

Open Source Usability

Professor Michael Terry

October 29, 2007

Talk Overview

- Who I am / research interests
- *General* overview of Human Computer Interaction (HCI)
- HCI at Waterloo
- Open source usability

Talk Protocol

- You should be actively asking questions...

Who I Am

- One of your professors specializing in Human Computer Interaction (HCI)
- Teach CS 349 (User Interfaces) and CS 489 (Human Computer Interaction)
 - CS 349 teaches mechanics of *constructing* interactive applications
 - CS 489 teaches process of discovering *what* to build, design process, and evaluation of designs

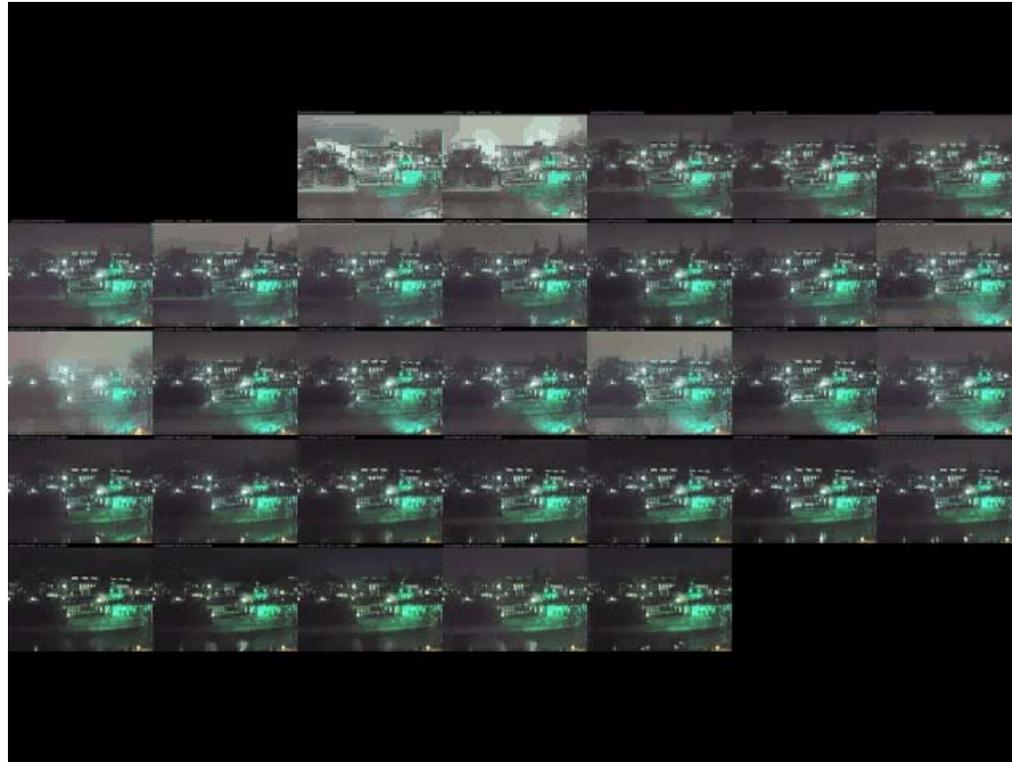
Research Interests

- Open source usability
 - This talk!
- Computational support for creative practices
- Novel time lapse visualization techniques

Time Lapse: Scales of Time

- Normally perceive world sequentially
 - Second-by-second, minute-by-minute
- Continuous barrage of stimuli to filter
- Easy to miss long-term trends and events
 - Example: Patterns of clapping
- Visualizations can help us discover these trends and events

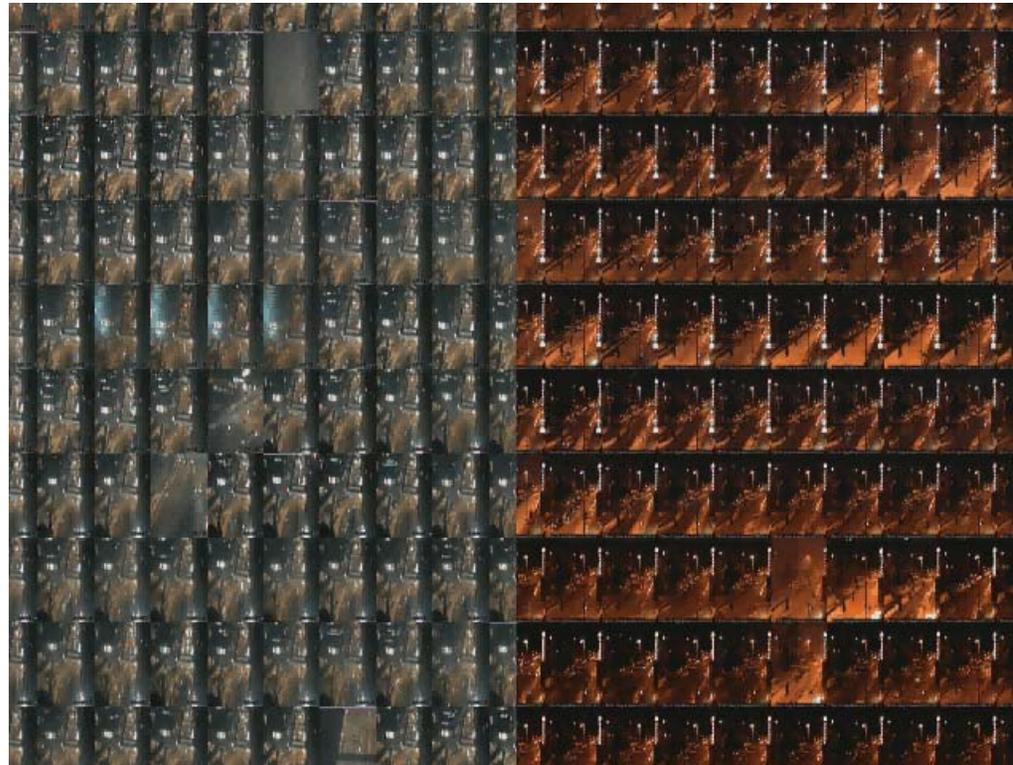
Time Maps



Time Maps



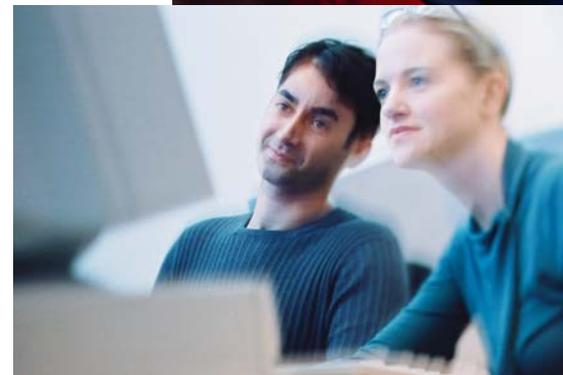
Time Maps



Switching Gears...

- Overview of HCI
 - Types of problems investigated
 - Goals of HCI

What is HCI?



HCI's Concerns

- *What* to build?
- For *whom*?
- In what *form*?
- What should be the *language of interaction*?
- *How well* does it meet its goals?
- What are the local and global *impact* of the choices made?

Example: Remote Communication



HCI is Interdisciplinary

- Computer scientists
- Psychologists
- Anthropologists
- Industrial designers
- Graphic designers
- Electrical engineers
- Artists

Goals of HCI

- Make computational artifacts that are “useful, usable, used” (Dix, Finlay, Abowd, Beale 2004)

Useful: Tool accomplishes what is required

Usable: Performs task “easily and naturally”

Used: People *want* to use it

More Goals of HCI

- Safe and forgiving
 - Mission critical systems, such as air traffic control, airplane cockpits, power plants, weapons systems...
- Universally accessible
 - Computers can compensate for physical and cognitive impairments
- Quick to learn to use
 - And to remember how it should be used...
- Efficient

Avoid Imprecision!

- People often cite “easy to use” and “usable” as laudable goals of interface design
- But how can we measure “easy to use” or “usable”?
- Terms mean many things to many different people
 - Not directly *measurable*
- Saying a system is “easy to use” or “usable” does not tell us *why* users perceive it to be that way
 - This makes it difficult to replicate successes and avoid similar mistakes

Be Specific and Prioritize

- Efficiency, learnability, desirability, cognitive load, physical exertion are all worthwhile measures for considering a system's design...
- But first concern should always be *usefulness* of system

Why HCI?

- Why should we be concerned with HCI?
- Isn't it enough to practice the Darwinian evolution of user interfaces?
 - Survival of the fittest?
 - Recombination of the best features into new interactive systems?
(Insert your own Mac OS X / Vista joke here...)
- Computation has the potential to be one of the most significant complements to our human intellect...

“There is a growing mountain of research. But there is increased evidence that we are being bogged down today as specialization extends. The investigator is staggered by the findings and conclusions of thousands of other workers — conclusions which he cannot find time to grasp, much less to remember, as they appear.”

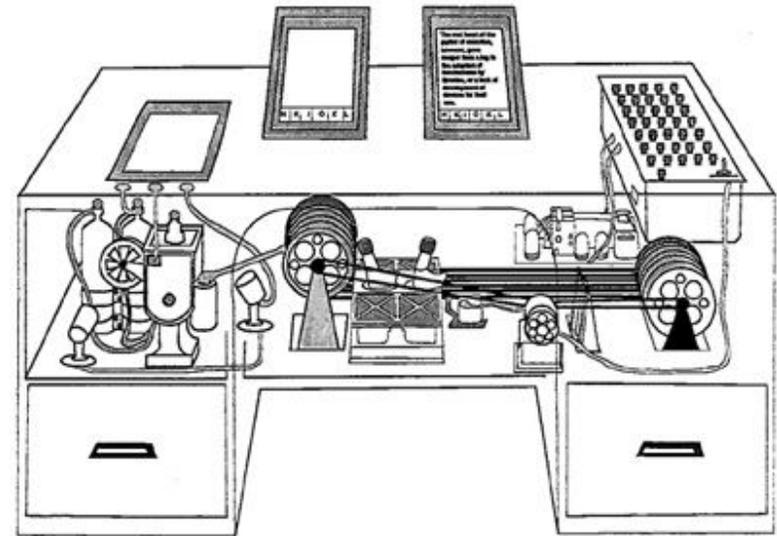
Vannevar Bush

- Headed “Office of Scientific Research and Development” in US
 - Manhattan project, other WWII science efforts
 - Oversaw 6000 scientists
- 1945 article in *The Atlantic* inspires computer scientists to present day
- Goal was to augment human intellect



Vannevar Bush

- “A memex is a device in which an individual stores all his books, records, and communications... It is an enlarged intimate supplement to his memory.”
- Proposes associative links between content
- Dual display setup!
- Direct annotation of stored content
- Proposes direct connection to nervous system



Today...

- Computers can fit on:
 - Desktop



Today...

- Computers can fit on:
 - Desktop
 - In our hands



Today...

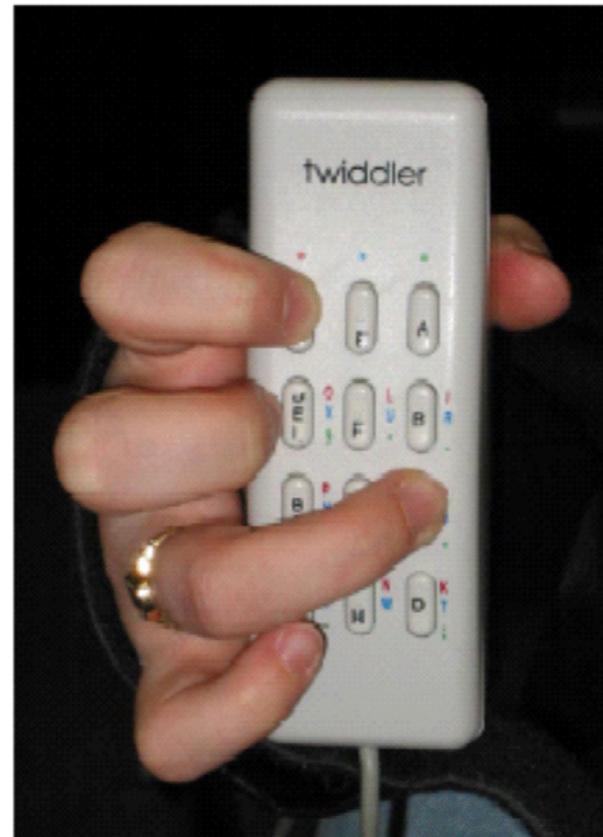
- Computers can fit on:
 - Desktop
 - In our hands
 - On our wrists



Today...



Thad Starner



On our glasses...

Today...



Pioneer's
Embedded
Display



MIT's Stitched
Number Pad

In our clothes

Today...

- Hard disks can store lifetimes of memories
 - See “MyLifeBits” by Gordon Bell at MS
 - 1 terabyte to hold our 20th century lives?
- Search engines put *volumes* of information at our fingertips
 - When is the last time you used a card catalog?
- Computation is *cheap*
 - \$100’s for basic computational machinery, ~\$1/GB for storage
- CPUs can process billions and billions of instructions every second
- Computation can be packaged in forms both large and small
- But...

It's not enough to have faster,
more capable machines...



Source: http://www.ctv.ca/servlet/ArticleNews/story/CTVNews/20061115/wii_PS3_061115/20061115?hub=Specials&pr=0

How can we *fruitfully*
and *reliably* augment
humans with all of
this computational power?

The Goals of HCI

- HCI seeks to understand how computation can become a better partner to more efficiently accomplish our tasks, better entertain us, help us communicate, keep us healthier...
- ...better achieve the things that matter to us.

HCI at Waterloo

- Ed Lank, Bill Cowan, and I are primary HCI researchers in CS
 - 1 Post Doc, 1 PhD, 5 MS students
- <http://hci.uwaterloo.ca>
- Research group investigates problems in
 - Motion kinematics
 - Human perception
 - Practices surrounding mobile gaming with Nintendo DS
 - Open source usability...

Gaming Practices of DS

- DS is portable
- Has ad-hoc wireless
- What unique gaming practices arise with this platform?
- How are practices different from other gaming platforms?



Gaming Practices of DS

- Work done by Christine Szentgyorgyi, Ed Lank and myself
- In-depth interviews, observations of 9 users, with 2 observations of larger gaming events
- Three primary findings:
 - Renegade gaming
 - Blockers to ad-hoc, pick-up games
 - “Private gaming sphere”

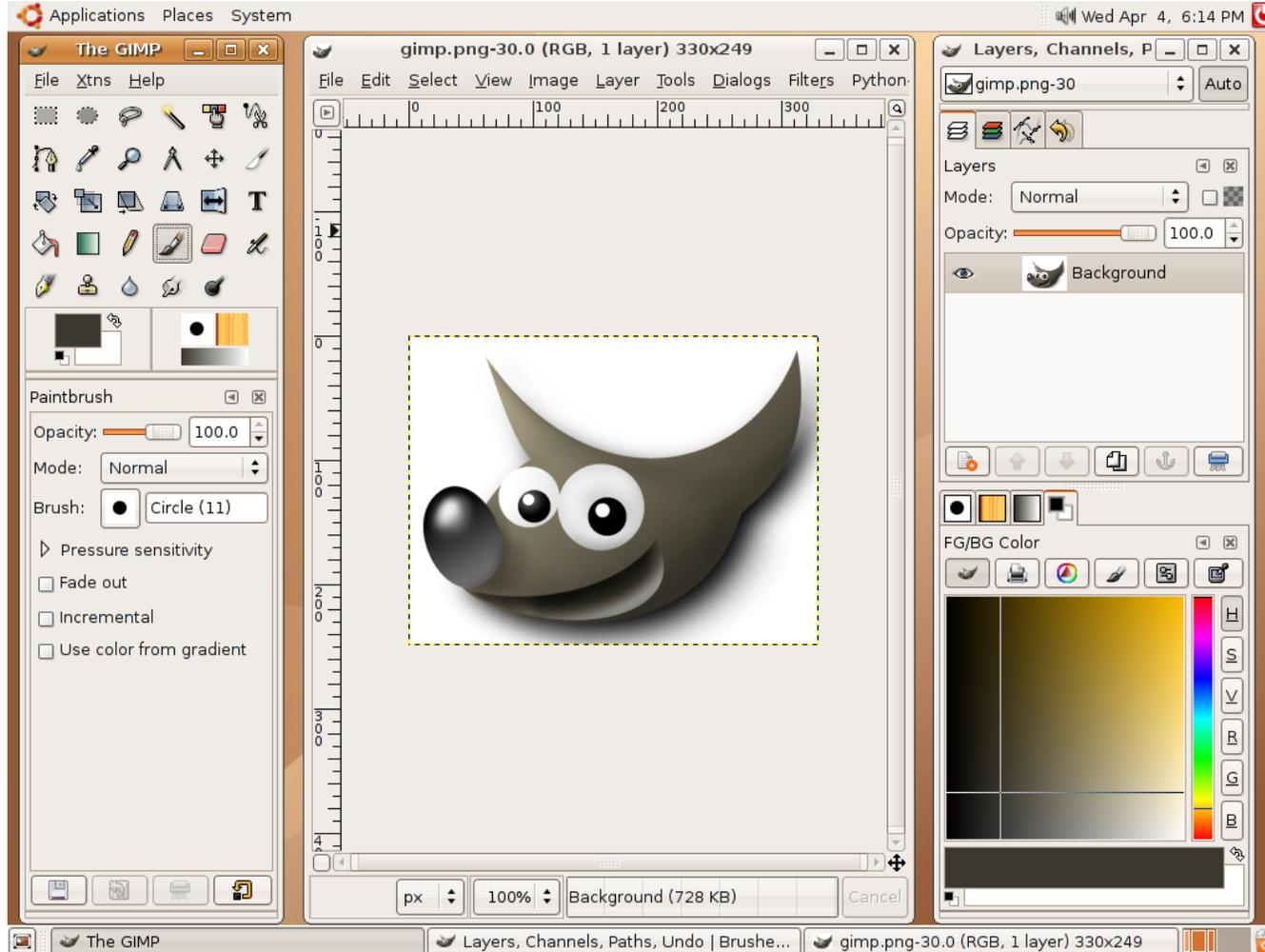




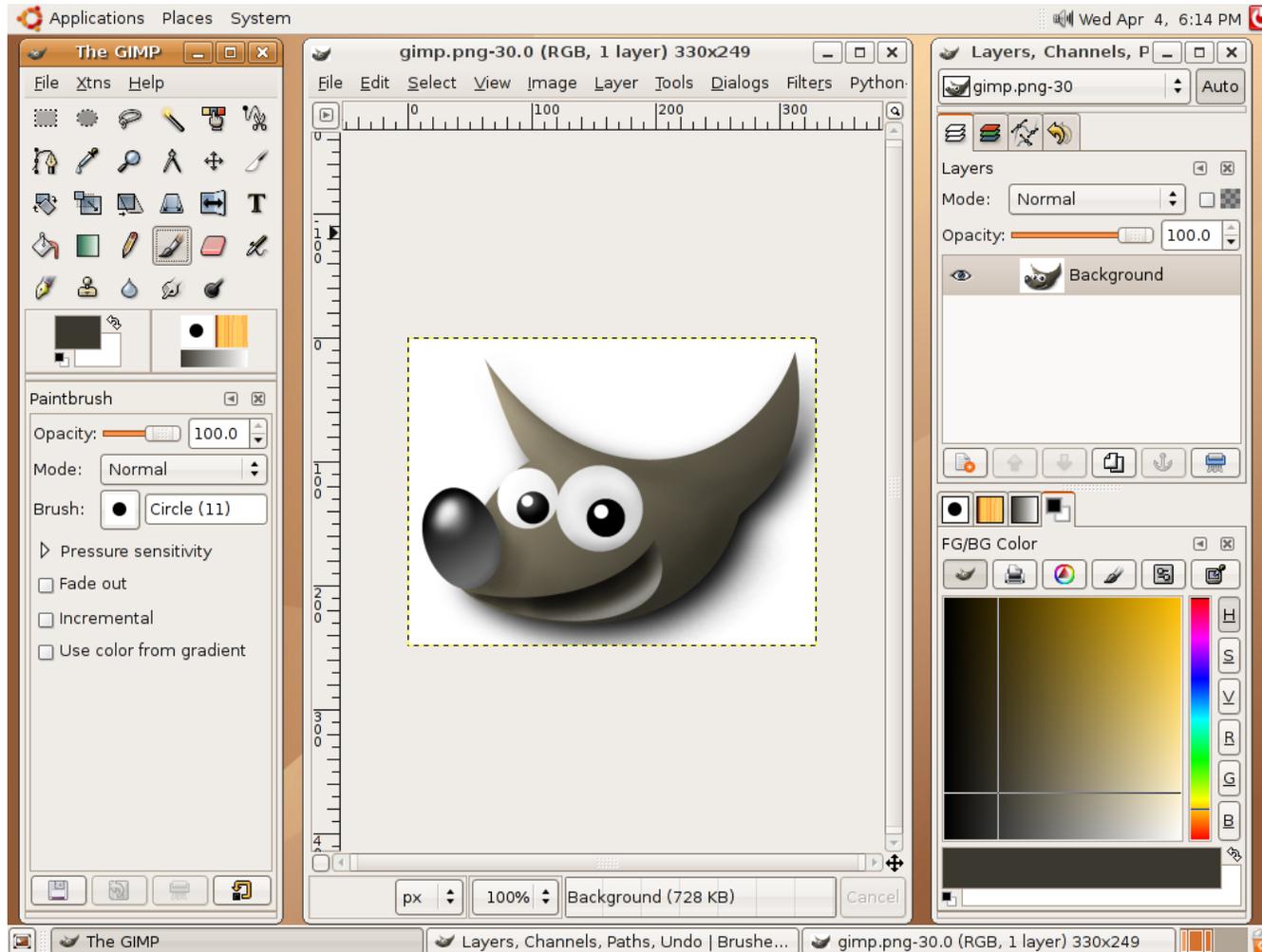
Shifting Gears...

- Open source usability!

GIMP

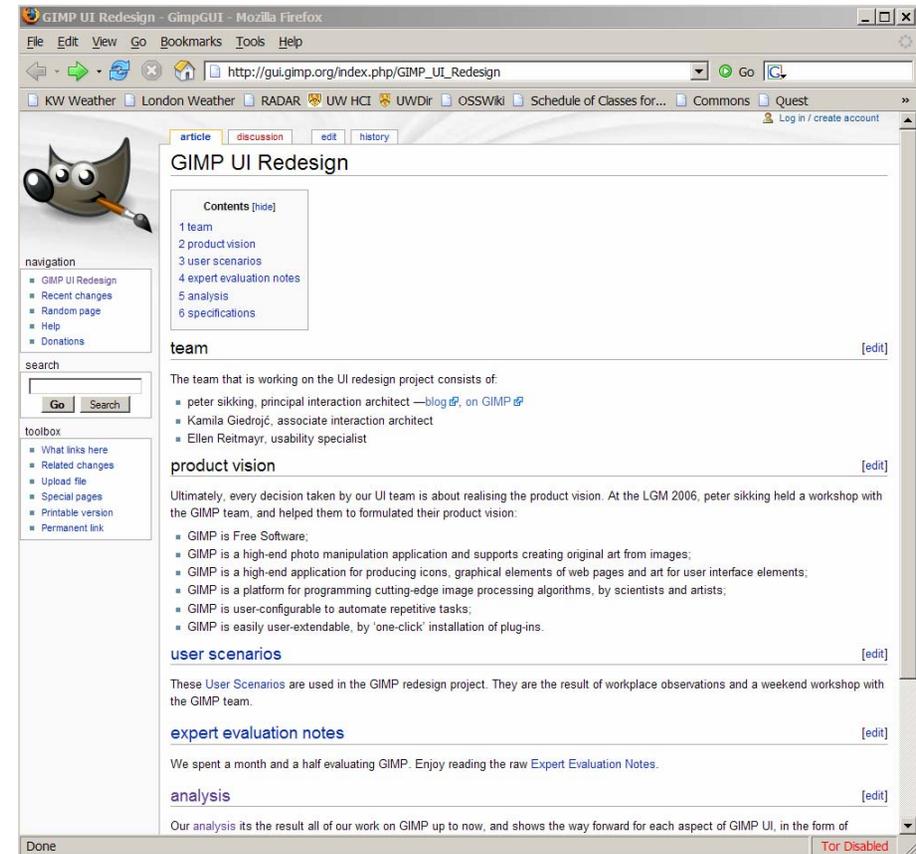


GIMP + Usability



Addressing Usability

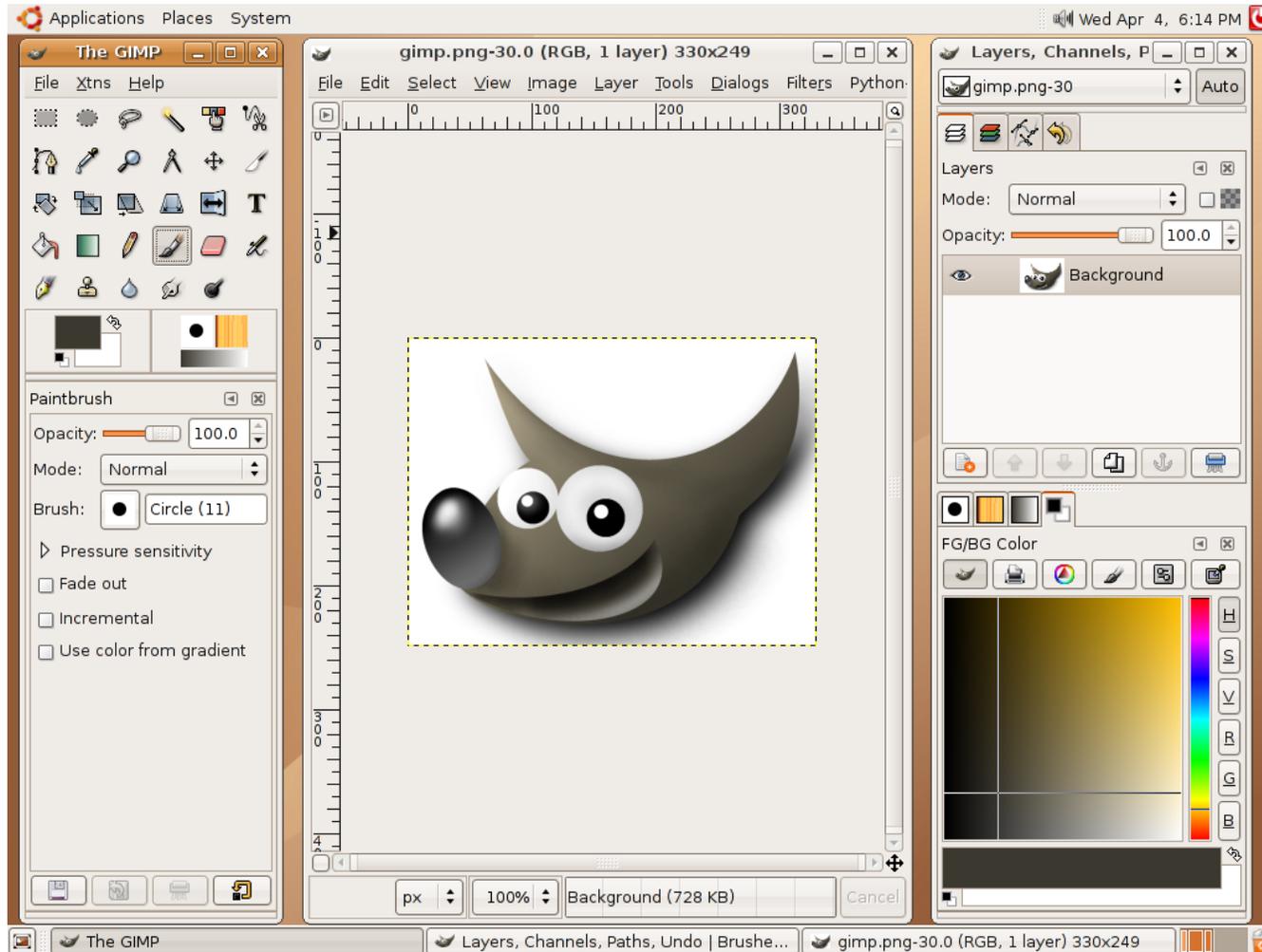
- GIMP usability efforts underway
 - Current focus is on interaction design
- But important questions remain regarding *broad, community-level use and composition of community*
- Who, What, Where, When, Why?



Usability Data vs. Usage Data

- *Usability data* describes areas in need of improvement
 - Points of confusion, inefficient workflows, missing features...
- *Usage data* quantifies the *community* of users
 - How the software is used and in what computing environment
- Usage data helps to:
 - Quantify impact (of design flaws, new changes)
 - Prioritize efforts
- Without usage data, speculation arises...

GIMP's Window Management

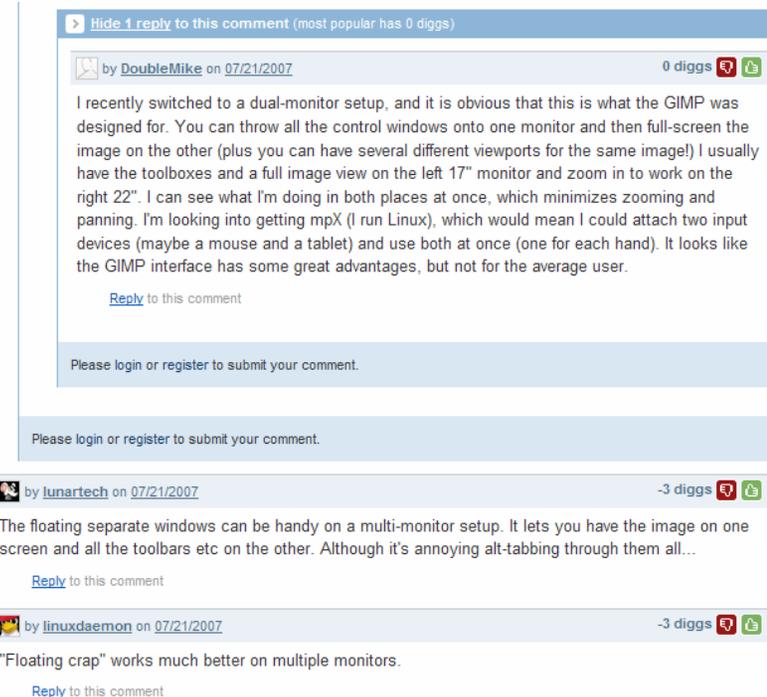


Example 1

“I recently switched to a dual-monitor setup, and it is obvious that this is what the GIMP was designed for. You can throw all the control windows onto one monitor and then full-screen the image on the other...”

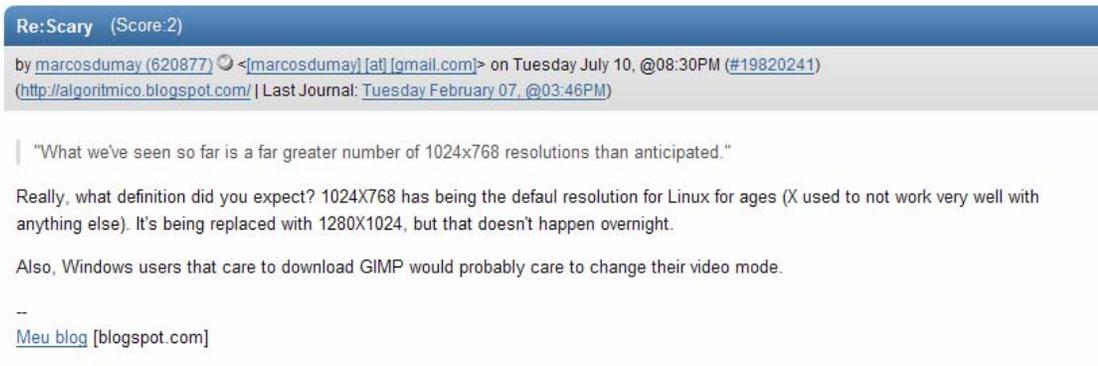
“The floating separate windows can be handy in a multi-monitor setup.”

“Floating crap’ works much better on multiple monitors”



The screenshot shows a Digg comment thread. At the top, there is a blue bar with the text "Hide 1 reply to this comment (most popular has 0 diggs)". Below this, the first comment is by "DoubleMike" on "07/21/2007", with "0 diggs". The comment text reads: "I recently switched to a dual-monitor setup, and it is obvious that this is what the GIMP was designed for. You can throw all the control windows onto one monitor and then full-screen the image on the other (plus you can have several different viewports for the same image!) I usually have the toolboxes and a full image view on the left 17" monitor and zoom in to work on the right 22". I can see what I'm doing in both places at once, which minimizes zooming and panning. I'm looking into getting mpX (I run Linux), which would mean I could attach two input devices (maybe a mouse and a tablet) and use both at once (one for each hand). It looks like the GIMP interface has some great advantages, but not for the average user." Below the comment is a "Reply to this comment" link and a light blue box with the text "Please login or register to submit your comment." The second comment is by "lunartech" on "07/21/2007", with "-3 diggs". The comment text reads: "The floating separate windows can be handy on a multi-monitor setup. It lets you have the image on one screen and all the toolbars etc on the other. Although it's annoying alt-tabbing through them all..." Below the comment is a "Reply to this comment" link and a light blue box with the text "Please login or register to submit your comment." The third comment is by "linuxdaemon" on "07/21/2007", with "-3 diggs". The comment text reads: "'Floating crap' works much better on multiple monitors." Below the comment is a "Reply to this comment" link and a light blue box with the text "Please login or register to submit your comment."

Example 2



“1024x768 has been the default resolution for Linux for ages... It’s being replaced with 1280x1024, but that doesn’t happen overnight. Also, Windows users that care to download the GIMP would probably care to change their video mode.”

Grounding the Discussion

- Open source projects lack infrastructure for collecting usage data
 - Many commercial software companies have this infrastructure in place
- Open source smolt, Debian's popularity contest collect *some* data about users' computing environments
 - But not data on software usage itself
- Enter *ingimp*...

What is ingimp?

- instrumented version of GIMP
- Tool to see into *community's* use of GIMP
- Snap-in replacement for GIMP
- First example of *open instrumentation*



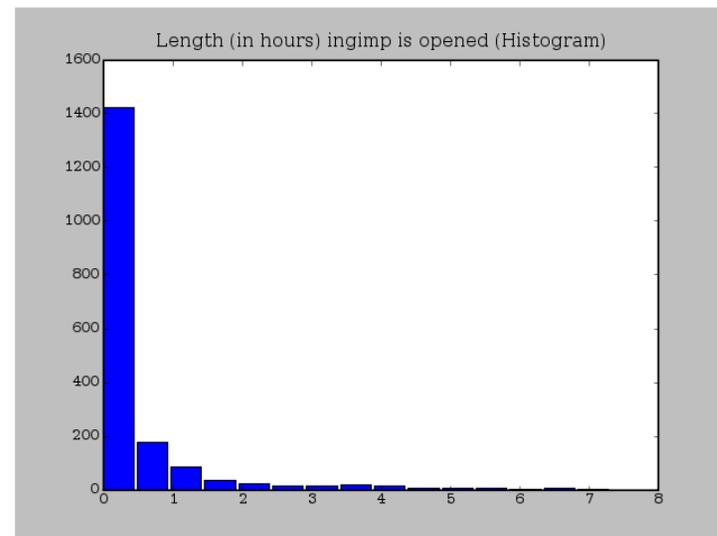
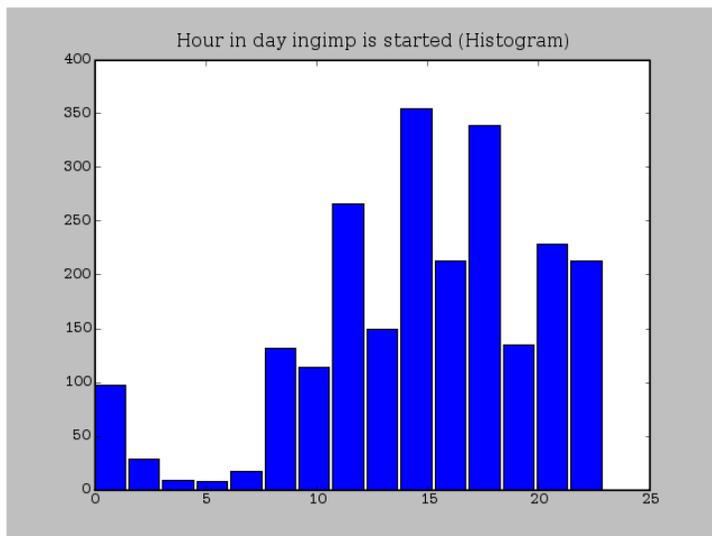
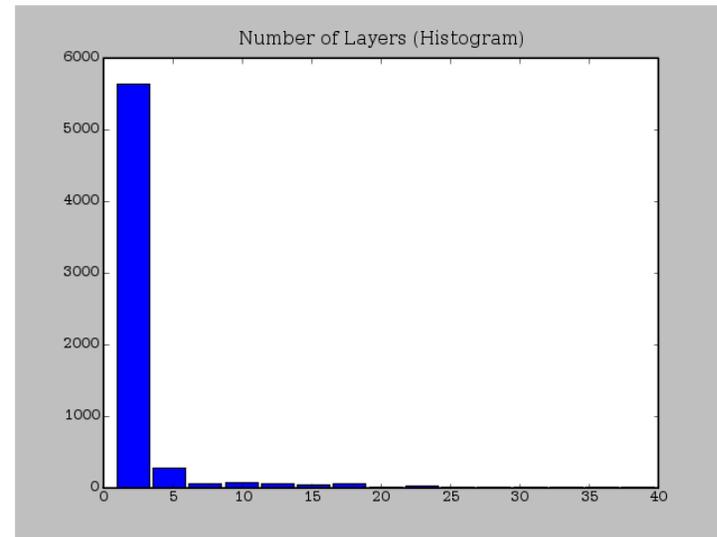
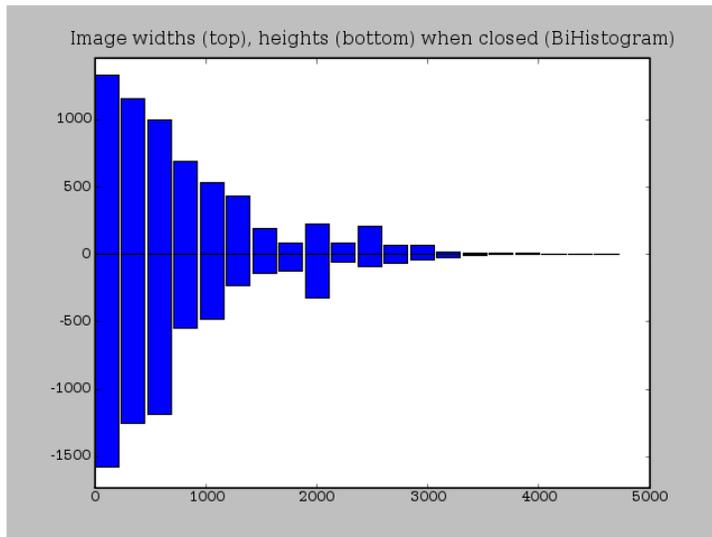
Open Instrumentation

- *All* collected data publicly available
 - *Anyone* can view/download/analyze the collected data
- Instrumentation itself is open
 - Anyone can inspect the source to see *exactly what data are collected and how*
 - *Highly transparent* process
- Openness creates new challenges and opportunities for instrumentation

ingimp

- Demo...
- (<http://www.ingimp.org>)

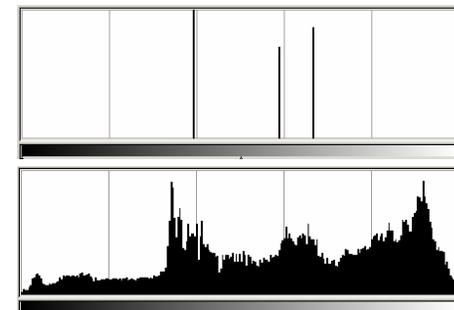
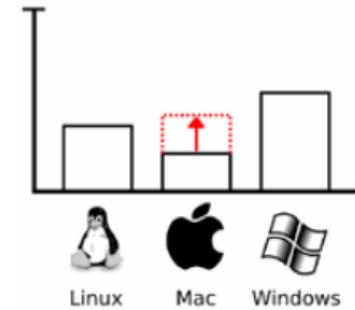
ingimp Website: Stats + Raw Data



Data Collected

Five primary types of data collected:

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



Activity Tags



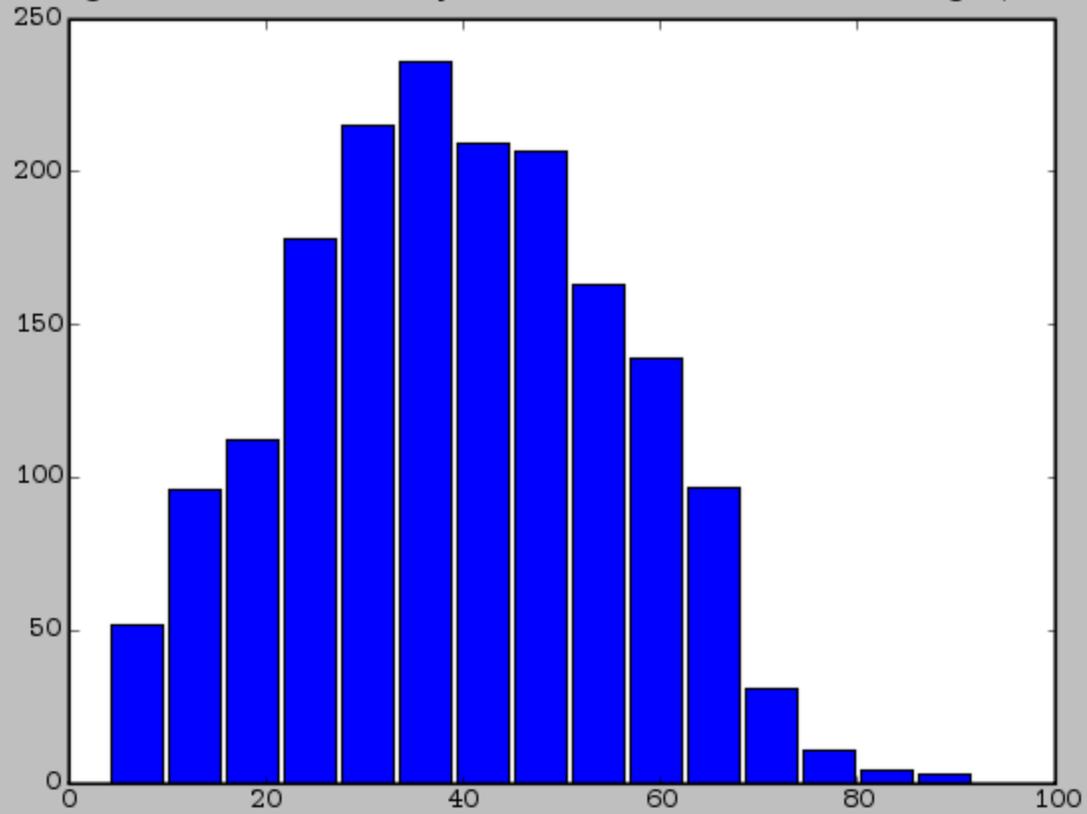
Value of Data

- Data reflects *in-situ* use of application
 - *Not* use in artificial environment
 - Reveals how people use GIMP on day-to-day basis
 - People quickly forget it's there
- Completely unique, *public* data store of *longitudinal use* data
- Data ground discussions about user base, design implications

Example: Monitor Use

- Percentage of multimonitor users? Most common resolutions?
- 96% of ingimp users (~550) have single monitor setup
- Most common resolutions
 - 1024x768 (33%)
 - 1280x1024 (33%)
 - ~90% are Windows users
- Impact of GIMP's window management scheme apparent...

Percentage of Monitor Used by Document Window, on Average (Histogram)



The Price of Instrumentation

- Instrumentation provides valuable usage data...
- ...but the “cost” is a partial compromise of an individual’s privacy...
- ...without any perceived, direct benefit to users
- Issues found in *any* application/service that monitors how people use software

Research Opportunities

- How can we motivate usage of instrumented software?
- Are there ways users can more directly (and quickly) benefit from instrumentation?
- How can we collect real-world usage data in a privacy-sensitive manner?

ingimp: Motivating Use

- *Personalized and personable* statistics



Personalized Statistics

- Many ingimp statistics *personalized* to enable comparisons
- Thinkalouds reveal subjects like this
 - Play to innate curiosities

Average Amount Command Used in Session

| Command Name | Average Number of Times Used When Used in Session | Standard Deviation | Command Name | Your Average Use When in a Session |
|--------------|---|--------------------|-----------------|------------------------------------|
| Add Anchor | 301.267716535 | 676.868776394 | Ink | 30.0 |
| Clone | 160.933333333 | 243.804719485 | Fuzzy Select | 28.0 |
| Dodge/Burn | 116.5 | 161.927452892 | Item visibility | 15.0 |
| Ink | 116.435483871 | 390.440676671 | Undo | 10.666666666 |
| | | | Move Floating | |

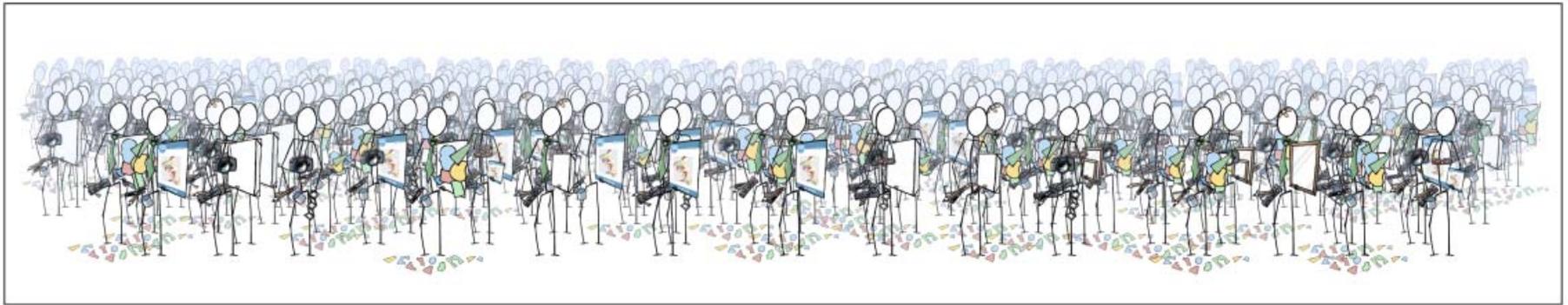
ingimp Personas

- Each user has personal online *persona*
- A personal information visualization
- Simplified design so other software projects can create them









ingroup

Interconnected 1000 Design Multiplication Program
March and July 2007

Personal Stats Findings

- Studies indicate people like the personas
- More importantly, people are using the statistics to *learn* about the software
 - Most frequently used commands to learn about software's functionality
- Tight feedback loop gives collected data *direct value* to *end-users*
- Suggests range of alternative uses of usage data
 - Help systems, tutorials, suggestive interfaces, recommender systems (e.g., Linton *et al*, 2000)

The Price of Instrumentation

- Motivating use / personal benefits
- Privacy issues

Privacy Considerations

- All collected data made publicly available
 - Not your usual usability lab setup...
- Set of rules need to be in place to honor privacy

Privacy Preserving Measures

- Disable logging at start-up
- ingimp designed to collect only application *usage* information
 - Motto: *No parameters*
- Key/mouse activity, but *no details of what key or where mouse used*
- Transparent data collection via human-readable XML log file format



```
<LogEntry entry_num="581" event_date="2007_05_03_11_52_55"
  elapsed_time_usec="1140515625.000000">
  <WindowKeyEvent window_name="GIMP Text Editor" window_role="gimp-text-editor"
    event_type="9" modifier_state="0"/>
</LogEntry>

<LogEntry entry_num="583" event_date="2007_05_03_11_52_55"
  elapsed_time_usec="1140656250.000000">
  <WindowKeyEvent window_name="GIMP Text Editor" window_role="gimp-text-editor"
    event_type="9" modifier_state="0"/>
</LogEntry>

<LogEntry entry_num="584" event_date="2007_05_03_11_52_55"
  elapsed_time_usec="1140812500.000000">
  <WindowKeyEvent window_name="GIMP Text Editor" window_role="gimp-text-editor"
    event_type="9" modifier_state="0"/>
</LogEntry>

<LogEntry entry_num="585" event_date="2007_05_03_11_52_56"
  elapsed_time_usec="1141234375.000000">
  <WindowKeyEvent window_name="GIMP Text Editor" window_role="gimp-text-editor"
    event_type="9" modifier_state="0"/>
</LogEntry>

<LogEntry entry_num="586" event_date="2007_05_03_11_52_56"
  elapsed_time_usec="1141375000.000000">
  <WindowKeyEvent window_name="GIMP Text Editor" window_role="gimp-text-editor"
    event_type="9" modifier_state="0"/>
</LogEntry>

<LogEntry entry_num="587" event_date="2007_05_03_11_52_56"
  elapsed_time_usec="1141531250.000000">
  <WindowKeyEvent window_name="GIMP Text Editor" window_role="gimp-text-editor"
    event_type="9" modifier_state="0"/>
</LogEntry>
```

Source Code

- *Anyone can inspect the source code*
- Users report an increased sense of security knowing source code available for inspection
- Other software may not need to open *all* source code to gain these benefits
 - Could selectively display only the logging source code

Deeper Privacy Considerations

- Anyone can inspect the source
 - But how many do?
- Consent agreement outlines privacy issues
 - But how many read it?
- Open source community is global
 - Localization takes time in resource-strapped open source project

Issues Not Unique

Microsoft Windows

Error signature

BCCode: 100000d1 BCP1: 994950A4 BCP2: 00000002 BCP3: 00000008
BCP4: 994950A4 OSVer: 5_1_2600 SP: 2_0 Product: 256_1

Reporting details

This error report includes: information regarding the condition of Microsoft Windows when the problem occurred, the operating system version and computer hardware in use, and the Internet Protocol (IP) address of your computer.

We do not intentionally collect your name, address, email address or any other form of personally identifiable information. However, the error report may contain customer-specific information in the collected information. While this information could potentially be used to determine your identity, if present, it will not be used for that purpose.

The data that we collect will only be used to fix the problem. If more information is available, we will use it when you report the problem. This error report will be sent using a secure connection to a database and will not be used for marketing purposes.

To view technical information about the error report, [click here](#).
To see our data collection policy on the web, [click here](#).

firefox.exe

firefox.exe has encountered a problem and needs to close. We are sorry for the inconvenience.

If you were in the middle of something, the information you were working on might be lost.

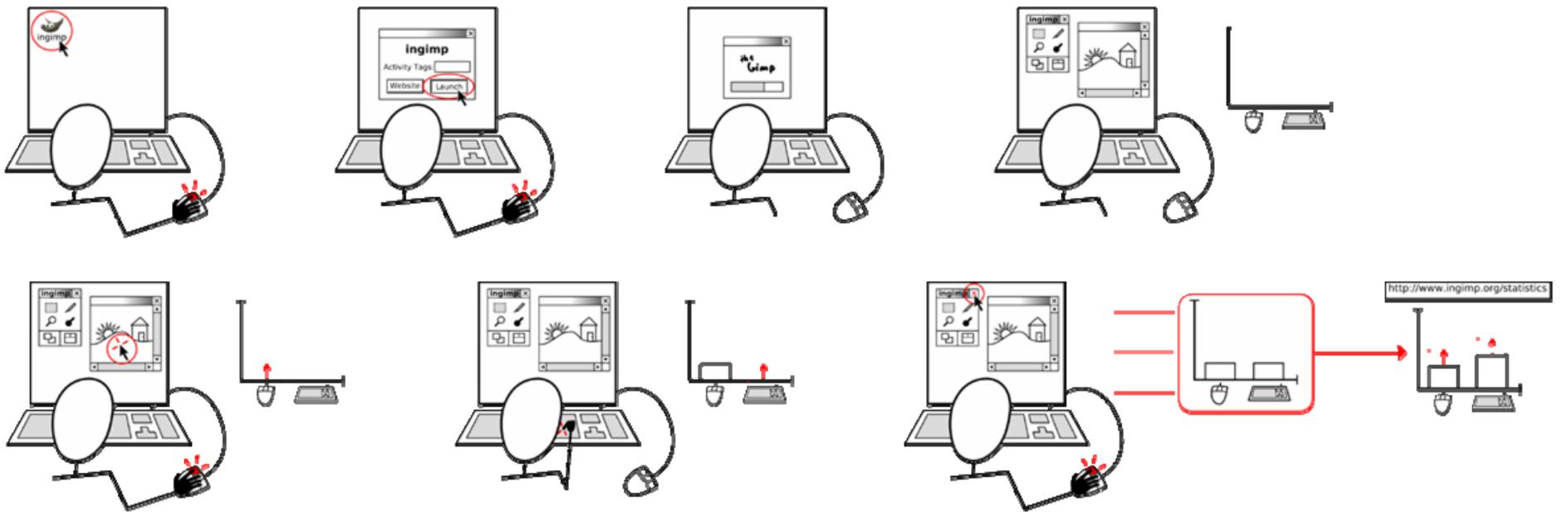
Please tell Microsoft about this problem.
We have created an error report that you can send to us. We will treat this report as confidential and anonymous.

For more information or report contains, [click here](#).

Crash Report:

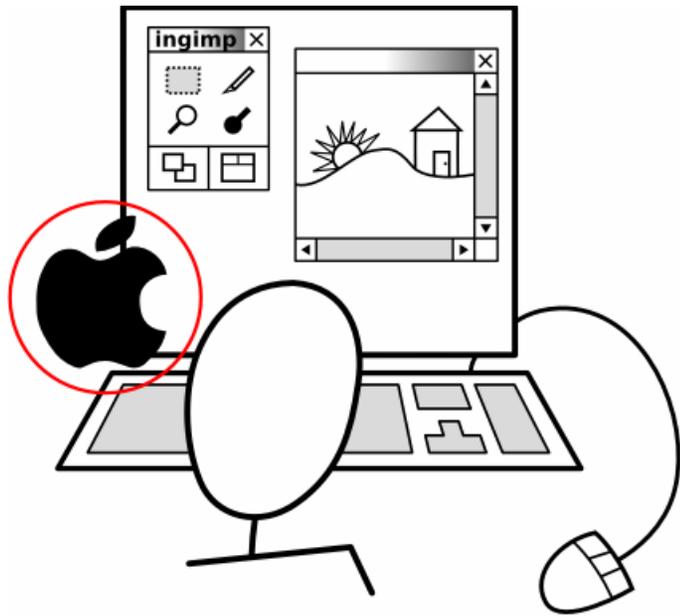
```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE AdobeCrashReport SYSTEM "AdobeCrashReporter.dtd">
<crashreport version="1" applicationName="InDesign" applicationVersion="4040"
build="4.0.4.658">
  <time year="2006" month="11" day="28" hour="17" minute="24" second="19" />
  <user host="iMac-main.local" />
  <system platform="macintosh" osversion="10.4.8" ram="2048" machine="Power Macintosh"
model="PowerMac8,1" cpuCount="1" cpuFreq="1800000000" busFreq="600000000" />
  <crash exception="Read Only Memory Exception" instruction="80430004">
    <backtrace crashedThread="0">
      <thread index="0">
```

? Mail using my default mail application

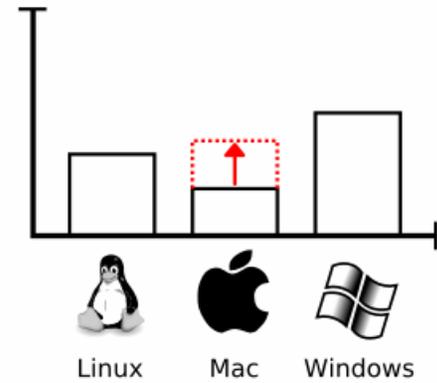


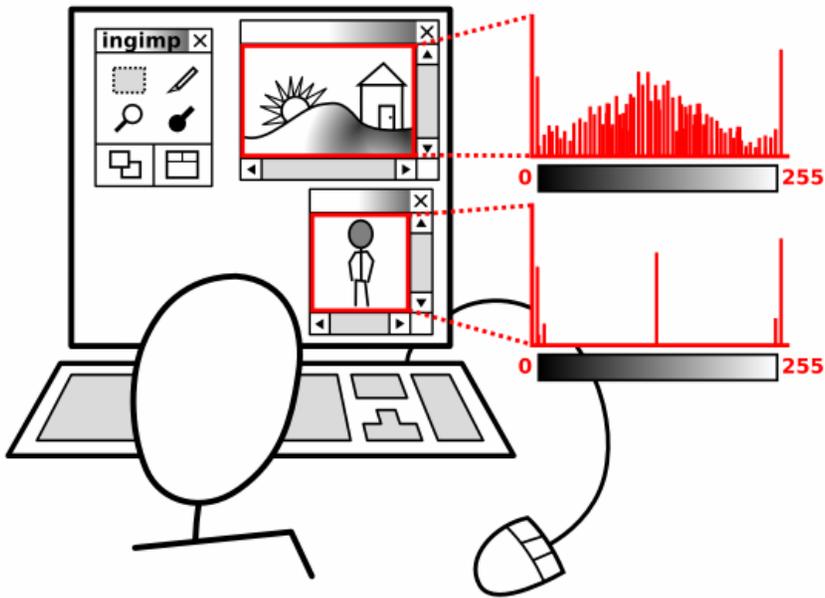
ingimp Pictograms

- Visual language indicating how software functions
 - *Not how to use the software*
- Iteratively designed with native, non-native speakers
- Overview “teaches” visual language
 - Individuals generally understand meaning after “reading” overview
- More specific logging information follows



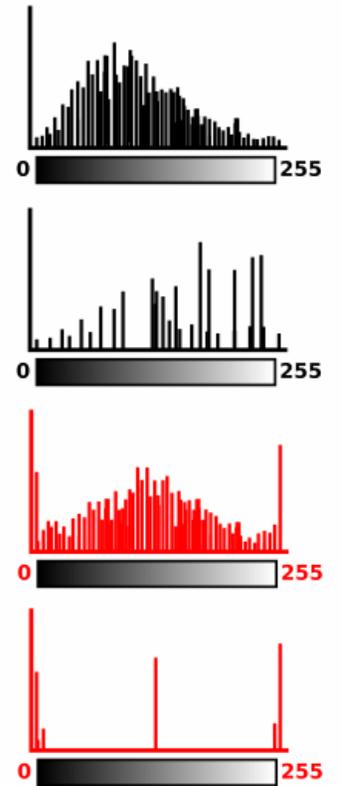
<http://www.ingimp.org/statistics>

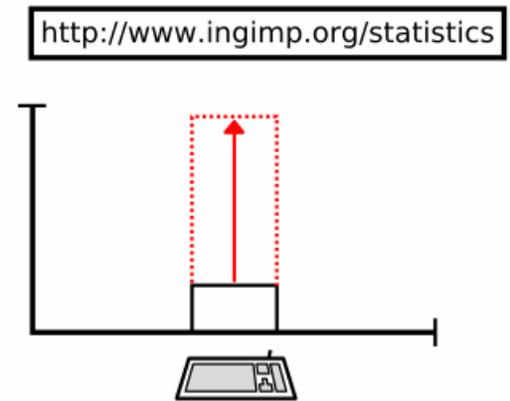


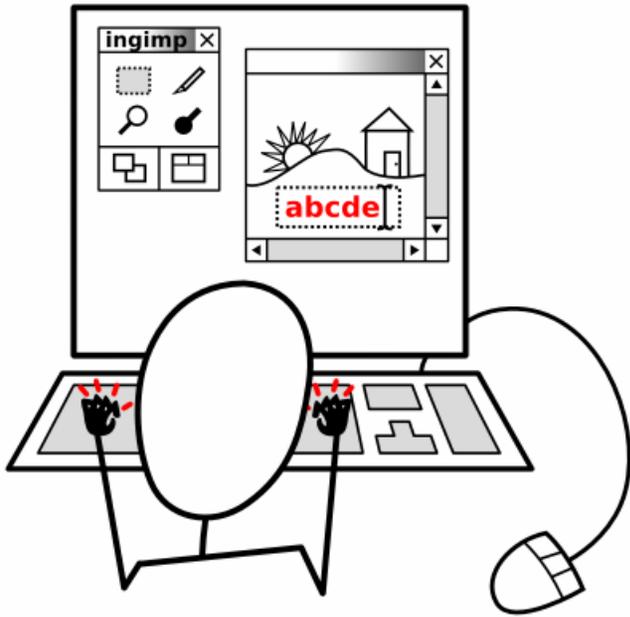


<http://www.ingimp.org/statistics>

HISTOGRAMS







<http://www.ingimp.org/statistics>



Pictograms: Broader Use

- No standard methods to communicate third party data collection practices
 - Logging, crash reports, ...
 - Text and data dumps current state-of-the-art
- Pictograms can complement text-based agreements
 - Afford scanning
 - Wordlessness eases localization burdens
 - Attention-getting to increase chance of viewing

Summing Up...

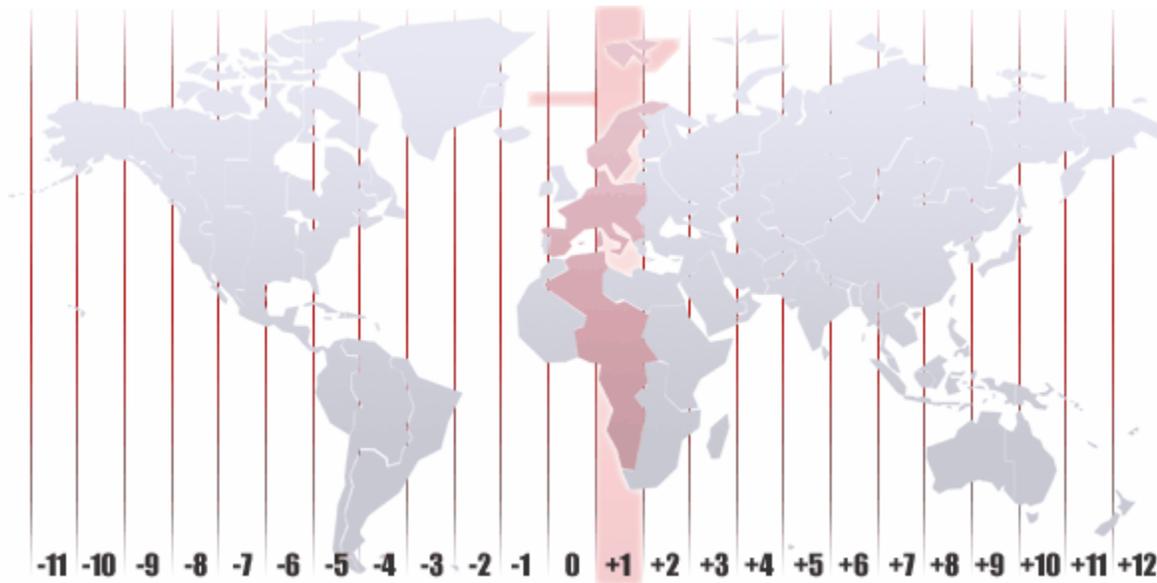
- ingimp:
 - Brand new initiative in *sustainable open source usability*
 - First example of *open instrumentation*
 - Creating a completely unique repository of *real-world, quantitative, longitudinal usage data*
 - Over 500 installations in first three months
 - Majority non-English speaking locales
 - Over 2000 logs with 100,000's of data points logged

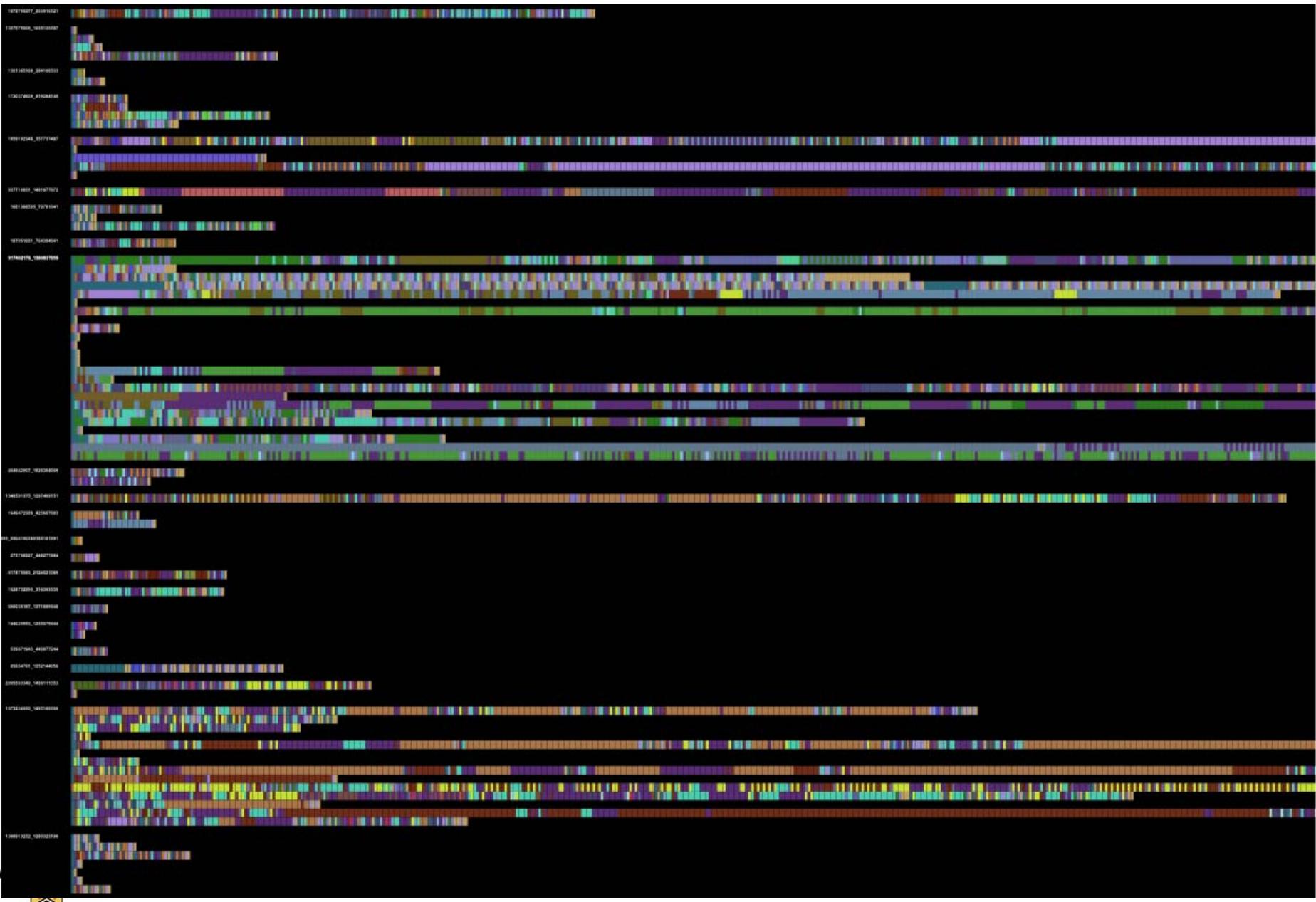
Take-Home Points

- *Personalized* and *personable* statistics help motivate use of instrumented applications
- Immediate availability of collected data can directly benefit end-users
- Transparency in logging process helps allay privacy concerns
- Pictograms make critical privacy information “come alive” without relying on lengthy text
- *More usable end-user open source software benefits everyone*
 - Industry, academia, education, developing countries...

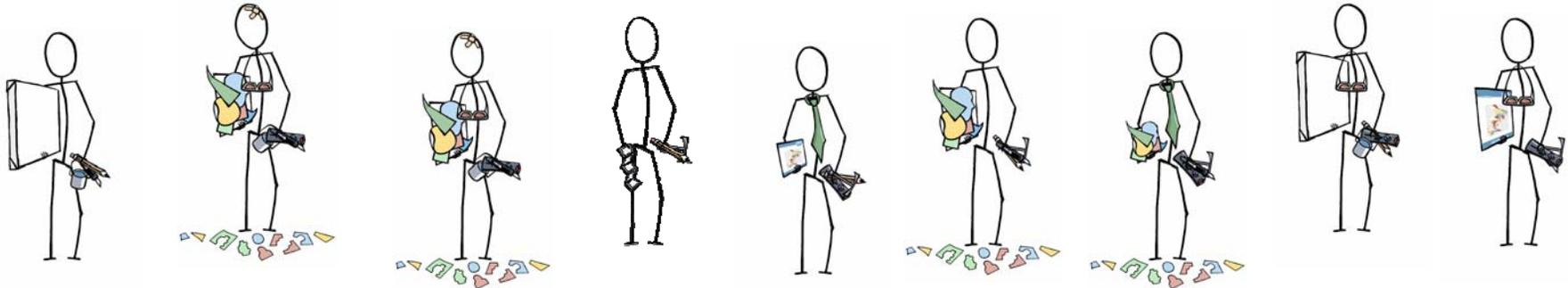
Future Directions

- Richer data analyses, visualizations...





Credits, Thanks, & Links



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

<http://www.ingimp.org>
mterry@cs.uwaterloo.ca

Your Assignment

- Paper prototype for system presenting illustrated consent agreements
- User needs to be able to see
 - Overview
 - Specific data collected
- User needs to be able to jump to specific data collection visualizations
- Little-to-no use of language
- Ideas
 - Something like an installation/customization wizard?
 - Some kind of animation?