ICAME42 | tu
crossing boundaries through corpora

*The **Language Technology and Data Analysis Laboratory (LADAL)***

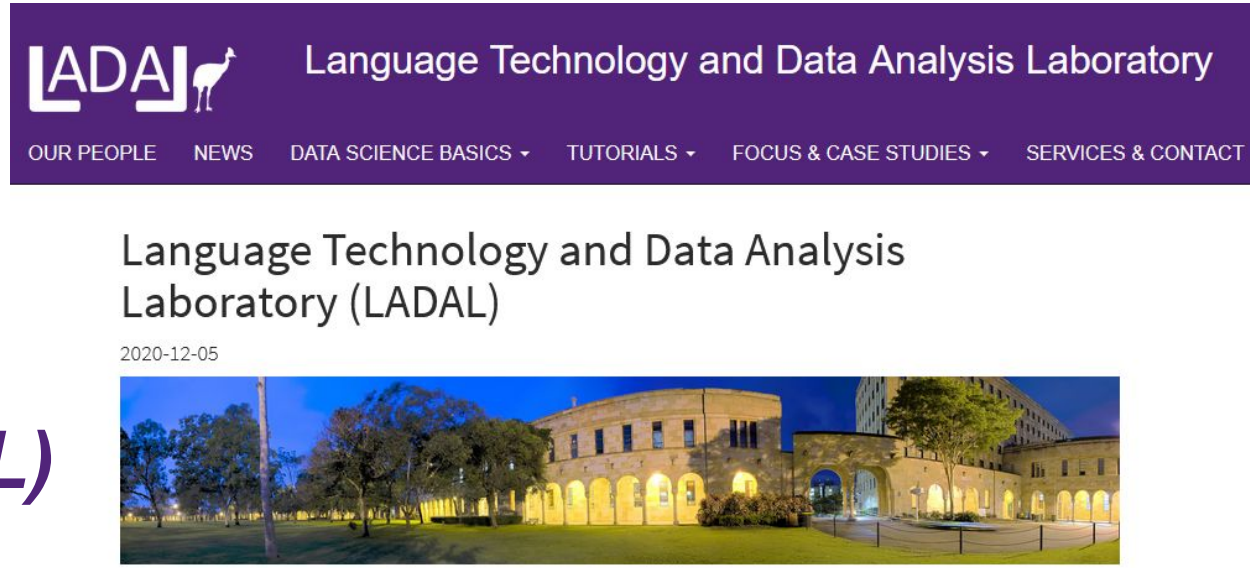
Martin Schweinberger

THE UNIVERSITY OF QUEENSLAND AUSTRALIA
CREATE CHANGE

NORGES ARKTISKE UNIVERSITET • UIT •
Logosymbol.

LADAL

AcqVA Aurora Lab



Background and Motivation

Consultation (statistics, designs | tools, data management)

- An accessible resource (website + repo) I can point to for people interested and motivated to work professionally with data
  - Not various tools but a single environment (R | RStudio)*
  - Quant + Comp Basics | Data Management | Data processing | Viz | Statz | TA | Case Studies
  - RNotebooks (Rmd): Formatted text | Code | Output**
  - Light(!): free, open source, transparent, collaborative & potentially crowd-sourced
  - Community building (events, webinars, e.g. LADAL Opening Webinar Series 2021:)

* No longer set on R (also Python and selected tools)

** Interactive | Screencasts & Video Tutorials | Exercises

Challenges

- Audiences with very different **levels of expertise** (from absolute beginners to expert users)
- Audiences with vastly **different interests**, expectations, and needs
- **Inertia**
  - people want to continue using the tools they are used to
  - drag & drop | point & click
    (e.g. SPSS; blackbox tools: Leximancer, Voyant tools)
- Establishing infrastructures requires **resources**
  - Fantastic support at UQ (SLC & Michael Haugh)!
  - Team: difficult to find qualified people
    (often only temporary or in-betweens)
  - A lot of work: RAs, staff, contributors & admin

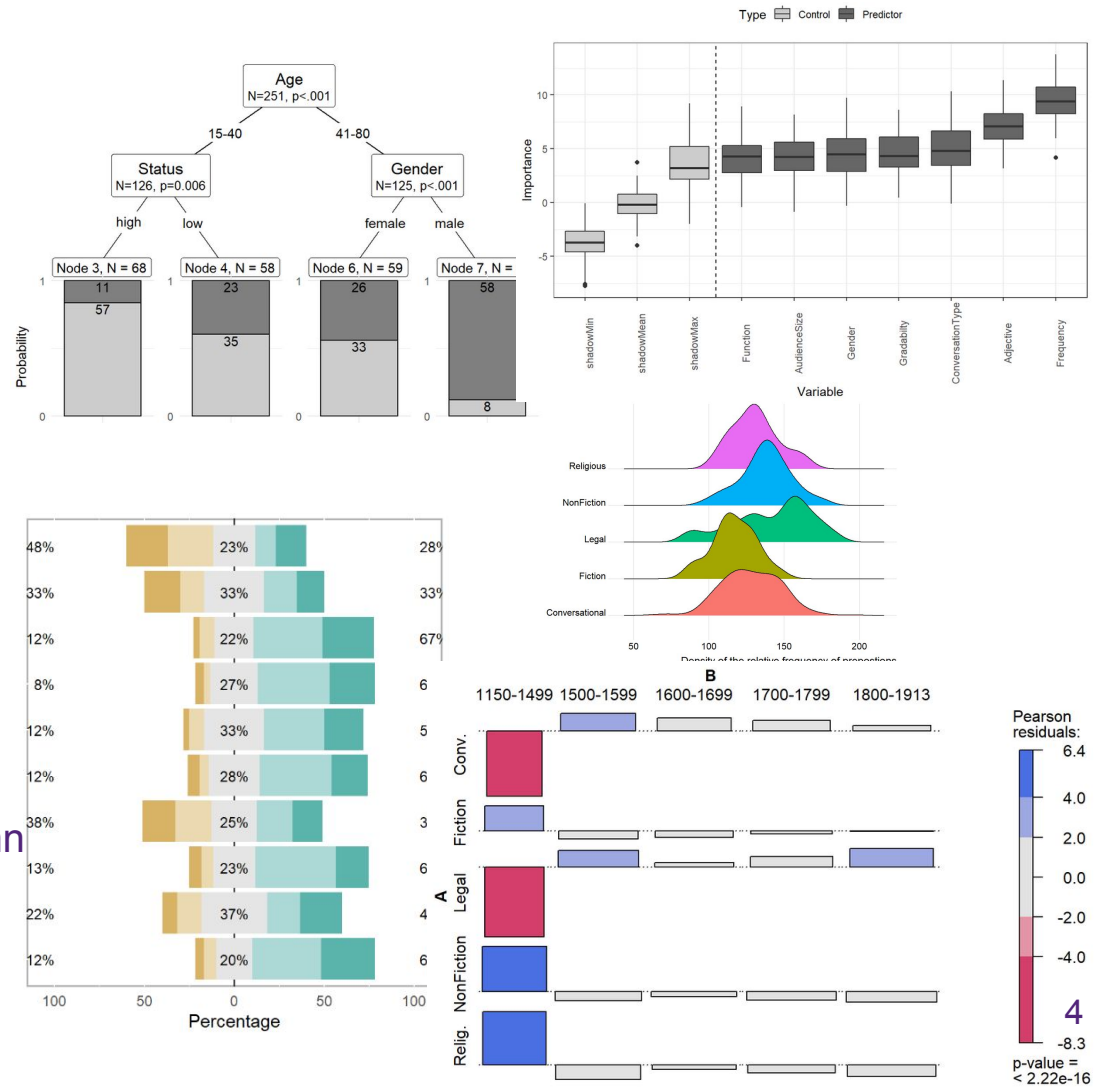
Language Technology and Data Analysis Laboratory (LADAL)

What we have achieved so far

- Self-guided study **materials (**online **tutorials)** on data extraction, processing, management, visualization, statz and learning to “code”
- Face-to-face **consultations** and practical hands-on **workshops**
- **Webinars** (e.g. LADAL Opening Webinar Series 2021, see https://slcladal.github.io/opening.html)
- **Collaboration**

  (e.g. Text Crunching Center at UZH, AcqVA Aurora Lab, Media Research Methods Lab, VARIENG, see https://slcladal.github.io/people.html)
- **Collaborators**

  (e.g. Stefan Gries, Erich Round, Gerold Schneider, Gregor Wiedemann Joseph Flanagan, etc. see https://slcladal.github.io/people.html)

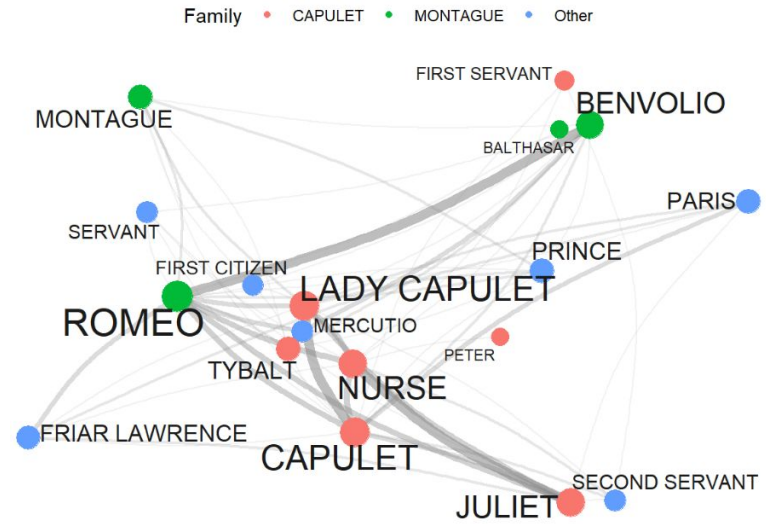


Language Technology and Data Analysis Laboratory (LADAL)

What we hope to do in the future

- Correct and update existing materials
- Add exercises (target materials for course content)
- Video tutorials & screencasts (add YouTube channel)
- Expands methods and case studies (Speech + Text)
- Activate crowd-sourcing and start community building

Network of persons in Shakespeare's Romeo and Juliet



ICAME42 | tu
crossing boundaries through corpora

*The **Language Technology and Data Analysis Laboratory (LADAL)***

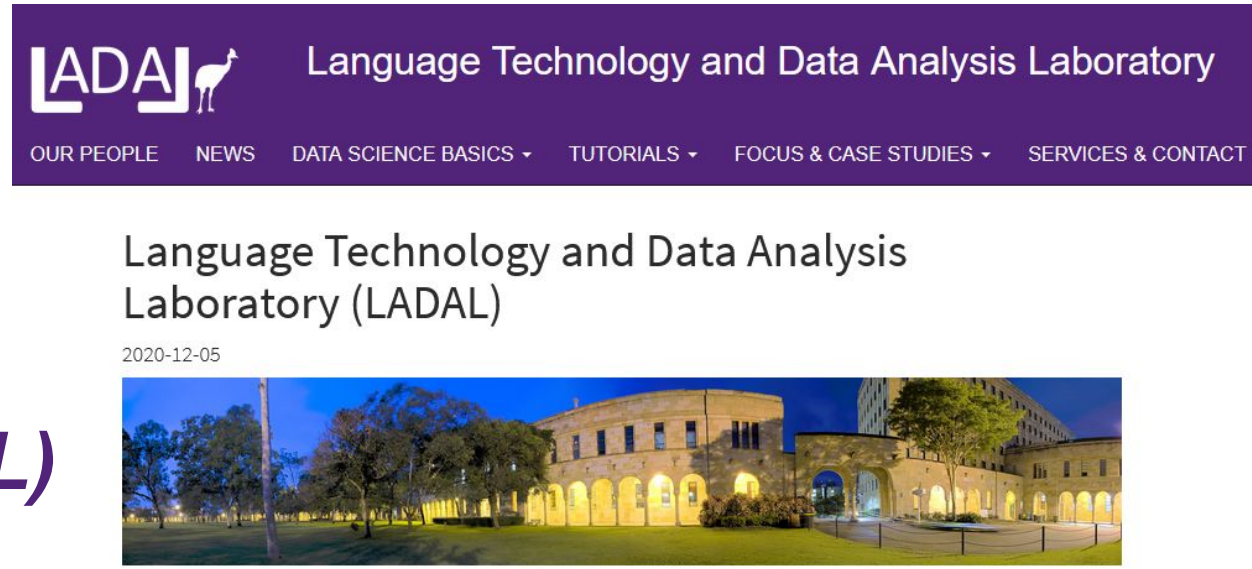
Martin Schweinberger

The University Of Queensland AUSTRALIA
CREATE CHANGE

Norges arktiske universitet · UIT
Logosymbol.

LADAL

AcqVA Aurora Lab



References

LADAL: The Language Technology and Data Analysis Laboratory, LADAL. School of Languages and Cultures, The University of Queensland, Access Aug. 13, 2021, url: https://slcladal.github.io/index.html.

AcqVQ Aurora Lab. UiT Aurora Center for Language Acquisition, Variation, and Attrition, The Arctic University of Norway, Tromsø, Access Aug. 13 2021, url: https://site.uit.no/acqvalab/.

PoLaR: Psycholinguistics of Language Representation (PoLaR) lab. UiT Aurora Center for Language Acquisition, Variation, and Attrition, The Arctic University of Norway, Tromsø, Access Aug. 13 2021, url: https://site.uit.no/polar/.

TCC: Text Crunching Center, TCC. Department of Computational Linguistics, University of Zurich, Access Aug. 13 2021, url: https://www.cl.uzh.ch/en/TCC.html.