Package manager Nextflow mlflow Guix Nix Know configuration Apptainer Dockerfile Constraint variability on Watermark Definition performances measures Files Dependency manager Krun npm Fix the JIT and GC events Make venv Java performance Reproducible Experimental state evaluation through variability reduction artifacts construction rigorous replay compilation. conda Maven Version manager Manually written procedure VMSH: Hypervisor-DVC agnostic Guest SVN Git Reproducible Reproduction Overlays for VMs artifacts access procedure **Artifacts repository** Reproducible Reproducible experiment environment Software Performances Figshare Zenodo Heritage Reproducibility Hyperfine Maven DockerHub Central JMH Reproducible Environment Measurement conclusions Infrastructure deployment deployment framework Energy Ansible Terraform PowerAPI System deployment E-Surgeon VM Addressing Rigorous statistical measurement bias analysis process **Analysis workflow** Apptainer Docker Snakemake Bayesian Data Analysis 🗅 Producing wrong data Benchmark design without doing anything in Empirical Software obviously wrong! Engineering Research Benchmark precision Rigorous benchmarking in and random initial reasonable time state