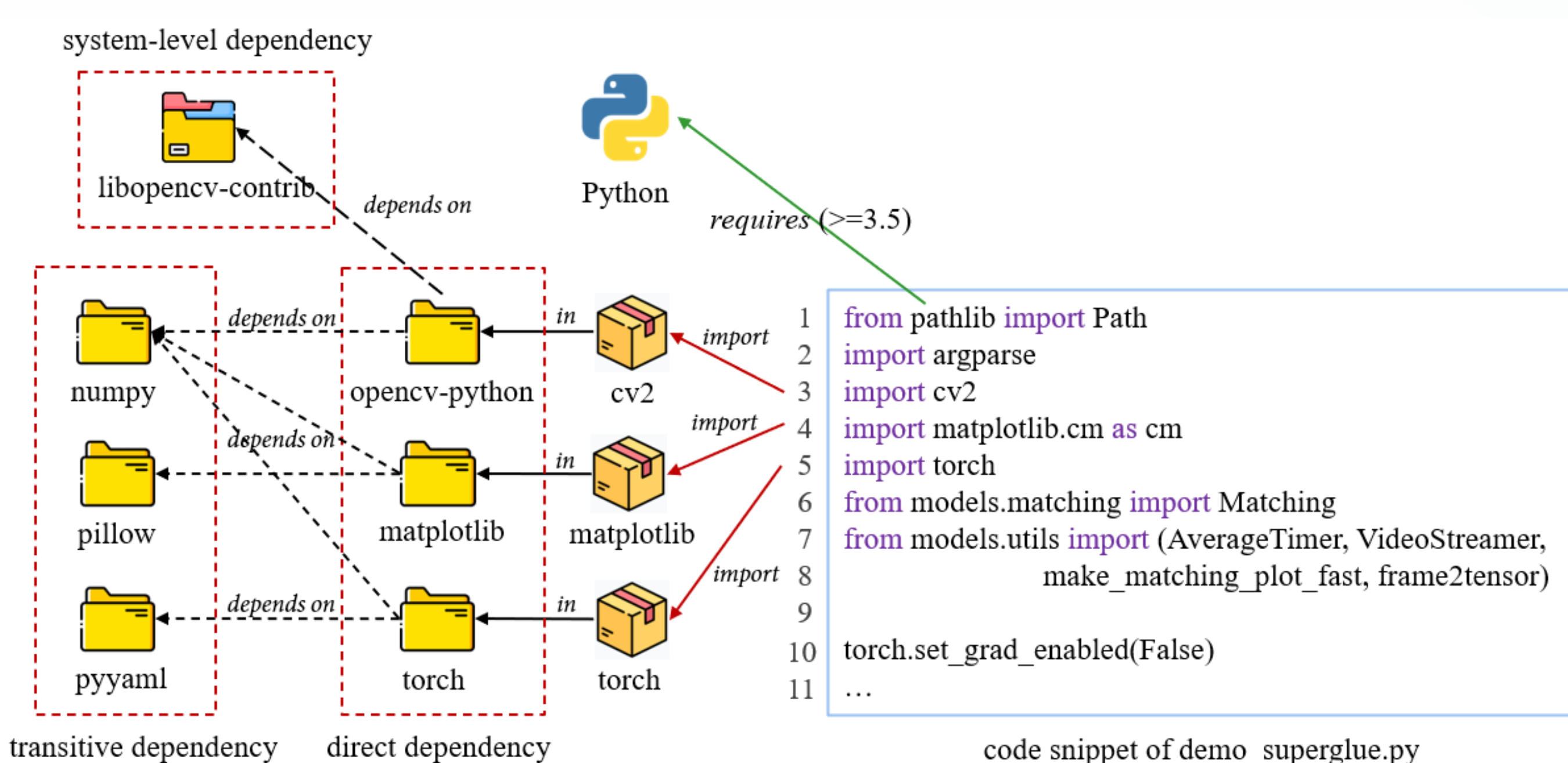# 基于知识的**Python**程序环境依赖推断

叶宏杰，陈伟，窦文生，吴国全，魏峻

Knowledge-Based Environment Dependency Inference for Python Programs

The 44th International Conference on Software Engineering (ICSE'22)

联系方式：陈伟，chenwei@otcaix.iscas.ac.cn
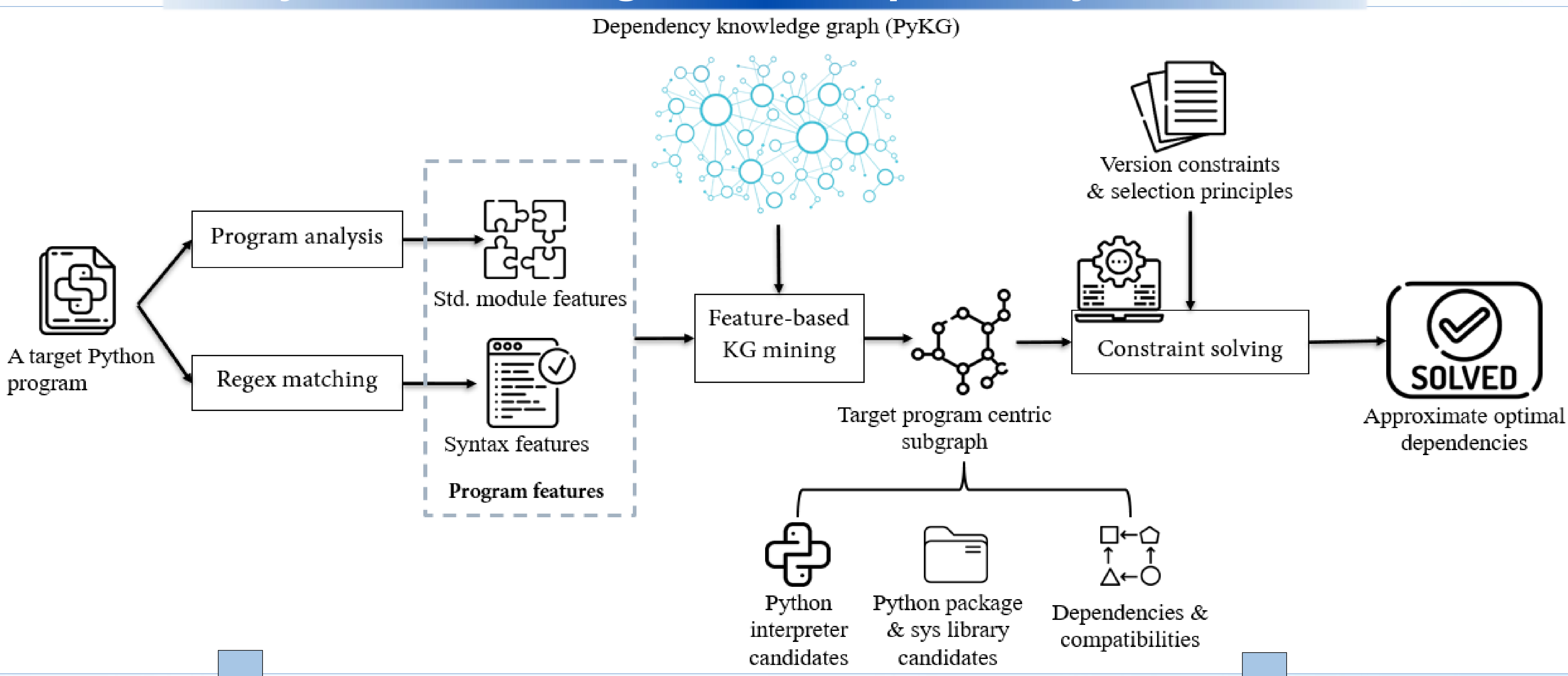
Dependencies are complex and diverse
- Third-party packages (TPP)
- System libraries
- The Python interpreter

Version constraints make things worse
- Python version constraint
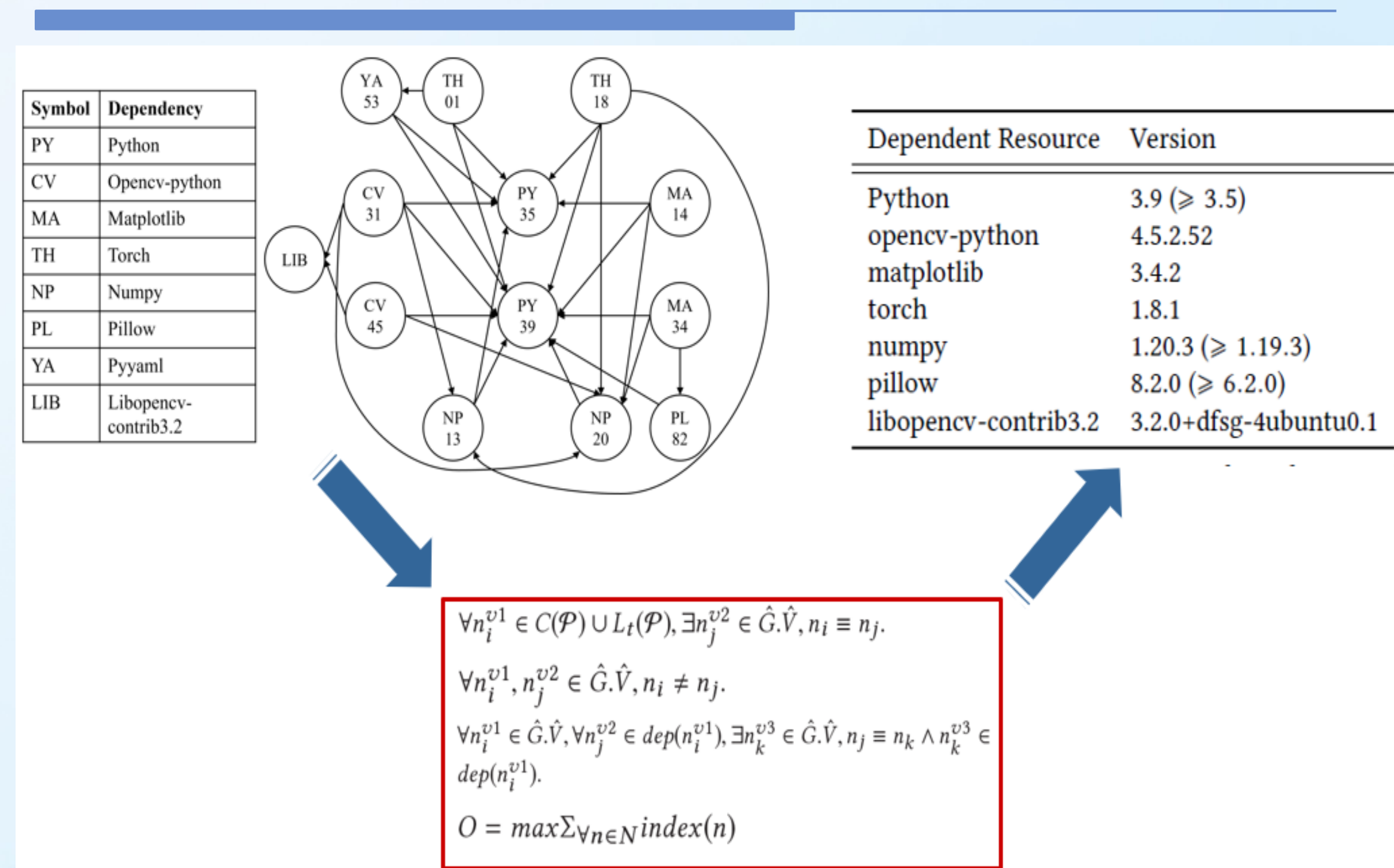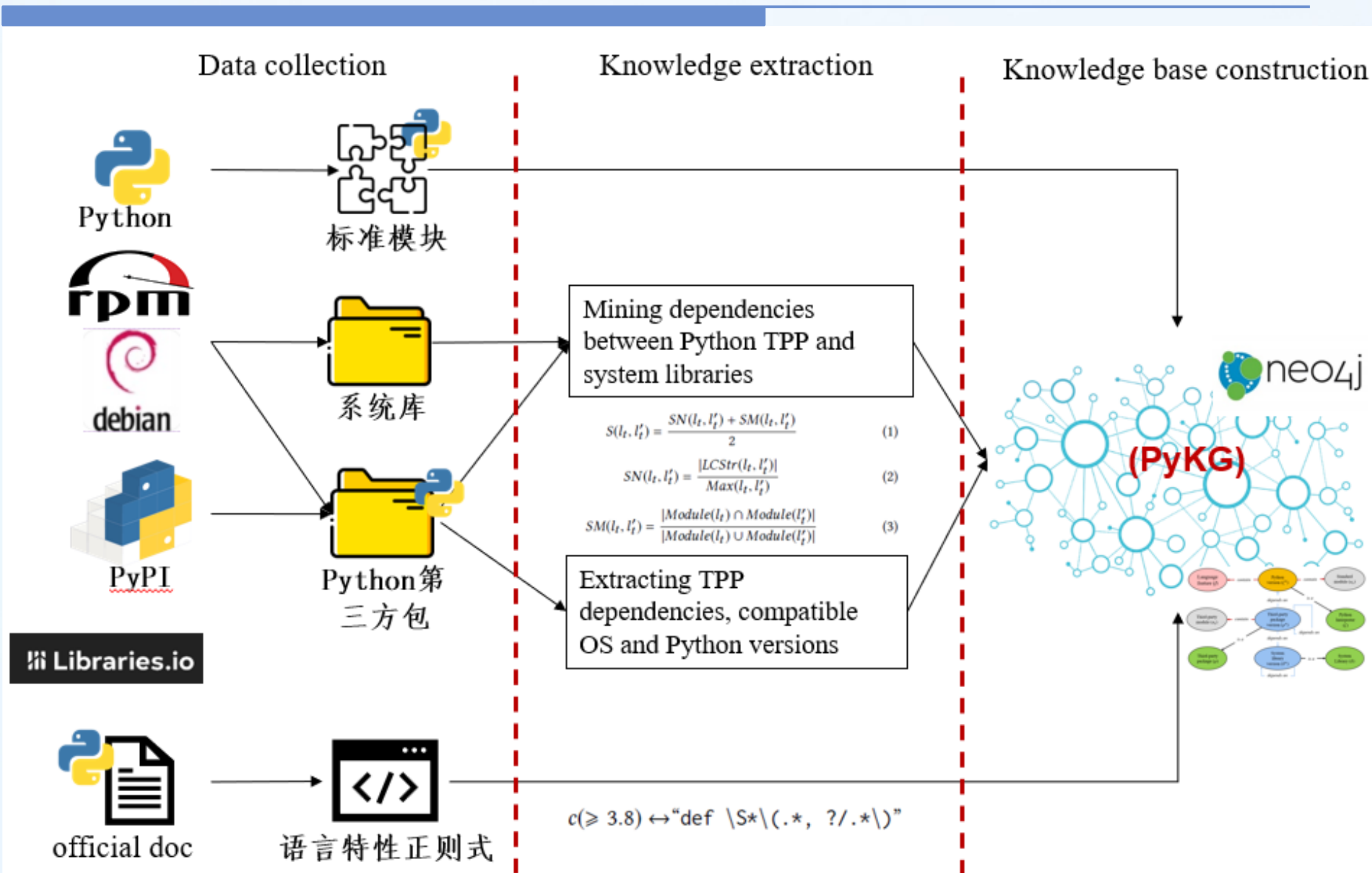- TPP version constraint

Program build and runtime failures are prevalent

PyEGo: Knowledge-Based Dependency Inference

PyKG: Knowledge Graph

Dependency Inference

Evaluation

	Tool	ACC	ADP	ATP	AT (sec.)
HG2.9K	**PyEGo**	**46.14%** (1334/2891)	3.71	**1.41**	**0.69**
	pipreqs	10.27% (297/2891)	**1.52**	1.52	2.18
	DockerizeMe	30.85% (888/2891)	7.32	5.99	13.45
SD	**PyEGo**	**62.00%** (62/100)	8.91	**4.01**	2.52
	pipreqs	45.00% (45/100)	**6.25**	6.25	**2.43**
	DockerizeMe	23.00% (23/100)	13.18	10.41	10.37

PyEGo outperforms the state-of-the-arts significantly
- ACC is **1.4x – 4.5x** of pipreqs and DockerizeMe
- The fewest third-party packages on average
- PyEGo runs **0.1x – 18.5x** faster than pipreqs and DockerizeMe