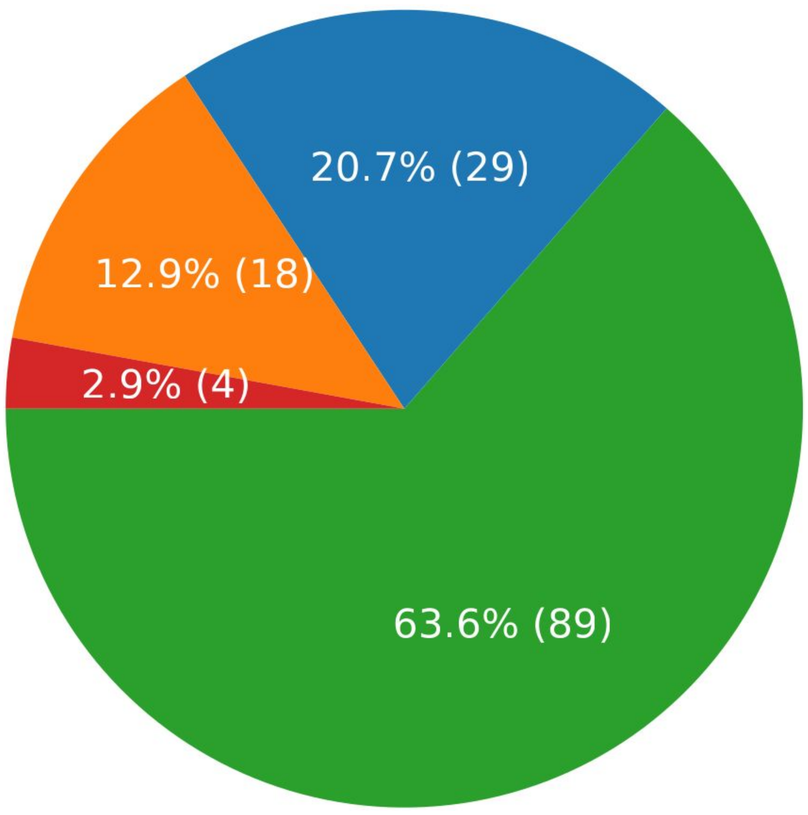
Automatic Discovery of Artifacts in Cybersecurity Literature

Marton Bognar, Arthur Bols, Jo Van Bulck
DistriNet, Department of Computer Science, KU Leuven

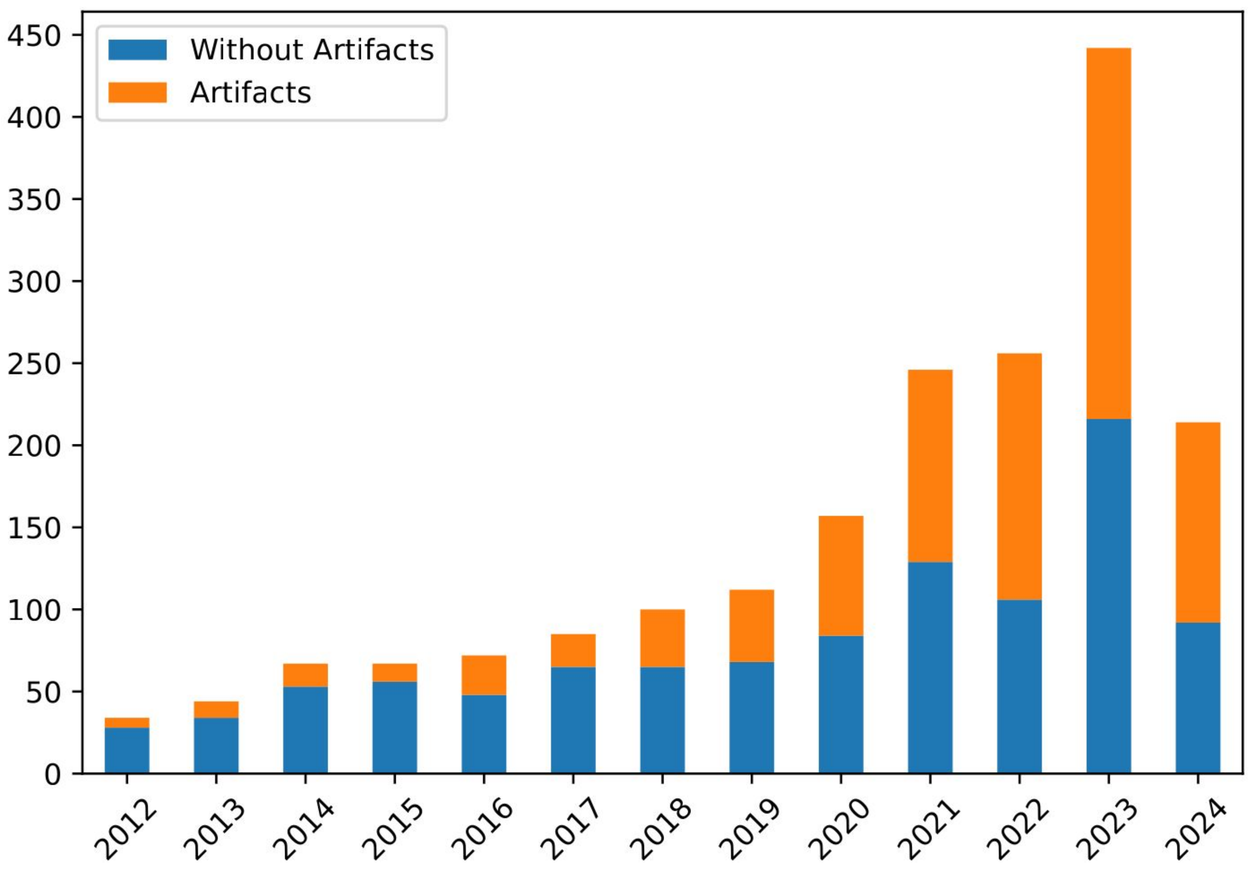
Goal:
- Automatically discover **artifacts** in papers
- Check artifact availability

Steps:
- Scrape PDFs → parse → **candidate URLs**

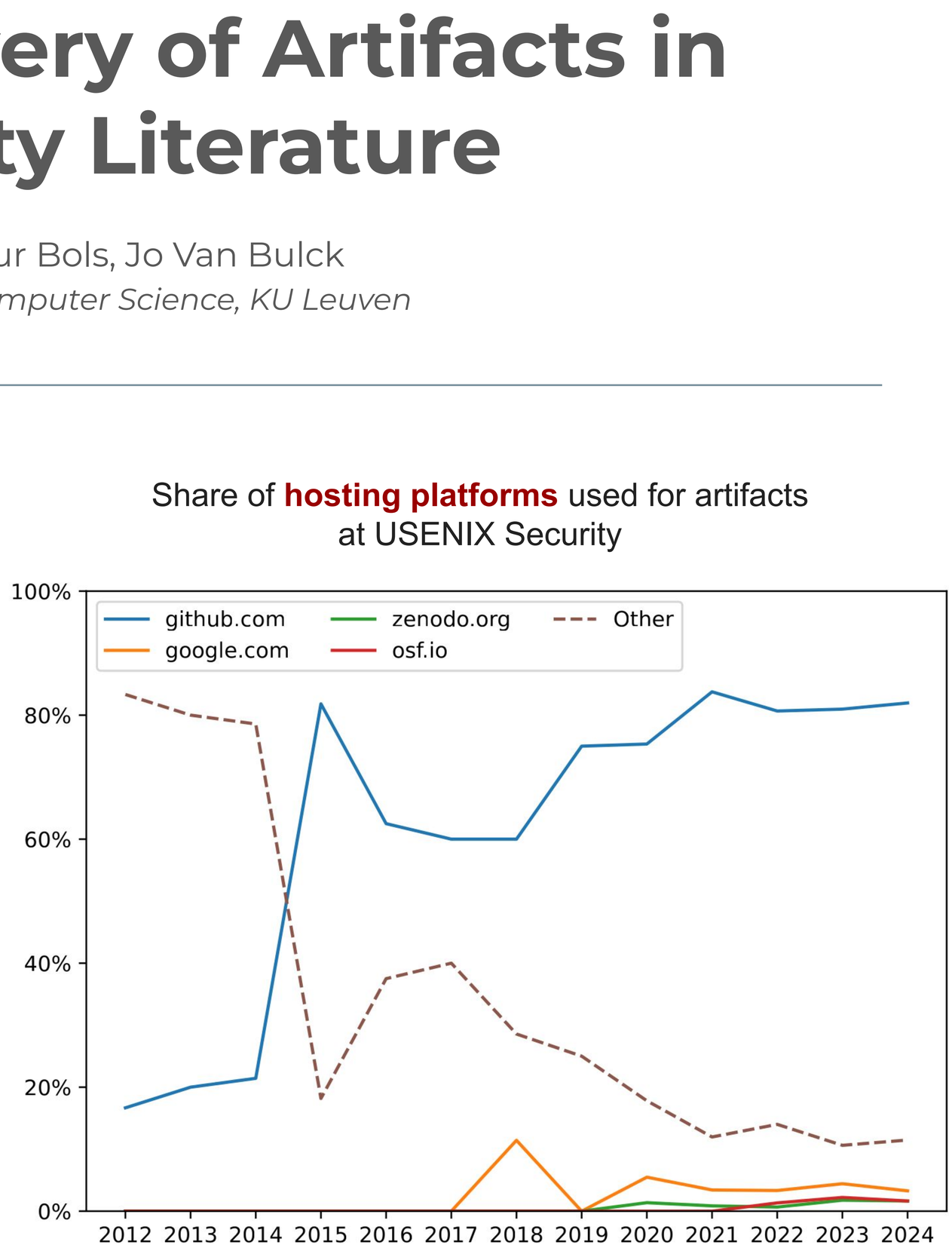
Detection rate of our tool.
Green: exact match, Blue: alternative link, Orange: link missing or unavailable, Red: link not detected



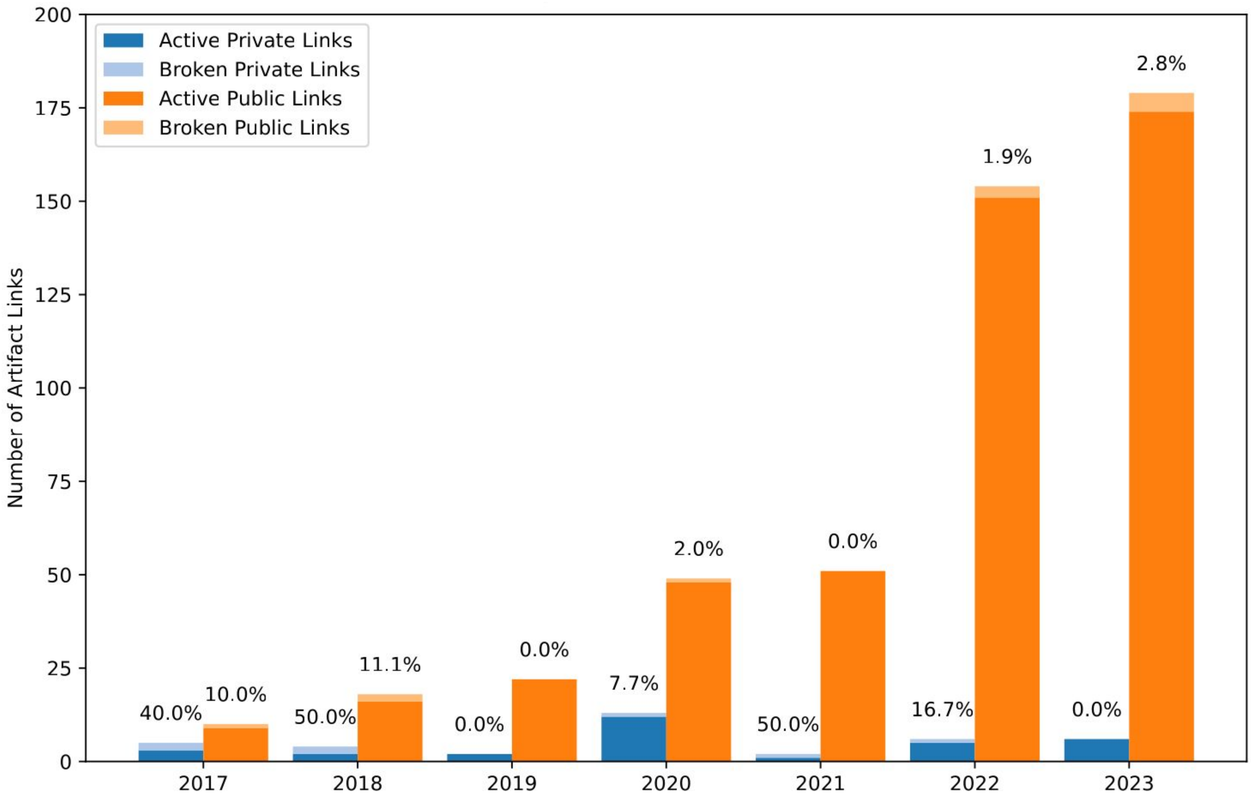
USENIX Security papers with and without **linked artifacts**



Share of **hosting platforms** used for artifacts at USENIX Security



Availability of artifact links hosted on **public services** and **private** (personal, institutional) websites



Future work:
- Investigating recent trends: **mandatory artifacts, stable archiving**
- Comparison to other CS fields
- More historical analysis, data repository
- Comparison with other tools (**LLMs**)