



Neha Moopen @NehaMoopen

Jul 9, 2022 · 16 tweets · [NehaMoopen/status/1545806638371799046](https://twitter.com/NehaMoopen/status/1545806638371799046)

The #LIBER2022 conference has ended, but during my long train commute 🇩🇰 -> 🇩🇪 -> 🇫🇷 - I finally started putting together some highlights from the pre-conference workshop I led on Data Science in Libraries. Let's go! @LIBERconference @LIBEReurope

1/15



We had about +10 participants, and most of them were motivated to join the workshop to explore how data science could be implemented in libraries.

2/15

When asked to provide their definition/idea of how data science could be applied in libraries
- participants' responses ranged from research data management and teaching information/data skills to analyzing library collections and research output.

3/15

The screenshot shows a Wooclap poll interface. At the top left is the Utrecht University logo. To its right is the URL www.wooclap.com/FBGMTZ. Below the URL is the poll question: "What is your definition/idea of data science in libraries?". On the far right is an info icon (a circle with an 'i').

The poll consists of five cards:

- Card 1: "Using scientific methods, processes, algorithms and systems to extract knowledge and insights" (3 likes, icons: grid, list, cloud)
- Card 2: "The application of Data Science to library topics" (2 likes, icon: list)
- Card 3: "Facilitating TDM and analytics on library collections" (1 like, icon: list)
- Card 4: "Teaching information resource focused classes to students regarding data collection, analysis" (1 like, icon: list)
- Card 5: "When DS touches metadata it's for the library" (1 like, icon: list)

Below the cards, there are additional options: "Analyzing aspects of the scholarly communication & research workflow landscape" (3 likes, icons: list, cloud), "support in data management" (no likes, icon: cloud), and "Text and data mining ?" (no likes, icon: list). On the right side of the poll area are navigation arrows (left and right) and a mail icon.

At the bottom of the screen, the Wooclap logo is on the left, followed by navigation buttons: Questions (2 / 2), Messages, a heart icon, a lock icon, a search icon (60 %), a refresh icon, and Exit. On the far right are participant counts (13 people), a help icon (i), and a settings icon.

After this, it was time to get into discussion groups. I had some fun with naming the subgroups after iconic Dutch treats 😊

4/15

Team Tompouce

Tompouce: An iconic Dutch dessert from the equally iconic Dutch department store HEMA. If you ever want to read about it (with Google Translate) - you can follow this link:
<https://www.hema.nl/geschiedenis/tompouce>



Team Stroopwafels

Stroopwafel: The one and only Dutch Syrup Waffle/Cookie. The best ones are baked fresh on market days or in busy areas. Today, you get the supermarket version - still good!



Team Poffertjes

Poffertjes: Dutch Mini Pancakes. Children are crazy for them and plenty of adults as well. I've always gotten them from street carts in the busier cities or markets.



And thanks to @jeroenson, the participants actually got to enjoy one of these iconic treats!

5/15

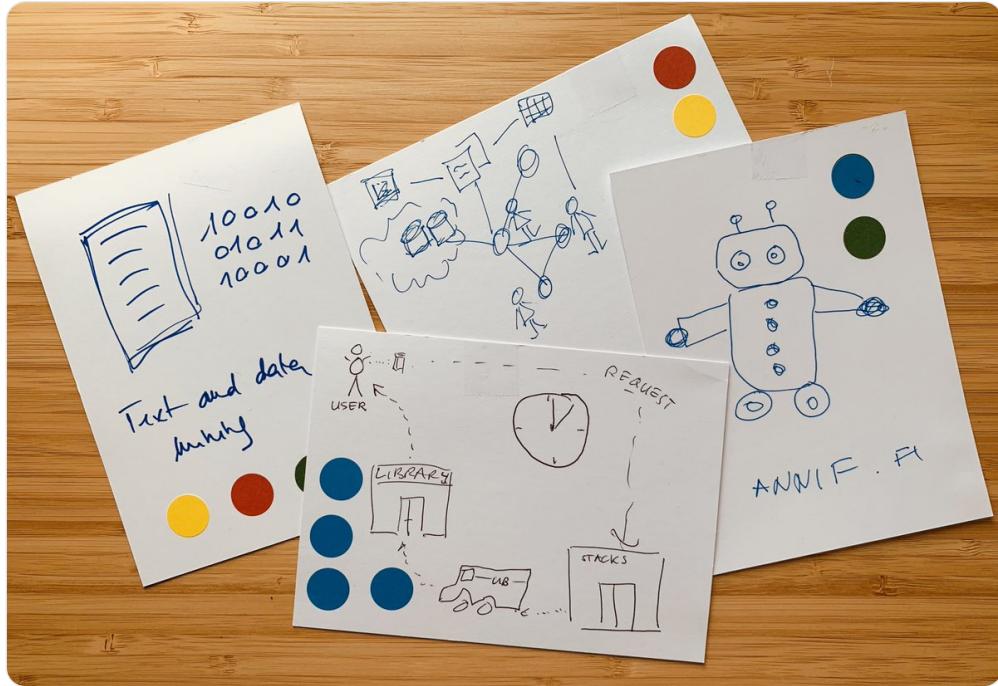


Within the subgroups, participants were asked to discuss examples/use cases of how data science could be applied in their jobs/libraries ...

6/15

But there was a catch! They had to draw out the examples/use cases and the discussion would start with trying to guess what everyone's artwork was describing 🖍

7/15

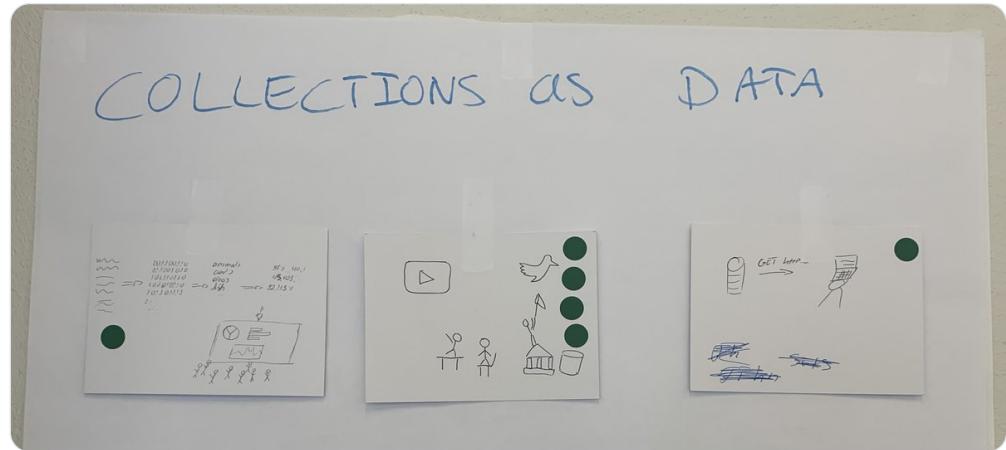


After the subgroups reported back on their discussions, I went on to present the categories/typologies that our working group has been using to cluster examples and use cases:

8/15

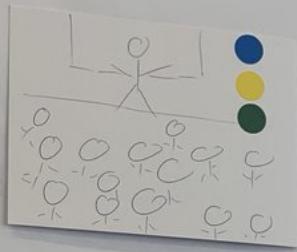
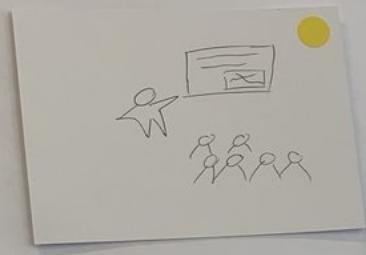
Collections as Data Data science activities that facilitate the use of library collections in computationally-driven research and teaching. This can include activities that ensure that data from library collections are high-quality, rich with information, reliable, suitable for analysis, and easily accessible for computational interactions.	Library Intelligence Data science activities geared towards the improvement of traditional library services and support for decision-making by library management.
Research Support Data science activities to support researchers through the research lifecycle. This can cover areas such as research data management, research data/software engineering, digital humanities, and (digital) information skills.	Research Intelligence Data science activities in compiling and visualizing data for decisions and benchmarking within the scientific community. Given the scale of the data available, Research Intelligence often requires the implementation of data pipelines and dashboard tools.

This led to our next subgroup activity, where participants sorted their examples/use cases into the categories presented - all while having a good discussion 💬

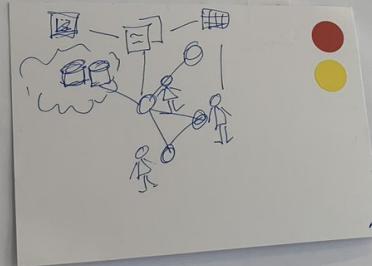


RESEARCH

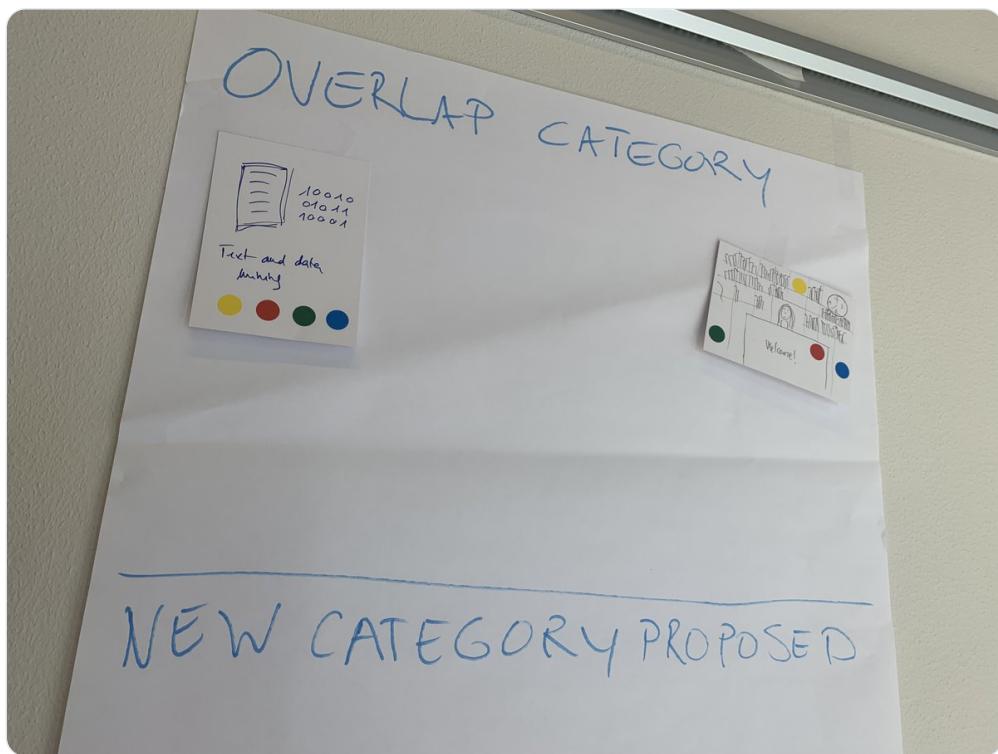
SUPPORT



RESEARCH INTELLIGENCE



We saw that we had examples/use cases for all categories, but there was also overlap across categories - as seen by multiple colored dots per flashcard + the extra 'overlap' category we created. This suggests that we might have to reconsider the definitions of the categories...



We were running short of time, so we wrapped up with a discussion on what participants would like to see in the landscape analysis which is being developed by the Data Science in Libraries Working Group 0.



LIBER Data Science in Libraries Working Group - LIBER Europe

<https://bit.ly/3nR1XjI>

We noted that clearer definitions, more concrete examples/use cases, and an overview of who is doing what + where would be valuable in the landscape analysis:

12/15

Landscape Analysis

- It would be useful to see various examples:
 - What kind of data libraries are working with- case examples
 - For what purposes the data is analyzed- examples
 - Development of tools to help data analysis- examples
 - Examples of collaboration- library-stakeholder(s)
- What tools to use when publishing data and other information resources for more transparency
 - What existing tools used
 - Course of action
 - Newly developed tools
- One challenge is becoming discipline specific, focusing on one area more than the others. One problem would be stepping on other territories or sometimes doing something that has already been done.
- Transparency of ongoing or past projects
 - Avoid overlap and nervousness in bringing up new projects.
- Guiding frame about collections. Structured data collection.
- What is meant by data "science"? Analytics, Engineering, etc.
- What is the landscape for?
- The range of the landscape
- An overview of people engaging with data (Carpentries, ...)
- Definitions - what is research intelligence? RI to integrate into the research support portfolio
- Highlight that there is other data stuff than data science. Data engineering, data analysis etc. Maybe everything isn't data science. Maybe it's only data science if we are actually doing science.
- A directory would be nice - who is where, and what are they working on/with
- Examples for making the case for a service area (e.g. research intelligence)

And that's it! If you've come this far, thanks for reading.

And if you find this topic interesting, please consider joining our working group on Data Science in Libraries. We're looking for new members!

13/15

Data Science in Libraries

Applying Data Science in Libraries: Join Our New Working Group - LIB...

<https://bit.ly/3yTqfjm>

Before signing off, I want to thank all the members of the working group + the workshop participants for their input.

I also want to thank The Two Jeroens & Joeri. Yes, I had to make that sound like a band name 😊

14/15

So that's @jeroenbosman (and @MsPhelps too) for inspiring the exercises we did in the workshop, @jeroenson for supplying the Emergency Stroopwafels, and Joeri Both (VU Amsterdam) for helping me out on-location since I was to do the workshop alone. Cheers!

15/15



@threadreaderapp unroll

• • •