lot of statistics
 - Population, consumer price index, average income ...
- Transport and mobility
 - Train delays, parking spots, ...
- Environment
 - Air quality, biodiversity, ...
- Geospatial data
 - Satellite data, maps, ...

Formats

- No single format for each and every type of data
 - Depends on application / systems / conventions
- Preference for open, structured file formats
 - CSV, GTFS, (Geo)JSON, KML, ODS, sqlite, XLSX, XML ...
- (RDF / Linked Data)
- Known / documented APIs
 - OGC GIS, OpenAPI/Swagger, SPARQL ...
- Availability and meaning often more important

Open data licenses



Quick tour of licenses

- Well-known licenses

- Various Creative Commons, Open Data Commons
- <https://creativecommons.org/about/cclicenses/>
- <https://opendatacommons.org/licenses/>

- Legal situation in Belgium

- Status: it's complicated (federated state)
- Quite a few slightly different custom licenses ...
- ... many of them are very similar to CC-licenses

Differences and similarities

- Not all licenses can be combined !
- What they have in common
 - No guarantees, non-exclusive, (mostly) irrevocable
 - Right to distribute
- Important differences between licenses
 - Can be used commercially or not
 - Attribution required or not
 - License of combined datasets

Open data business models



Additional services and data

- Combining and transforming data
- Providing Service Level Agreements
- Additional services
 - E.g. user-friendly search, visualizations, reports...
- Examples
 - Law: <https://lex.be> , <https://jura.kluwer.be>
 - Companies: <https://opencorporates.com>

Example: MapTiler with OpenStreetMap data

[Home](#)

Contact

Algemeen adres

FOD Beleid en Ondersteuning

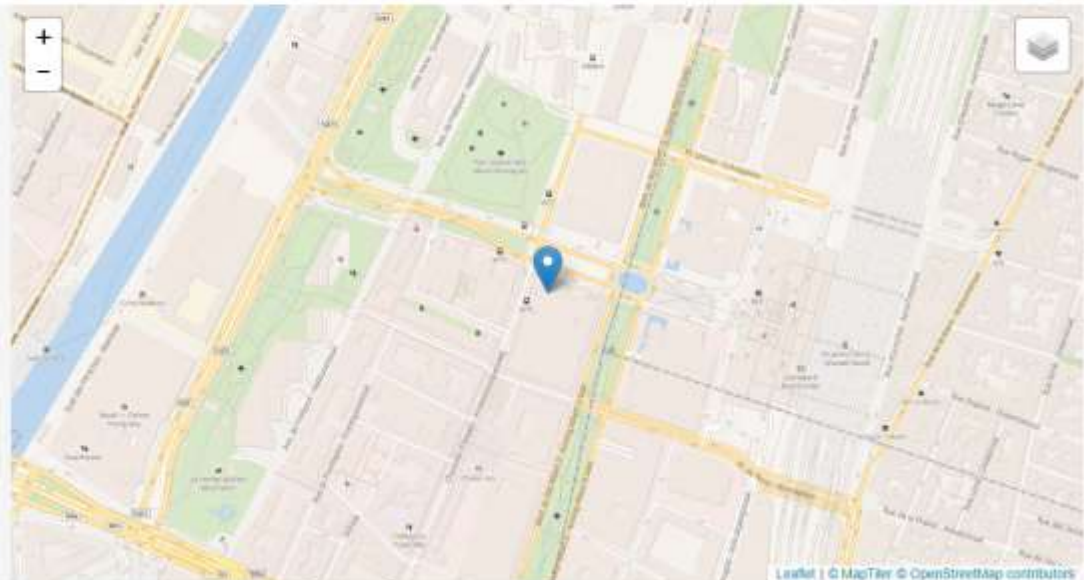
Simon Bolivarlaan 30, bus 1
WTC III
1000 Brussel

info@bosa.fgov.be

+32 (0)2 740 74 74, van maandag tot
vrijdag van 8u tot 17u

🗺️ Routebeschrijving

Bekijk de ligging en bereikbaarheid [via](#)
[Google Maps](#) of in [dit document](#).



Where to find open data ?

- Websites and (open data) portals
 - <https://data.gov.be>
- If you can't find it... just ask !
 - opendata@belgium.be
 - @BartHanssens
 - #data.gov.be:matrix.org

Open data portals

NL EN FR DE

Other official information and services: www.belgium.be 

 **Data.gov.be**

[HOME](#) [DATA](#) [APPS](#) [NEWS](#) [INFO & FAQ](#) [API/RSS](#) [CONTACT](#)

Data, tools and resources. More than 10,000 datasets.



Agriculture and
Fisheries



Culture and
Sports



Economy and
Finance



Education



Energy



Environment



Health



International



Justice



Population



Public Sector



Regional



Science and
Technology

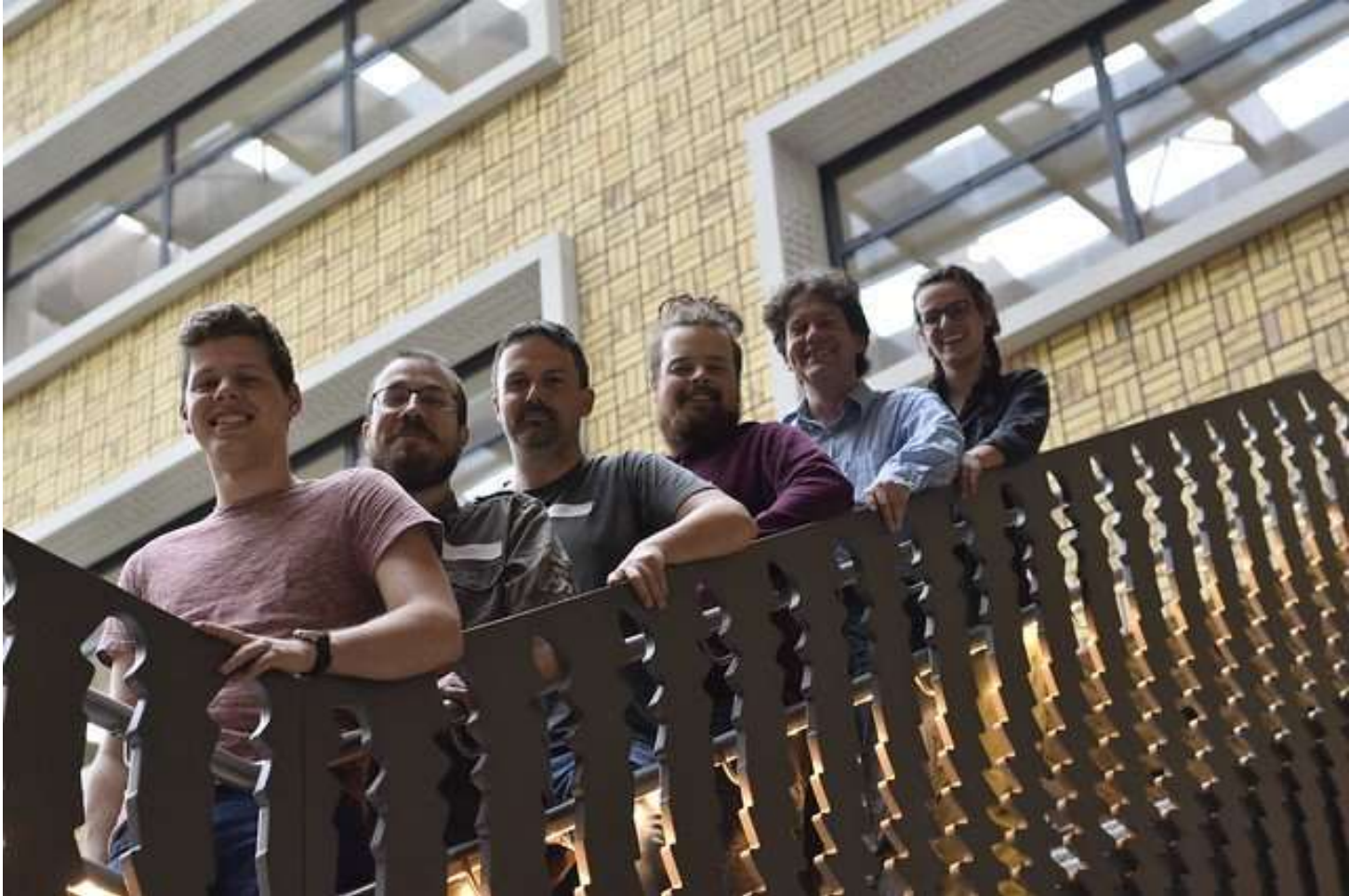


Transport

Data.gov.be and the Open Data Task Force

- Open Data Task Force: FPS BOSA DG DT + DAV/ASA
- Portal is “by-product” of full (meta)data export
 - See also <https://github.com/fedict/dcat>
- “No Wrong Door” principle
 - Forward questions to Regions / Cities if needed

Combining open source and open data



Open Summer of Code

- Organized by OpenKnowledge Belgium
 - + partners from government and private sector
 - <https://summerofcode.be/>
- Students get paid to build open source projects
 - Often projects which also use open data
 - Mostly IT, but also business, design, marketing, ...
 - 3 x 4 day in July
 - With coaches, presentations on soft skills, ...

Example: BeST Addresses

- Addresses + geo location from the 3 Regions
 - Published and updated weekly by BOSA DG DT
 - <https://opendata.bosa.be/>
- 2019 oSoC project
 - Python tools developed by students
 - Data also available on <https://openaddresses.io>
 - <https://github.com/oSoc19/best>

Open data communities

- OpenKnowledge Belgium
 - Local “chapter” of OpenKnowledge
 - <https://be.okfn.org>
- OSGeo Belgium
 - Open data and open source
 - <https://www.osgeo.org/local-chapters/osgeo-belgium/>
- Open Justice
 - <https://openjustice.be>

Other community initiatives



Data-driven projects

- Mozilla CommonVoice
 - <https://commonvoice.mozilla.org/en/about>
- Telraam traffic counters
 - <https://telraam.net/en>
- Global low-power data transfer / IoT
 - <https://www.thethingsnetwork.org>

Share your open data stories !

- Are you working on a thesis / paper on open data ?
- Did you create a stunning app / service ?
- Are you organizing an open data-driven event ?
- Let us know, so we can:
 - Tweet and make some fuzz
 - Promote your work
 - Convince more organizations to open data

Questions ?



Thank you !

@BartHanssens

<https://data.gov.be>

opendata@belgium.be

[#data.gov.be:matrix.org](https://matrix.org)

BOSA.be