Perspectives Workshop: New Frontiers for Empirical Software Engineering

Vic Basili
Nachi Nagappan
Dieter Rombach
Andreas Zeller

Software Engineering

"the application of *engineering*
to software"

or

"the application of a *systematic, disciplined, quantifiable approach*
to the development, operation,
and maintenance of software"

Software Engineering

systematic

disciplined

quantifiable

Empirical Software Engineering

"You can't control what you can't measure"

Tom DeMarco

systematic

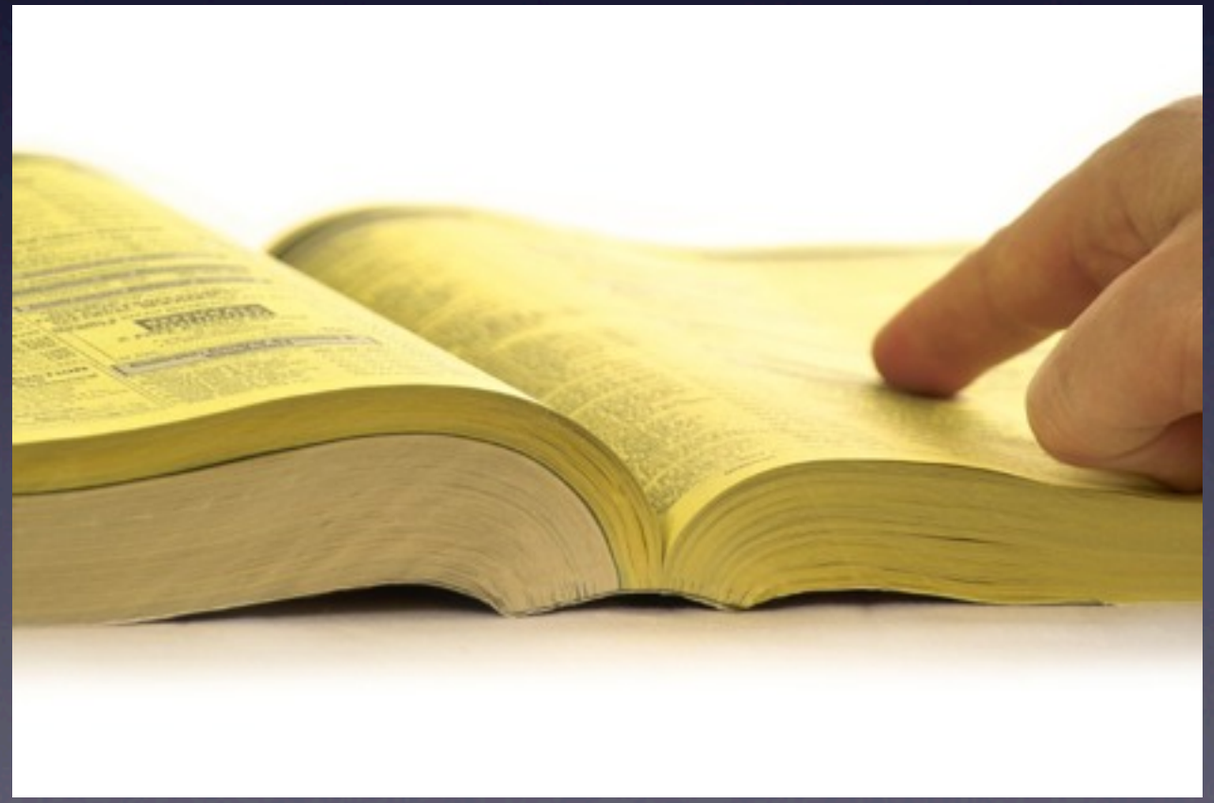
disciplined

quantifiable

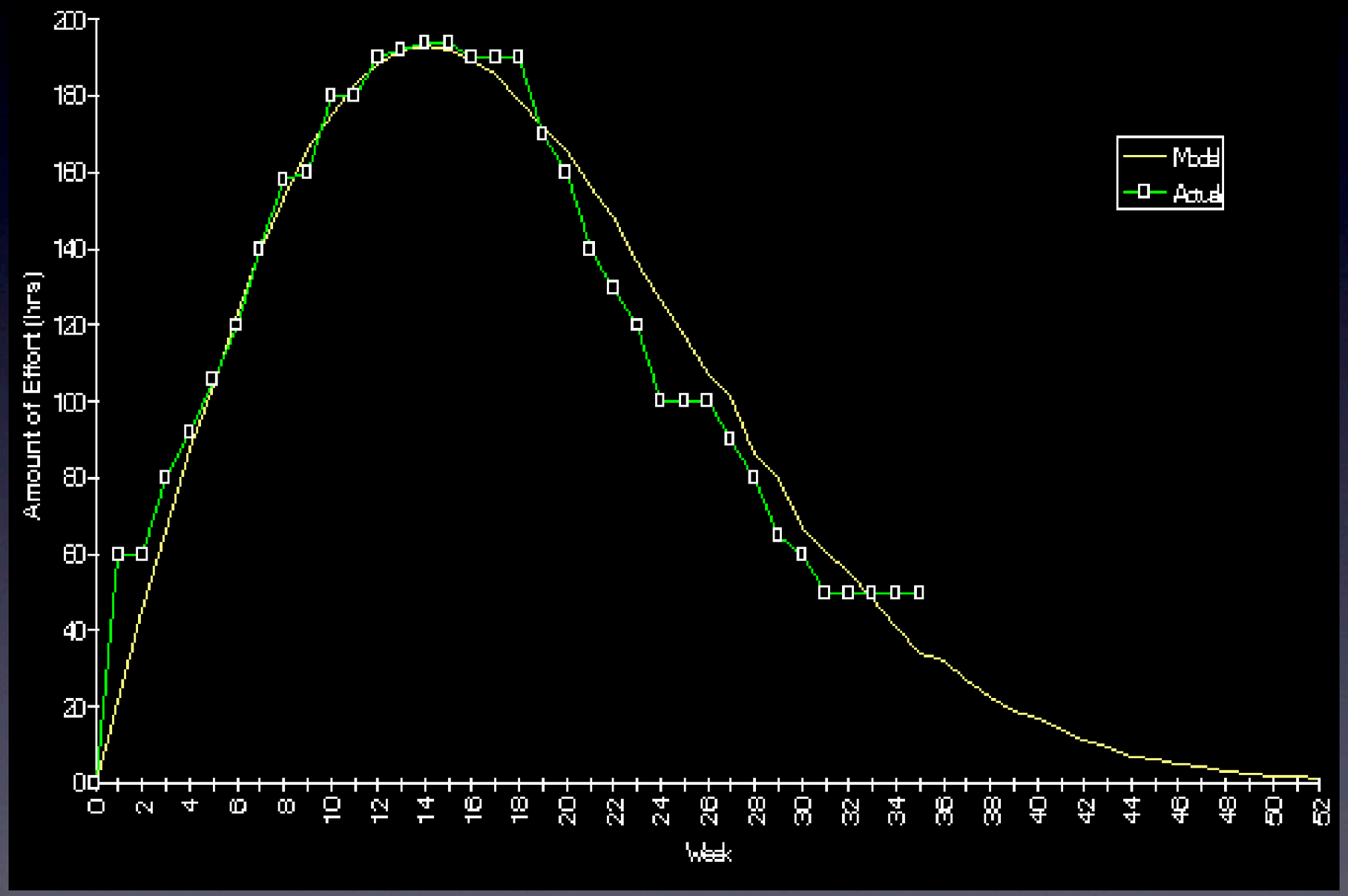
Empirical Software Engineering

"You can't control what you can't measure"

Tom DeMarco



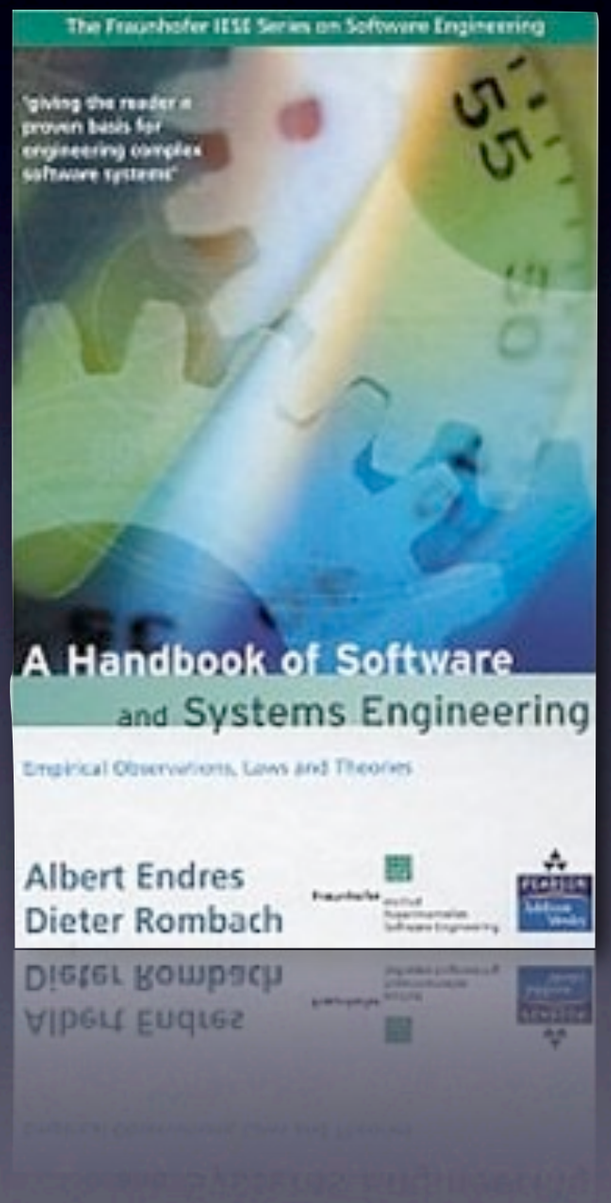
Studies

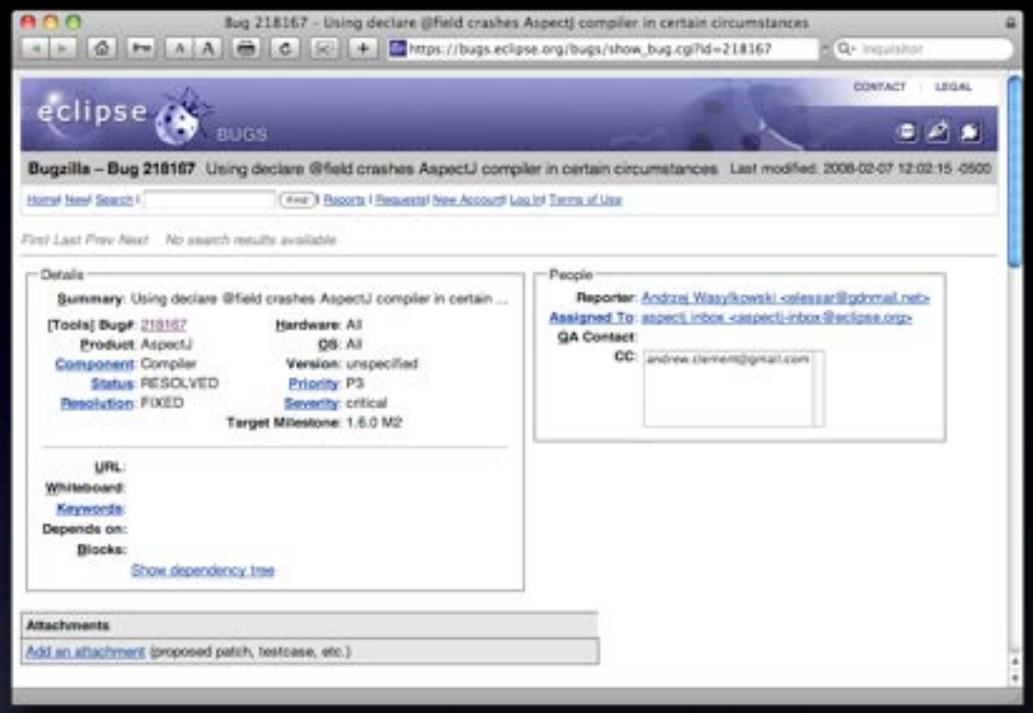


Rosenberg, L. and Hyatt, L. "Developing An Effective Metrics Program"
European Space Agency Software Assurance Symposium, Netherlands, March, 1996

Empirical SE 1.0

- Data collected manually
- Few data points available
- Few studies available

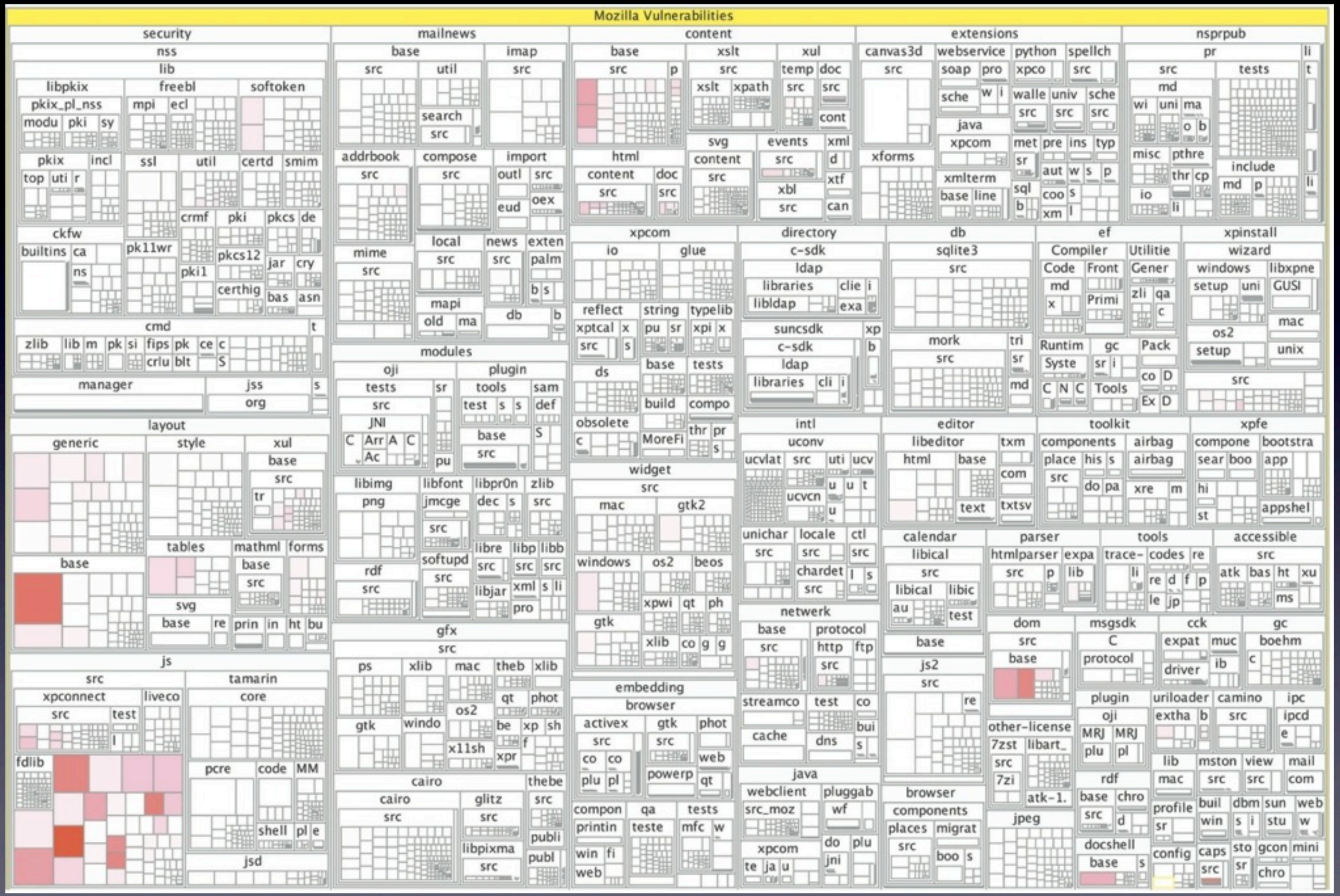




Bugs

Versions

Map *bugs* to *code*

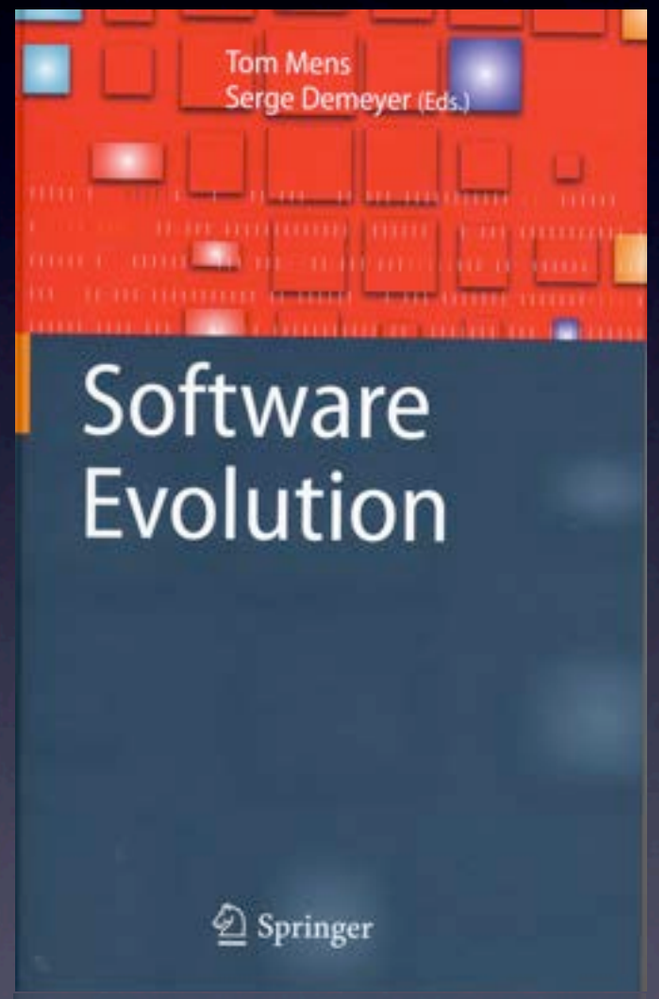

Mozilla Vulnerabilities
security
nss
lib
libpkix
pkix_pl_nss
modu
pki
sy
pkix
top
uti
r
incl
ckfw
builtins
ca
ns
freebl
mpi
ecl
softoken
ssl
pk11wr
util
crmf
pki1
pki
pkcs12
certhig
certd
pkcs
jar
bas
smim
de
cry
asn
cmd
zlib
lib
m
pk
si
fips
crlu
pk
blt
ce
c
S
t
manager
jss
org
s
layout
generic
base
style
svg
base
re
xul
base
src
tr
tables
mathml
base
src
forms
prin
in
ht
bu
mailnews
base
src
util
search
src
imap
src
addrbook
src
compose
src
import
outl
src
eud
oex
mime
src
local
src
news
src
exten
palm
b
s
mapi
old
ma
db
b
modules
oji
tests
sr
src
JNI
C
Arr
A
C
Ac
pu
plugin
tools
test
s
s
sam
def
S
base
src
libimg
png
libfont
jmcge
src
softupd
src
libpr0n
dec
s
libre
src
libjar
libp
src
libb
src
xml
s
li
pro
zlib
src
rdf
src
gfx
src
ps
xlib
mac
theb
xlib
qt
phot
gtk
windo
os2
be
xp
sh
f
x11sh
xpr
cairo
cairo
src
glitz
src
libpixma
src
thebe
src
publi
publ
content
base
src
p
html
content
src
doc
src
xslt
src
xslt
xpath
svg
content
src
xul
temp
src
doc
src
cont
events
src
xbl
src
xml
d
xtf
can
xpcom
io
reflect
xptcal
x
src
s
ds
obsolete
c
glue
string
pu
sr
base
build
MoreFi
typelib
xpi
x
tests
compo
thr
pr
s
widget
src
mac
gtk2
windows
os2
beos
xpwi
qt
ph
gtk
xlib
co
g
g
embedding
browser
activex
src
co
co
plu
pl
gtk
src
web
powerp
qt
phot
compon
printin
win
fi
web
qa
teste
tests
mfc
w
directory
c-sdk
ldap
libraries
libldap
clie
exa
i
suncsdk
c-sdk
ldap
libraries
cli
i
xp
b
intl
uconv
ucvlat
src
ucvcn
uti
u
ucv
u
t
u
unichar
src
locale
src
chardet
src
ctl
src
l
s
netwerk
base
src
protocol
http
src
ftp
streamco
cache
test
dns
co
bui
s
java
webclient
src_moz
xpcom
te
ja
u
pluggab
wf
do
jni
plu
db
sqlite3
src
mork
src
tri
sr
md
editor
libeditor
html
base
text
txm
com
txtsv
calendar
libical
src
libical
au
libic
test
base
js2
src
re
browser
components
places
src
migrat
boo
s
extensions
canvas3d
src
xforms
webservice
soap
sche
pro
w
i
java
xpcom
xmlterm
base
line
python
xpco
walle
src
met
sr
sql
b
pre
aut
coo
xm
univ
src
ins
w
s
l
spellch
src
sche
src
typ
p
nsprpub
pr
src
md
wi
uni
ma
o
b
misc
pthre
thr
cp
io
li
tests
include
md
p
li
t
li
ef
Compiler
Code
md
x
Front
Primi
Runtim
Syste
C
N
C
gc
sr
i
Tools
Pack
co
D
Ex
D
toolkit
components
place
src
his
s
do
pa
airbag
airbag
xre
m
parser
htmlparser
src
p
expa
lib
tools
trace-
li
codes
re
re
d
f
p
le
jp
accessible
src
atk
bas
ht
xu
ms
dom
src
base
other-license
7zst
src
7zi
libart_
atk-1.
jpeg
msgsdk
C
protocol
cck
expat
muc
driver
ib
gc
boehm
c
plugin
oji
MRJ
plu
MRJ
pl
rdf
base
src
chro
d
docshell
base
s
uriloader
extha
b
lib
mac
profile
sr
config
camino
src
mston
src
view
src
buil
win
dbm
s
i
sun
stu
caps
src
sto
sr
ipc
ipcd
e
mail
com
web
w
gcon
chro
mini
xpfe
compone
sear
boo
hi
st
bootstra
app
appshel
xpinstall
wizard
windows
setup
uni
libxpne
GUSI
mac
os2
setup
unix
src

Where do bugs come from?

- Earlier bugs?
- Developer experience?
- Specific domains?
- Lack of testing?
- Code complexity?

Empirical SE 2.0

- Data collected automatically
- Thousands of data points
- Can be widely automated
- Explosion in papers & topics



Invitees

- Leaders in Empirical Software Engineering
- Representatives from
  - Program Analysis + Testing
    leads to better mining of software archives
    primary target of defect prediction activities
  - Process Analysis and Modeling
    search for features that correlate with defects
  - Emerging Domains
    posing new challenges for Software Engineering

Empirical Research and Program Analysis

- Better analysis via empirical findings
- Better benchmarks (real bugs + changes)
- Better cost-benefit-risk studies
- Dealing with partial information

Empirical Research with Web 2.0

- How do Web 2.0 techniques influence SE?
- How can they help in empirical studies?
- Need new data collection methods
- Need infrastructure for shared research

Software Engineering and Climate Science

- SE in Climate Science is different
- Domain experts who develop programs
- Build high quality software w/ lots of QA
- Development = Scientific work

Software Engineering and Health IT

- Socio-technical networks for patients
- Rate my doctor; recommender systems
- How to architecture such systems?
- Privacy and reliability issues

Specific Outcomes

as reported by attendees

- Special Issues of IEEE Software
  Climate Change Science and Software (2011)
  Social Networking and Software (2013)
- "Several proposals" on
  data mining and process work
  leveraging empirical data for analysis
- Extended Editorial Boards (ESE, IST)
- PC Chairs met at Dagstuhl (ICST)

Retro-Perspectives

- Outcome was substantial
- Outcome was mostly indirect and political
- Organizer's role was matchmaking
- Driven by participants, not by vision
- Driven by opportunities, not issues

Empirical SE 2.0

- Data collected automatically
- Thousands of data points
- Can be widely automated
- Explosion in papers & topics

Empirical Research and Program Analysis

- Better analysis via empirical findings
- Better benchmarks (real bugs + changes)
- Better cost-benefit-risk studies
- Dealing with partial information

Empirical Research with Web 2.0

- How do Web 2.0 techniques influence SE?
- How can they help in empirical studies?
- Need new data collection methods
- Need infrastructure for shared research

Software Engineering and Climate Science

- SE in Climate Science is different
- Domain experts who develop programs
- Build high quality software w/ lots of QA
- Development = Scientific work

Software Engineering and Health IT

- Socio-technical networks for patients
- Rate my doctor; recommender systems
- How to architecture such systems?
- Privacy and reliability issues

Perspectives Workshop: New Frontiers for Empirical Software Engineering

Vic Basili
Nachi Nagappan
Dieter Rombach
Andreas Zeller

Specific Outcomes

as reported by attendees

- Special Issues of IEEE Software
  Climate Change Science and Software (2011)
  Social Networking and Software (2013)
- "Several proposals" on
  data mining and process work
  leveraging empirical data for analysis
- Extended Editorial Boards (ESE, IST)
- PC Chairs met at Dagstuhl (ICST)

Retro-Perspectives

- Outcome was high
- Outcome was mostly indirect and political
- Organizer's role was matchmaking
- Driven by participants, not by vision
- Driven by opportunities, not issues