Open Source as a Test Bed for Evaluating Creativity Support Tools

Michael Terry

University of Waterloo

Canada

Who I Am

- Assistant Professor in the…
- David R. Cheriton School of Computer Science at the…
- University of Waterloo, in…
- Canada
- Where it is currently…

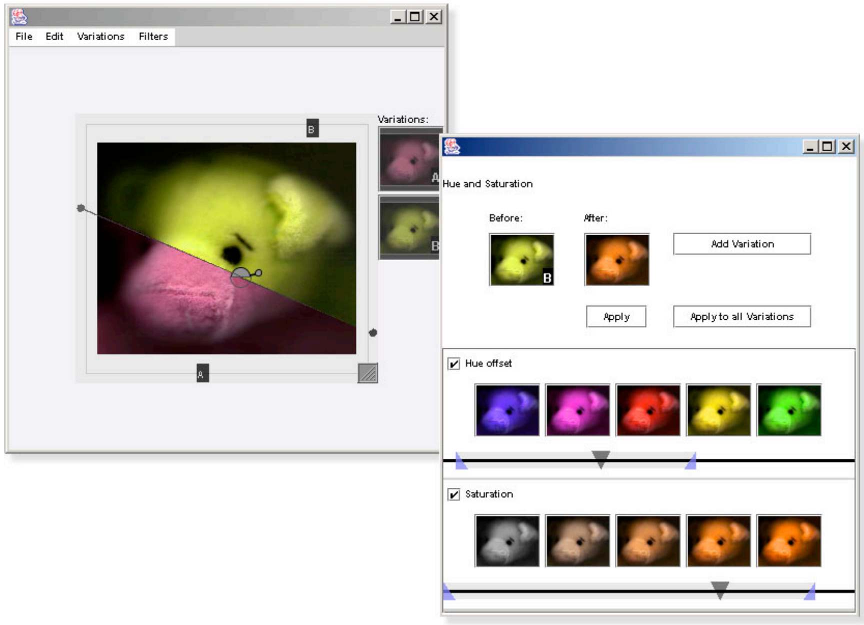
Very cold and snowy…

Research Past, Present, Future

- Creativity support tools
  - Side Views (UIST ’02)
  - Parallel Pies (CHI ’04)
  - PhD work at Georgia Tech, Beth Mynatt

- Open source usability
  - ingimp (CHI ’08)

- Open source as test bed for researching novel interaction tools



Talk Outline

- Approaches to supporting creativity
- Barriers to long-term evaluation
- Open source as substrate for widespread, longitudinal testing
- ingimp: Design, architecture, current results, how *you* can use it today
- Future possibilities for open source and creativity support tools

“By ‘augmenting human intellect’ we mean increasing the capability of a man to approach a complex problem situation, to gain comprehension to suit his particular needs, and to derive solutions to problems…

[We seek] more-rapid comprehension, better comprehension, the possibility of gaining a useful degree of comprehension in a situation that previously was too complex, speedier solutions, better solutions, and the possibility of finding solutions to problems that before seemed insoluble…”

Douglas Engelbart’s “Augmenting Human Intellect,”
1962 SRI Report

Achieving Engelbart’s Goal

- General categories of tools to support creativity (adapted from Shneiderman, Creativity & Cognition ’99)
  - *Creative activities* support tools
  - Domain-specific tools *with scaffolding*
  - Domain-specific tools with limited or no scaffolding
  - Domain-independent *process* support tools
  - Collaboration support tools (person-to-person)
  - Research tools

Approaches to Supporting Creativity

- Tool *itself* intended to lead to creative outcomes
  - Cohen's Aaron
- Support tools for *creative activities*
  - Brainstorming software
  - Concept mapping tools
- Domain-specific tools *with scaffolding for task*
  - MasterWriter (songwriting software)
  - Final Draft (screenwriting software)

Approaches to Supporting Creativity

- Domain-specific tools with limited or no creativity scaffolding
  - Make it easier to manipulate complex data
  - Word processors
  - Image manipulation applications

- Domain-independent *process* support tools
  - Undo
  - Version control systems
  - “What-if” tools
  - Simulations
  - Visualizations

Approaches to Supporting Creativity

- Collaboration support tools (person-to-person)
  - Change tracking
  - Communication tools

- Research tools
  - Search engines
  - Digital libraries
  - Dissemination tools

Measuring Success

- There are a lot of ways computation can support creativity…
- …but across all creativity tools, one basic requirement must be met for tool to be considered long-term success:

Used

Long-Term Use

- Creativity happens in the messy world
  - Long-term use studies (e.g., Computer Clubhouses) always illuminating
- Would like to understand tool use when:
  - Tool has been widely adopted
  - Experts have become experts with the tool
  - Reappropriation has occurred
- What are its long-term strengths and weaknesses across a variety of contexts?

Barriers to Long-Term Adoption

- Basic usability issues in tool itself…
- May not be useful…
- May not mesh with cultures of practice…
- May be inefficient…
- May be learnability issues…
- ...We want to understand what makes or breaks a tool in the long term, outside confines of the laboratory

Evaluation Methodologies

- Report from NSF Workshop on Creativity Support Tools, 2005:
  - Controlled studies
  - Field studies
  - Surveys
  - Deep ethnography

- Talk focus: How can we more easily determine *long-term*, *wide-spread* use?

Barriers to Long-Term Studies

- Costly to implement production-quality creativity support tools
  - Engineering, support issues

- Creativity support tools often need to exist in a larger *tool context*
  - Larger suite of tools, larger application
  - Example: History tools need to function in larger context. Same with previewing tools

- Costly to recruit subjects, collect data

Lowering the Costs to Building

- Three primary ways to lower building costs:
  - Scripting architectures (e.g., VBScript in MS Office)
  - Plug-ins (e.g., Eclipse, Photoshop, Audacity)
  - Open source software (full control over augmenting app)

- Methods save time by allowing one to *build* on the work of others
  - Don’t need to create larger tool context

Reasons to Go Open Source

- Open source becoming *very attractive* research platform:

- End user open source software continues to mature (Firefox, Eclipse, Audacity, OpenOffice…) giving options in a variety of task domains

- Lots of options for extending/instrumenting applications (plug-ins, scripting languages, source code itself)

- Access to enthusiastic, early-adopter user base

Success Stories

- Mylyn (formerly Mylar) (Mik Kersten and others at UBC)
  - Task-management plug-in for Eclipse
  - Conducted study with 99 developers
  - Now a successful, widely used plug-in
- Statistical bug isolation (Liblit *et al*)
  - Instrumented applications help discover bugs in the wild
  - Also benefited from enthusiastic user base
  - (See “The Open Source Proving Grounds,” Ben Liblit, 2005)

Advancing the OSS for Research

- Existing examples were one-off solutions
- How can we further lower barrier to using open source as a test bed for evaluating experimental creativity support tools?
- Enter open instrumentation and ingimp…

Open Instrumentation

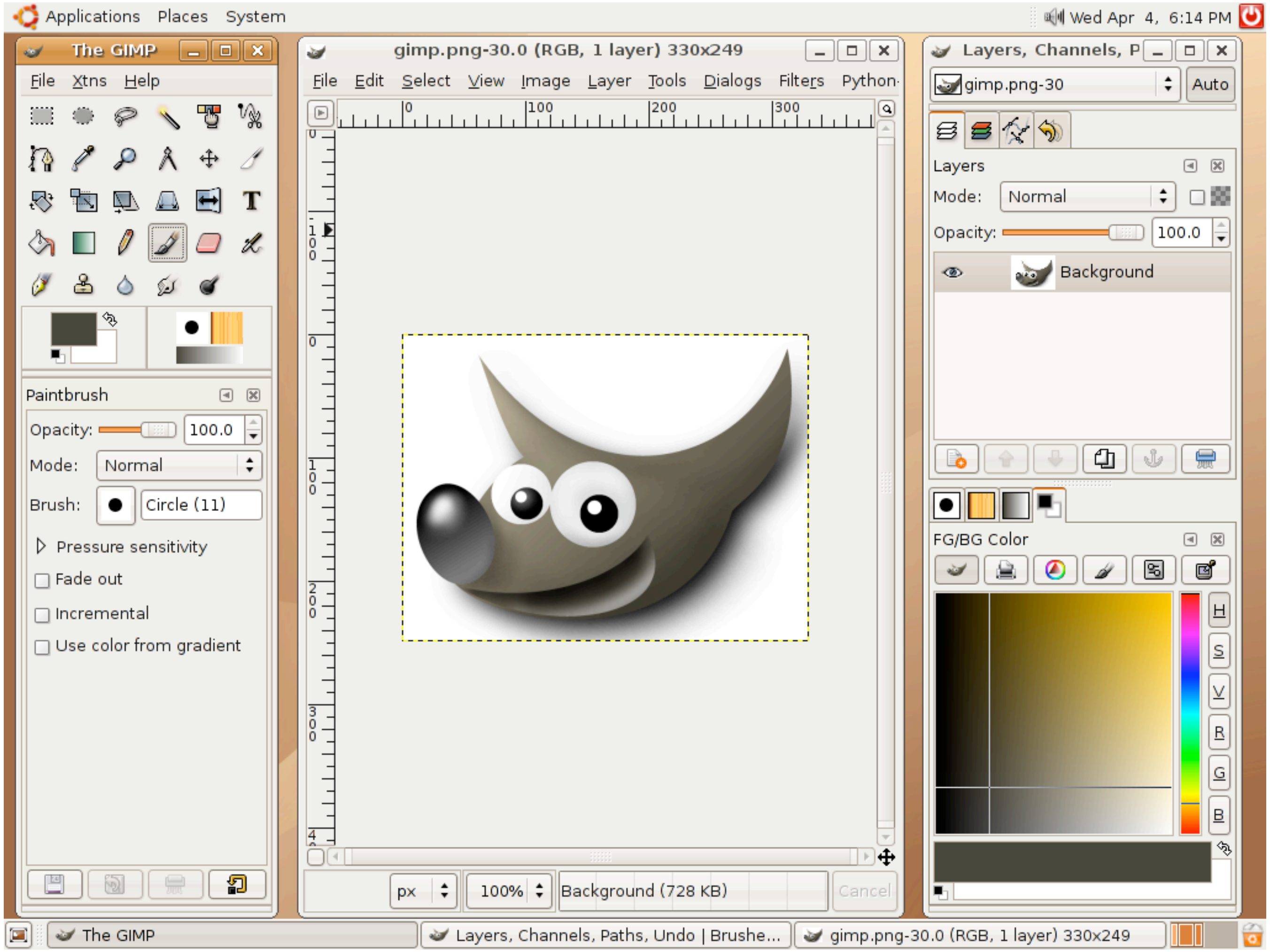
- Instrument an open source application to create a *data collection framework*

- Provide well-defined ways for third parties to make use of and extend the data collection framework

- Make data collection and dissemination public

- ingimp first example of this concept (Terry *et al*, CHI ’08)

GIMP



Paintbrush

Opacity: 100.0

Mode: Normal

Brush: Circle (11)

▷ Pressure sensitivity

□ Fade out

□ Incremental

□ Use color from gradient

Layers

Mode: Normal

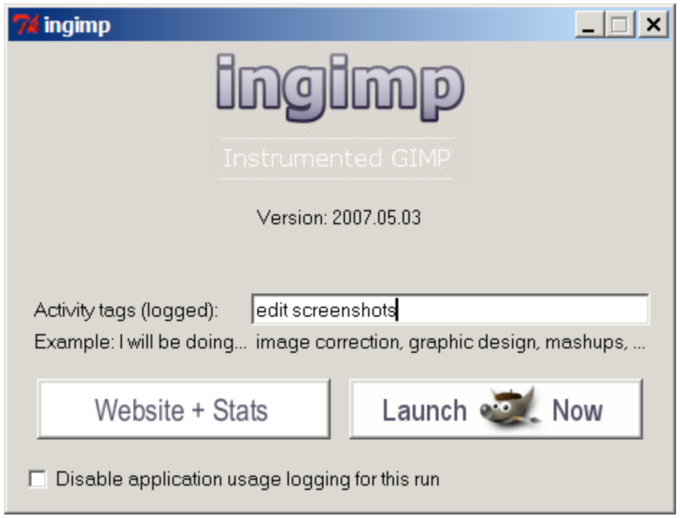
Opacity: 100.0

Background

px 100% Background (728 KB)

ingimp

- instrumented version of GIMP
  - http://www.ingimp.org
- Collects *usage* data
- All usage data publicly available on website
- Provides a *framework* to enable researchers to more easily collect / tag their own specific data



Version: 2007.05.03

Activity tags (logged): edit screenshots

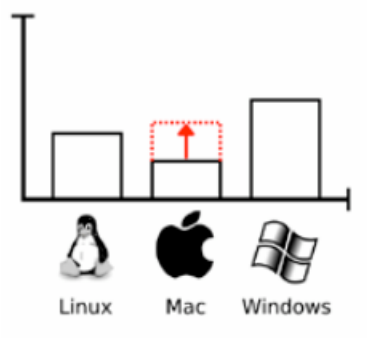
Example: I will be doing... image correction, graphic design, mashups, ...

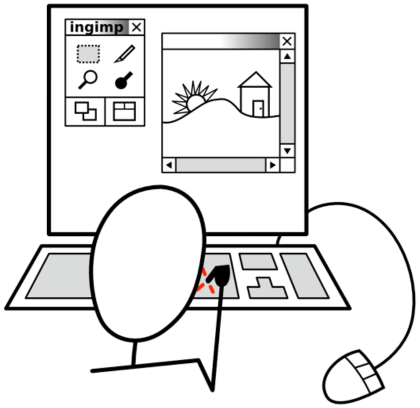
□ Disable application usage logging for this run

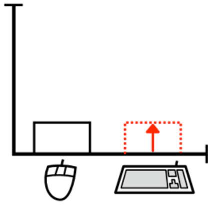
University of Waterloo

Data Collected

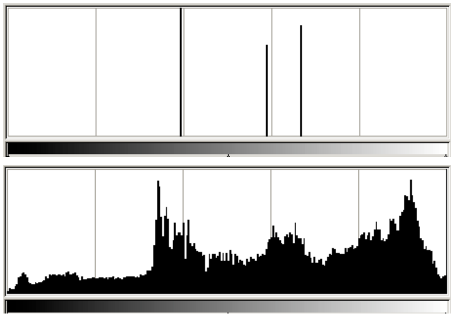
Five primary types of data collected:





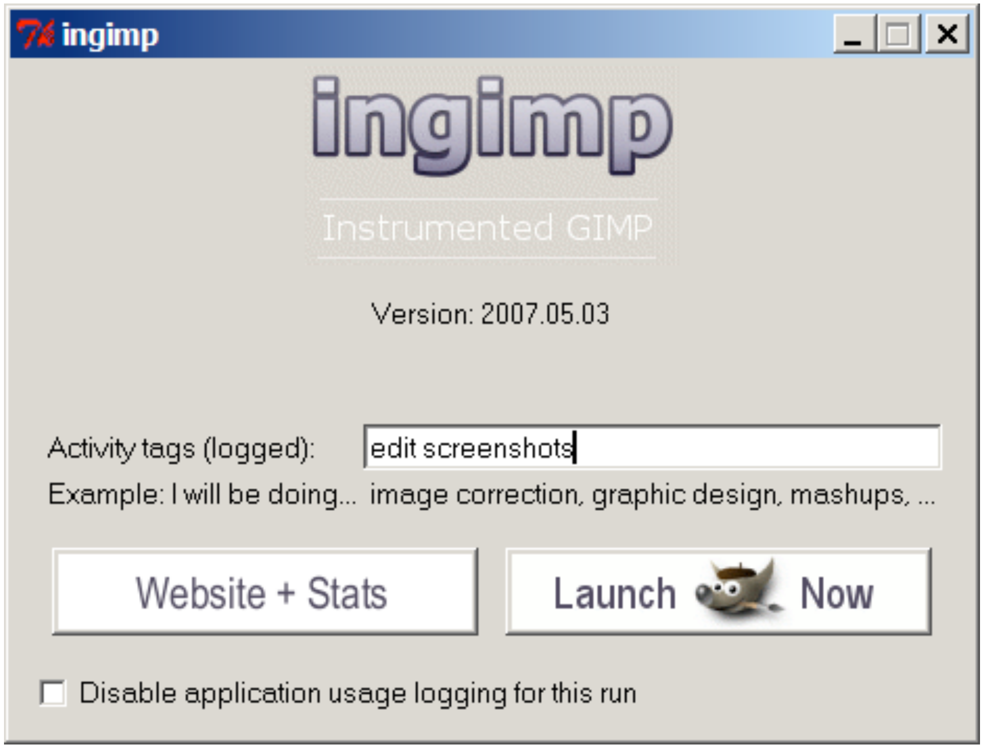


1. System characteristics
2. Command use
3. Interface events
4. Document characteristics
5. Activity tags



University of Waterloo

Activity Tags



Version: 2007.05.03

Activity tags (logged): edit screenshots

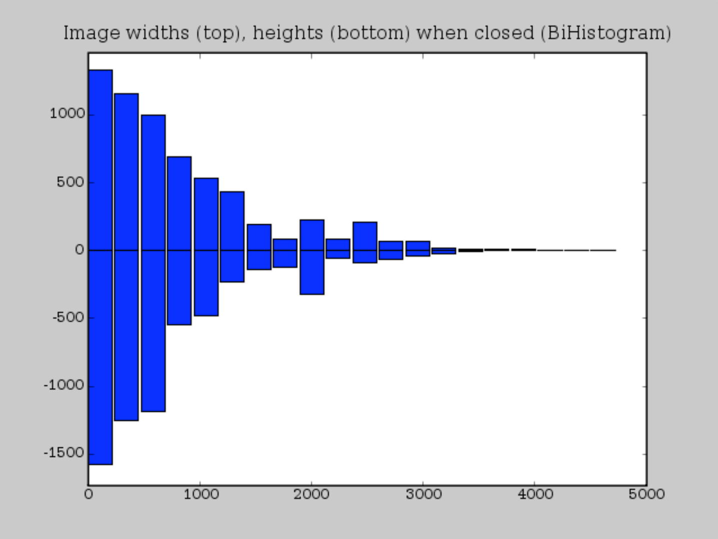
Example: I will be doing... image correction, graphic design, mashups, ...

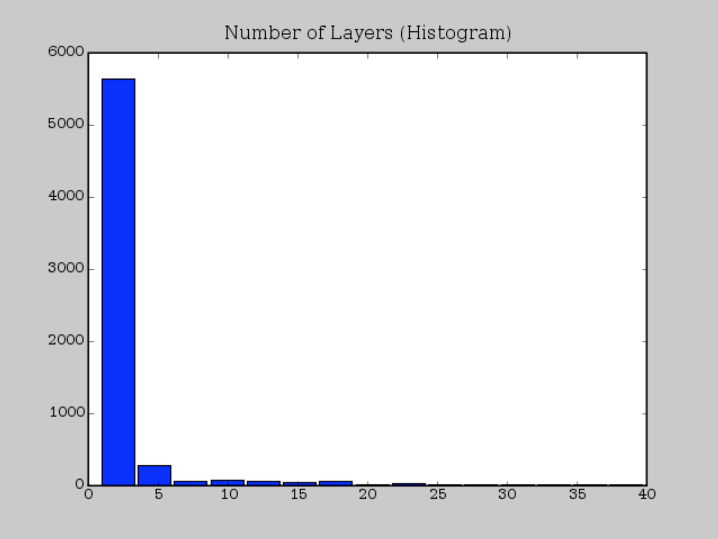
☐ Disable application usage logging for this run

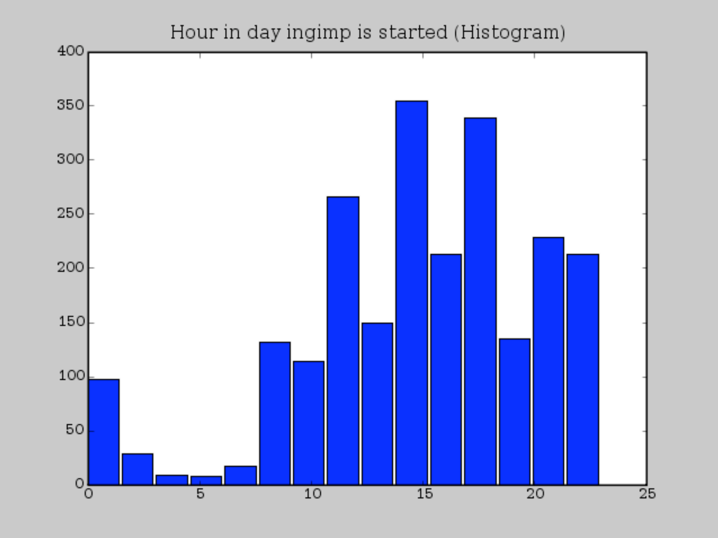
What We Can Learn From Data

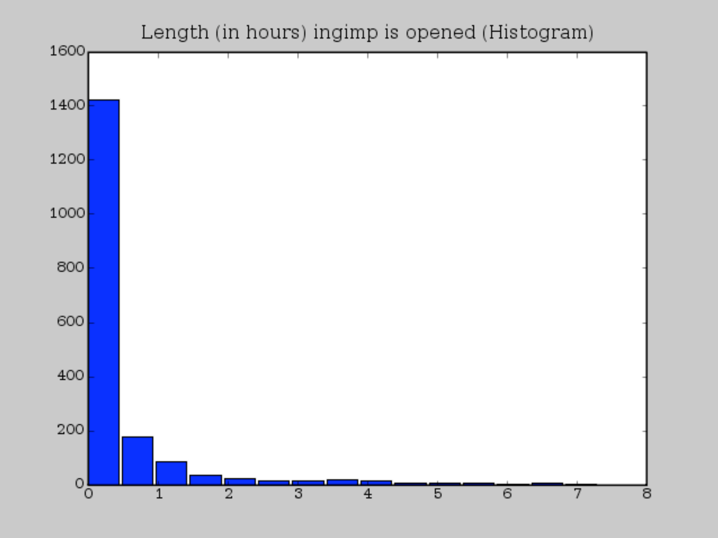
- Tool adoption / abandonment
- Frequency of use
- Some of the surrounding context of use
- How tools used (problem solving context)
- Information about *types* of users
- Users' descriptions of their tasks

ingimp Website: Stats + Raw Data









ingimp’s Design for Researchers

- Provides basic data collection infrastructure
- Tags: Free-form descriptions
- Way to add extra log data without code
- Source code for further instrumentation
- Mechanism to contact users
- Existing user base

Basic Infrastructure

- Data collection infrastructure built, tested, deployed
  - Get lots of logging data for “free”

- Most new commands should automatically be recorded by system
  - Little-to-no extra logging effort required on your part to collect basic data regarding new tool’s use

- Have spent significant time considering privacy issues in architectural design

Tags

- Tags can be used as basic means of understanding users' tasks, getting feedback
  - "Panoramic stitching, animation, website design, Garmin icon editing, opensim avatar texturing, cel shading, testing to see what it looks like.... AGAIN no real work this session so should be omitted from any staistical analysis just assigning jpg files to use ingimp, gel analysis, cropping, overlays, Toning down 'strobe party' advert banner, I will be doing photographic image enhancement and rotoscoping…"
- Reveal diversity of users, their willingness to participate

Extra Logging Capabilities

- Environmental variable can be set to include extra data in log file
- Intended to allow third parties to include information such as experimental conditions of a laboratory study
- No code needed to include the extra data

Source Code

- Source code availability allows additional data logging possibilities
  - Easier to extend existing logging framework than start from scratch

- *However*: ingimp server will reject log files not matching existing log format (for security reasons)
  - But can store log files locally

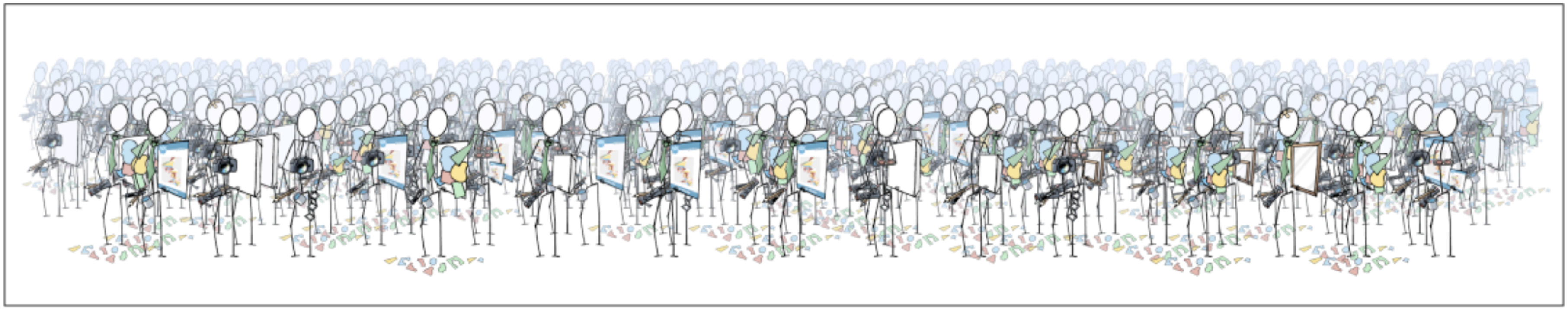
Mechanisms to Contact Users

- When closing app, server can send back a message and URL to open in web browser
- Can be used to direct users to surveys, interviews, updates…
- At present, must coordinate with us to make use of this feature

Additional Uses of Infrastructure

- *Distributed experimental evaluations* become possible
  - Create a set of alternative tools
  - Randomly activate just one tool for each user
  - Compare usage between tools
  - Similar to existing strategies used by web apps (e.g., Amazon)

- We’re doing this right now with illustrated consent agreements
  - User randomly assigned to one of 3 conditions
  - Timings taken of use of consent agreements



ingimp

Instrumented GNU Image Manipulation Program

Userbase as of July 2007

- Over 700 installations in first 6 months
- Users are from around the world
  - Most users not native English speakers
- Over 4,000 log files collected
- Already using it to field test new research ideas
  - Illustrated consent forms

Future Work

- Open data analysis (coming in next 4 months?)
  - MediaWiki modified to allow statistical queries against data set via *wiki text*
  - Opens data analysis to entire community
- Java, .NET-based infrastructure for easier access to extending GIMP platform (coming ???)

Conclusion

- OSS lowers barriers, costs associated with long-term, widespread testing of creativity support tools in real-world contexts
- Open instrumentation opens app to research possibilities
- OSS community an *eager* party to the testing of new research ideas
  - Early adopters, easily reached
- ingimp an open, extensible logging infrastructure that can serve as substrate for widespread evaluations
  - Example of open instrumentation for other projects

- ingimp contributors: Matthew Kay, Brad Van Vugt, Brandon Slack, Francois Marier, Jaime Ruiz, Tom Ayre, Terry Park, Tom Hazelton
- Work partially supported by a grant from NSERC

http://www.ingimp.org

mterry@cs.uwaterloo.ca