Appendix E: Different systems for data extraction and analysis

Table 1 Different systems for data extraction and analysis

	System 1	System 2	System 3
Description	Ubuntu 22.04.02 LTS, Linux 5.15.0	macOS 15.4 with arm64	Ubuntu 22.04.03 LTS, Linux 5.15.0 with 128 CPUs per task and 64 GB memory
Python version	3.10.12	3.12.0	3.10.12
Django version	4.1.7		
Requests version	2.28.2		
Pandas version	1.5.3	2.1.1	2.1.1
Matplotlib version		3.8.0	
Seaborn version		0.13.0	
Sklearn version		1.3.1	1.3.1
Numpy version		1.26.0	1.26.0
Scipy version		1.11.3	1.11.4
Statsmodels version		0.14.0	
Joblib version			1.3.2
Yaml version		6.0.1	
Quarto version	1.3.361	1.5.41	
Jupyter notebooks version	6.5.3	7.0.4	
Shap version		0.46.0	

The following libraries were used with a random state and seed of 42 wherever applicable:

- Data extraction: django, requests
- Data preparation: pandas
- Data visualization: matplotlib, seaborn
- Feature selection: pandas, numpy, sklearn
- Hyperparameter tuning: sklearn, numpy
- Feature Importance: shap
- Statistical tests: numpy, scipy, statsmodels
- Authoring system: Quarto with Jupyter notebooks