PSDI

PHYSICAL SCIENCES DATA INFRASTRUCTURE

The Semantic Web is Dead – Long Live the Semantic Web

The Future of Semantics in the Physical Sciences (and beyond!)

CHEMBL 15 Year Symposium

2nd October 2024

Dr Samantha Pearman-Kanza
University of Southampton

https://www.psdi.ac.uk/

Presentation Outline

- About Me & PSDI
- Inception, Perceptions & Misconceptions of the Semantic Web
- The Semantic Web in the Physical Sciences –Barriers & Challenges
- Mitigations & Suggestions for Best Practice
- The Future of Semantics in the Physical Sciences (and beyond)
- Thoughts for the Future

https://www.pinterest.co.uk/pin/539306124105554567/

PHYSICAL SCIENCES DATA INFRASTRUCTURE

About Me

- Computer Scientist turned Web Scientist
- Senior Enterprise Fellow at University of Southampton
- Pathfinder Lead on Process Recording for PSDI
- Researcher on Process Recording for PSDI & AIChemy Hub
- Lab Horizons Columnist: CompSci Cat
- Research Interests: Semantic Web Technologies, IoT, Research Data Management, Digitisation, Lab of the Future, Paperless Labs, Re-use of Technology
- @SamiKanza

The Semantic Web is so much more than just a concept or a technology, it's a philosophy, a way of life, and above all,

What is PSDI?

PSDI
PHYSICAL SCIENCES
DATA INFRASTRUCTURE

PSDI
PHYSICAL SCIENCES
DATA INFRASTRUCTURE

Physical Sciences Data Infrastructure

An Integrated Data Infrastructure for the Physical Sciences

PSDI aims to accelerate research in the physical sciences by providing a data infrastructure that brings together and builds upon the various data systems researchers currently use.

- UK funded project through the UKRI Digital Research Infrastructure theme (DRI) via EPSRC
- Developed out of a community statement of need
- Lead institutions: University of Southampton and Science and Technology Facilities Council (STFC)
- Still predominantly in requirements gathering and developing proof of concept
- Launch of resources in early 2025

Aim of PSDI

Data is a major driver of research in Physical Sciences

Driven by **community** needs, **PSDI** will provide

A data infrastructure that

connects existing

experimental and computational facilities

within Physical Sciences and beyond

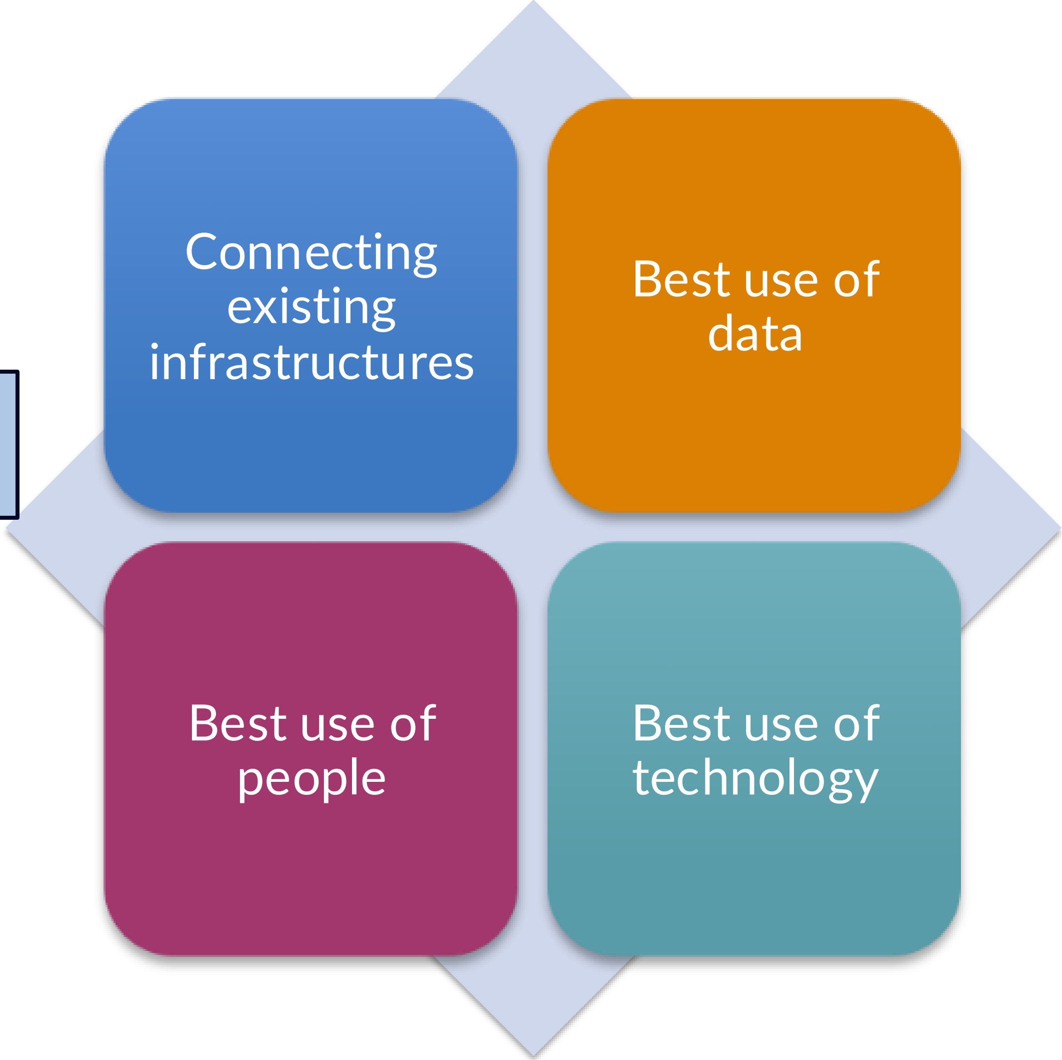
- A platform for data collection, sharing, aggregation, integration and curation
- Supporting analysis across experimental, simulation and reference data
- Combining and enhancing existing data infrastructures
- Sustaining data resources beyond lifespan of individual research projects

Transitioning physical science to a fully digitally enabled research discipline

PSDI Principles & Focus Areas

Our current (but growing) focus is on:

- Process Recording
- Metadata & Semantic web technologies
- Data Management Plans & Data Skills
- Access to trusted data resources
- Scientific Workflows



Domain Exemplars:

- Biomolecular simulation
- Catalysis
- Machine learnt interatomic-potentials
- Materials
- NMR Crystallography

Full recommendations at: https://www.psdi.ac.uk/the-pilot/recommendations

Inception, Preconceptions & Misconceptions of the Semantic Web

PSDI
PHYSICAL SCIENCES
DATA INFRASTRUCTURE

Conclusions

- **The Semantic Web is coming!**
  - Joint development between DARPA/EU/and W3C communities
  - Languages and tools are available to play with
    - Http://www.daml.org/
  - W3C interest group available for those wishing to join the discussion
    - Www-rdf-logic@w3c.org (live or archived)
  - Ongoing DoD and commercial projects
- **Come join us**
  - Submit ontologies/marked up pages
  - Develop tools or help test ours
- **Get in on the next big thing early!**

www.daml.org

WI-2001: Maebashi, Japan

MIND

https://www.wi-consortium.org/wicweb/pdf/wi-hendler.pdf

SCIENTIFIC AMERICAN
Main Menu | Interview | Bookmarks | Feedback
Current Issue | Explore! | Ask the Experts | Marketplace | Search the Site
FEATURE ARTICLES

The Semantic Web

A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities

by TIM BERNERS-LEE, JAMES HENDLER and ORA LASSILA

The entertainment system was belting out the Beatles' "We Can Work It Out" when the phone rang. When Pete answered, his phone turned the sound down by sending a message to all the other *local* devices that had a *volume control*. His sister, Lucy, was on the line from the doctor's office: "Mom needs to see a specialist and then has to have a series of physical therapy sessions. Biweekly or something. I'm going to have my agent set up the appointments." Pete immediately agreed to share the chauffeuring. At the doctor's office, Lucy instructed her Semantic Web agent through her handheld Web browser. The agent promptly retrieved information about Mom's *prescribed treatment* from the doctor's agent, looked up several lists of *providers*, and checked for the ones *in-plan* for Mom's insurance within a *20-mile radius* of her *home* and with a *rating* of *excellent* or *very good* on trusted rating services. It then began trying to find a match between available *appointment times* (supplied by the agents of individual providers through their Web sites) and Pete's and Lucy's busy schedules. (The emphasized keywords indicate terms whose semantics, or meaning, were defined for the agent through the Semantic Web.)

In a few minutes the agent presented them with a plan. Pete didn't like it—University Hospital was all the way across town from Mom's place, and he'd be driving back in the middle of rush hour. He set his own agent to redo the search with stricter preferences about *location* and *time*. Lucy's agent, having *complete trust* in Pete's agent in the context of the present task, automatically assisted by supplying access certificates and shortcuts to the data it had already sorted through.

Almost instantly the new plan was presented: a much closer clinic and earlier times—but there were two warning notes. First, Pete would have to reschedule a couple of his *less important* appointments. He checked what they were—not a problem. The other was something about the insurance company's list failing to include this provider under *physical therapists*: "Service type and insurance plan status securely verified by other means," the agent reassured him. "(Details?)"

Lucy registered her assent at about the same moment Pete was muttering, "Spare me the details," and it was all set. (Of course, Pete couldn't resist the details and later that night had his agent explain how it had found that provider even though it wasn't on the proper list.)

Expressing Meaning

Pete and Lucy could use their agents to carry out all these tasks thanks not to the World Wide Web of today but rather the Semantic Web that it will evolve into tomorrow. Most of the Web's content today is designed for humans to read, not for computer programs to manipulate meaningfully. Computers can adeptly parse Web pages for layout and routine processing—here a header, there a link to another page—but in general, computers have no reliable way to process the semantics: this is the home page of the Hartman and Strauss Physio Clinic, this link goes to Dr. Hartman's curriculum vitae.

The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users. Such an agent coming to the clinic's Web page will know not just that the page has keywords such as "treatment, medicine, physical, therapy" (as might be encoded today) but also that Dr. Hartman *works* at this *clinic* on *Mondays*, *Wednesdays* and *Fridays* and that the script takes a *date range* in *yyyy-mm-dd format* and returns *appointment times*. And it will "know" all this without needing artificial intelligence on the scale of 2001's Hal or Star Wars's C-3PO. Instead these semantics were encoded into the Web page when the clinic's office manager (who never took Comp Sci 101) massaged it into shape using off-the-shelf software for writing Semantic Web pages along with resources listed on the Physical Therapy Association's site.

The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation. The first steps in weaving the Semantic Web into the structure of the existing Web are already under way. In the near future, these developments will usher in significant new functionality as machines become much better able to process and "understand" the data that they merely display at present.

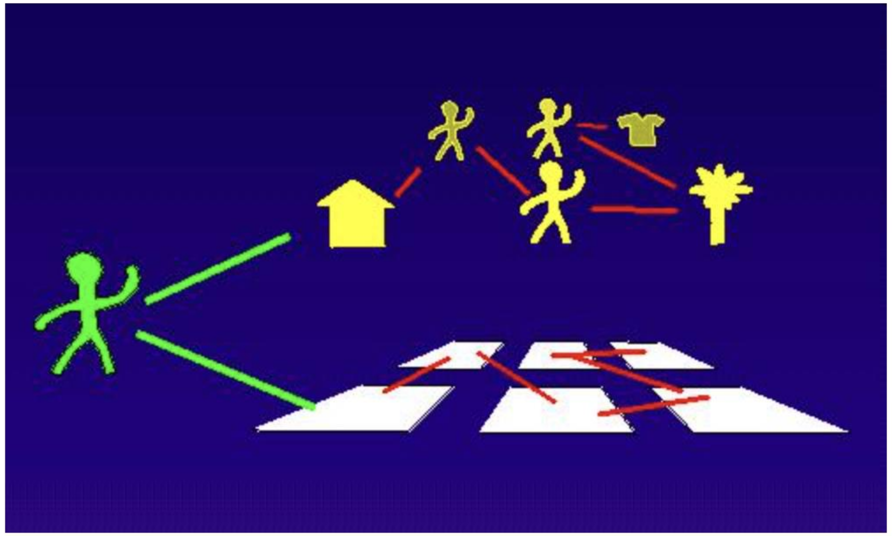
The essential property of the World Wide Web is its universality. The power of a hypertext link is that "anything can link to anything." Web technology, therefore, must not discriminate between the scribbled draft and the polished performance, between commercial and academic information, or among cultures, languages, media and so on. Information varies along many axes. One of these is the difference between information produced primarily for human consumption and that produced mainly for machines. At one end of the scale we have everything from the five-second TV commercial to poetry. At the other end we have databases, programs and sensor output. To date, the Web has developed most

Our story actually begins 30 years ago

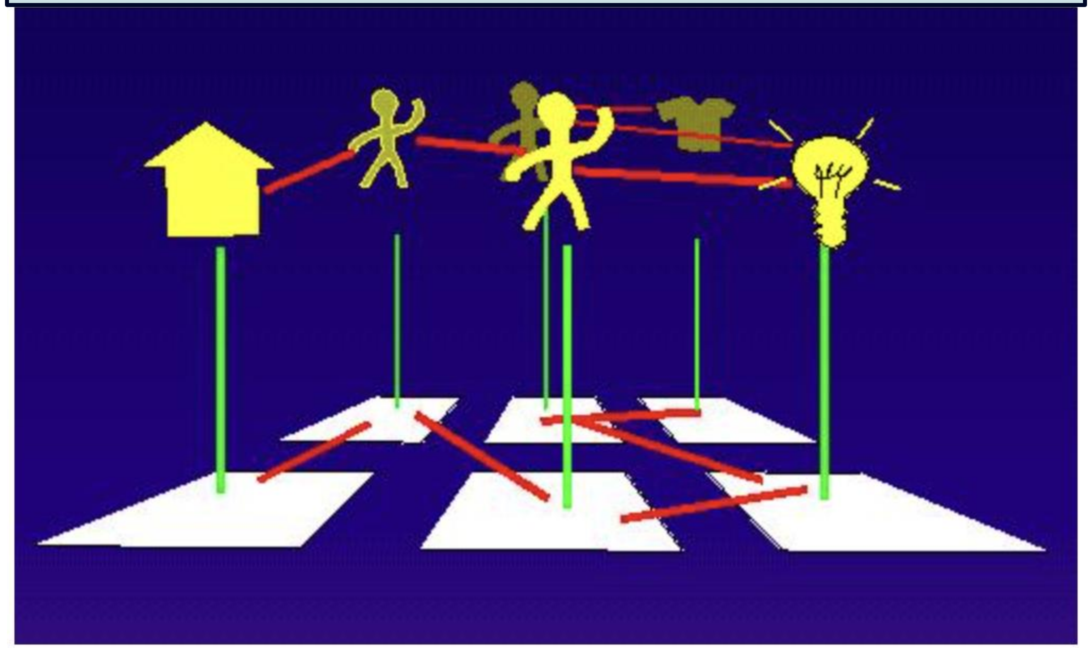
For example, a document might describe a person. The title document to a house describes a house and also the ownership relation with a person.

Adding semantics to the web involves two things: allowing documents which have information in machine-readable forms, and allowing links to be created with relationship values. Only when we have this extra level of semantics will we be able to use computer power to help us exploit the information to a greater extent than our own reading.

An important effect of developing security protocols on the web is the abstract space of web information is linked to reality. By taking verifiable responsibility for web statements, a party guarantees an isomorphism between the web and reality.



This means that machines, as well as operating on the web information, can do real things. For example, a program could search for a house and negotiate transfer of ownership of the house to a new owner. The land registry guarantees that the title actually represents reality.



1994 – First World Wide Web Conference – Tim Berners Lee shares his vision of the Semantic Web

Debunking Common Misconceptions

The concept of machine-understandable documents does not imply some magical artificial intelligence allowing machines to comprehend human mumblings. It relies solely on the machine's ability to solve well-defined problems by performing well-defined operations on well-defined data. So, instead of asking machines to understand people's language, the new technology, like the old, involves asking people to make some extra effort, in repayment for which they get major new functionality – just as the extra effort of producing HTML mark-up is outweighed by the benefit of having content searchable on the web.

The Semantic Web will enable machines to COMPREHEND semantic documents and data, not human speech and writings.

Berners-Lee, T. and Hendler, J., 2001. Publishing on the semantic web. *Nature*, *410*(6832), pp.1023-1024.

Berners-Lee, T., Hendler, J. and Lassila, O., 2001. The semantic web. Scientific american, 284(5), pp.34-43.

The Semantic Web – Critiques

- 2014 – *"Going through all the effort of putting semantic markup with no guarantee of a payoff for yourself was a stupid idea."* https://bibwild.wordpress.com/2014/10/28/is-the-semantic-web-still-a-thing/

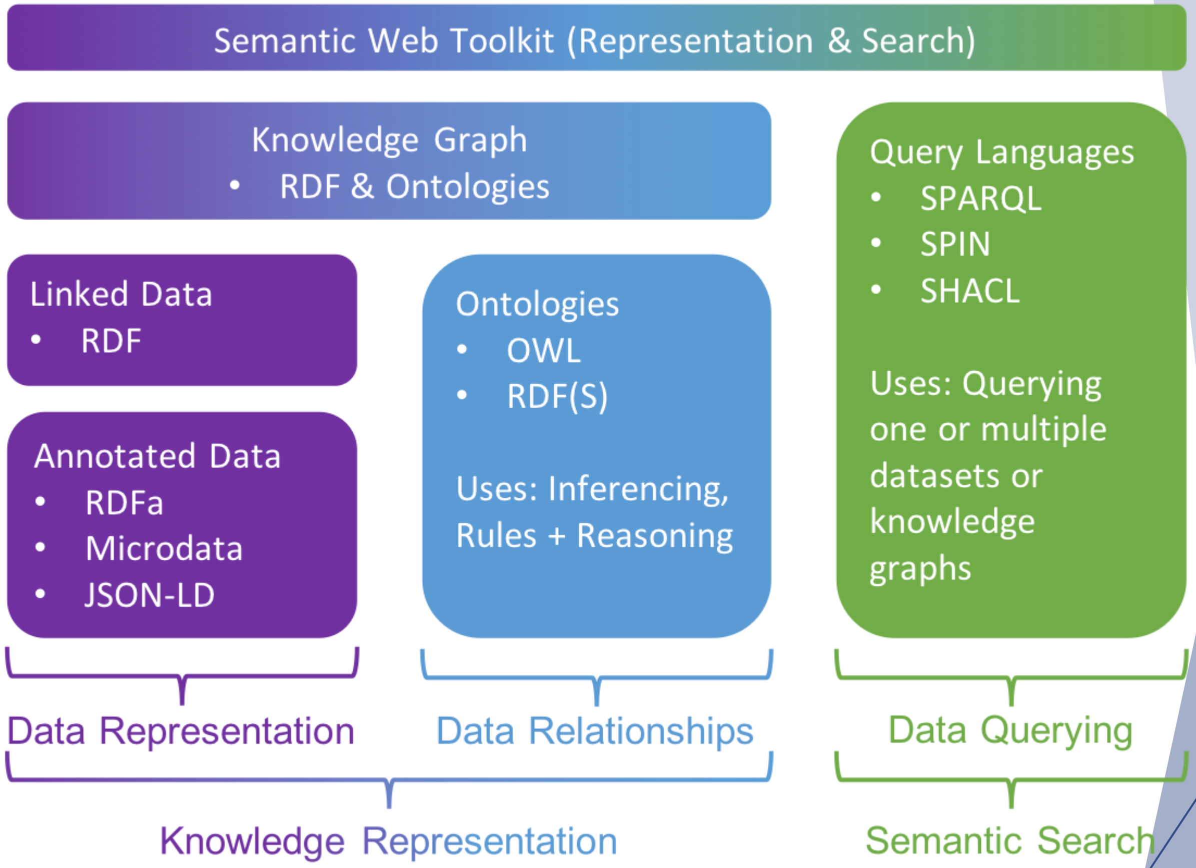
- 2016 - *"Semantics is hard to understand"* https://www.linkedin.com/pulse/why-semantic-web-has-failed-kurt-cagle/

- 2017 - *"The Semantic Web was a great idea and still is. But I'm not seeing a future for it as it is. It needs to evolve and integrate it's ideas with artificial intelligence"* https://hackernoon.com/semantic-web-is-dead-long-live-the-ai-2a5ea0cf6423

- 2022 – *"in some ways, the Semantic Web was on life-support since its inception, and it continued to survive only with the medical intervention of academic departments who had no need to produce useable software or solve serious industry needs."* https://terminusdb.com/blog/the-semantic-web-is-dead/

So what is the Semantic Web really?

- The Web of Linked Data
- A way to bring context and meaning to data
- A set of common standards for data representation, integration, and search
- Used to create knowledge graphs and metadata



https://www.sciencedirect.com/science/article/pii/B978012801238311520X?via%3Dihub

The Semantic Web in the Physical Sciences –Barriers & Challenges

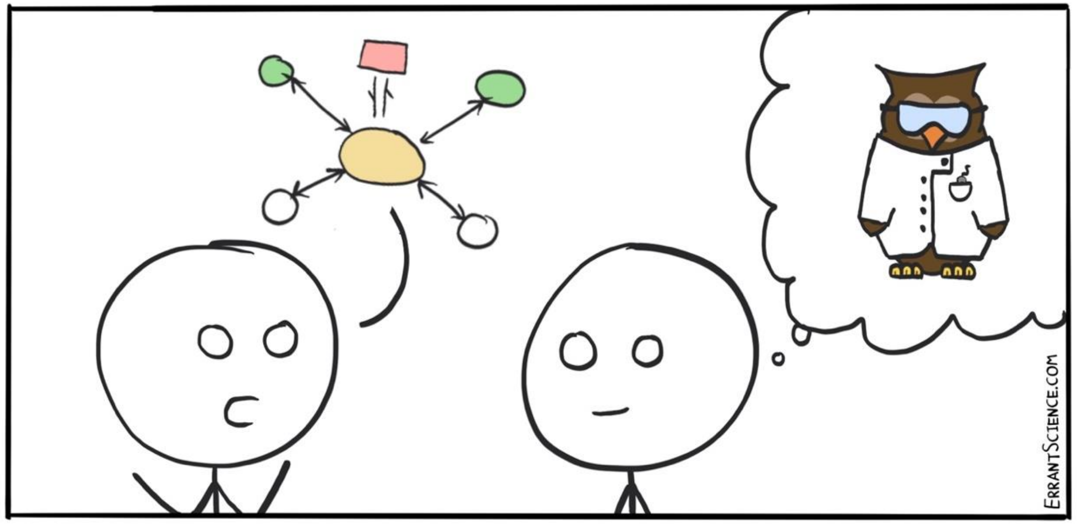
PSDI
PHYSICAL SCIENCES
DATA INFRASTRUCTURE

- Ontologies – Generic, Physical Sciences, Reactions
- Data & Ontology Integration
- Creation of Linked Data
- Scalability & Suitability
- Software Cost/ Availability
- Adoption & Usability

Image created using https://openart.ai/

Generic Ontology Barriers

- Too many ontologies (and yet not enough?)
- "Lets just make a new one" mentality
- Lack of standards for Ontologies
- Ontology projects frequently don't consider the full data lifecycle
- Ontologies aren't always well maintained
- Ontologies aren't always FAIR



Physical Sciences Ontology Barriers

- Even finding one ontological approach to classify the physical sciences is a challenge
- There are many ontologies competing in the same space, and yet large gaps still remain in others e.g. **reactions**
- Creating and maintaining ontologies to accurately represent aspects of the physical sciences e.g. chemical entities and their properties is complex, and requires domain knowledge

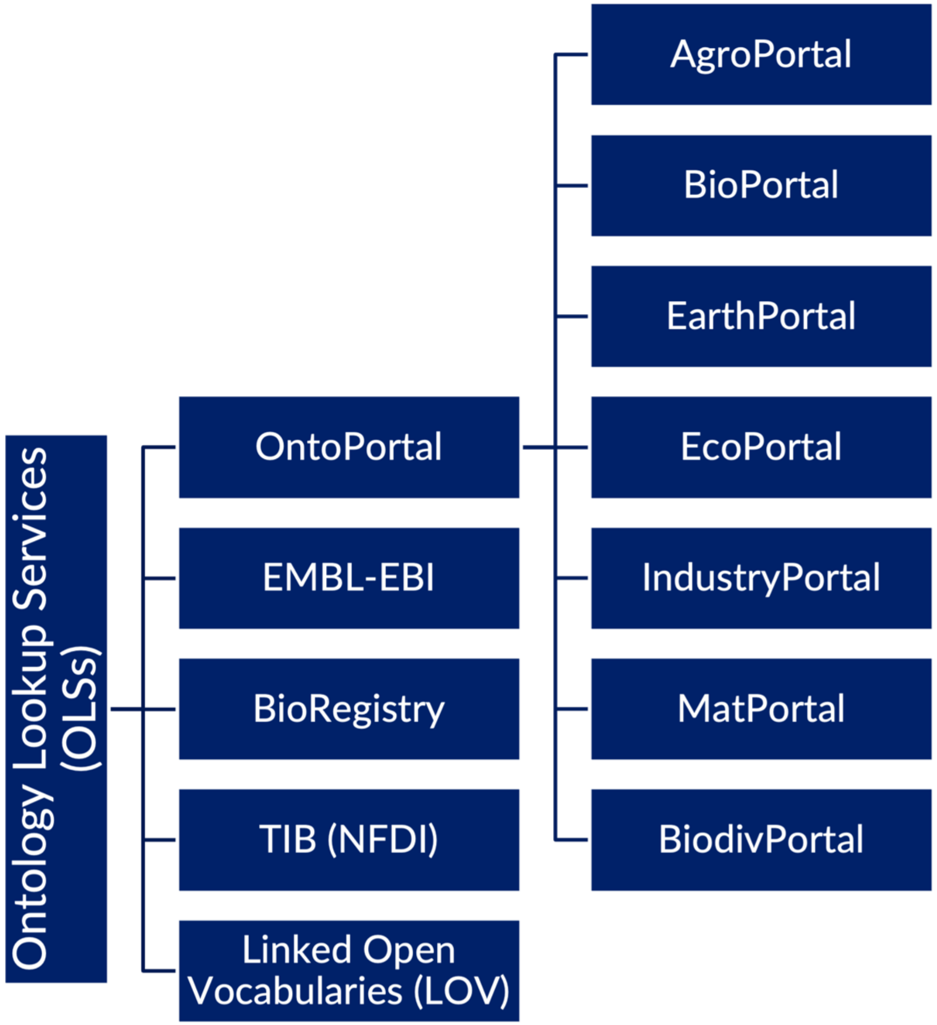


Diagram created by Oscar Robinson for PSDI Internship Project Semantic Landscape o the Physical Sciences

Data & Ontology Integration Barriers

- Physical Sciences data comes from disparate sources and is available in many formats, making it hard to standardize
- Harmonizing/Aligning ontologies across the physical sciences is complex, time consuming, and technically challenging

Image created using https://openart.ai/

Linked Data Creation/Conversion

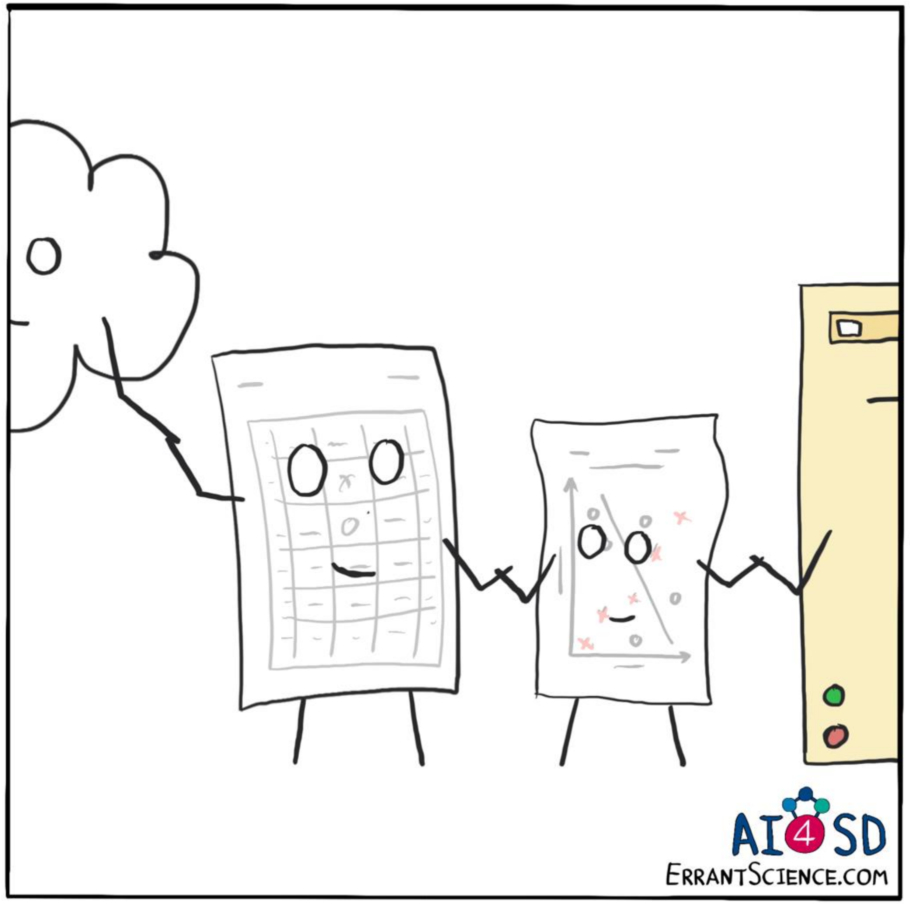
- Tools to convert structured data into linked data are still lacking
- R2RML even with additional libraries is repetitive
- Automated tools are very clunky and require a lot of human editing
- LLMs do a passable job in some instances but miss domain nuance
- Writing custom scripts is time consuming and are rarely fit for re-use

https://imgflip.com/memegenerator/

CONVERTING POOR QUALITY DATA TO LINKED DATA DOESN'T NECESSARILY FIX IT

Scalability & Suitability

- Some data doesn't lend itself well to being represented semantically (e.g. numerical classifications)
- Issues with reading, writing an d deleting data with certain (free) versions of triple stores
- Large datasets lead to even larger knowledge graphs



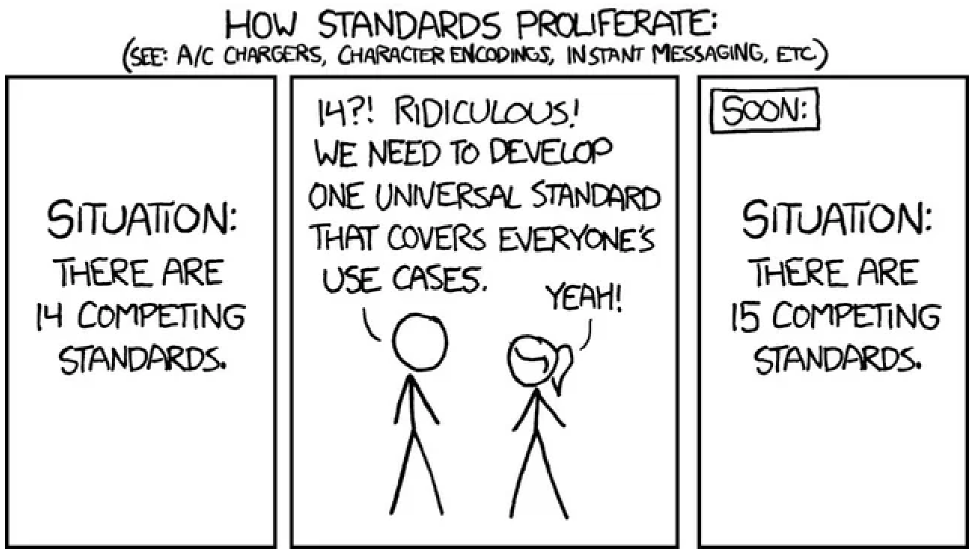
Potential for exponential grown of data that is unmanageable

Software Cost & Availability

https://imgflip.com/memegenerator/

Adoption & Usability

- People don't like change
- Time consuming to implement
- Chicken and egg - can't see the value until you have put in the effort, but why put in the effort if you don't understand the value?
- Too tempting to "start again"



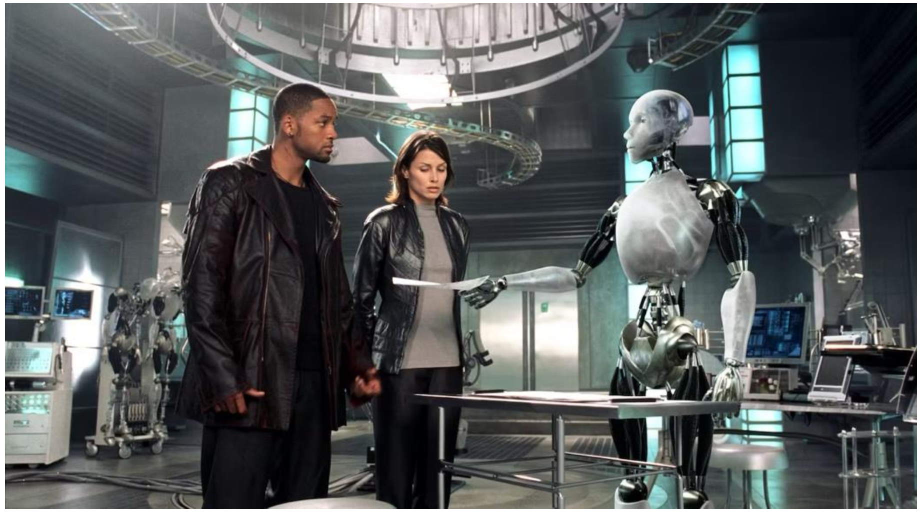
Implementing these technologies without proper DOMAIN & TECHNICAL knowledge is a recipe for DISASTER

Mitigations & Suggestions for Best Practice

https://memegenerator.net/instance/68158697/hermione-leviosa-swish-flick

You need to ask the right questions

- What is your use case for semantics?
- Do you need an ontology?
- Do you need a knowledge graph?
- How much data do you need to represent semantically?
- What data would benefit from being in a semantic format?
- What do you want your data to look like and how are you going to use it?
- What domain and technical knowledge does your project require?



https://www.inverse.com/culture/sci-fi-movies-january-2023-hbo-max-i-robot

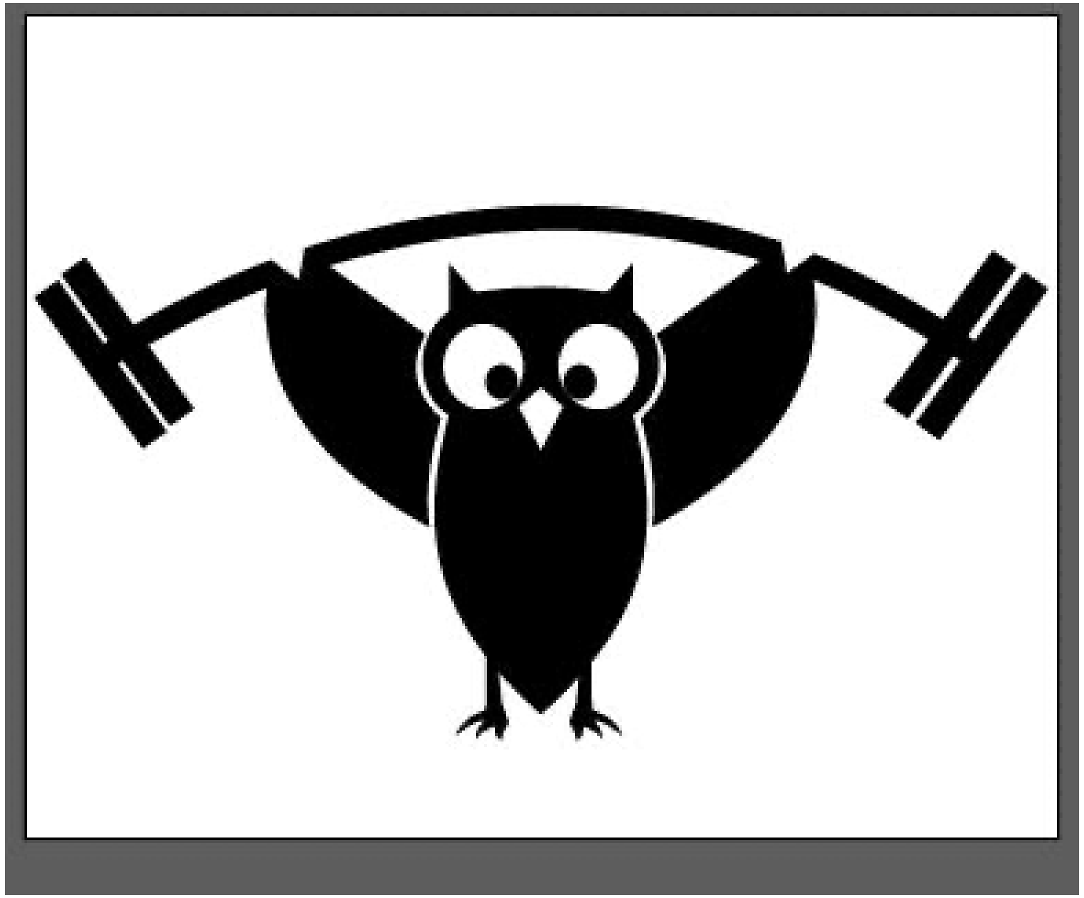
Consider your Conversion in Advance

- Converting data into linked data is non-trivial – what are you adding?
- Do you need semantic data or metadata?
- Conceptualising an Ontology vs using it in a dataset are two different things
- You need to consider the entire semantic data lifecycle of a project

https://www.woorank.com/en/blog/implement-semantic-web

PSDI PHYSICAL SCIENCES DATA INFRASTRUCTURE

Think about the WHY and the WHAT



https://cuk00jan.wordpress.com/2014/02/14/who-is-alpha-owl/

- Think about WHY you are making Ontologies/Linked Datasets and WHAT they are for
- Play to the technologies and standards strengths
  - Ontologies vs Taxonomies
  - JSON-LD vs Triples
  - OWL vs SHACL

Re-use and extend Ontologies where possible

- There are lots of ontologies out there! See if there's any you can re-use before making a new one (Search Papers/OLS/Google/LLMS)
- Re-use doesn't just mean ontologies! You can re-use design patterns as well!
- Consider trying to create an extension to an existing ontology if appropriate

https://www.fosteropenscience.eu/content/cartoonreusable-data

Modularize your Ontologies

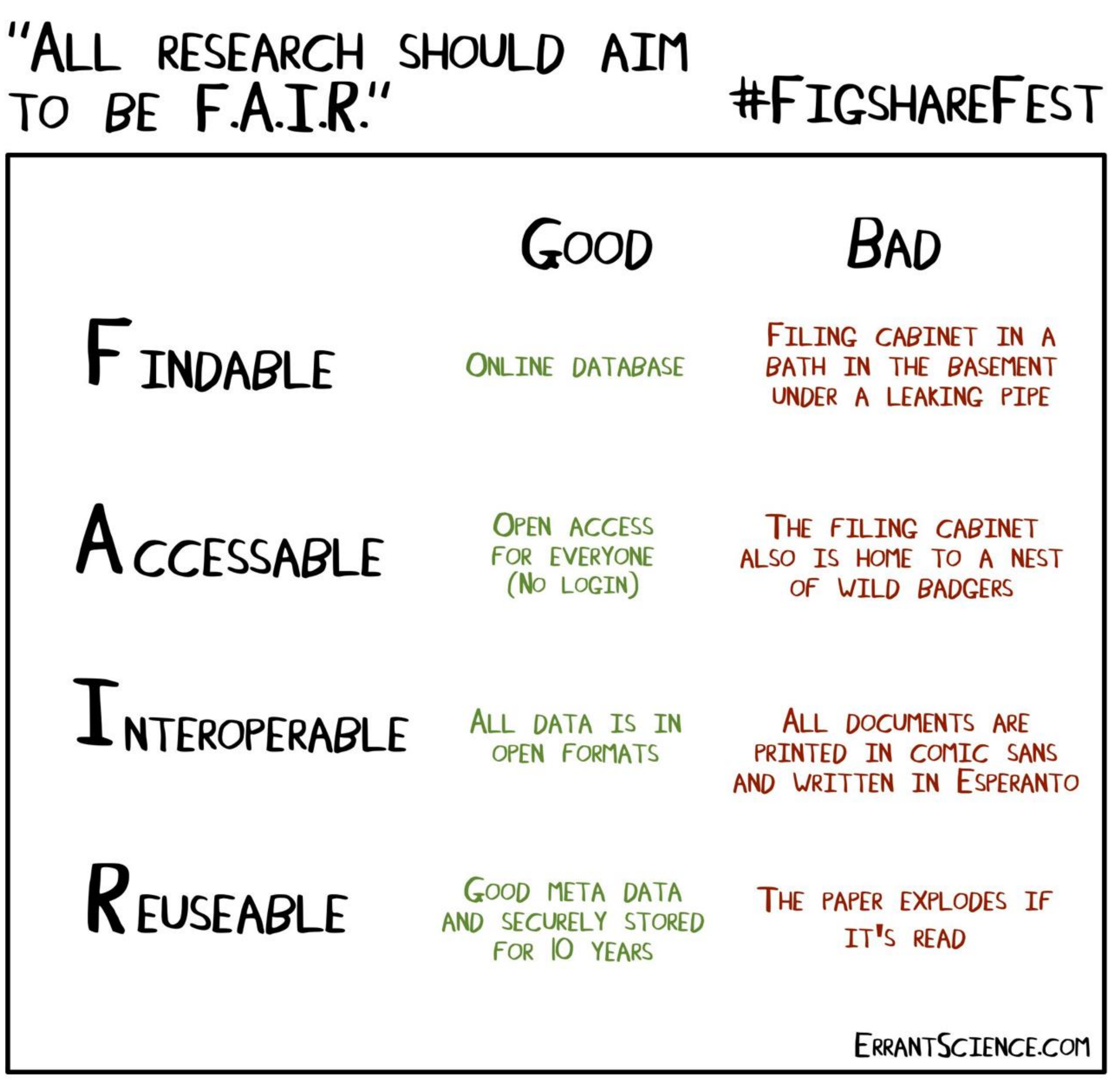
- Break your ontologies into smaller related modules for ease of use and improved re-use, and

<table>
<tr><th rowspan="2">Relation to time / Granularity</th><th colspan="5">Continuant</th><th rowspan="2">Occurrent</th></tr>
<tr><th colspan="3">Independent</th><th colspan="2">Dependent</th></tr>
<tr><td>Complex of organisms</td><td colspan="2">Family, community, deme, population</td><td rowspan="4">Environment</td><td rowspan="2">Organ function (FMP, CPRO)</td><td>Population phenotype</td><td>Population process</td></tr>
<tr><td>Organ and organism</td><td>Organism (NCBI Taxonomy)</td><td>(FMA, CARO)</td><td rowspan="2">Phenotypic Quality (PaTO)</td><td rowspan="2">Biological process (GO)</td></tr>
<tr><td>Cell and cellular component</td><td>Cell (CL)</td><td>Cell component (FMA, GO)</td><td>Cellular function (GO)</td></tr>
<tr><td>Molecule</td><td colspan="2">Molecule (ChEBI, SO, RnaO, PrO)</td><td colspan="2">Molecular function (GO)</td><td>Molecular process (GO)</td></tr>
</table>

https://www.researchgate.net/publication/268412593_Ontology-Based_Knowledge_Representation_of_Experiment_Metadata_in_Biological_Data_Mining/figures?lo=1

Don't forget your standards – FAIR is a Four Letter Word

- Even your standards should have standards
- Consider:
  - Design patterns
  - Upper-Level Ontologies
  - Consistency
- There's no point having beautiful Semantic data if no-one can find, access or re-use it
- Making it Semantic DOESN'T mean its USEFUL – GIGO still applies



The Future of Semantics in the Physical Sciences

https://imgflip.com/i/4x69cr

Established and Evolving Use Cases of Semantic Web in Physical Sciences (and beyond)

PSDI
PHYSICAL SCIENCES DATA INFRASTRUCTURE

Well Established Use Cases

- Defining common shared vocabularies for re-use
- Data Linkage & Integration to create Semantic Knowledgebases/Databases
- Inferencing
- Semantic Search
- Machine Readable Data
- Unlocking the potential of AI/ML

Evolving Use Cases

- Extension to existing Ontologies/New Ontologies
- Semantic Metadata
- Automated Metadata Extraction/Semantic Annotations
- Semantic ELNs & ELN Interoperability
- Validation using SHACL
- Leveraging Ontologies/Structured Data in LLMs

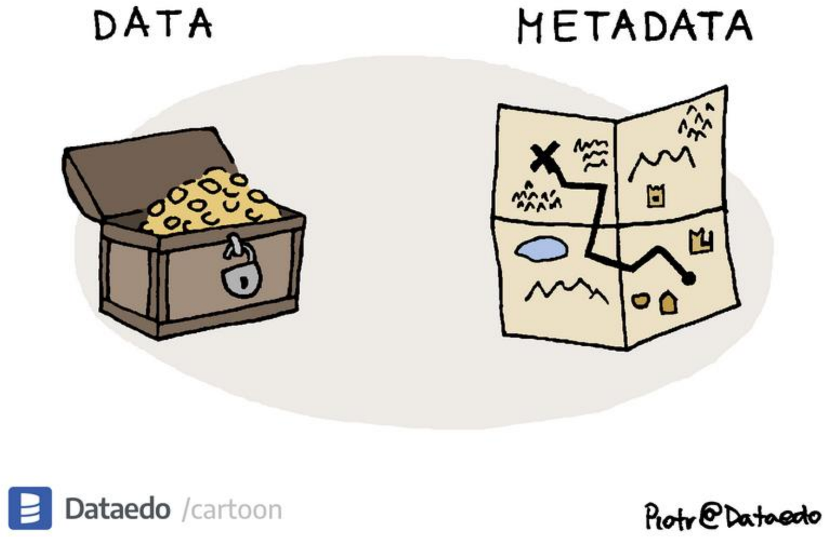
Extensions to Existing Ontologies/New Ontology Development

- Reactions is a key area – but there will be others
- We need the community to raise these issues and convene to solve them

- PSDI, NFDI (NFDI4Chem & NFDI4Cat), University of Cambridge (Alexei) are in discussions to convene a joint meeting to discuss *“current state and future of "ontologies" describing in-depth chemical reactions capturing electronic states and 3D geometries - or even dynamics - of reactants”*
- Discussions will be commenced at the Ontologies4Chem Workshop in December – slot reserved for PSDI/NFDI4Cat/NFDI4Chem discussions on areas of interest

Semantic Metadata

- Noticeable shift towards describing scientific research data using metadata and ontologies – intrinsic links with FAIR
- Creating Semantic Metadata
- Using JSON-LD (either with DCAT or schema.org)
- Using RO-Crate with JSON-LD

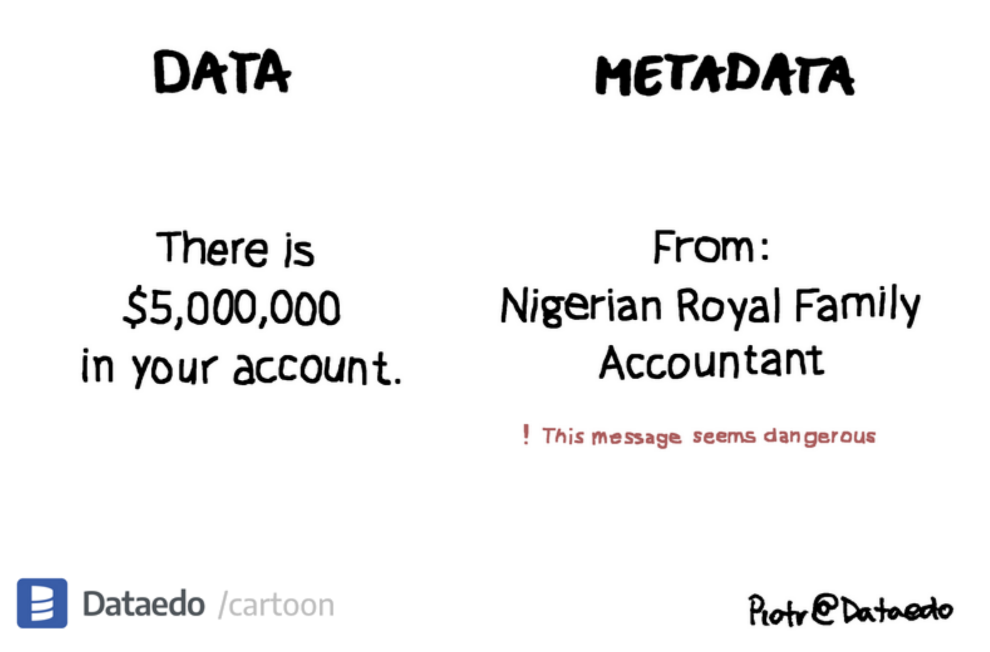


- Key Challenges
  - Lack of integration with repositories
  - Barrier to entry re knowledge
  - Too many ontologies to choose from
  - Inconsistency of ontologies
  - Ontologies may look like they have the term you are referring to but when you dig deeper you find out its something different entirely

Automated Metadata Extraction/Semantic Annotation

- Researchers are starting to consider how to automate semantic annotation/metadata extraction – either through custom tools or leveraging LLMs
- Number of libraries/tools available that look to extract chemistry-based information and generic information - see https://github.com/MADICES/MADICES-2024/discussions/9 for more information

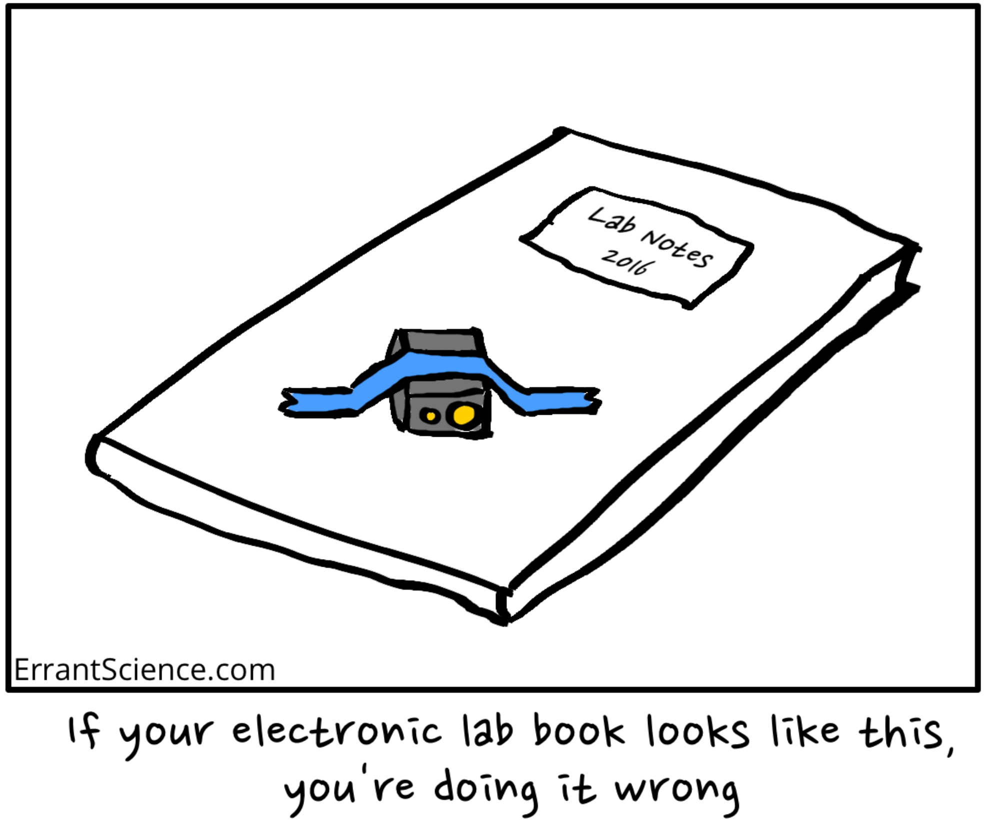
- Key Challenges
  - Alignment with ontologies is HARD – Domain knowledge is KEY
  - Datasets are often isolated from papers that contain relevant information needed for their metadata
  - Authors use names such as "chemical1" – which are hard to tie back to real chemical identifiers!



PSDI
PHYSICAL SCIENCES
DATA INFRASTRUCTURE

Semantic ELNs

- ELNs are starting to consider the powers of Semantics
- Work on ELN FileFormat (using RO-Crate) for interoperability between ELNS - https://github.com/TheELNConsortium/TheELNFileFormat
- Helmholtz-Zentrum Hereon to facilitate the digitalization efforts within several institutes. Herbie builds upon semantic technologies like OWL and SHACL to offer a user-friendly interface to the RDF ecosystem.



PSDI
PHYSICAL SCIENCES DATA INFRASTRUCTURE

Using SHACL for Validation

- QuaRe tool - https://github.com/uniba-mi/quare
- Assessing the FAIRness of Software using RDF and SHACL
- Hummel, T., Martin, L. and Henrich, A., 2024. Assessing the FAIRness of Software Repositories Using RDF and SHACL. In *Knowledge Graphs in the Age of Language Models and Neuro-Symbolic AI* (pp. 160-175). IOS Press.

README GPL-3.0 license

QuaRe: Validate your GitHub Repositories against Quality Criteria

Summary • Installation and Prerequisites • Usage • Repository Representation Ontology • Developer Information • Citation • License

Summary

QuaRe is a single-page application that allows users to test if GitHub repositories of interest comply with certain quality criteria that they should fulfill according to the type of project in the repository.

Installation and Prerequisites

Thanks to Docker, only Docker and Docker Compose are required for using the tool.

Usage

After cloning or downloading this repository, simply run `docker compose up` in a command line from the root folder of the repository to start the tool. The frontend can then be accessed via http://localhost:3000. (If necessary, the backend can be accessed via http://localhost:5000.)

The frontend currently provides two pages, namely the Validation page and the Specification page which can be selected using the navigation bar.

Leveraging Ontologies/Structured Data in LLMs

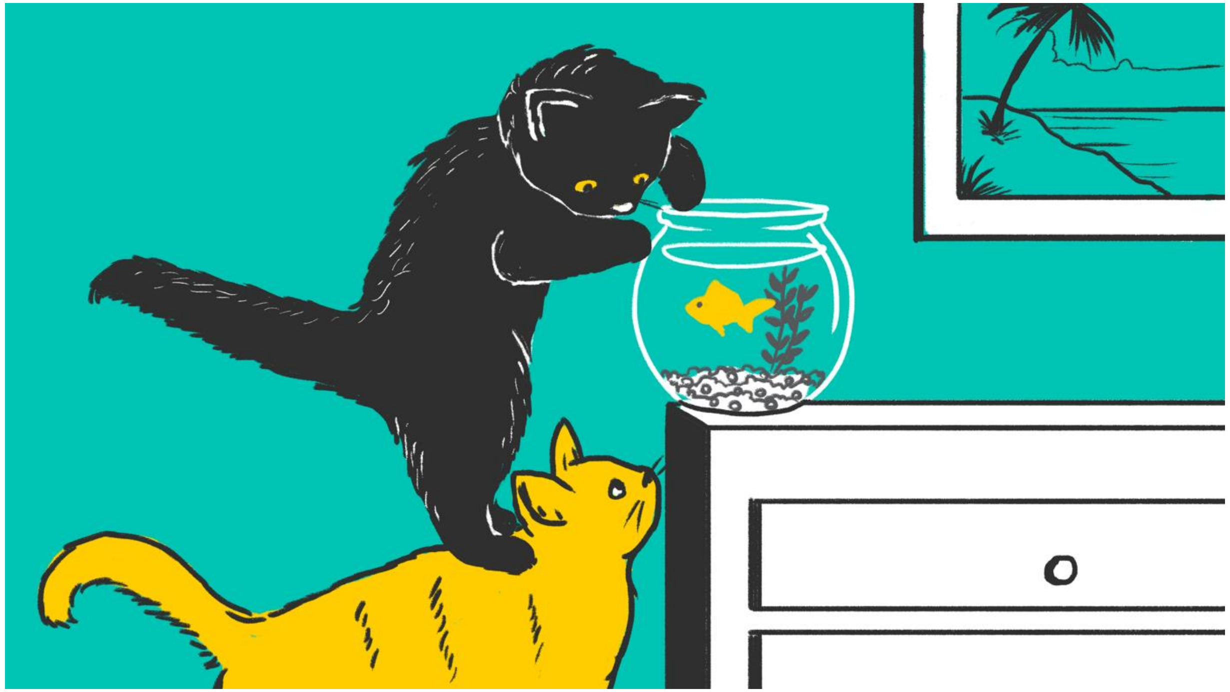
- LLMs & Ontologies were a new ”hot topic” at recent Semantic Conferences
- Countless papers are being published in this area, and naturally some are emerging in the physical sciences – specifically chemistry

https://ffineis.github.io/blog/2017/09/13/case_for_big_govt.html

- Key Message: Structured datasets and controlled vocabularies to describe the terms within them, WILL ALWAYS BE NEEDED.
- However good your AI is, do you really think it won’t be improved by consistent, structured data embedded with meaning and context?
- Even LLMs aren’t immune to GIGO

Thoughts for the Future

- We're all in this together and we need to work together!
- A lot of the technology is there, we just need to use it properly!
- This is as much a human endeavor as a technical one!



https://www.thedesigngym.com/skip-collaboration-collaborations-sake/

Relevant Publications & Talks

- Kanza, S., Willoughby, C., Gibbins, N., Whitby, R., Frey, J.G., Erjavec, J., Zupančič, K., Hren, M. and Kovač, K., 2017. Electronic lab notebooks: can they replace paper?. Journal of cheminformatics, 9(1), pp.1-15. https://doi.org/10.1186/s13321-017-0221-3
- Kanza, S., Stolz, A., Hepp, M. and Simperl, E., 2018. What does an ontology engineering community look like? A systematic analysis of the schema. org community. In The Semantic Web: 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, Proceedings 15 (pp. 335-350). Springer International Publishing. https://doi.org/10.1007/978-3-319-93417-4_22
- Kanza, S., 2018. What influence would a cloud based semantic laboratory notebook have on the digitisation and management of scientific research? (Doctoral dissertation, University of Southampton). https://eprints.soton.ac.uk/421045/
- Kanza, S., Gibbins, N. and Frey, J.G., 2019. Too many tags spoil the metadata: investigating the knowledge management of scientific research with semantic web technologies. Journal of cheminformatics, 11(1), p.23. https://doi.org/10.1186/s13321-019-0345-8
- Kanza, S. and Frey, J.G., 2019. A new wave of innovation in Semantic web tools for drug discovery. Expert Opinion on Drug Discovery, 14(5), pp.433-444. https://doi.org/10.1186/s13321-019-0345-8
- Kanza, S. and Frey, J.G., 2020. Semantic technologies in drug discovery—potential, practical, possibilities. http://dx.doi.org/10.1016/B978-0-12-801238-3.11520-X
- Kanza, S., 2021. AI3SD Video: Semantic Web in Scientific Research–Possibilities & Practices. http://dx.doi.org/10.5258/SOTON/P0078
- Kanza, S., Willoughby, C., Bird, C.L. and Frey, J.G., 2022. eScience Infrastructures in Physical Chemistry. Annual Review of Physical Chemistry, 73, pp.97-116. https://doi.org/10.1146/annurev-physchem-082120-041521
- Kanza, S., Willoughby, C., Knight, N.J., Bird, C.L., Frey, J.G. and Coles, S.J., 2023. Digital research environments: a requirements analysis. *Digital Discovery*. https://doi.org/10.1039/D2DD00121G
- Kanza, S. P. (2023, August 13). We don't talk about Semantic Web Technologies - Presentation at ACS Fall 2023. ACS FALL 2023 Harnessing the Power of Data (ACS FALL 2023), San Francisco, CA & Hybrid. Zenodo. https://doi.org/10.5281/zenodo.10149599

PSDI
PHYSICAL SCIENCES
DATA INFRASTRUCTURE

Acknowledgements

- **PSDI Southampton Team:** Simon Coles, Jeremy Frey, Nicola Knight, Cerys Willoughby, Aileen Day,Louise Saul, Raymond Whorley, Philip Leadbitter, Jo Grundy, Mark Anderson, Victoria Hooper, Beckie Isla, Katy Dolin *(University of Southampton)*
- **PSDI Internship Projects:** Oscar Robinson & Ashley Doel *(University of Southampton)*
- **Data Revival Team:** Samuel Munday & Jeremy Frey *(University of Southampton)*
- **Semantic Web Research Teams**
  - **SEED:** Pistoia Alliance, Pfizer, Biovia, Dotmatics, Sanofi, BenchSci, AstraZeneca, Elsevier, GSK, Bayer, Merck, SciByte, University of Southampton, Bristol-Myers Squibb, ChemAxon, IDBS, Linguamatics, Arxspan
  - **BSP:** Hugo Mills & Hannah Wickenden & Derek Scuffell (Syngenta), Jeremy Frey & Nicholas Gibbins *(University of Southampton)*
  - **S3W:** Susan Halford & Faranak Hardcastle *(University of Bristol)*, Nicholas Gibbins & Mark Weal *(University of Southampton)*, Cathy Pope *(University of Oxford)*
- **PhD Supervisors:** Jeremy Frey & Nicholas Gibbins *(University of Southampton)*

PSDI & Personal Details - Questions

www.psdi.ac.uk

@PSDI_UK

@PSDI_UK

linkedin.com/company/psdiuk

linktr.ee/samanthakanza

Mailing List: *https://www.jiscmail.ac.uk/PSDI*