UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES
Département d'astronomie

ISDC

Low-latency Smart ISDC

Volodymyr Savchenko

Low-latency alerts & Data analysis for Multi-messenger Astrophysics

14/01/2021, Paris (remote)

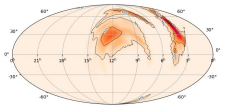
INTEGRAL/ISDC

Launched in 2002, currently till March 2023 (might overlap with O4).

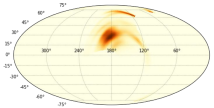
I work for the INTEGRAL Science Data Center, advocated and lead INTEGRAL multi-messenger activities.

Working primarily project-focused and supporting synergies for the cause, occasionally leading to some reusable developments.

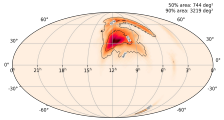
Relying on **open standards** and schemas like those of **IVOA** to facilitate organized action. Let's collaborate on tools! Build tools for reuse.



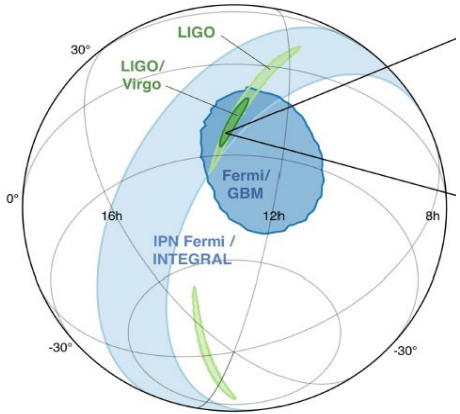
GBM + LIGO/Virgo



GBM + LIGO/Virgo + INTEGRAL
no-detection constrain

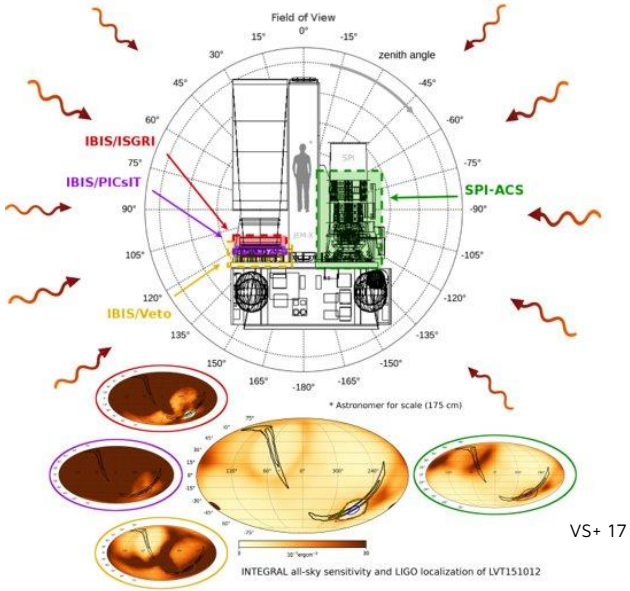


GBM + LIGO/Virgo
Improved
confirms INTEGRAL



For INTEGRAL: VS+ 17

Multi-mission localizations. Kevin H. nazgul



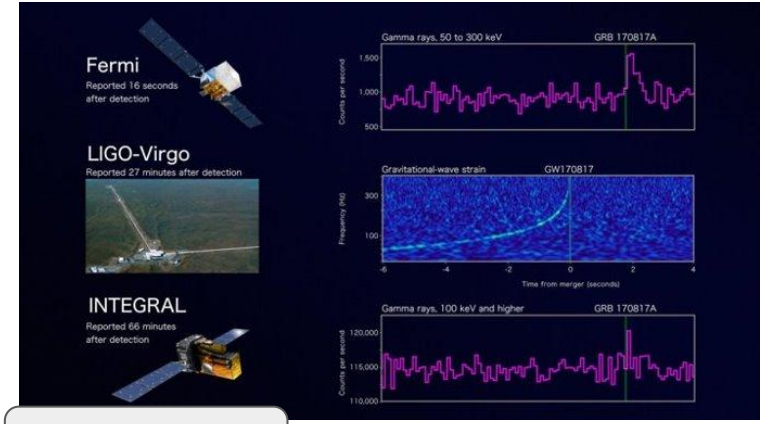
VS+ 17

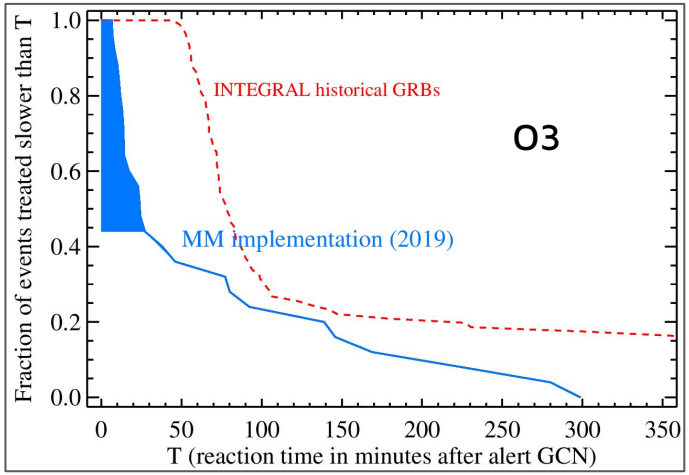
INTEGRAL is so INternational that combining its own instrument data is a challenge. Need experts in every instrument to work together! (part of ISDC mandate is to make them do so)

This is a challenge!

See talk of CF for introduction to INTEGRAL and review of scientific impact of my work.

INTEGRAL/ISDC and the world at lower latency





Data available with **20 s delay**. Extra 10 min time to quantify the background (interplay between **latency and quality**)

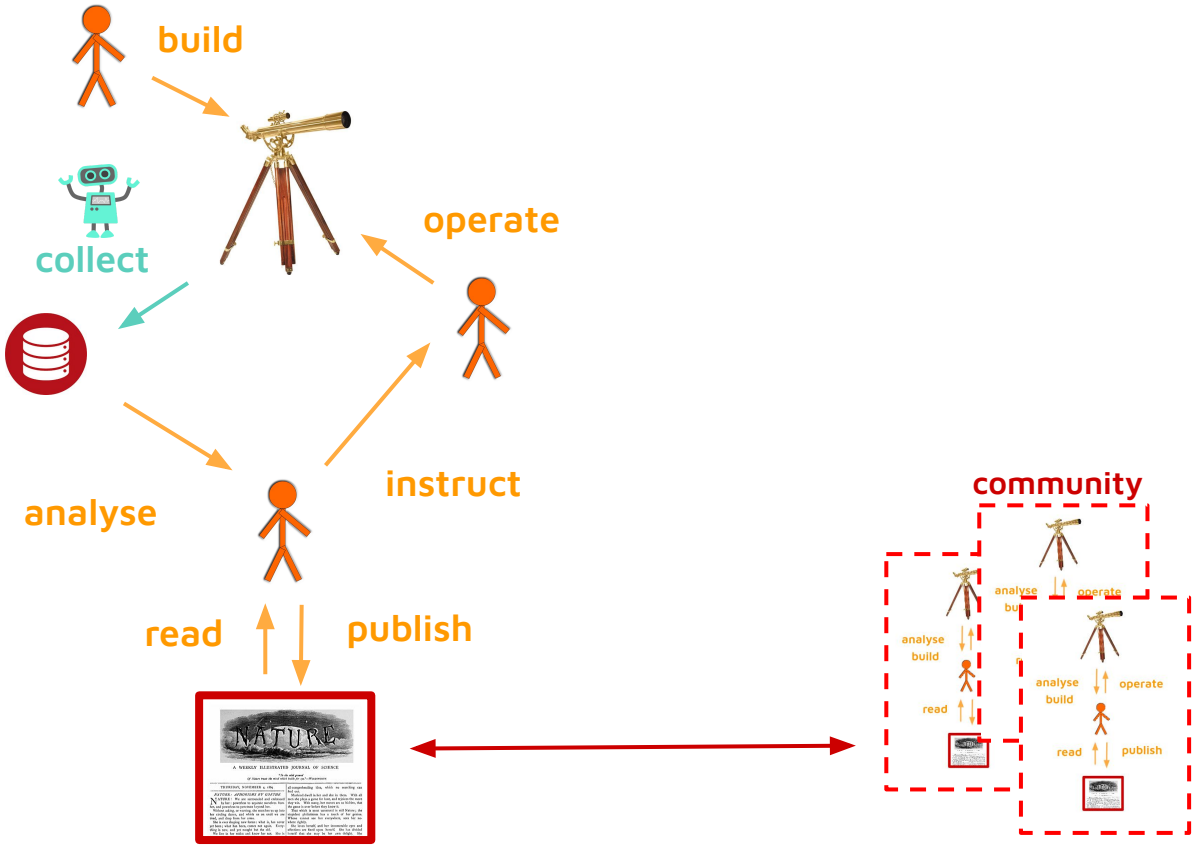
More dirty results faster? IBAS (especially ACS) does not wait and produces plenty of false alerts.

Is this about building better automation/pipelines? In part, but also why this is hard, especially with available manpower.

Complementary approach is to **share more data**, and we will very soon share **20 s latency** data with the community.

High latency, Low bandwidth, High creativity, High fidelity

- Reaction to sky: **slow**
- Reaction to papers: **slow**
- Trials (p-hacking): **uncontrolled**
- Publishing: **slow**
- Scalability: **bad**
- Creativity: **high**
- Communication: **clear and nuanced but imprecise and slow**

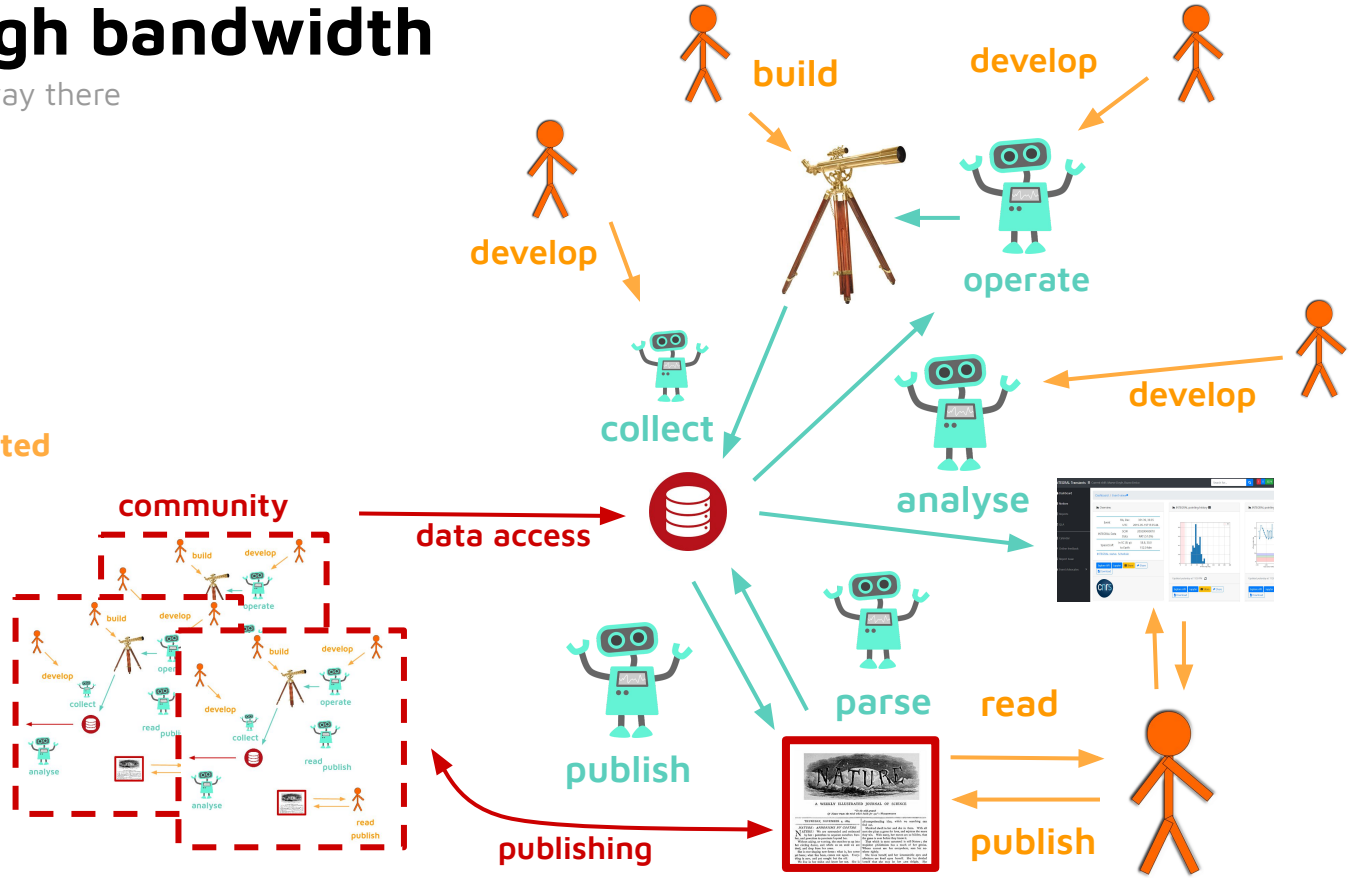


Human reaction and processing **is slow**, even if it's within even one person. But people are **smart**

Low latency, high bandwidth

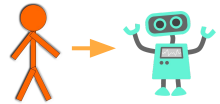
"Ideal" picture: most reality halfway there

- Reaction to sky: **fast**
- Reaction to papers: **fast**
- Trials (p-hacking): **controlled**
- Publishing: **fast**
- Scaling: **good**
- Creativity: **low**
- Communication: **precise, limited**



- **Making smart robots is hard**: always lacking **developers who are also research scientists.**
- If all is automated, **scientists have hard time seeing what's going on**, since **they do not speak robot**
- Robots are **fast**, but **lack creative reaction** in **new situations**.

Development space: help scientists make robots



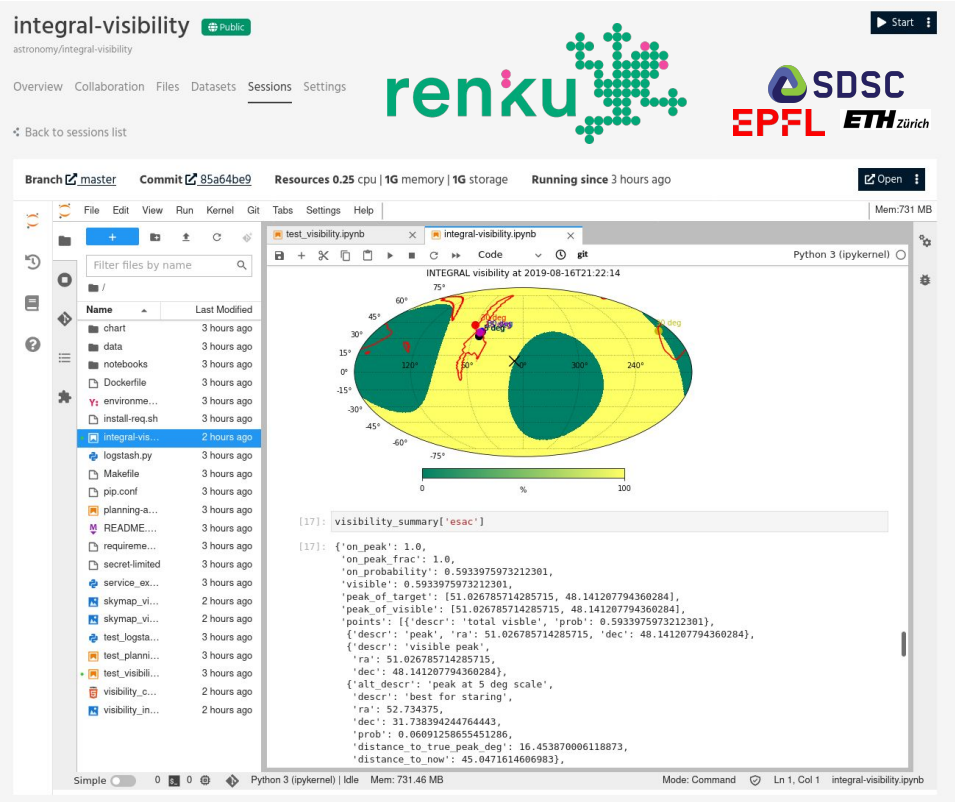
There are much more scientists who can make a jupyter notebook than write organized code.

Google-collab, binder, ESA DataLabs

We use Swiss-made **Renku** platform, to:

- **Continuous integration and testing**
- Supports in **publishing of data and code** (e.g. in zenodo)
- Support in annotation for scientists and robots reuse! **FAIR!**

Develops and integrates metadata in a Knowledge Graph (see later)



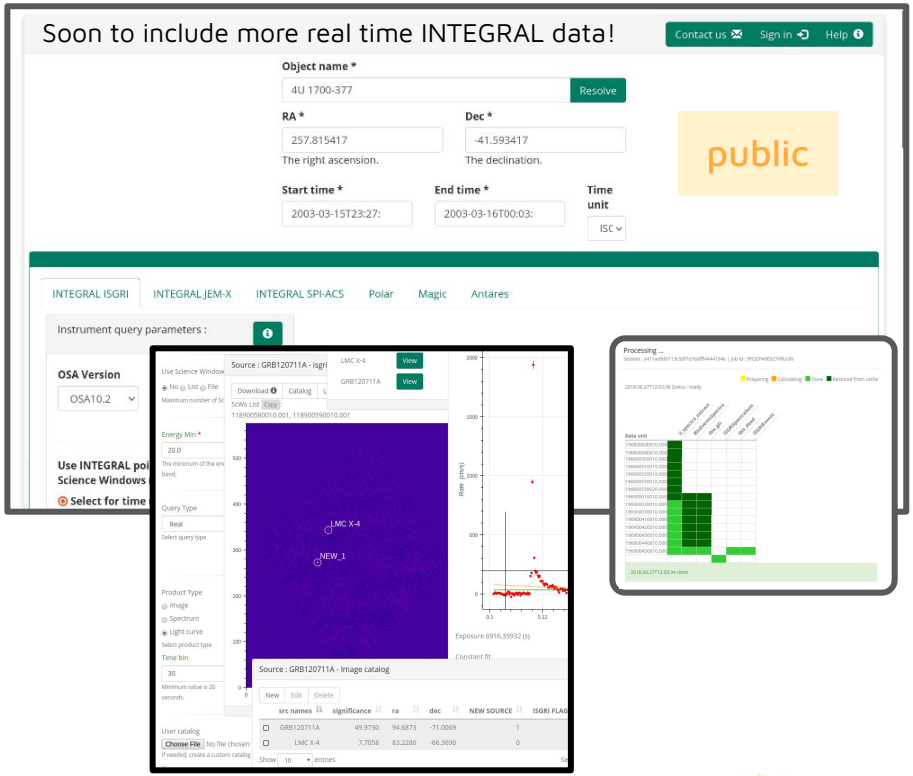
https://github.com/oda-hub/

Tools for exploring research data

See talk of AN later today for more extended review of MM(O)DA

https://www.astro.unige.ch/mmoda

Deploy your own with **helm charts/kubernetes** https://github.com/oda-hub



Integrating with Knowledge Graphs

https://marketplace.eosc-portal.eu/services/astronomical-online-data-analysis-astrooda

EUROPEAN OPEN SCIENCE CLOUD

Add “creativity”: Linked Open Data Knowledge Graphs

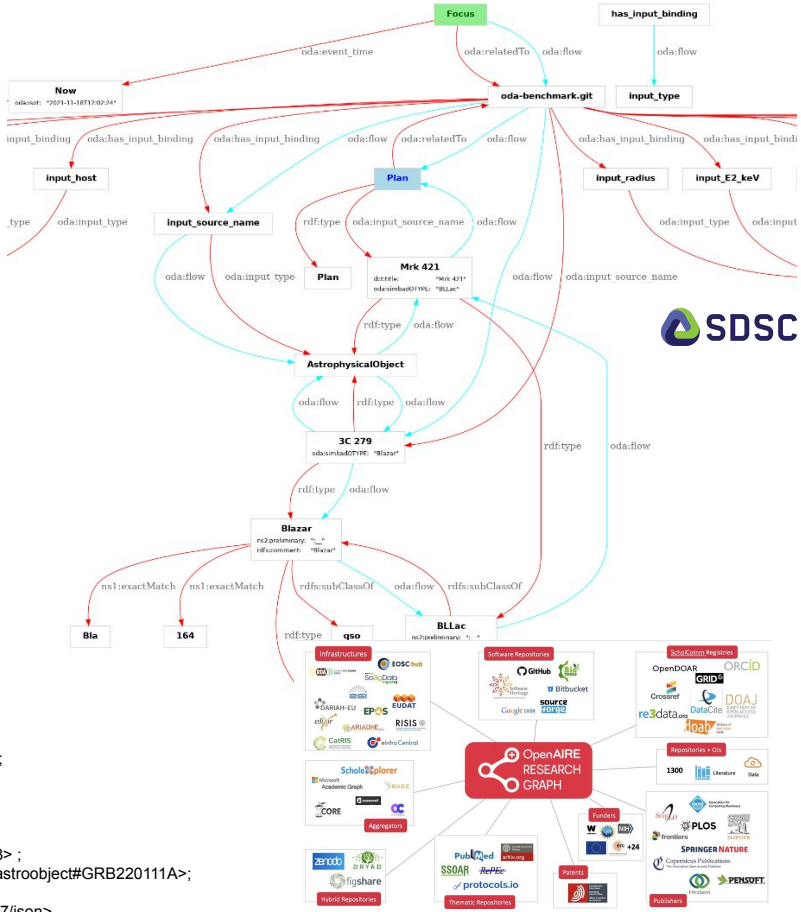
People know a lot, and form free associations. Robots have much information too. E.g. much insight is reported in GCN Circulars but only accessible to people.

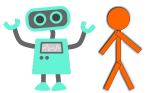
- **Global linked identifiers** - **URI** (ivo://, http://, ..): building common vocabularies. URIs point to documents, workflows, data, astro objects
- Explain possible **relations** in **ontologies**
- Embedding and following references, to express connections between different URIs

We develop tools:

- **scrobbling** gcns, atels, tns, arxiv: I wish people would put more references to structured data like URL to json, but we try to **consume graciously**
- for **annotating** and **publishing** integrate **code/workflows** with **data**: making sure we **produce cautiously**
- Using **graph relations** to rank and **optimize publication production**

With this, we recently wrapped up INTEGRAL FoV calibration. We publish next INTEGRAL MM (GWTC3, recent IceCube) catalog with RDF metadata.





JSON-LD

XML

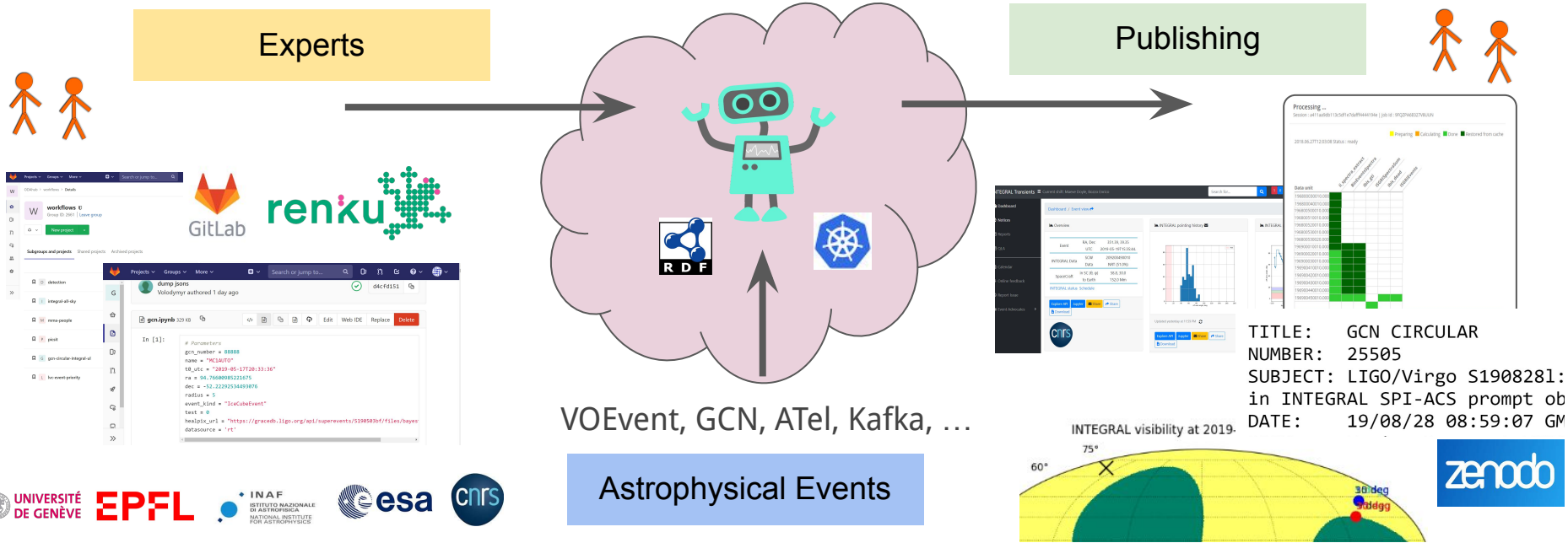
```
paper:gcn31435 paper:NUMBER "31435" ;
        paper:SUBJECT "GRB 220111A: BALROG localization …";
        paper:balrog_dec 6.380345e-01;
        paper:balrog_ra 1.498846e+02;
        paper:gbm_trigger_id 663621714 ;
        paper:grb_isot "2022-01-11T19:21:49.430000" ;
        paper:location <https://gcn.gsfc.nasa.gov/gcn3/31435.gcn3> ;
        paper:mentions_named_event <http://odahub.io/ontology/astroobject#GRB220111A>;
        paper:source "GCN" ;
        paper:url_json <https://grb.mpe.mpg.de/grb/GRB220111807/json> .
```

https://linked-open-data.space/ - small test example

FAIR INTEGRAL Astrophysical Transient Analysis

Multi-messenger transient astronomy puts especially high demands on **confusion-free low-latency interoperability**, which we addressed by developing two key components:

- **Environment** to develop, test, and integrate data reduction, theoretical models, statistical methods, spacecraft operation tools
- **Knowledge-Graph-enabled engine** finds combinations of data, adapters, statistical methods, publishers, planning, and **disseminates provenance-tracked standard results** along with public data and code



UNIVERSITÉ DE GENÈVE EPFL INAF ISTITUTO NAZIONALE DI ASTROFISICA NATIONAL INSTITUTE FOR ASTROPHYSICS esa cnrs

Conclusions

- **Low-latency** is hard to reconcile with **high-quality** and especially **high-creativity**, and costs **development effort**
- Sharing is key, and low-latency, real time SPI-ACS data (20 s latency, in 8 s blocks) soon available on https://www.astro.unige.ch/mmoda/ will be announced by a GCN.
- We make open tools for data centers (other projects also do) focused on added value analysis, aligned with software industry developments and EOSC:
  - Online analysis tools and tools and components: to help non-developers **explore**
  - Development, testing, benchmarking, integration tools: to help non-developers **contribute**
  - Building KG-enabled knowledge base, integrating papers, workflows, data, and learning from this graph
- Let's interoperate based on standards and tools! https://github.com/oda-hub

Sustainable goals of equitable access, transparency in decision making, see talk