Antonio Martini

Professor, University of Oslo

17-10-2023

DO WE KNOW OUR TD BETTER THAN AI?

Who is Antonio Martini?

- ◉ Italian
  - No kebab pizza! ☺
  - 12 years in Scandinavia– survived many winters!

- ◉ Previously
  - Software Developer
  - PhD in Software Engineering
  - Postdoc
  - Independent Consultant
  - Principal, Strategic Researcher at CA Technologies (now Broadcom)
  - 12 years of research and practice on technical debt

- ◉ Hobbies
  - Board games, strategy computer games, pool, etc.
  - Football, volleyball, beach volley, fencing
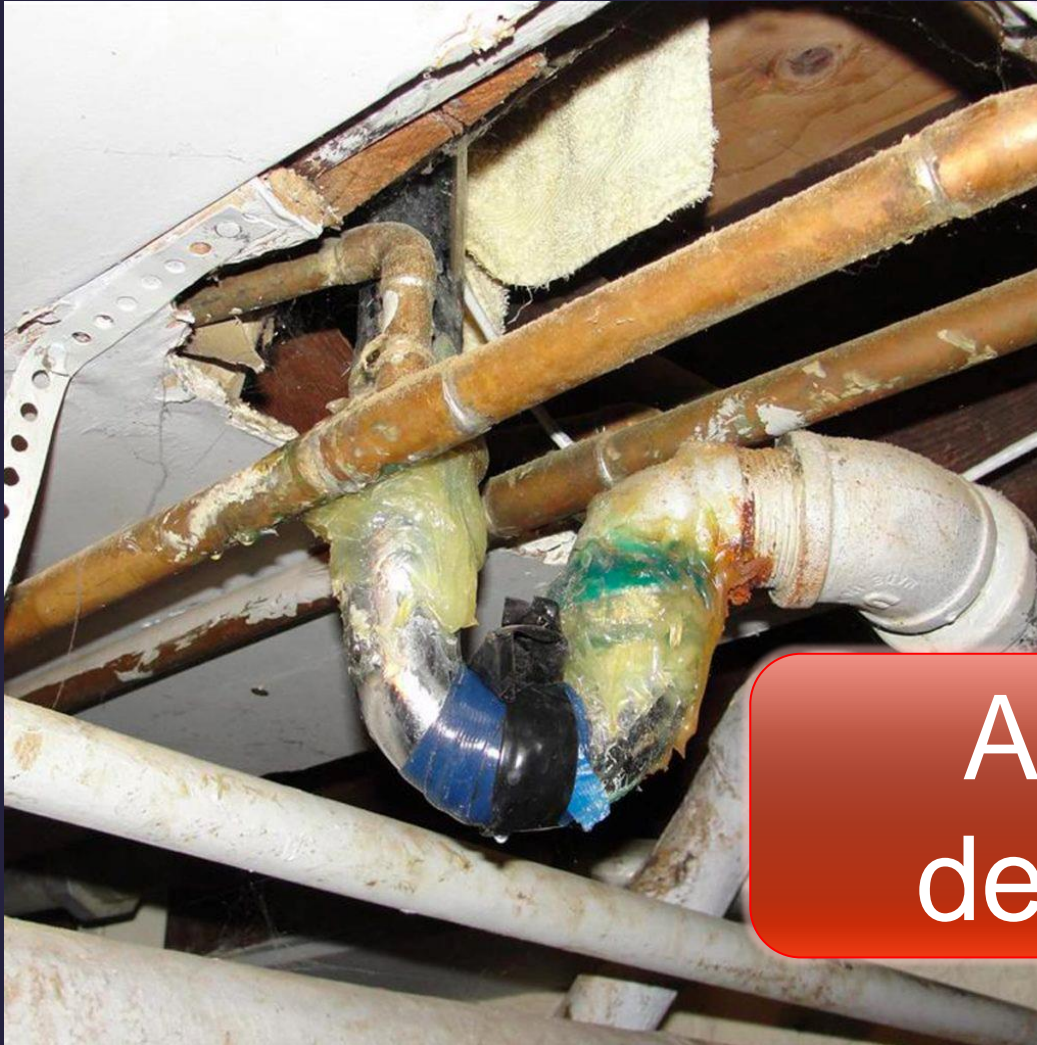 - Piano, Drumset, etc.
  - Travel!
  - …and no time for them! ☺

What is technical debt?

What the users see

Wonderful features!

What the developers see

Awful design!

TechDebtOps – A 3-year project funded by the Research Council of Norway

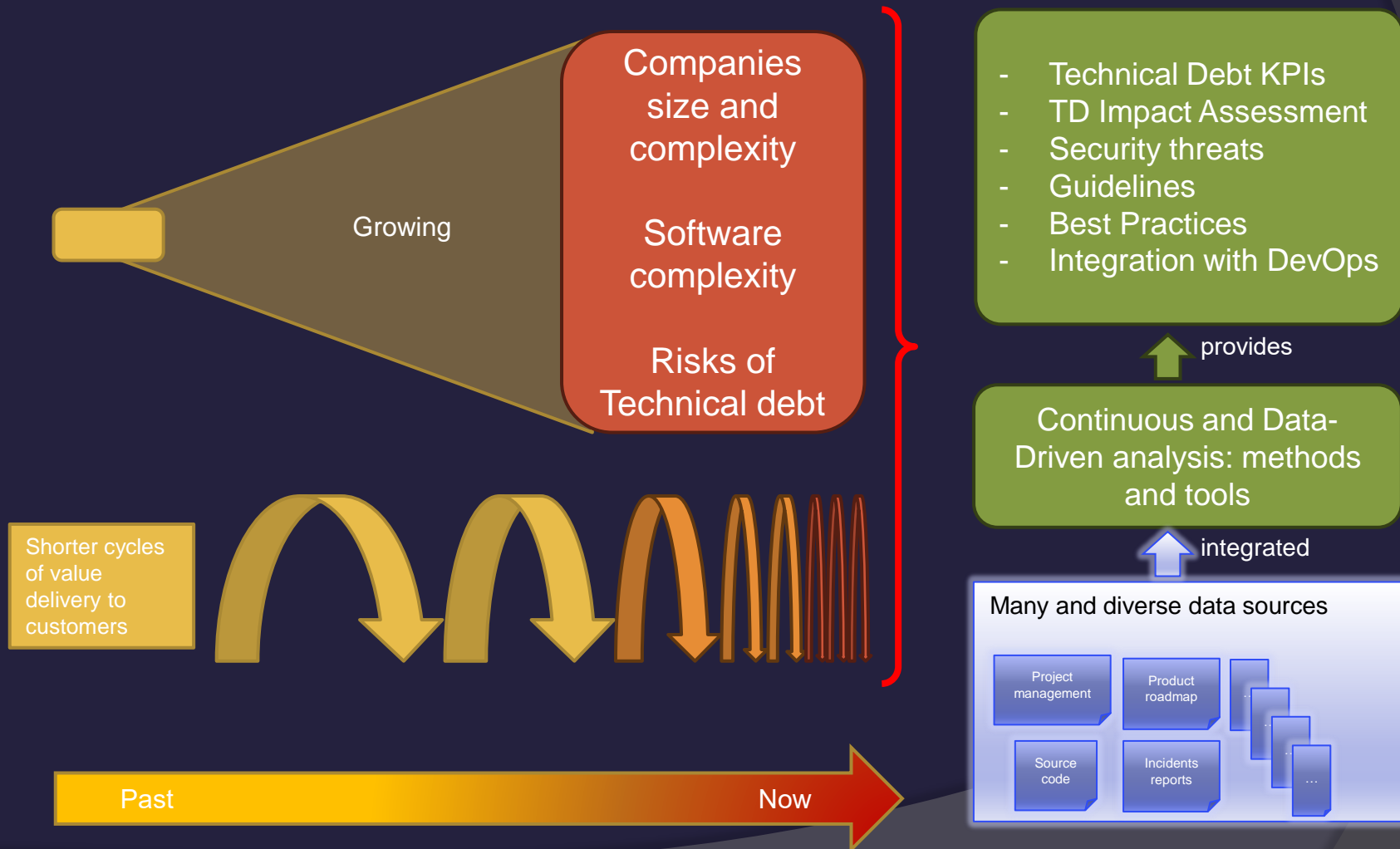**Data-driven continuous management of technical debts for sustainable software development**

Challenges and goals of technical debt

Growing

Companies size and complexity

Software complexity

Risks of Technical debt

- Technical Debt KPIs
- TD Impact Assessment
- Security threats
- Guidelines
- Best Practices
- Integration with DevOps

provides

Continuous and Data-Driven analysis: methods and tools

integrated

Many and diverse data sources

Project management

Product roadmap

Source code

Incidents reports

Shorter cycles of value delivery to customers

Past

Now

AI trends

- Current trends in AI
  - Large Language Models (LLMs)
  - Transformers-based (fast-)training

Generative AI for software development

- Why LLMs in software development?
  - Many past (and limited) approaches are often based on one specific data source only
    - E.g. Code repo analysis
  - We need more context!
    - A lot of the context is encoded in documents using natural language

(semi-) automatic tracking TD: ML-based recognition of TD in github issues *

* D. Skryseth, K. Shivashankar, I. Pilán and A. Martini, "Technical Debt Classification in Issue Trackers using Natural Language Processing based on Transformers," *TechDebt 2023*

Classification of TD in github issues

* D. Skryseth, K. Shivashankar, I. Pilán and A. Martini, "Technical Debt Classification in Issue Trackers using Natural Language Processing based on Transformers," *TechDebt 2023*

We can monitor technical debt in a more context-dependent way
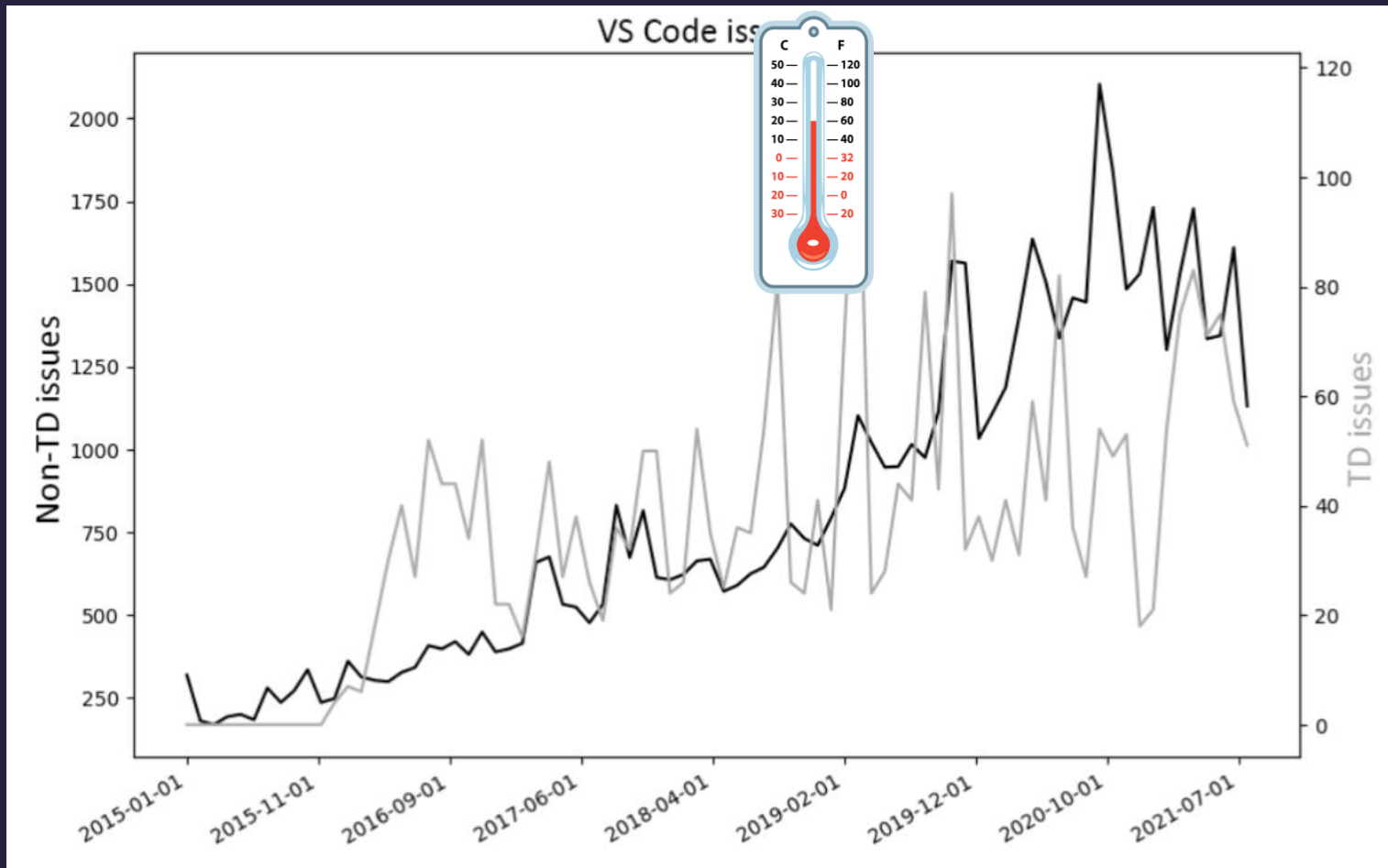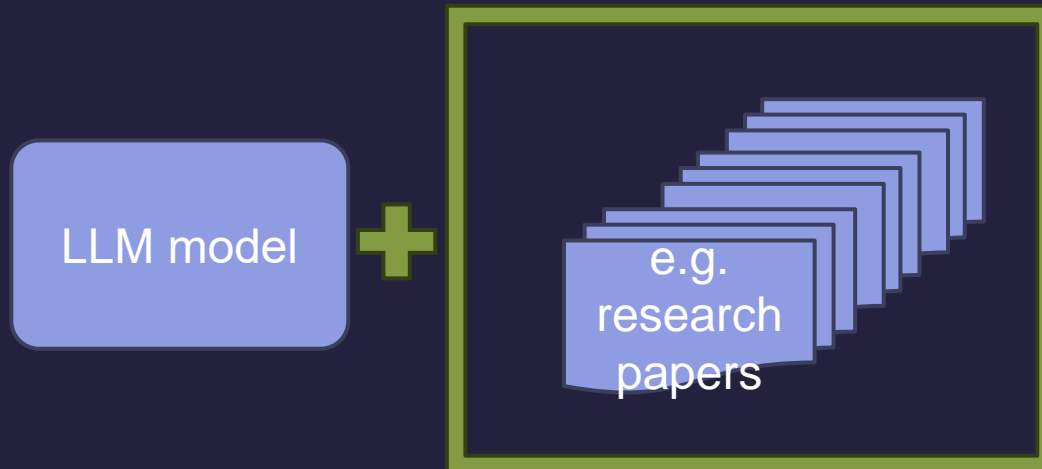

* D. Skryseth, K. Shivashankar, I. Pilán and A. Martini, "Technical Debt Classification in Issue Trackers using Natural Language Processing based on Transformers," *TechDebt 2023*

Technical debt guidelines with AI

- LLMs fine-tuning or retrieval augmented generation
  - reliable sources
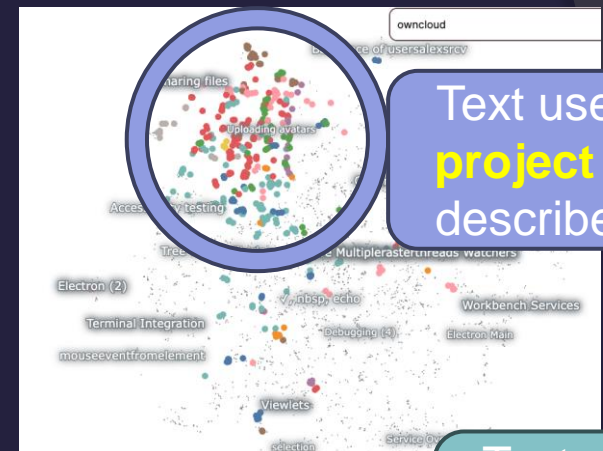

LLM model **+** e.g. research papers

TD visualization with AI
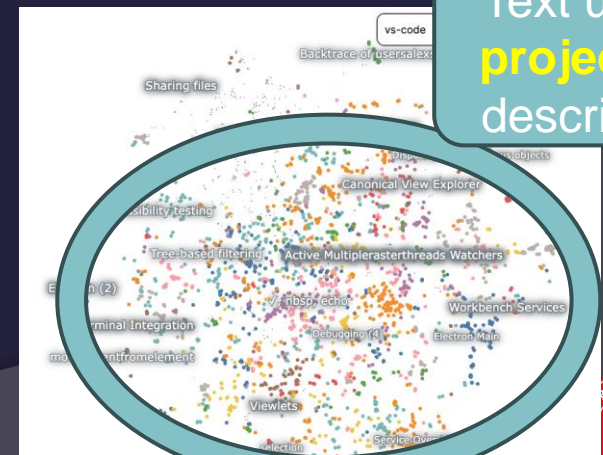
⊙ Different context = different TD

Clustering of topics used when describing TD issues
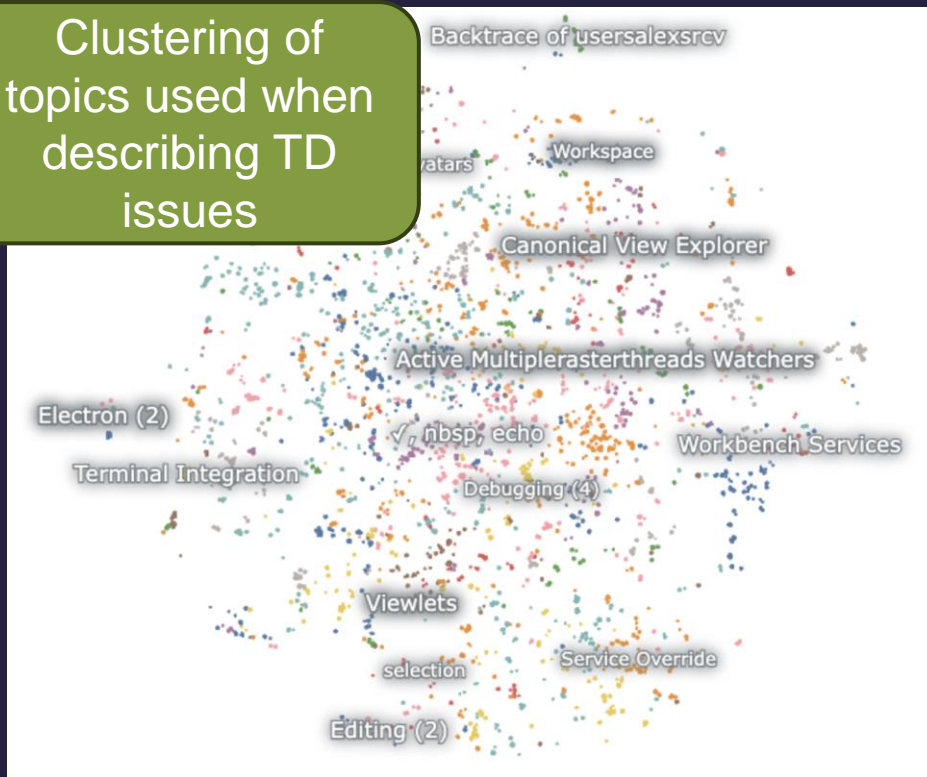
Text used in **project A** to describe TD

Text used in **project B** to describe TD

Questions?

Comments?

- References:
  - For further information:
    - antonio.martini@ifi.uio.no