

# SemanticLayerTools

Malte Vogl, ModelSEN  
MPI for the History of Science, Berlin, Germany



# Background

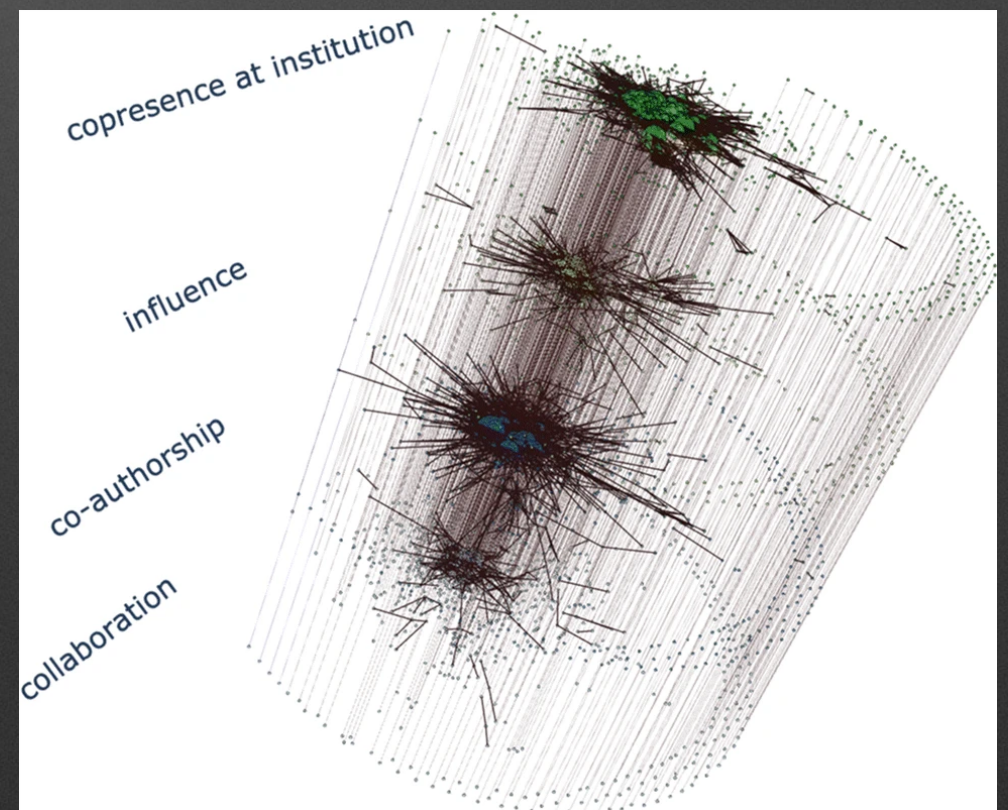
## ModelSEN project

- focus in History of Science
- How to model knowledge processes
- Common basis: bibliometric data
- Analyzed with historical background knowledge

## How to get the experts in the loop?

### Build software package

- to ease rinse and repeat
- make process reproducible
- fully understand impact of assumptions on historical interpretation



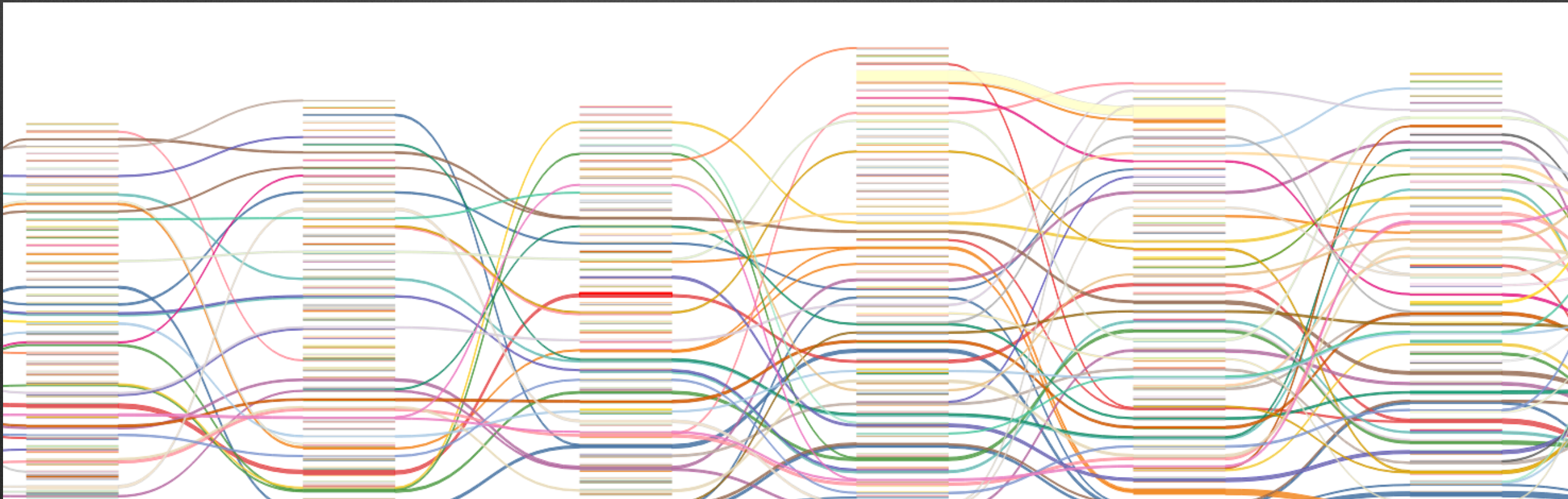


# Approaches I

Define each publication as multilayer network with 3 layers

- Authors
- Papers
- ngrams
- Connect papers by same authors, or using same ngrams
- Run infomap clustering to see flow across time between important modules

Example: NASA ADS corpus on astrophysics



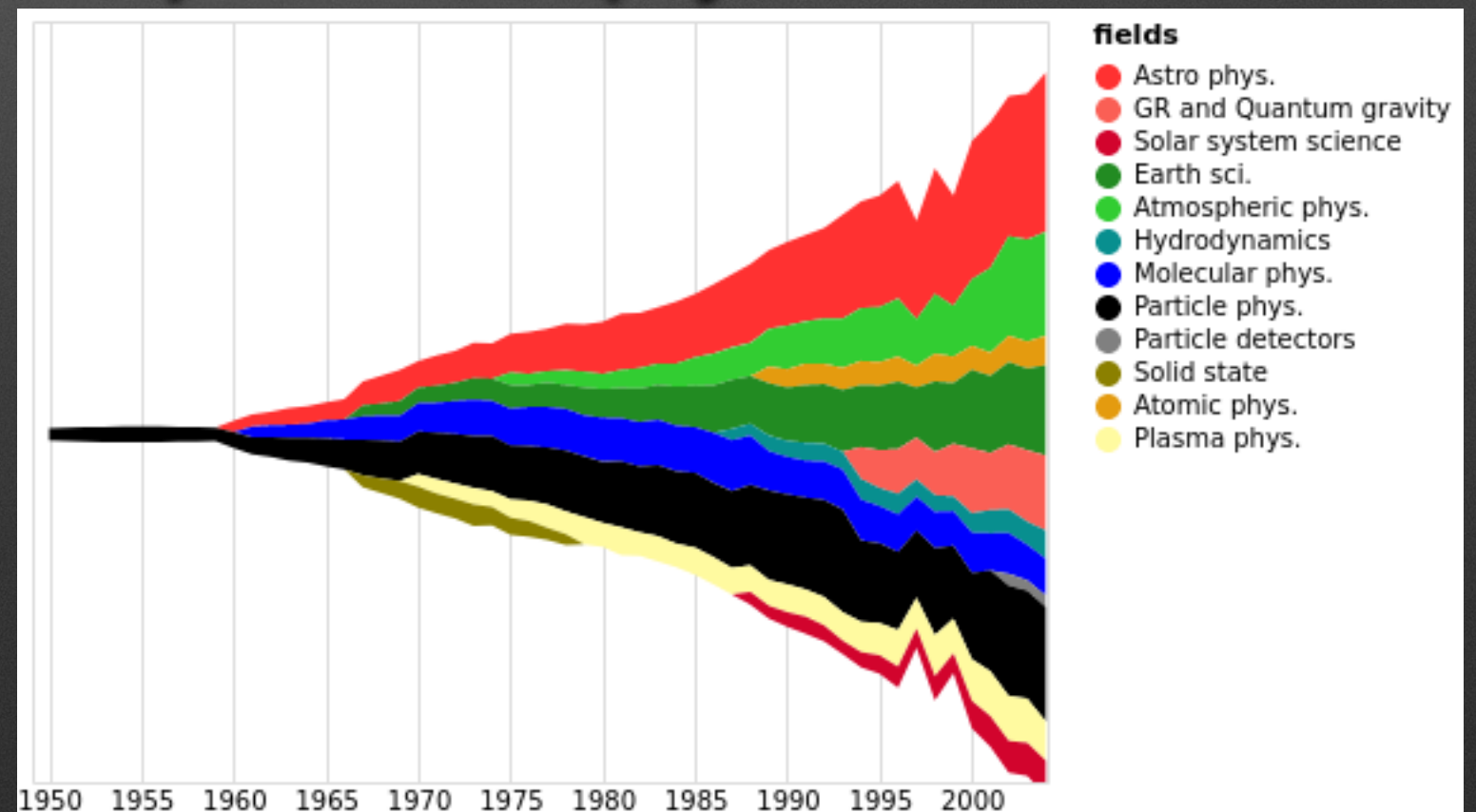


# Approaches II

Find "scientific fields" by using (co)citation links for timesliced corpora

- Uses temporal clustering approach of Leiden algorithm
- Writes reports of found clusters in relation to
  - important authors and institutions
  - and NMF topics

Example: Again NASA ADS corpus on astrophysics





# Thanks !

Project: <https://modelsen.mpiwg-berlin.mpg.de/de/>  
Source: <https://gitlab.gwdg.de/modelsen/semanticlayertools>  
Docs: <https://semanticlayertools.readthedocs.io>