

GlamMapping Trove

Arianna Betti (U of Amsterdam)

Thom Castermans (TU Eindhoven)

Bettina Speckmann (TU Eindhoven)

Hein van den Berg (VU Amsterdam / U of Amsterdam)

Kevin Verbeek (TU Eindhoven)



Visual Analytics for Library (Users)

Why so little applied solutions and/or suboptimal ones?

- **Sociological** factors
 - Sociology of research cooperation across strongly diverse fields in academia & their different practices
- **Organisational** factors
 - in-out academia interaction complexity (academic researchers vs business and GLAM sector - cultural heritage)
- **Public policy-related** factors
 - trends in governmental public sector funding policy (local & across the globe), nature of GLAM sector & its managerial organisation

Visual Analytics for Library (Users)

Our recent investigation

- **Methods**
 - research on visualisation tools
 - expert user testing & peer-group evaluations
 - market analysis
- **Timeframe**
 - several projects 2010-2015
- **Techniques**
 - interdisciplinary research meetings
 - internal evaluation
 - workshops and interviews

Libraries, users & visualisation

Libraries: interest to visualise their data on maps

- improve **accessibility** of information for society
- improve citizens' **awareness** for their information

Users (incl. researchers) want to access library data in a way which is

- fast
- immediate
- flexible
- insightful
- attractively designed
- **non-resource-hungry** (little money, little time, & low technoskills)

Visual Analytics for Library (Users)

Why so little applied solutions and/or suboptimal ones?

- **Technological** factors
 - scalability
 - speed
 - openness
 - insightfulness
 - flexibility
 - user-friendliness

Libraries, users & visualisation

Innovation implementation: interrelation of factors

- GLAM **users** are typically **conservative** & innovation-resistant
 - too high threshold of new technological solutions, esp. if not out-of-the-box & in need of (interdisciplinary) development/adaptation
- **Scarce funding** for **libraries** (staff & non-staff resources)
 - choices must be made
 - priority is given to digitisation

Enter GlamMap

Evaluation by users

- Free like Google Maps but
 - More flexible, insightful, more beautiful

Evaluation by libraries

- Unrivalled combination of
 - scalability
 - speed
 - flexibility
 - openness
 - user-friendliness
 - insightfulness

Libraries, users & visualisation

Preliminary conclusions

- Use of viz tools like GlamMap in library (project) portals might
 - ▣ contribute to spreading awareness of sophisticated visualisation solutions
 - ▣ result in increased demand for more insightful, state-of-the-art visualisation technologies

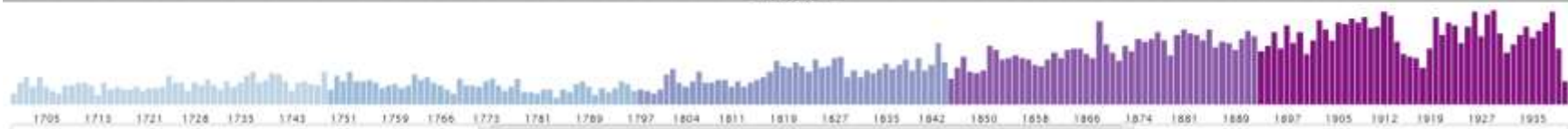
GlamMap: state of affairs

- development in strategic partnerships
 - *Visual Analytics for the World's Library Data* (2015-2019): U of Amsterdam, TU Eindhoven & OCLC
- give priority to develop fully customised solutions for software and services for specific groups of users
 - researchers
 - libraries & other GLAM actors (e.g. Rijksmuseum)
- regular releases of the well-researched and well-tested components of GlamMap as free software
- developing and testing of new innovative features in a closed environment
- most recent focus of development: **scalability**

GlamMap Demo

Title (1902 books)

Author	Title	Place	Year
Baumgarten, Alexander Gottlieb	Actiois logica, aucta et in systema redacta a Jo. G. Tolneri	Halle	1773
Beattie, James	Essay on the nature and immutability of truth	London	1773
Beattie, James	Verhandeling over de natuur en onveranderlykheid der waarheid	Utrecht	1773
Cudworth, Radolphus	Systema intellectuale hujus universi	Leiden	1773
Dulams, Louis	La logique ou l'art de raisonner	Paris	1773
Genesius, Antonius	Elementorum artis logico-criticae libri V	Bassano del Grappa	1773
Genesius, Antonius	---	Coimbra	1773
Helvétius, Claude Adrien	De l'homme, de ses facultés intellectuelles ...	London	1773
Herberth, Eardo	Elementa logicae electricae	Fulda	1773
Maki, Paulus	Compendiana logicae institutio	Vienna	1773
Ploucquet, Gotthard	Sammlung der Schriften, welche den logischen Calcul ... betreffen	Tübingen	1773
Sauro, Jean	Filosofia de	Ferris	1773



Scalability

Scalability

- Performance
- Visual design



Importing Trove

Importing

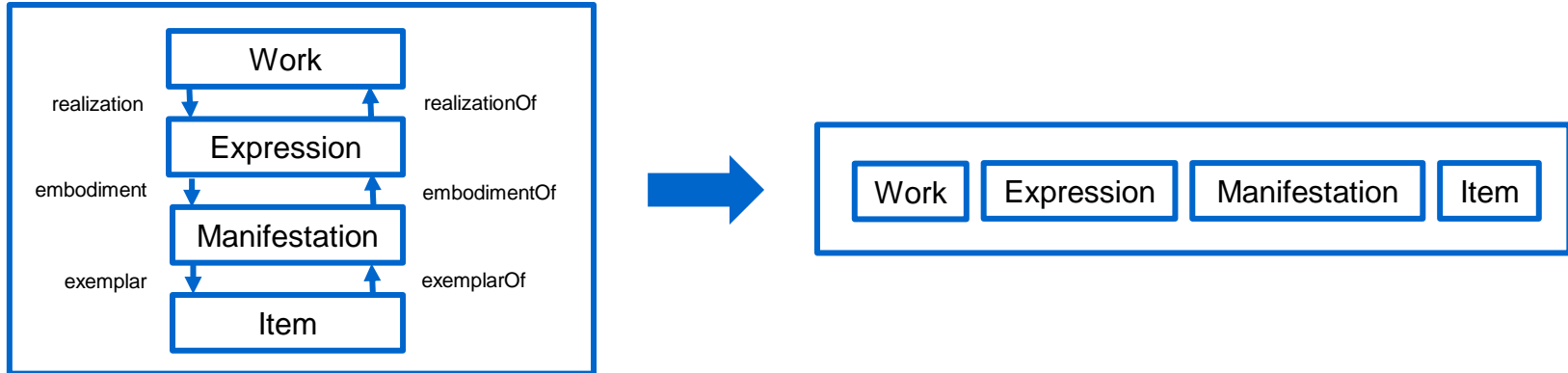
- Cannot use Trove data on the fly
- Import in our database
- 60,000,000 records



Accessing Trove data

Accessing

- FRBR hierarchy
 - Works / Expressions / Manifestations / Items
 - Aeneid / Fitzgerald's translation / paperback edition / copy thereof
- Access requires grouping and “joins” ➔ **flatten**



Geocoding

Geocoding

- Translate city names into coordinates
- Use online service: [GeoNames.org](https://www.geonames.org)
 - ▣ Limited to 30,000 per day / 2,000 per hour
- **Solution**: group same names for efficiency

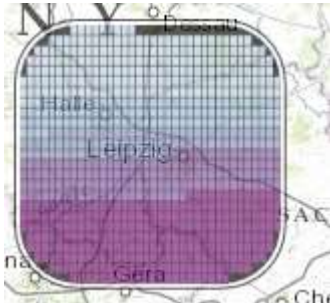
“London”



Symbol changes

Drawing performance

- Drawing individual shapes takes time
- 60,000,000 items → not visible & very slow
- **Solution:** combine items in same category
- Ignore rounded corners, no longer relevant



Visual scalability

Visual scalability

- Symbols with many items may cover entire map
- **Solution:** “compress” symbols with many items



Future challenges

What we are working on

- Filtering on big data sets
- Faster clustering
- Improved geocoding



Import your data

Explore your own data

- Your data can be imported into GlamMap
- GlamMap.net/glamdev
- Must sign up to import data

(Temporary) restrictions

- Requires a specific format (details on webpage)
- Cannot be too big

Use Case: History of Ideas

Arthur O. Lovejoy (1873-1962): *The Great Chain of Being* (1936)

- studies how concepts change through time across contexts
- largely interpretive & text-based, crucially relying on sources search
 - large-scale amount of diverse types of textual sources scattered in libraries or archives across the planet, possibly very many *unknown*
 - Identification of source by metadata a major task

Glammapping Trove aids strongly sources identification

- quick visual overview of most/all author's works, their translations/editions or contemporary works by other authors elsewhere
- easy identification of both known and unknown sources
- results easier to navigate than list-based