

# semanticClimate's knowledge graphs

Peter Murray-Rust

*University of Cambridge semanticClimate*

FOSDEM, Brussels BE, 2025-02-01

NIPGR team  
Volunteers  
TIB Germany

We'll guide you through  
the presentation

And show some videos



AMI, semanticClimate



Chuff the OKF Okapi



Free and  
Open Source  
(F/OSS)

# WE'RE SERIOUS!



What are we going to see?

- Overview of semanticClimate
- Intern presentations
- Live Demos

This is NOT a “student project”. We’re semantifying “**the survival guide for humanity\***”

Because nobody else yet is doing it



<https://www.bbc.com/news/science-environment-65000182>

*Not available in any Indian languages*

# semanticClimate

## Builds a **community**

Automatically collects **climate information** ,  
makes it **semantic**, **annotates** it:

**for learning, re-use and advocacy by everyone**

# Sustainable Tech is created by Community



- Youth Learning by Doing; Internship Programme hosted at the National Institute of Plant Genome Research (NIPGR)



# Mentors and Support



Dr. Gitanjali Yadav



Dr. Renu Kumari



Ms. Shweeta N Hegde



Dr Ambreen Hamdani

"Doctor WHO\*"



**Simon Worthington:** TIB Hanover (German National Library of Science and Technology:  
Leibniz Information Centre for Science and Technology and University Library)

2020-current; and 100+ participants  
6-monthly internships (thanks NIPGR)

<https://blogs.ch.cam.ac.uk/pmr/2009/06/06/the-doctor-who-model-of-open-source/>



They joined in January 2025

## *IPCC Chapters analysis*



AMI



**Moobashara Jawed**

AR6/WG1/Chapter 2: Changing  
State of the Climate System



**Anudev S**

AR6/WG2/Chapter 6: Cities,  
settlements and key infrastructure



**Mebin Joseph**

AR6/WG2/Chapter 5: Food, Fibre  
and Other Ecosystem Product



**Shabnam Barbhuiya**

AR6/WG1/Chapter 10: Linking  
Global to Regional Climate Change



**Sharon Nayak**

AR6/WG2/Cross-Chapter Paper 7:  
Tropical Forests



**Madhana Kumar**

AR6/WG2/Chapter 4 : Water

Sharon will be  
presenting!



## **"Learning-by-Doing"**

- Driven by public presentations!
- new joiners often have no previous coding

# Presentations from OKF

## “TheTechWeWant 2024”



Shweata and Parijat describe IPCC; how to make it semantic

Enshittification alert! There may be YouTube ads



- [https://www.youtube.com/watch?v=o50Jd1w6xKw&t=16842s&ab\\_channel=OpenKnowledgeFoundation](https://www.youtube.com/watch?v=o50Jd1w6xKw&t=16842s&ab_channel=OpenKnowledgeFoundation) (Shweata N Hegde)
- [https://www.youtube.com/watch?v=V7Sxo0IXe8g&t=7800s&ab\\_channel=OpenKnowledgeFoundation](https://www.youtube.com/watch?v=V7Sxo0IXe8g&t=7800s&ab_channel=OpenKnowledgeFoundation) (Parijat Bhadra)

# Software freedoms

0. **to run the program** for any purpose
1. to study how the program works (**source code** )
2. **to redistribute copies** so you can help your neighbor.
3. To distribute copies of your **modified versions**

[https://en.wikipedia.org/wiki/The\\_Free\\_Software\\_Definition](https://en.wikipedia.org/wiki/The_Free_Software_Definition)



<https://blogs.ch.cam.ac.uk/pmr/2014/07/01/visit-of-richard-stallman-rms-to-cambridge/>



# Our mission

- **To create verifiable computable climate knowledge**
- **we take the following as verifiable sources of truth**
  - **IPCC framework (PDF or web-HTML)**
  - **UNFCCC reports (COPs)**
  - **1000/week scholarly climate publications**



## What's in the IPCC material?

### **framework (PDF or web-HTML)**

- 3 Working Groups (WG1, WG2, WG3)
- several Special Reports
- Synthesis Report
- 15000+ bibliography
- 700-term Glossary
- 1500 acronyms

We make all of these semantic (machine parsable into structured, annotated data).

Total:

- 20,000 pages of PDF => HTML
- 1000 figure captions

# Knowledge Graphs



our approach is very simple

no OWL,  
no Protege

RDF triple = node-edge-node

Wikidata is a  
enhanced  
triple Store

# Graph Visualization and Analysis

many tools...

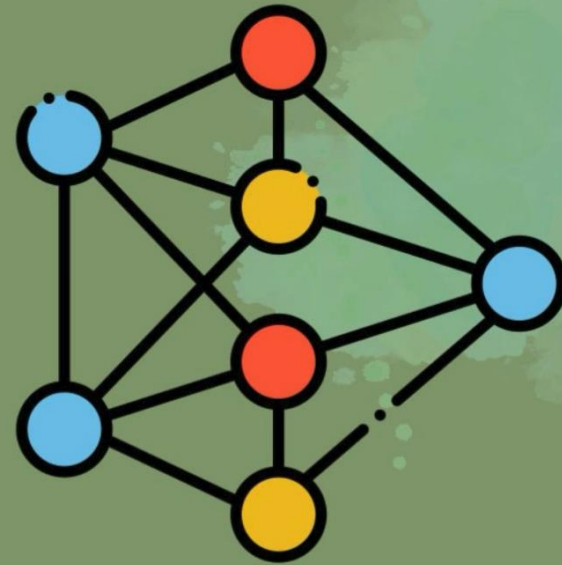
- networkx
- D3

query languages...

- SPARQL
- GraphQL

# Sharon demonstrates network and a ToC graph

Networks of  
IPCC  
Working  
Groups

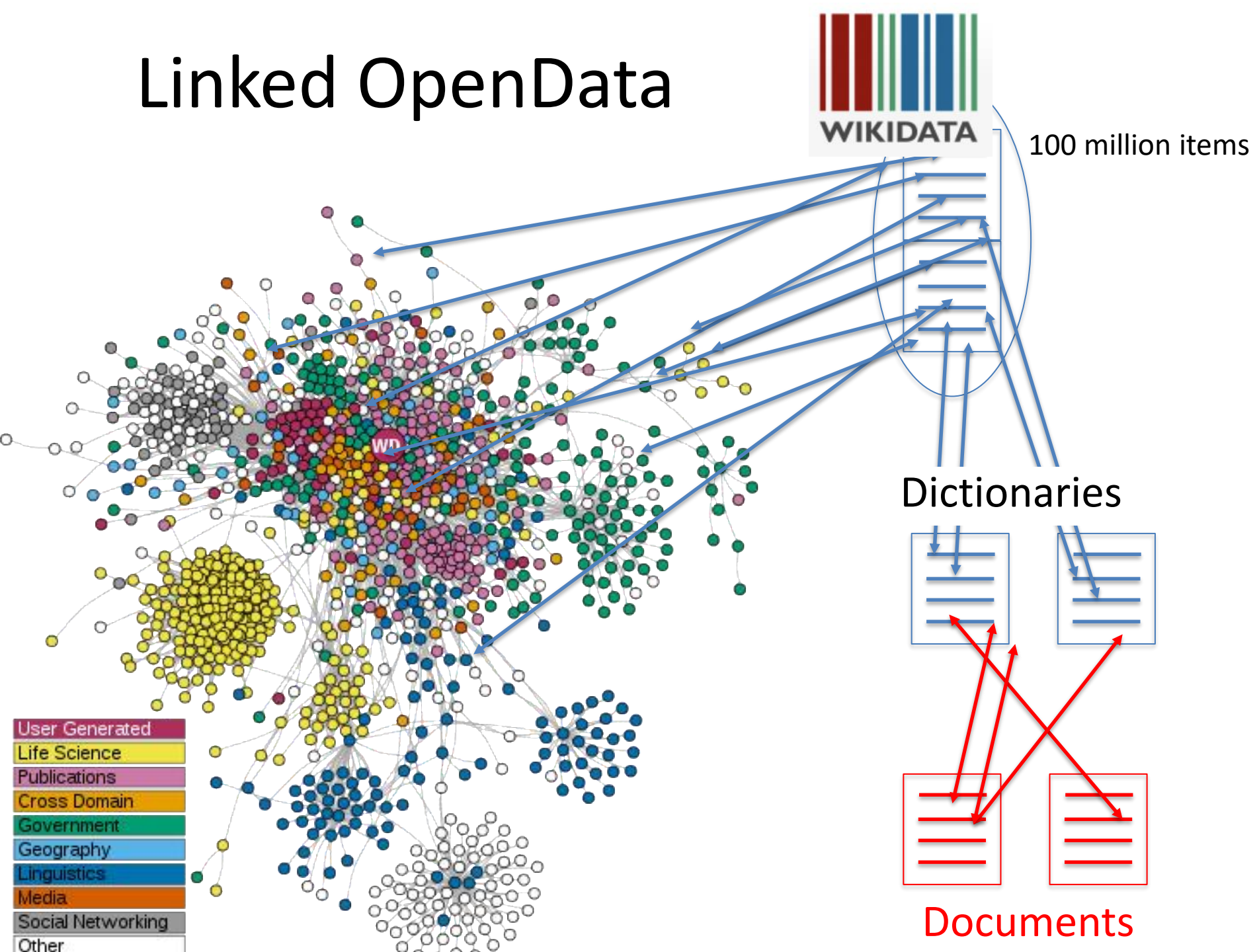


By Team  
semanticClimate

~ Sharon



# Linked OpenData



# Renu's Resources

<https://semanticclimate.github.io/p/en/posts/resources/>

- [Glossary/Encyclopedia \(700-2000 IPCC terms\)](#)  
(SimonWorthington)
- [Knowledge Graph of the IPCC glossary](#)  
(ShweataHegde)
- [Wikidata Knowledge Graph of a section of SYR \(IPCC Synthesis\)](#) (Egon Willighagen)

## **OUTREACH**

OKFN ORDEM (Shweata and Ambreen) Feb 15 2025

we'd love to

- Scale up knowledge graphs
  - Find new users
  - make our terminology available
  - find new volunteers (climate, IT, languages)
  - Build tested libraries
- 
- Have a look at our resources!

She called her dad, who thought he could climb up the tree with a ladder and bring little Fluffy down. **The others sat next to her and fondled her.**

# Hallucinations!

---

And then he got to the, they pulled his into the house on the window, **and he slapped out his infl Sexuality,**

---

everybody in the truck, the whole family, just waving and yelling. My goodness. **That was pretty, extremely barbaric.**

---

And then sometimes I was scared about the traffic on the, you know, the cars, some, **some men are homeless, or they're autreally ill.**

---



Polly,  
stochasticParrot

Should we trust HuggingFace?



AMI, semanticClimate





I don't know where their data comes from  
And they don't have an advisory board

We must create our own OKF-Open framework!

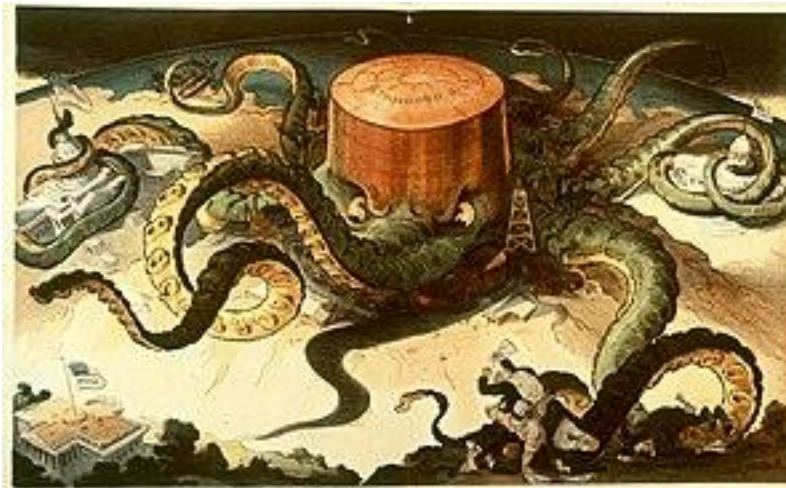


the data are private,  
run by a **private for-profit company**

	<b>Hugging Face</b>
<b>Company type</b>	Private
<b>Industry</b>	Artificial intelligence, machine learning, software development
<b>Founded</b>	2016; 9 years ago
<b>Headquarters</b>	Manhattan, New York City
<b>Area served</b>	Worldwide
<b>Key people</b>	Clément Delangue (CEO) Julien Chaumond (CTO) Thomas Wolf (CSO)
<b>Products</b>	Models, datasets, spaces
<b>Revenue</b>	▲ US\$15 million (2022)
<b>Number of employees</b>	170 (2023)
<b>Website</b>	<a href="https://huggingface.co">huggingface.co</a> 

“We are on a mission to democratize *good* machine learning, one commit at a time.”

# ROBBER BARONS



1904 depiction of an acquisitive and manipulative [Standard Oil](#) (founded by [John D. Rockefeller](#)) as an all-powerful octopus

Those practices included

- [unfettered consumption and destruction of natural resources](#),
- [influencing high levels of government](#),
- [wage slavery](#),
- [squashing competition by acquiring their competitors](#) to create [monopolies](#) and/or [trusts](#) that control [the market](#), and
- \* [schemes to sell stock at inflated prices](#) to unsuspecting [investors](#).<sup>[2]</sup>

# Tiktok's enshittification

The company manually allocates surplus to creators, and they can take it away again, too.



Cory Doctorow  · [Follow](#)

17 min read · Jan 21, 2023

Here is how platforms die: first, they are good to their users; then they abuse their users to make things better for their business customers; finally, they abuse those business customers to claw back all the value for themselves. Then, they die. I call this enshittification, and it is a seemingly inevitable consequence arising from the combination of the ease of changing how a platform allocates value, combined with the nature of a "two-sided market", where a platform sits between buyers and sellers, hold each hostage to the other, raking off an ever-larger share of the value that passes between them.

<https://pluralistic.net/2023/01/21/potemkin-ai/#hey-guys>

Engage your brain!  
Control AI or it will  
control you!



AMI, semanticClimate

- [https://x.com/okfn\\_okapi](https://x.com/okfn_okapi)



Thanks to:

- \* Yusuf Hamied Travel Fellowship (Cambridge India)
- NIPGR Hosting and internships
- Indian National Young Academy of Scientists
- Egon Willighagen
- Wikimedia
- The brave people who have challenged injustice