How we (do not) share our data in humanities

Eleri Aedmaa
Institute of Estonian and General Linguistics

UNIVERSITY OF TARTU

RDA Meets Estonian Researchers
8 November 2017, Tartu, Estonia
Data in humanities

Platforms

Challenges

Solutions
Personal experience

SHOW ME THE DATA!!!!
What is our data?

✔ text, e.g. books, newspapers, documents, blogs, tweets
✔ audio, e.g. recordings of speech, sounds, music
✔ video, e.g. recordings of interviews, fieldwork
✔ image, e.g. photos of findings
✔ metadata, e.g. when and where the photos of findings were taken
Main resources

- text and speech corpora
- dictionaries and other lexical resources
- databases
- archives
- repositories
Organisations and platforms – some examples

- For linguistic data – Center of Estonian Language Resources – text and speech corpora, lexical resources, databases, language processing tools by University of Tartu, Tallinn University of Technology, Institute of the Estonian Language and Estonian Literary Museum

- University of Tartu Library, National Library of Estonia – databases

- The National Archives of Estonia – e.g. millions of digital images with free of charge access in Saaga
When humanists share data?

- required by a funder as an outcome of the project
- advised or asked by the reviewers/supervisors
- personal interest and skills
- presence of a good and an efficient platform
Why humanists do not share data?

- little community
- no funding and time
- no skills
- no demand
- data includes sensitive information
- fear
- documentation
How to push data sharing in humanities?

✔️ explain the importance and benefits of data sharing

✔️ educate students/researchers/colleagues about sharing possibilities (e.g. repositories)

✔️ make it easy for the researchers to share their data

✔️ make it difficult to publish without data sharing
Conclusion

✔ Data in humanities is diverse and might be difficult to digitize and share.

✔ Humanists lack confidence, skills, money and reasons to share their data.

✔ Besides creating platforms and other conditions for the data sharing, it is important to explain the necessity of data sharing.