

Mining Software Repositories

Software Evolution – L6T1

Dr. Vadim Zaytsev aka @grammarware, March 2021



Empirical Software Evolution

- Software is an observable phenomenon
 - like weather
- Empirical bottom-up:
 - seek hypotheses
- Empirical top-down
 - test hypotheses

Mining Software Repositories..

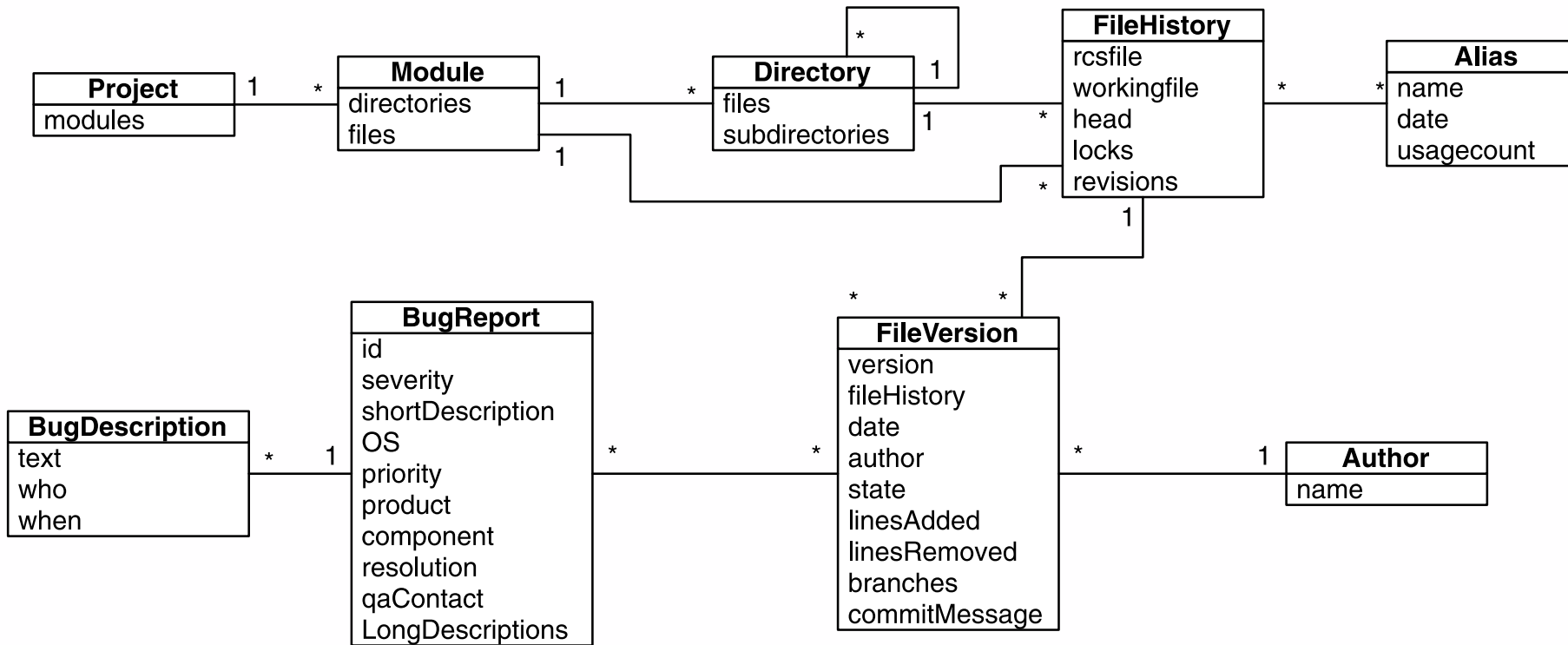
- Developer effort
- Social network
- Change impact
- Hotspot analysis
- Defect prediction



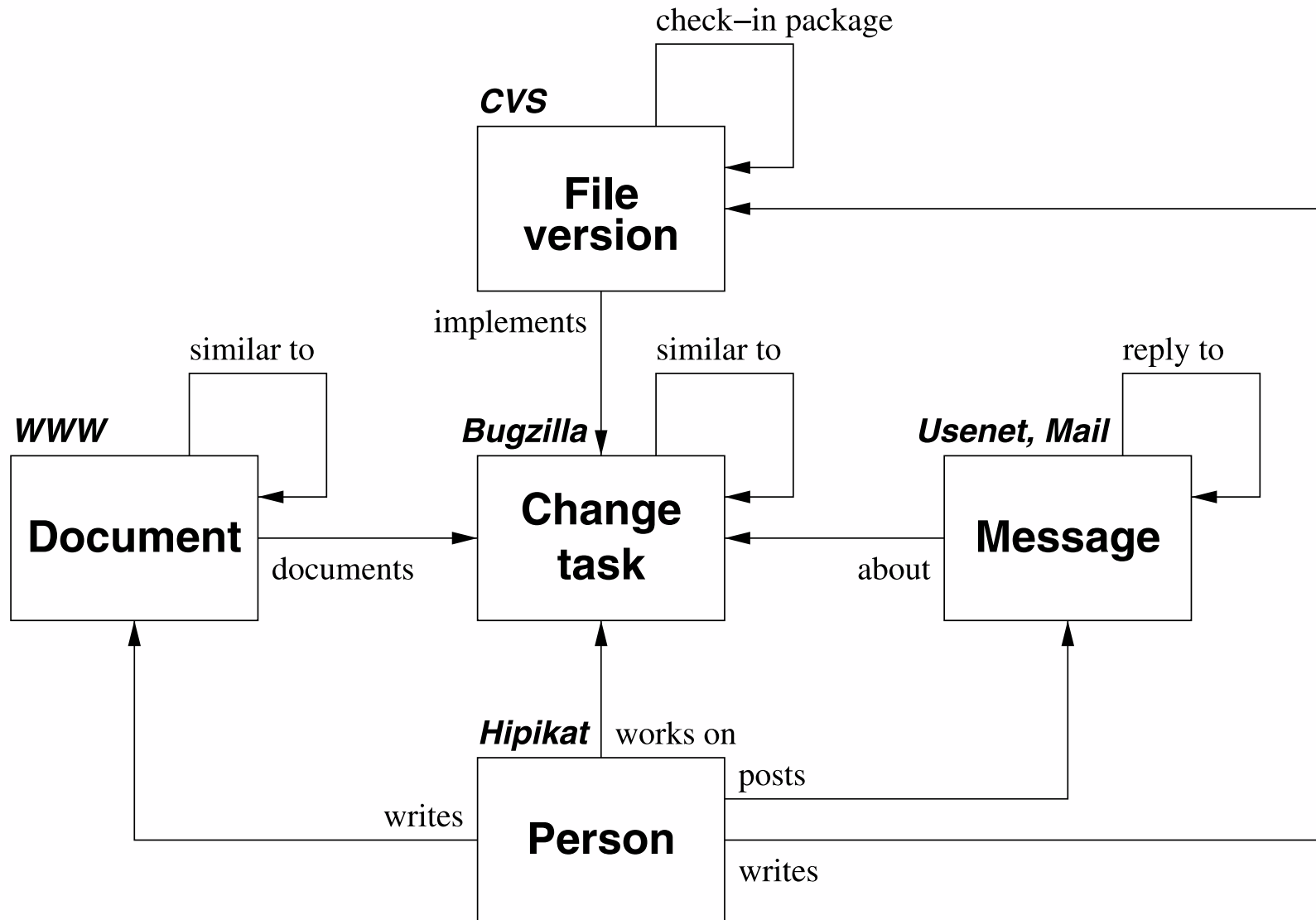
Mining Software Repositories

- Data retrieval
 - source control, bug tracker, docs, emails, chats...
- Data modelling
 - models, database
- Data analysis
 - tools, panels

Release History Database Model



Hipikat Model



Questions Answered by MSR

- What kind of projects live longer?
 - <https://doi.org/10.1145/3379597.3387511>
- Does copy&paste programming lead to license violations?
 - <https://doi.org/10.1145/3379597.3387455>
- Do users leave projects where bugs don't get fixed fast?
 - <https://doi.org/10.1145/3379597.3387456>
- Is *software* just *code* with some extras?
 - <https://doi.org/10.1145/3379597.3387442>

Conclusion

- MSR is a hot topic
- Good for comprehension and analytics
- Next: understanding and prediction
- Q&A Sessions @ Canvas
 - ⇒ v.zaytsev@utwente.nl
 - ⇒ <https://discord.gg/n7VQAPNBPD>