Dynamic curation of artifacts and experiments is changing the way digital libraries will operate

Bruce Childers¹, Jack Davidson², Wayne Graves³, Bernie Rous³, David Wilkinson¹

¹University of Pittsburgh
²University of Virginia
³Association for Computing Machinery

Contact: childers@cs.pitt.edu
Emerging Trends

- Mandates, accountability, leverage
- ACM/IEEE communities: Artifact Evaluation (AE), Replicated Computational Results (RCR)
  - Evaluate artifacts (SW, data, scripts, etc.)
  - Incentivize: better practices & more access
  - ACM TOMS, RTSS, ICSE, PLDI, PPoPP, CGO, OOPSLA…
  - “Badging” adopted by ACM (this year)
- Hands-on, Third-party examination
  - Installing, running, modifying, rerunning

RCR: http://toms.acm.org/replicated-computational-results.cfm
AE: http://www.artifact-eval.org
Emerging Trends

- Support for evaluation, e.g., VMs, workflows, etc.
- "Active curation platforms"
  - Direct access to artifacts/experiments
  - End-to-end: Innovate, submission, publish, repeat, archive
  - Different for different communities
    - From algorithms, to scripts, to complex workflows, to real systems!
    - Seems unlikely that single solution will serve all needs

- DLs need to be prepared for this trend
  - Increasingly diverse objects, ways reviewed (e.g., AE), delivered and manipulated
Pilot Study with ACM DL

- Understand technical capabilities & integration
  1. Identify & catalog capabilities/audiences
  2. Find & develop use cases to test/evaluate
  3. Apply exemplar platforms to use cases
  4. Prototype integration/interfaces w/ACM DL
  5. Pilot studies with use cases, platforms, & DL
  6. Inform guidelines/practices for authors, developers & publishers on how to integrate the platforms

- Pilot studies, insight on interfaces & integration, and insight on practices/standards
Pilot 1: Algorithm Comp. (Scenario)

- **A vs. B comparison on different data sets**
  - **Scenario**
    - Author *publishes paper with a “wrapped” artifact in DL*
    - Reader wants to *repeat A vs. B experiments* from paper
    - Reader may want to *try different data sets*
    - Reader accesses the artifact, downloads it, sets it up & runs it
    - Simple “access and run locally” scenario, minimal DL requirements
  - **DL provides**
    - Artifact, wrapper for the artifact, links to full setup to re-run
    - Deployment description – resource requirements to run experiments (software and hardware), what the wrapped artifact does
A vs. B comparison on different data sets

- **Application:** *SC16 student cluster competition*
- **Platform:** *Collective Knowledge (CK) to wrap & run app.*
- **DL:** *Local execution, driven from command-line*
  - Access CK-wrapped application from DL, download to run locally
  - DL holds application, CK, & the CK wrapper
  - Wrapper pulls data sets for local execution and runs experiments

- **Community:** *SIGHPC*


http://dl.acm.org/citation.cfm?doid=2807591.2807619
Pilot 2: Share & Modify (Scenario)

- Change existing experiments w/new parameters
  - Author deploys artifact: *active curation platform in the cloud*
  - Reader accesses paper from the DL
  - Paper’s DL landing page delivers “active content” extracted from platform, which reader can manipulate
  - Reader *examines experiments, changes them, tries new ones*
    - Provenance, new results, crowd sourced contribution fed back to DL
  - More complex situation with independently hosted, online artifact that can be examined and modified simply without local deployment
Pilot 2: Share & Modify (Implementation)

- Change existing experiments w/ new parameters
  - Application: *Portuno access control* (large design space)
  - Platform: OCCAM
  - DL: *Interactive page, modify, & run from the page*
    - Cloud-hosted through active curation platform (OCCAM)
    - Author adds artifact to an active curation platform
    - Platform and DL are integrated to deliver content to DL/hand-off to active curation platform for deeper examination of artifact

- Community: SIGSAC


http://dl.acm.org/citation.cfm?doid=2613087.2613099
Pilot 3: Artifact Derivation (Scenario)

- Modify (source changes) to an artifact
  - Author deploys an artifact with paper through the DL
  - Reader plays with the artifact through platform
  - **Makes source changes & re-run on original author’s data sets**
  - **Reader deploys the modified artifact back to the DL**
  - Redeploy with changes, compare experiments before/after
  - Provenance of derivation, source changes, experimental runs
Pilot 3: Artifact Derivation (Scenario)

- Modify (source changes) to an artifact
  - DL provides
    - Access to artifact, similar to pilot 2
    - Ability to redeploy the modified artifact with changes, compare experiments before/after change
    - Provenance of derivation, source changes, experimental runs
  - Possibly most complex case, illustrating both integration of platform and source modification/redeployment
    - May simplify to show making changes on the active curation platform, and then extracting changed results in the DL?
Pilot 3: Artifact Derivation (Impl.)

- Modify (source changes) an existing artifact
  - Application: **DRAM address remapping**
    - Modify existing memory simulator to have address remapping
    - Simulator: SST framework using Prospero and DRAMsim2?
      - SST: Sandia’s simulation toolkit (widely used at DOE, well supported)
      - Artifact is wrapped to run in active curation platform
  - DL: similar to Pilot 2, but incorporating changed artifact
  - Platform: **commercial**
  - Community: **SIGMICRO**

A permutation-based page interleaving scheme to reduce row-buffer conflicts and exploit data locality, Zhao Zhang, Zhichun Zhu and Xiadong Zhang, ACM/IEEE Int’l. Symp. on Microarchitecture, 2000

http://dl.acm.org/citation.cfm?doid=360128.360134
Status

- Study is underway... Building out the pilots now

- Outcomes
  - Technical insight into how to approach integration
  - Demos to excite community (to contribute)
  - Feedback on policies & procedures
    - ACM Task Force on Software, Data and Reproducibility in Publications
    - ACM SIG GB Task Force Replication and Independent Verification

- Engaging the community: Your feedback???