user feedback in software evolution

Andrew J. Ko with Parmit Chilana

based on


finding and fixing everyday software defects

- helping developers diagnose failures
- helping users diagnose failures
- detecting usability defects automatically

where do defects come from?
what makes debugging difficult?
what do software teams need to fix defects?
finding and fixing everyday software defects

software team

user community
finding and fixing everyday software defects

software team

user community
finding and fixing everyday software defects

**users** have more opportunities than ever to influence what defects are found and fixed

- open development processes
- perpetual beta status
- open bug reporting
- web-based user feedback
- crowdsourced help in support forums
Users have more opportunities than ever to influence what defects are found and fixed.

Who contributes such feedback?
What do they write about?
What feedback is addressed?
Why is some feedback ignored?
How can teams get better feedback?
How can teams use feedback better?
research from the past two years
who contributes such feedback?
what do they write about?
what feedback is addressed?
why is some feedback ignored?

research in progress
how can teams get better feedback?
how can teams use feedback better?
open bug reporting in the Mozilla project

research from the past two years

who contributes such feedback?
what do they write about?
what feedback is addressed?
why is some feedback ignored?
outline

0. a brief introduction to bug reporting
1. who contributes such feedback?
2. what do they write about?
3. what feedback is addressed?
4. why is some feedback ignored?
a user’s **expectations** are violated

“I wanted this to be a new tab, not a new window!”
user visits bugzilla.mozilla.org
users are encouraged to look for an existing bug report that describes their issue

<table>
<thead>
<tr>
<th>Change in last day(s)</th>
<th>Component</th>
<th>Severity</th>
<th>Op Sys</th>
<th>Target Milestone</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Layout</td>
<td>normal</td>
<td>All</td>
<td>---</td>
<td>[INLINE] Too much space (for descenders) under image (&lt;img&gt;) e.g. in table cells</td>
</tr>
<tr>
<td>0</td>
<td>Startup and Profile System</td>
<td>critical</td>
<td>Windows XP</td>
<td>---</td>
<td>customized toolbar always reset to default on restart, bookmark engines lost, unable to add search engines (localstore.rdf corruption, upgrade or crash). SEE URL FIELD FOR SUPPORT LINK</td>
</tr>
<tr>
<td>0</td>
<td>Widget: Win32</td>
<td>major</td>
<td>Windows 98</td>
<td>---</td>
<td>Mousewheel doesn't work with some driver/Win32 OS combination (scrolling)</td>
</tr>
<tr>
<td>0</td>
<td>Plug-ins</td>
<td>critical</td>
<td>All</td>
<td>---</td>
<td>Application shortcut keys (keyboard commands such as f11, ctrl to fail to operate when plug-in (flash, acrobat, quicktime) has focus)</td>
</tr>
<tr>
<td>0</td>
<td>Layout</td>
<td>enhancement</td>
<td>All</td>
<td>---</td>
<td>(Warning 56k) Alt text is not displayed as a tooltip over &lt;img&gt;</td>
</tr>
<tr>
<td>0</td>
<td>Plug-ins</td>
<td>major</td>
<td>Mac OS X</td>
<td>Future</td>
<td>[OSX] Java applets draw on wrong tabs</td>
</tr>
<tr>
<td>0</td>
<td>Layout</td>
<td>critical</td>
<td>All</td>
<td>mozilla1.9beta1</td>
<td>Splitting Absolutely positioned frames not implemented - Missing page of content when printing or print previewing this site</td>
</tr>
<tr>
<td>0</td>
<td>General</td>
<td>enhancement</td>
<td>All</td>
<td>Firefox1.0</td>
<td>Options for where to open URLs from other applications (reused)</td>
</tr>
</tbody>
</table>
if they don’t find one, they submit a new bug report...

Summary

Options for where to open URLs from other applications (reuse tab, new window)

A sentence which summarises the problem. Please be descriptive and use lots of keywords.

Bad example: mail crashed
Good example: crash if I close the mail window while checking for new POP mail

Details

This is a dupe of http://bugzilla.mozilla.org/show_bug.cgi?id=75138 (for Mozilla). There’s a great discussion of it over there. However, that RFE has been listed for over a year now and is still marked for “future”.

Expand on the Summary. Please be as specific as possible about what is wrong.

Bad example: Mozilla crashed. You suck!
Good example: After a crash which happened when I was sorting in the Bookmark Manager, all of my top-level bookmark folders beginning with the letters Q to Z are no longer present

Reproducibility

Happens every time.

Steps to Reproduce

1.
2.
**Bug 172962 - Options for where to open URLs from other applications (reuse tab, new tab, new window)**

- **Status:** RESOLVED FIXED
- **Keywords:** conversion, fixed-aviary1.0
- **Product:** Firefox
- **Component:** General
- **Version:** unspecified
- **Platform:** All All
- **Importance:** P3 enhancement with [64 votes](#) (vote)
- **Target Milestone:** Firefox1.0
- **Assignee:** Dan M
- **QA Contact:**

**Reported:** 2002-10-06 15:23 PDT by [Duey](#)

**Modified:** 2009-11-24 21:21 PST ([History](#))

**CC List:**
- Add me to CC list
- 119 users ([edit](#))
  - 2005@bdam.de
  - alex@spamcop.net
  - ali@ebrahim.org
  - ask@swva.net
  - asqueella@gmail.com
  - beltzner@mozilla.com
  - benjamin@madbox.us

**Flags:**
- bugs: blocking-aviary1.0PR
- jst: blocking-aviary1.0
  - blocking1.9.0.19
  - wanted1.9.0.x
  - blocking1.8.1.next
  - wanted1.8.1.x
  - blocking1.8.0.next
  - wanted1.8.0.x
  - in-testsuite
  - inditus
reporter
who created
the report
date
when the
report was
created
assignee
the developer
assigned to
resolve this
report
resolution
the reason
the report
was closed
FIXED
DUPLICATE
INCOMPLETE
WONTFIX
WORKSFORME
INVALID
Friday, October 1, 2010
screenshots, patches, test cases, mockups, etc.

attachment
attacher
date

### Attachments

<table>
<thead>
<tr>
<th>Attachment</th>
<th>Details</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>the strings</strong> (1.17 KB, patch)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004-08-30 16:41 PDT, Mano</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>base code to control where URLs are opened</strong> (19.04 KB, patch)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004-09-19 16:54 PDT, Dan M</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>optional addition to divert window.open to a new tab</strong> (5.36 KB, patch)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004-09-19 16:58 PDT, Dan M</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>pref UI</strong> (6.74 KB, patch)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004-09-19 17:00 PDT, Dan M</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>windows-only native hookup of external URL control</strong> (5.03 KB, patch)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004-09-19 17:03 PDT, Dan M</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>mac-only native hookup of external URL control (Aviary branch)</strong> (7.35 KB, patch)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004-09-30 13:13 PDT, Dan M</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>linux-only native hookup of external URL control (Aviary branch)</strong> (6.97 KB, patch)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004-09-30 17:18 PDT, Dan M</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**See Also:**

- blocking1.9.3: ---
- status1.9.3: ---
- blocking1.9.2: ---
- status1.9.2: ---
- blocking1.9.1: ---
- status1.9.1: ---

Friday, October 1, 2010
This is a dupe of [http://bugzilla.mozilla.org/show_bug.cgi?id=75138](http://bugzilla.mozilla.org/show_bug.cgi?id=75138) (for Mozilla). There's a great discussion of it over there. However, that RFE has been listed for over a year now and is still marked for "future".

I think this feature would be important. I, myself, prefer links to open new windows. Others would prefer them to use an existing window. Now that you can also browse with tabs, other people might like a combination of the two (open a new tab in an existing window).

Right now it's pretty much random for me. Some applications open links in a new window while others seem to always use the top most window. However, I think it's important that opening new links be consistent. I can't get all of my applications to do the same thing, so it's logical that I should be able to set my browser to act only one way.

I would like to open up this enhancement for Phoenix since (I think) the developers of Phoenix can add features independent of Mozilla. However, I'll understand it if you feel this isn't that important. At least you will have considered it.

Reproducible: Always

Steps to Reproduce:

comments and commenters

anyone interested in this bug can add information to the report

---

Friday, October 1, 2010
a typical report lifespan

original report
comment
comment
comment
screenshot
comment
comment
patch 1.0
comment
patch 2.0
code review
patch 3.0
closed
1. **who** contributes bug reports?

2. **what** do they write about?

3. **what are the outcomes** of their reports?

4. **why** do their reports have these outcomes?
who contributes bug reports?

obtained all 496,766 reports except for private security patches

15 years of reports, including Netscape years

152,877 unique e-mail addresses

64% addresses only authored, attached to, or commented on 1 report

who's e-mail addresses were these?
who’s e-mail addresses were these?

1% **CORE** developers, drivers, super reviewers, module owners, peers
who was behind these 152,877 addresses?

1% **CORE** developers, drivers, super reviewers, module owners, peers

1% **ACTIVE** developers assigned bug reports
who was behind these 152,877 addresses?

1% **CORE** developers, drivers, super reviewers, module owners, peers

1% **ACTIVE** developers assigned bug reports

80% **REPORTERS**

reported and commented on bug reports

responsible for 54% of reports
who was behind these 152,877 addresses?

1%  **CORE** developers, drivers, super reviewers, module owners, peers

1%  **ACTIVE** developers assigned bug reports

80% **REPORTERS**
reported and commented on bug reports
responsible for 54% of reports

18% **USERS**
only commented on bug reports
# of active contributors by type, per 6 month period

Reporters and users fluctuate spike before a release.
1. **who** contributes bug reports? 
   mostly non-developer, one-time contributors who were active pre-release (REPORTERS)

2. **what** do they write about?

3. **what are the outcomes** of their reports?

4. **why** do their reports have these outcomes?
who contributes bug reports?

mostly non-developer, one-time contributors who were active pre-release (REPORTERS)

what do they write about?

what are the outcomes of their reports?

why do their reports have these outcomes?
2. what do they write about?

selected a sample of 50 REPORTER reports

inductively analyzed titles and descriptions
2. **What** do they write about?

The most salient feature was the **source of expectation**.

What group or person believed the behavior was defective?
2. **what** do they write about?

iterated through 3 samples of 100 reports, converging towards a set of 7 sources

- runtime logic
- specifications
- standards
- prior
- individual
- community
- genre
2 what do they write about?

developer intents
- runtime logic
- specifications
- standards

user expectations
- prior
- individual
- community
- genre
runtime violations

errors, warnings, assertion violations, crashes, hangs, and language-defined invalid states

“...scary deadlock assertions exiting mozilla after referencing nsInstallTrigger...”
specification violations

an agreed upon functional requirement among the application developers

“There’s an incorrectly placed PR_MAX in the code for pref width distribution of colspanning cells.”

application language

Andrew J. Ko – IBM Research Symposium, Sept. 30, 2010
standards violations

industry-wide functional specifications, reaching beyond the application’s developer community

“'codebase' attribute of the HTML 4.0 OBJECT element is not supported…”
violations of a reporter expectations

a reporter’s personal perspective about what the system should do

“Every time I Sort By Name by Bookmarks Firefox sorts and closes my Bookmark menu... Why does it do this??”
violation of a community’s expectations

a reporter’s belief about a “typical” user’s expectations

“The preference to not show the tab bar when only one tab is open could be set to false by default. This would at least alert a new user to the possibility that tabs exist. The old tabbed browsing preferences could be returned.”
violation of genre conventions

inconsistency with the behavior of a similar application

“Firefox does not limit the slideshow horizontal size to the window width. The same source works correctly in IE.”

industry application language

industry user community
app user community
app user
inconsistency with prior behavior

community expectation that behavior of previous versions would be preserved

“The latest version of Firefox only imports one certificate from each file. I used to import all certificates previously.”

industry user community

app user community

app user
1. **who** contributes bug reports?
   - mostly non-developer, one-time contributors who were active pre-release (REPORTERS)

2. **what** do they write about?
   - expectations from developer and user communities of varying population scope

3. **what** are the outcomes of their reports?

4. **why** do their reports have these outcomes?
1. **who** contributes bug reports?
   mostly non-developer, one-time contributors who were active pre-release (**REPORTERS**)

2. **what** do they write about?
   expectations from developer and user communities of varying population scope

3. **what** are the **outcomes** of their reports?

4. **why** do their reports have these outcomes?
3 what are the outcomes of their reports?

most DEVELOPER reports are fixed

- duplicate
- incomplete
- wontfix
- worksforme
- invalid
- fixed

Friday, October 1, 2010
what are the outcomes of their reports?

only 13% of REPORTER reports were fixed
what are the outcomes of their reports?

half of fixed reports were reported by just ~8,000 (6% of) REPORTERs
what are the outcomes of their reports?

most **REPORTER** reports were duplicate, worksforme, or invalid

- **reporter**
  - duplicate
  - incomplete
  - wontfix
  - worksforme
  - invalid
  - fixed

- **active**
  - duplicate
  - incomplete
  - wontfix
  - worksforme
  - invalid
  - fixed

- **core**
  - duplicate
  - incomplete
  - wontfix
  - worksforme
  - invalid
  - fixed
were the duplicates useful?

73% of REPORTERs’ duplicates referred to fixed reports

70% of REPORTERs’ duplicates referred to issues known for > 1 month

66% of REPORTER duplicates of fixed reports were created after a patch was attached

most REPORTER reports identified issues that were already known already patched
# of REPORTER reports by resolution (per 3 months)

dropped since the alpha version...
% of REPORTER resolution types (per 3 months)

↑ likelihood of invalid or incomplete

↓ likelihood of fixed
% fixed reports by contributor type (per 3 months)

Proportion by REPORTERs has dropped since 1.0.
what was the effect of the expectation source in report resolution?

sampled 1,000 REPORTER reports

selected the most salient expectation cited in each report

classified each as one of the 7 sources

redundant coding agreement was ~75%

sources of expectations

prior
runtime logic
individual
community
specifications
standards
genre
what was the effect of the **expectation source** on **REPORTER** report resolutions?

**expectation source** had a significant association with **resolution** $\chi^2(7, N=1000) = 35.8, p<.001$
what was the effect of the expectation source on REPORTER report resolutions?

**genre** and **individual** expectations least likely to be fixed

![Bar chart showing the distribution of report resolutions for different expectation sources.](chart.png)

- **INCOMPLETE**
- **EXPIRED**
- **WONTFIX**
- **INVALID**
- **WORKSFORME**
- **DUPLICATE_NOTFIXED**
- **DUPLICATE_FIXED**
- **FIXED**
what was the effect of the expectation source on REPORTER report resolutions?

multiple reporters filing duplicates from any source of expectation sometimes led to fixes
1. **Who** contributes bug reports?
   - mostly non-developer, one-time contributors who were active pre-release (REPORTERS)

2. **What** do they write about?
   - expectations from developer and user communities of varying population scope

3. **What are the outcomes** of their reports?
   - mostly not fixed, unless grounded in community expectations or in large numbers

4. **Why** do their reports have these outcomes?

---

Andrew J. Ko – IBM Research Symposium, Sept. 30, 2010
1. **who** contributes bug reports?
   mostly non-developer, one-time contributors who were active pre-release (REPORTERS)

2. **what** do they write about?
   expectations from developer and user communities of varying population scope

3. **what are the outcomes** of their reports?
   mostly not fixed, unless grounded in community expectations or in large numbers

4. **why** do their reports have these outcomes?
why do their reports have these outcomes?

a qualitative analysis of reports with REPORTER and USER comments

100 fixed
100 incomplete
100 invalid
100 worksforme
100 duplicate
100 wontfix

+ 40

reports with USER comments
(5% of all reports)
fixed reports
13% of REPORTER reports

terse, productive

an obvious shared understanding of process
usually a single REPORTER followed by a patch
some involved diagnosis by the REPORTER
before a patch could be written
incomplete, invalid, worksforme reports

38% of REPORTER reports

problems were ambiguous and ego-centric

1/3 identified issues already resolved in a recent build

“have you tried the latest nightly build?”

2/3 two thirds identified problems caused by exotic configurations:

“so i trashed the preferences and all was fine again. thank you all for your time. everyday mozilla is getting better, thank to people like you!” (104347:7)
duplicate reports

42% of REPORTER reports

most were about widely experienced problems
with nightly builds

88/100 were marked duplicate on the same day

most had only 1 comment, reminding reporters to check for duplicates

only 12/100 had attachments, such as logs and screenshots
wontfix reports
3% of REPORTER reports

narrow expert feature requests

“it would be nice if I could...”

half explained the resolution, saying the feature was not broadly useful to “regular” users

the rest were denied because the request was supported through other means (e.g., plug-ins)

some REPORTERs expressed frustration

“If you don't change Thunderbird, then Firefox on Mac must be changed, it must be done the same way. (383036:3)”
reports with **USER** comments
5% of all reports

regarded **contentious** Firefox design choices

- bookmarks, location bar, file handling, keyboard shortcuts, tabs, security, history

most **REPORTER** and **USER** contributions expressed

- agreement ("me too!")
- frustration ("this is ridiculous!")
- confusion about the process ("why was it closed?")
contentious reports with **USERs**
5% of all reports, 40 sampled

**REPORTERS** trivialized engineering work

Forcing my users to retype the filename (presuming they even know what it should be) is just plain oppressive, IMHO... The organizations (large choral groups) for which I'm creating sites use the right-click save extensively to download (instead of play in their browsers) audio files for rehearsing music. ... **It's just *code*, guys. Figure it out.**

(299372:49:reporter)
contentious reports with USERs
5% of all reports, 40 sampled

REPORTERS viewed it as a service

Over two months ago I gave complete information on when and how I got the error AND spent a great deal of time isolating the messages that caused it ... Which part of that is just saying "me too"? For crying out loud, I'm a nursing student, not a programmer. Do you do your own x-rays before going to the doctor?
(252697:10:user)
contentious reports with **USERs**
5% of all reports, 40 sampled

**REPORTERS** expected service for payment

Mozilla "Foundation", you have cash, you have the resources. FIX IT.
**PEOPLE DONATED MONEY TO HAVE YOU *FIX THIS KIND-OF SHIT*...**This is the *EXACT* sort-of situation that shows why open-source *FAILS*.
contentious reports with **USERS**
5% of all reports, 40 sampled

developers tried to explain the process

**OK, calm down everyone.** How many times do I have to say it? It's what we want to try out to start with. That is not code for "we've made a decision" or "your arguments all suck and we're going to ignore them"... **That's what the trunk is for. Experimentation.** We want to do a UI experiment.
1. **who** contributes bug reports?
   - mostly non-developer, one-time contributors who were active pre-release (REPORTERS)

2. **what** do they write about?
   - expectations from developer and user communities of varying population scope

3. **what are the outcomes** of their reports?
   - mostly not fixed, unless grounded in community expectations or in large numbers

4. **why** do their reports have these outcomes?
   - REPORTERS understood little about the context of resolution decisions, leading to missing information and egocentric requests
how well did this dialog between users and developers work?
the ~8,000 successful REPORTERs act like closed source beta testers
most active before a release
effective at writing good reports
trained over several years
the ~8,000 successful REPORTERs act like closed source beta testers
most active before a release
effective at writing good reports
trained over several years

all other REPORTERs were far less effective
most of their issues should have been triaged by a tech support
misunderstandings about process sometimes led to friction
bug reporting is a **skill**

- 🔺 open bug reporting communities appear to **cultivate** this skill
- ✗ few contributors acquire these skills, perhaps because are early negative experiences
- ✗ bug reporting is an ineffective place to gather **user feedback**

---

**REPORTERs**
- beta testers
- in release
- good reports
- years

were far less

- could have technical support

about process

---

Andrew J. Ko – IBM Research Symposium, Sept. 30, 2010
future work

do these trends occur in non-FLOSS projects via user feedback?

**do explicitly user-centered**

software teams treat user feedback differently?

are there better ways of gathering and **aggregating user feedback**

from help forums, new help tools?

what kind of data would help teams **interpret** user feedback when triaging bugs?