

A Literate Experimentation Manifesto

Jeremy Singer

University of Glasgow, UK



I have a dream ...

*that one day, our children will
be able to repeat our
experiments*

Three example problems



Green Shoots

- Publishers
- Academics
 - starting to address the problem ...



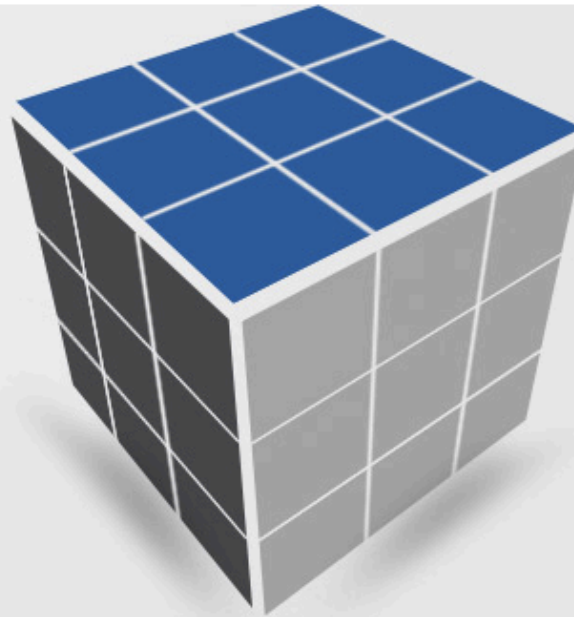
executable paper grand challenge

knowledge enhancement in the computational sciences

- Home
- News
- About the Challenge
- Meet the Winners
- Meet the Judges
- Abstract Requirements
- Finalists
- Rules

Meet the Winners

» Learn more



Winners

[Meet the Winners - Click here](#)

News Overview

winners of the Elsevier sponsored Semantic Web Challenge

» [Know more](#)

December 14, 2010 - Press Release: Elsevier Launches Executable Paper Grand Challenge

» [Know more](#)

Elsevier

versions of Adobe Acrobat and Distiller ONLY. The acm.job requirements to create an ACM compliant PDF with Adobe A

Thumbnail Image ACM is requesting a thumbnail image to identify your paper representation of your paper). This thumbnail image is "optical" dpi, 100 pixels wide in .jpg format. If you do submit a thumbnail caption (20-30 words) will be required.

Supplemental Material Supplementary material is "optional" and may include: A technical details; Source code for software; Source code for automated or anything else relevant. The SIGPLAN chair encourages authors to submit supplementary material to enable others to reproduce their work. [information.](#)

Please continue.

4. ACM Templates & Formats

Welcome to the Evaluate Collaboratory!



Welcome to the collaboratory on Experimental Evaluation of Software and Systems

This web site serves as a resource and a hub for everybody interested in understanding the current state of practice in experimental evaluation.

Currently, we provide the **Canon** (a **bibliography** on experimental evaluation), focusing on experimental evaluation.

<http://evaluate.inf.usi.ch/>

ker: Research Archive

This tracker is used to archive patches against specific versions of Jikes RVM that either reproduce a published search infrastructure that is not included in the main cvs.

Patches found in this archive may (or may not) be actively supported by their author, but none of them are. The role of the core team is simply to make this archiving service available to the research community. The goal is to produce reproducible research results and encouraging researchers to extend/improve on previous enhancements.

<http://jikesrvm.org/Research+Archive>

Search: [Advanced](#)

Page 1

	Summary	Status	Opened		
Signee:	<input type="text" value="Any"/>	Status: <input type="text" value="Any"/>	Category: <input type="text" value="Any"/>	Group: <input type="text" value="Any"/>	Submitter
<input type="text"/>	Artifact ID: <input type="text"/>	<input type="button" value="Filter"/>	<input type="button" value="Reset"/>	<input type="button" value="Permalink"/>	
0401	LeakChaser, a Java memory leak detection tool	Open	2011-03-21		
1791	enable replay compilation for Dacapo 9.12 bach	Open	2010-12-07		

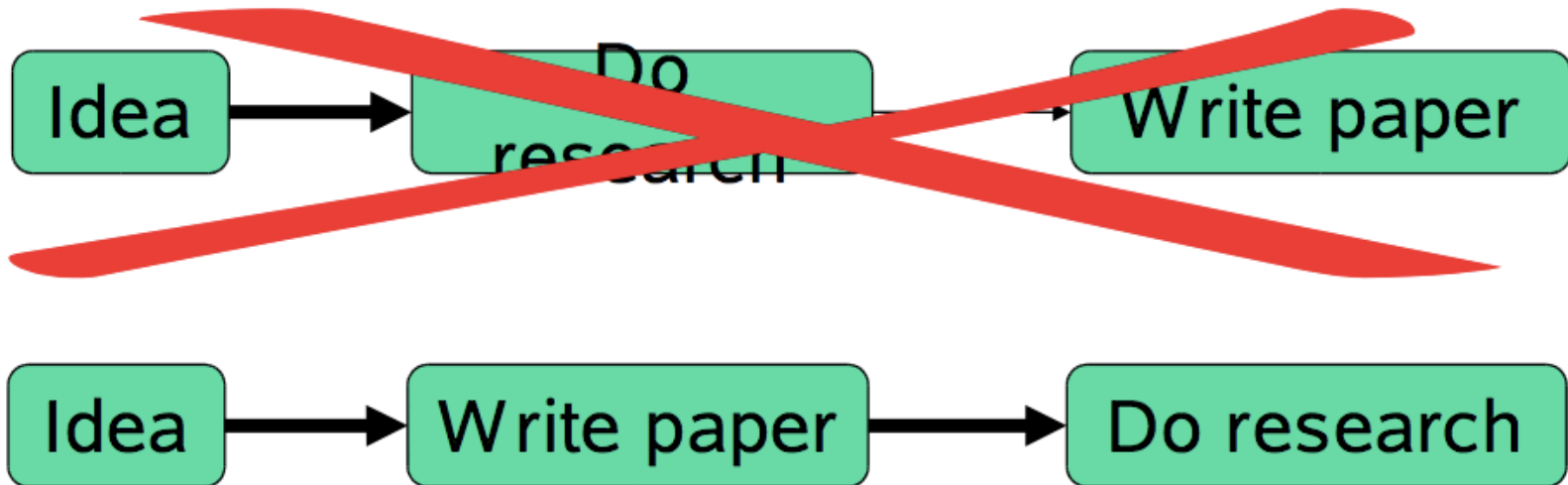
The default order



The SPJ order



Writing papers: model 2



Same problem occurs

- Experiment and paper are:
 - logically interdependent
 - not physically related
- Two representations can become inconsistent!

Literate Experimentation Slogans

- write up your experiments at *design time*!
- weave together the *description* with the *meta-description* of your experimental process



Literate Programming



- Invented by Donald Knuth
- WEB etc ...
- not popular – but some steps in right direction (Javadoc etc)

Knuth quote

Let us change our traditional attitude to the construction of programs: Instead of imagining that our main task is to instruct a computer what to do, let us concentrate rather on explaining to human beings what we want a computer to do.

change Knuth quote for LitExp

*Let us change our traditional attitude to the construction of **systems software experiments**: Instead of creating an experiment to demonstrate system-level improvements, let us concentrate rather on explaining to human beings how we design the experiment to work.*

Basic implementation

- weave together
 - **bash script** that manages experiment execution
 - **LaTeX** high-level description of experiment
- Relies on standard Unix tools
 - tested on Linux and Mac OS.

Simple example

- Trivial example presented here
- DaCapo benchmarks
 - two versions
 - some benchmarks retained across versions
 - are they the same?

Go through script /paper

Still needed

- <http://literateexperiments.org> portal
- Library of template experimental scripts
- Utility scriptlets
- emacs mode
- IDE integration
- Community support to archive experiments with papers
- Support from publishers / conference organizers

Conclusions

- Make life easier
 - for yourself
 - for others
- LaTeX + bash = literacy ??
- More complex example in paper

Literate Experimentation Slogans

- write up your experiments at *design time*!
- weave together the *description* with the *meta-description* of your experimental process

