# Fighting SPAM:

## *Whitelisting Revisited*

David Erickson
derickso@stanford.edu

Martin Casado
casado@cs.stanford.edu

Nick McKeown
nickm@stanford.edu

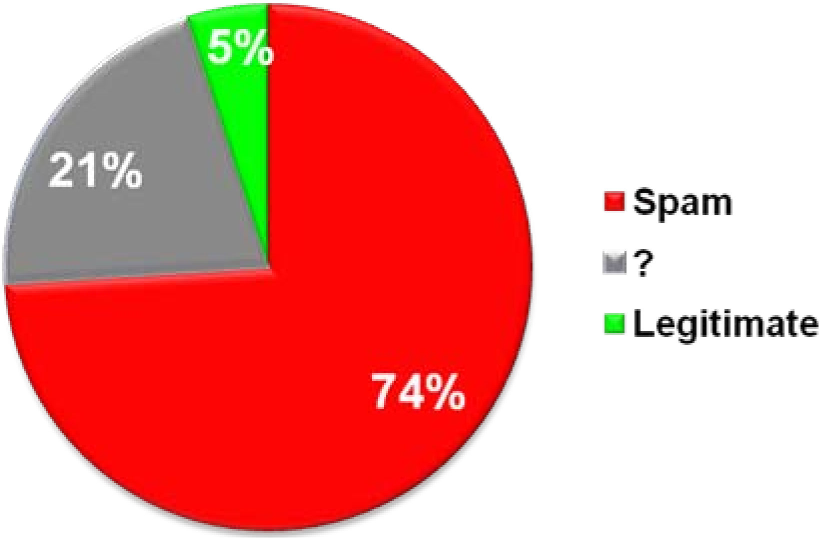
*Member project of the*

Stanford Clean Slate Program

http://cleanslate.stanford.edu

# Motivation

- In 2007, 74-95% of all email was SPAM



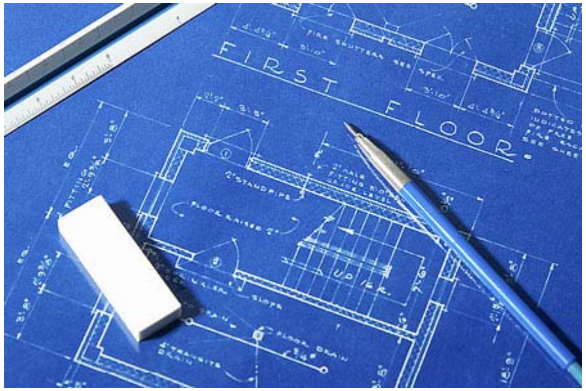
- 1.2% of employee time
  - $713 per year per employee
  - $200 billion cost to companies worldwide

# Whitelisting

- What is it?
  - – Email must match a whitelist entry to be delivered
  - – Entries contain email addresses / domains
- Often paired with challenge-response
  - – Shifts some burden from user to sender
  - – Has its own list of complaints
- Is it feasible?
  - – Lots of opinions, little data

# Methodology

- Built an operational system
  - – Default Off Email (DOEmail)
- Heavily instrumented
  - – Email and user behavior
- Running for nearly 2 years
  - – ~800,000 emails processed to date
- Real users
  - – 120+ accounts have received email

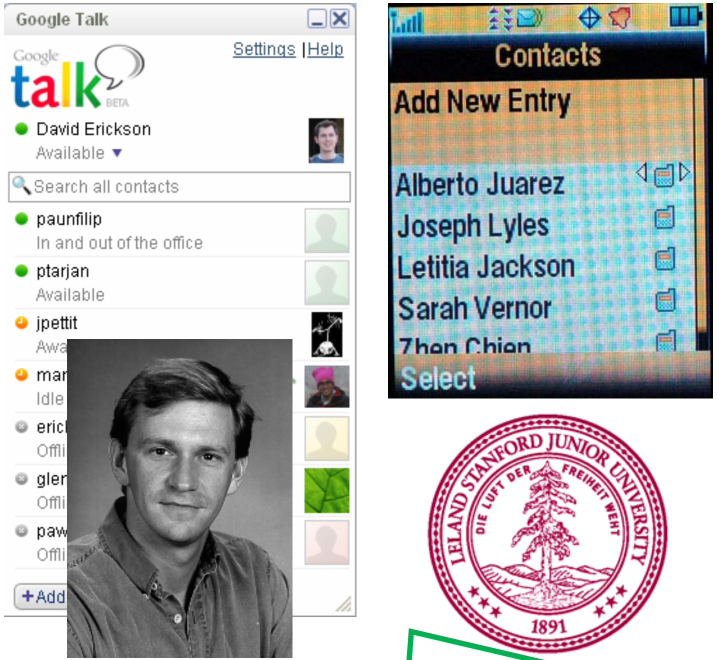


# Default Off Email

- Create an account
  - E.g. dericksо@doemail.org
- Forward existing email
- Set destination for cleaned email
- Install Mozilla Thunderbird
  - And use our custom add-on!
  - … or use the web interface
- Populate white/black lists

# Sender Categories

## Whitelist



Google Talk

Settings | Help

David Erickson
Available

Search all contacts

paunfilip
In and out of the office

ptarjan
Available

jpettit
Awa

mar
Idle

eric
Offli

gler
Offli

paw
Offli

+Add

Contacts

Add New Entry

Alberto Juarez

Joseph Lyles

Letitia Jackson

Sarah Vernor

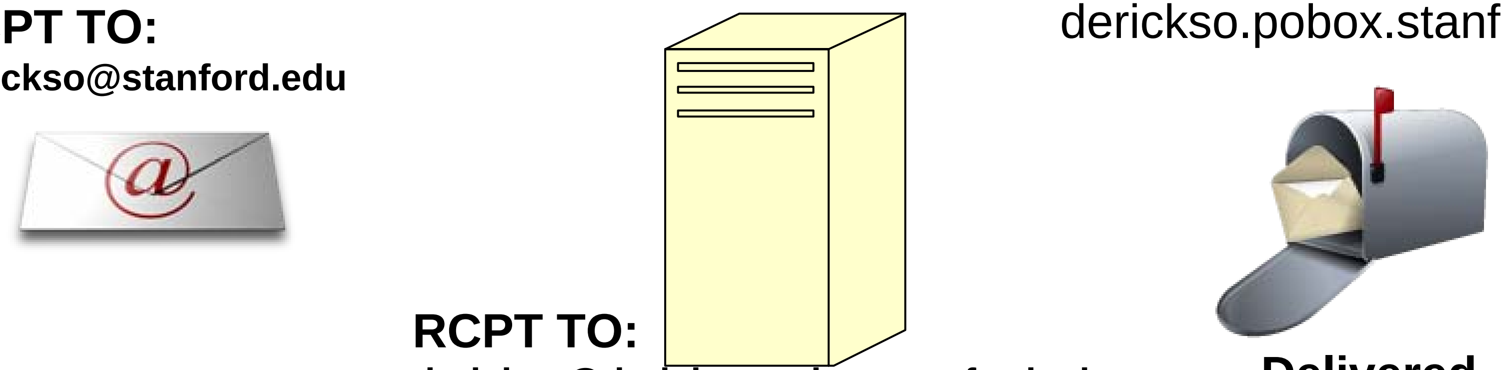
Zhen Chien

Select

## Blacklist

## Unknown

# Stanford Integration



stanford.edu

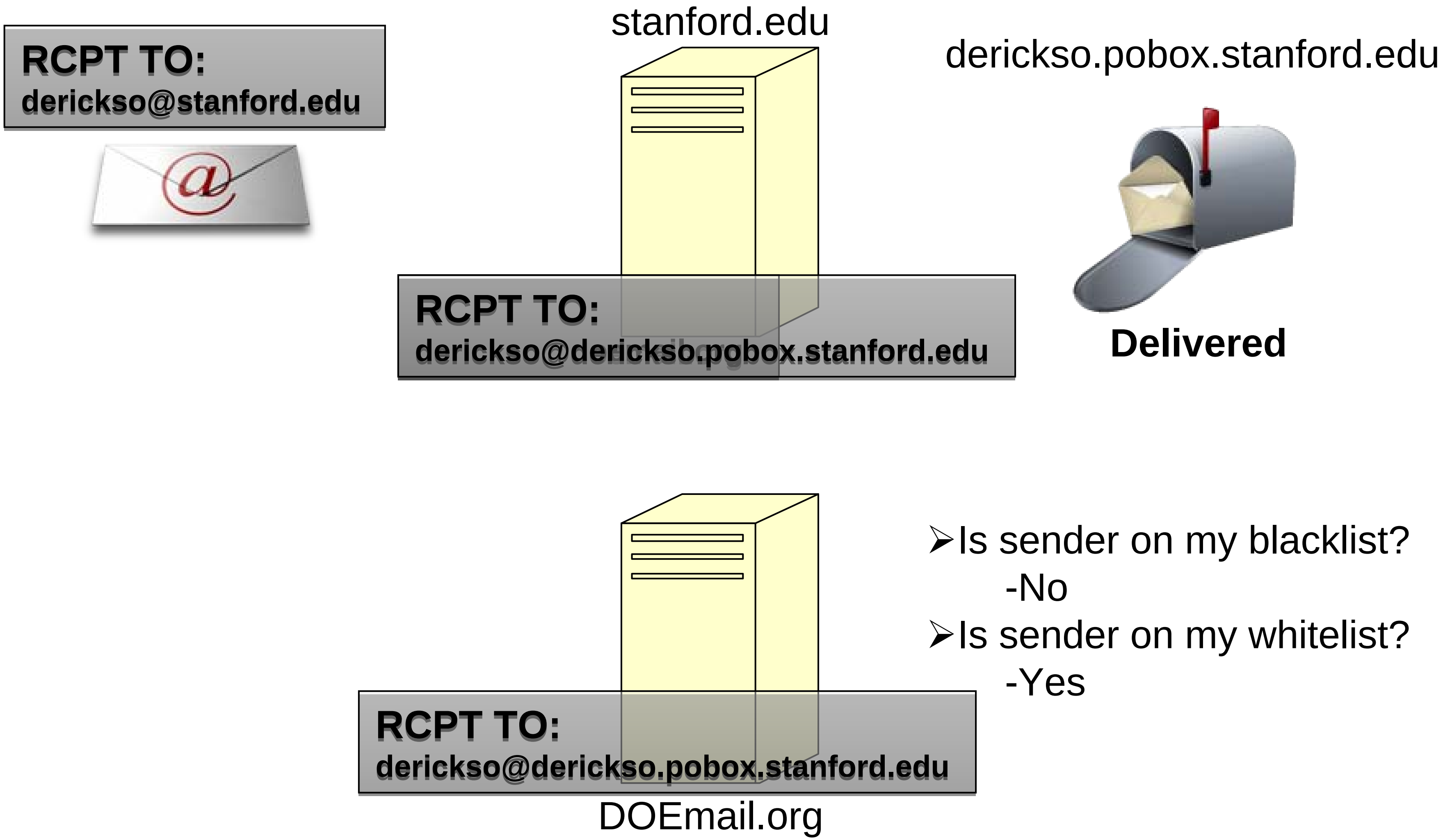
RCPT TO:
**dericksо@stanford.edu**

derickso.pobox.stanford.edu

**RCPT TO:**
**derickso@derickso.pobox.stanford.edu**

**Delivered**

# Stanford w/DOEmail Whitelist



**RCPT TO:**
**derickso@stanford.edu**

stanford.edu

**RCPT TO:**
**derickso@derickso.pobox.stanford.edu**

derickso.pobox.stanford.edu

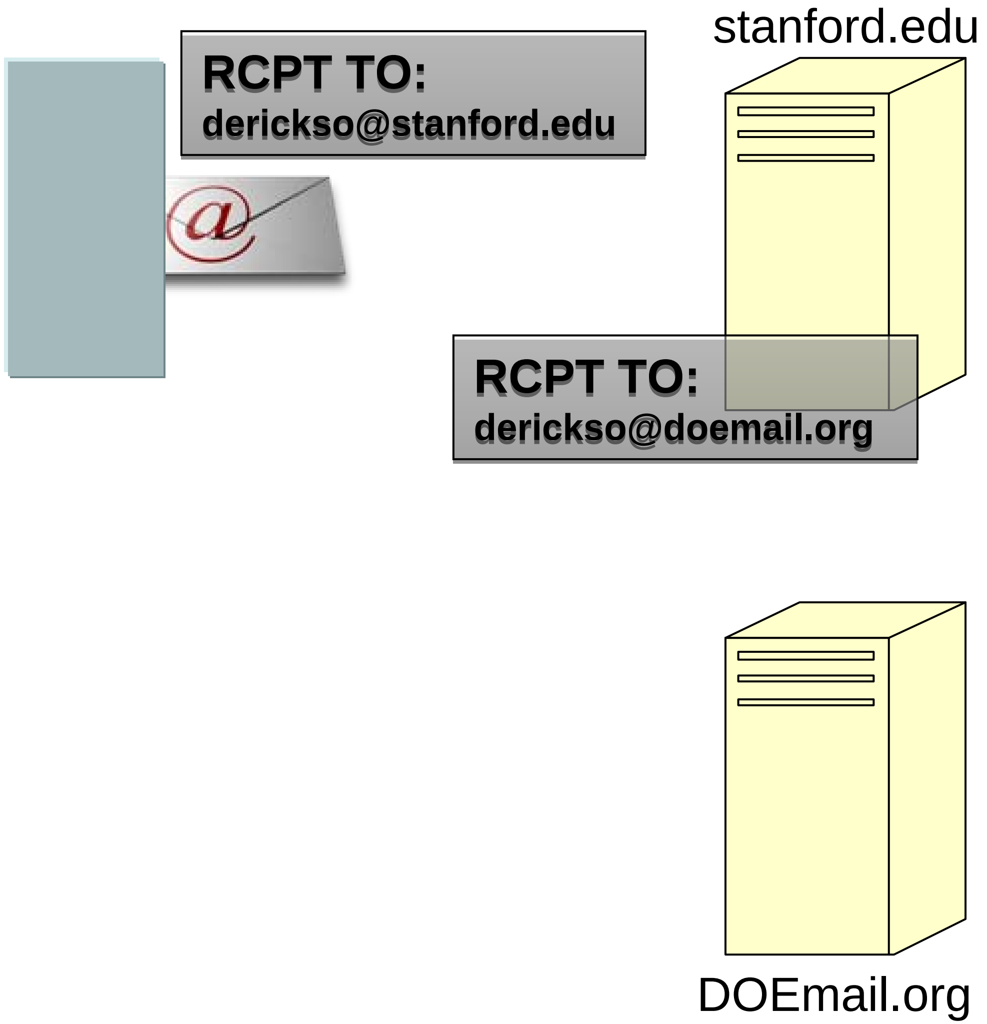
**Delivered**

- ➢Is sender on my blacklist?
  -No
- ➢Is sender on my whitelist?
  -Yes

**RCPT TO:**
**derickso@derickso.pobox.stanford.edu**

DOEmail.org

# Stanford w/DOEmail Unknown



stanford.edu

RCPT TO:
derickso@stanford.edu

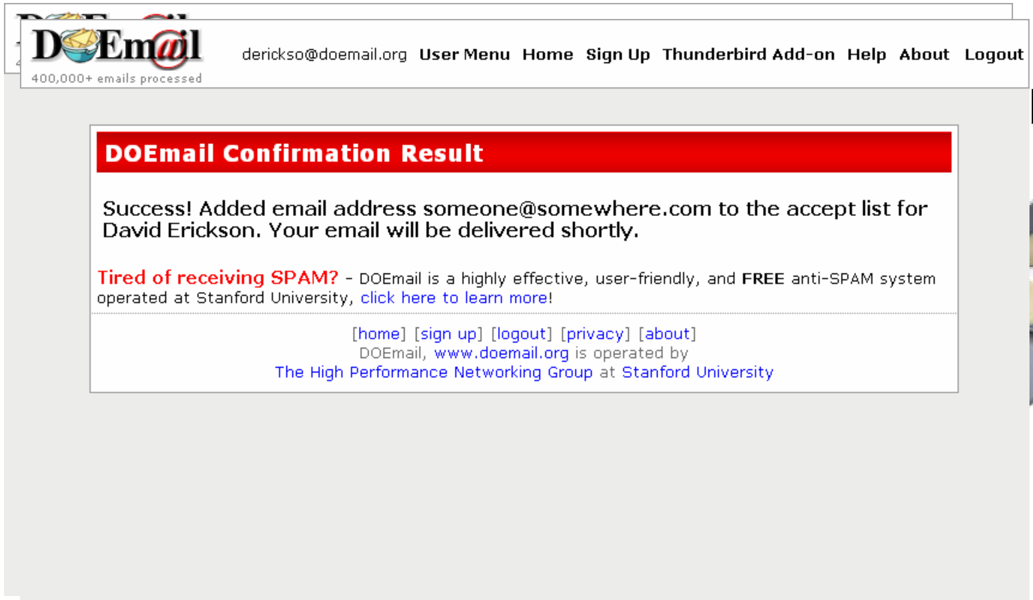
RCPT TO:
derickso@doemail.org

DOEmail.org

derickso.pobox.stanford.edu

- Is sender on my blacklist?
  -No
- Is sender on my whitelist?
  -No

# Stanford w/DOEmail Unknown



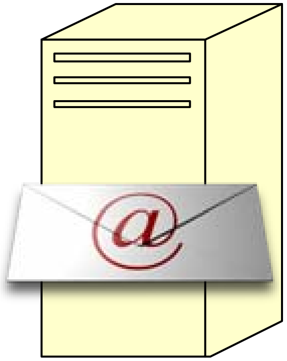
**DOEmail Confirmation Result**

Success! Added email address someone@somewhere.com to the accept list for David Erickson. Your email will be delivered shortly.

Tired of receiving SPAM? - DOEmail is a highly effective, user-friendly, and **FREE** anti-SPAM system operated at Stanford University, click here to learn more!

[home] [sign up] [logout] [privacy] [about]
DOEmail, www.doemail.org is operated by
The High Performance Networking Group at Stanford University

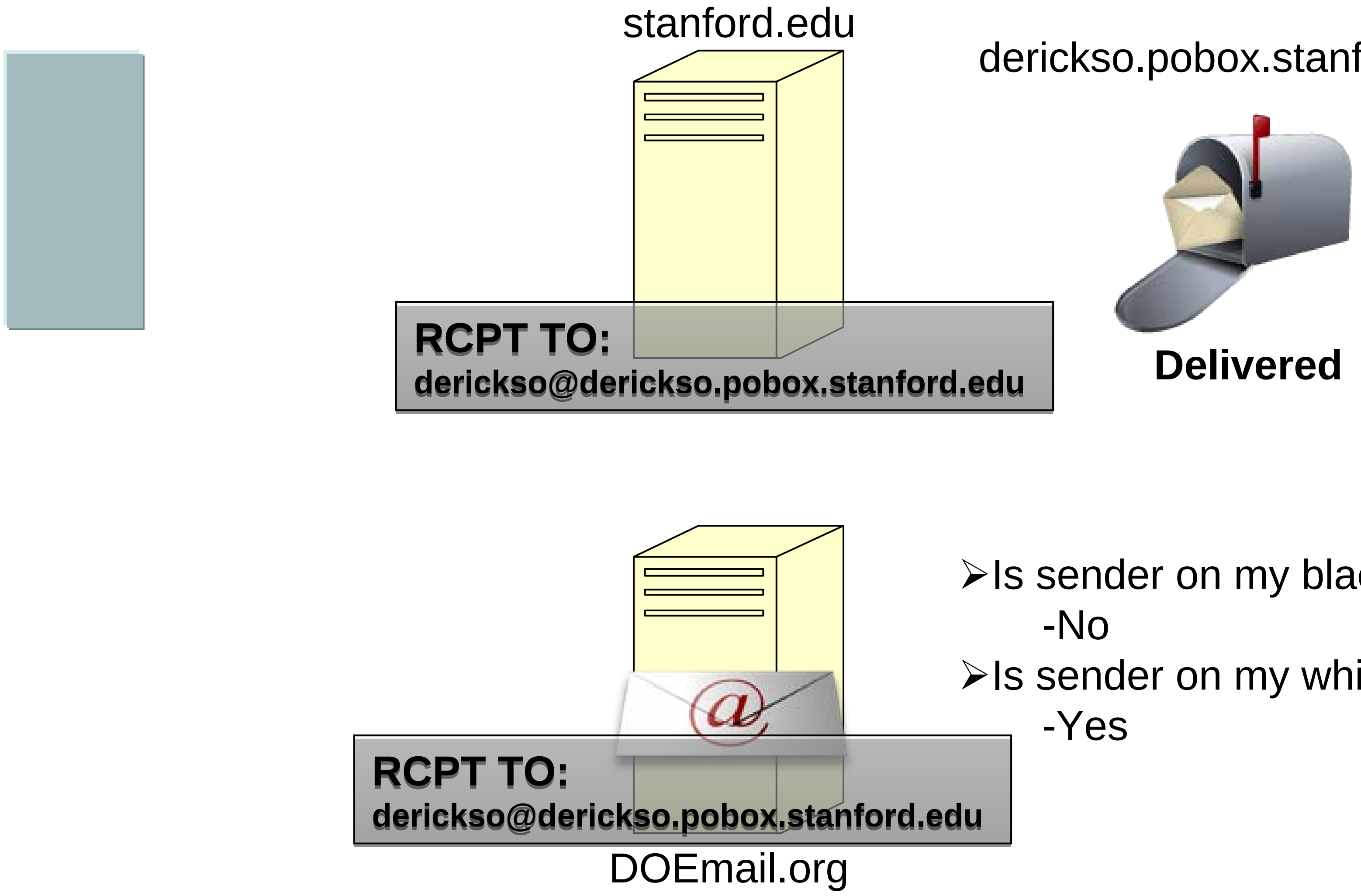
box.stanford.edu



DOEmail.org

- Is sender on my blacklist?
  -No
- Is sender on my whitelist?
  -No

# Stanford w/DOEmail Unknown



stanford.edu

derickso.pobox.stanford.edu

**RCPT TO:**
**derickso@derickso.pobox.stanford.edu**

**Delivered**

- Is sender on my blacklist?
  - -No
- Is sender on my whitelist?
  - -Yes

**RCPT TO:**
**derickso@derickso.pobox.stanford.edu**

DOEmail.org

# Tools

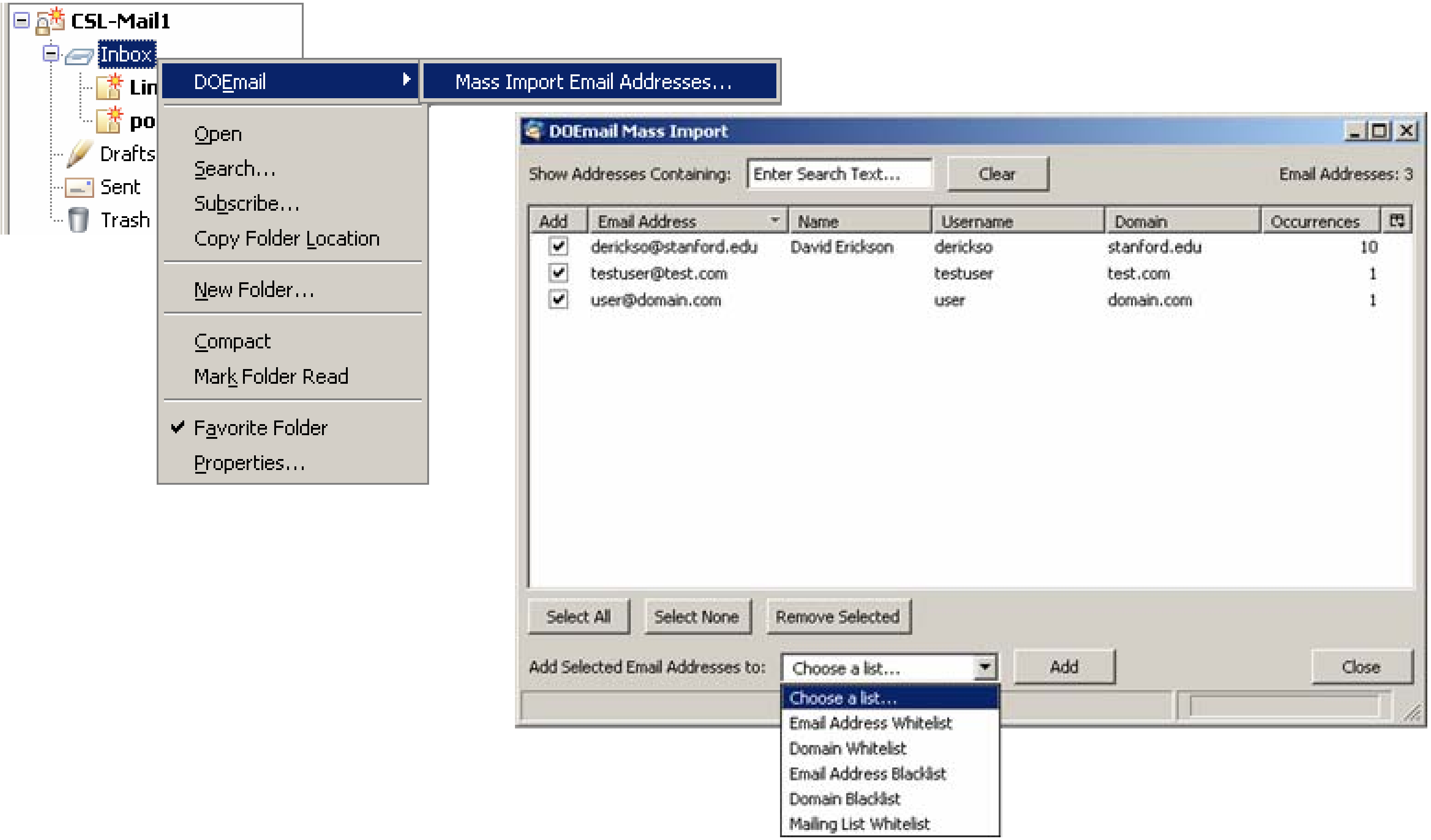
- Mozilla Thunderbird and Web Interfaces
- Import your whitelist
- Whitelist your email recipients
- Detect mailing lists
- View and manage pending email
- Monitor your statistics

mozilla
Thunderbird™

W3C® World Wide Web
consortium

# Thunderbird Add-on

- Import email addresses and domains from existing mail folders



Show Addresses Containing: Enter Search Text...

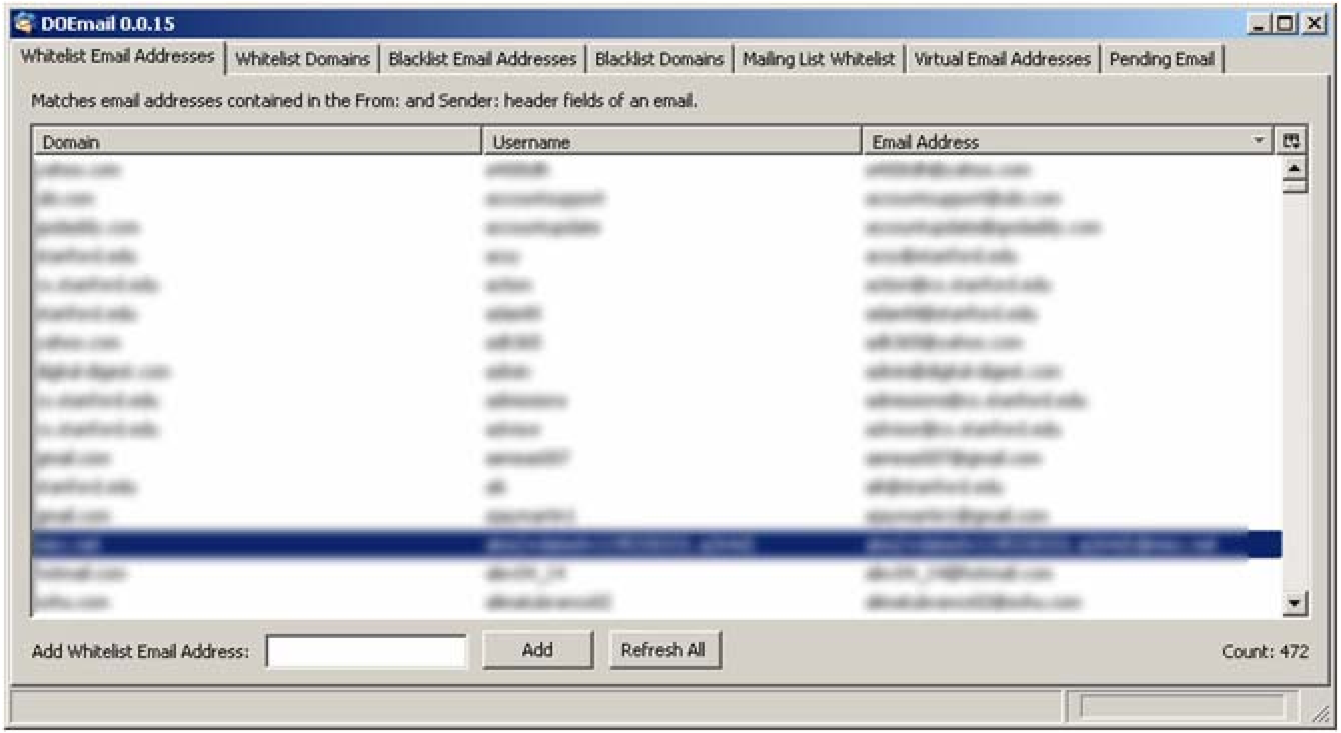
Email Addresses: 3

| Add | Email Address | Name | Username | Domain | Occurrences |
|---|---|---|---|---|---|
| ☑ | dericksо@stanford.edu | David Erickson | derickso | stanford.edu | 10 |
| ☑ | testuser@test.com | | testuser | test.com | 1 |
| ☑ | user@domain.com | | user | domain.com | 1 |

Add Selected Email Addresses to: Choose a list...

# Thunderbird Add-on

- Manage white and blacklists



Matches email addresses contained in the From: and Sender: header fields of an email.

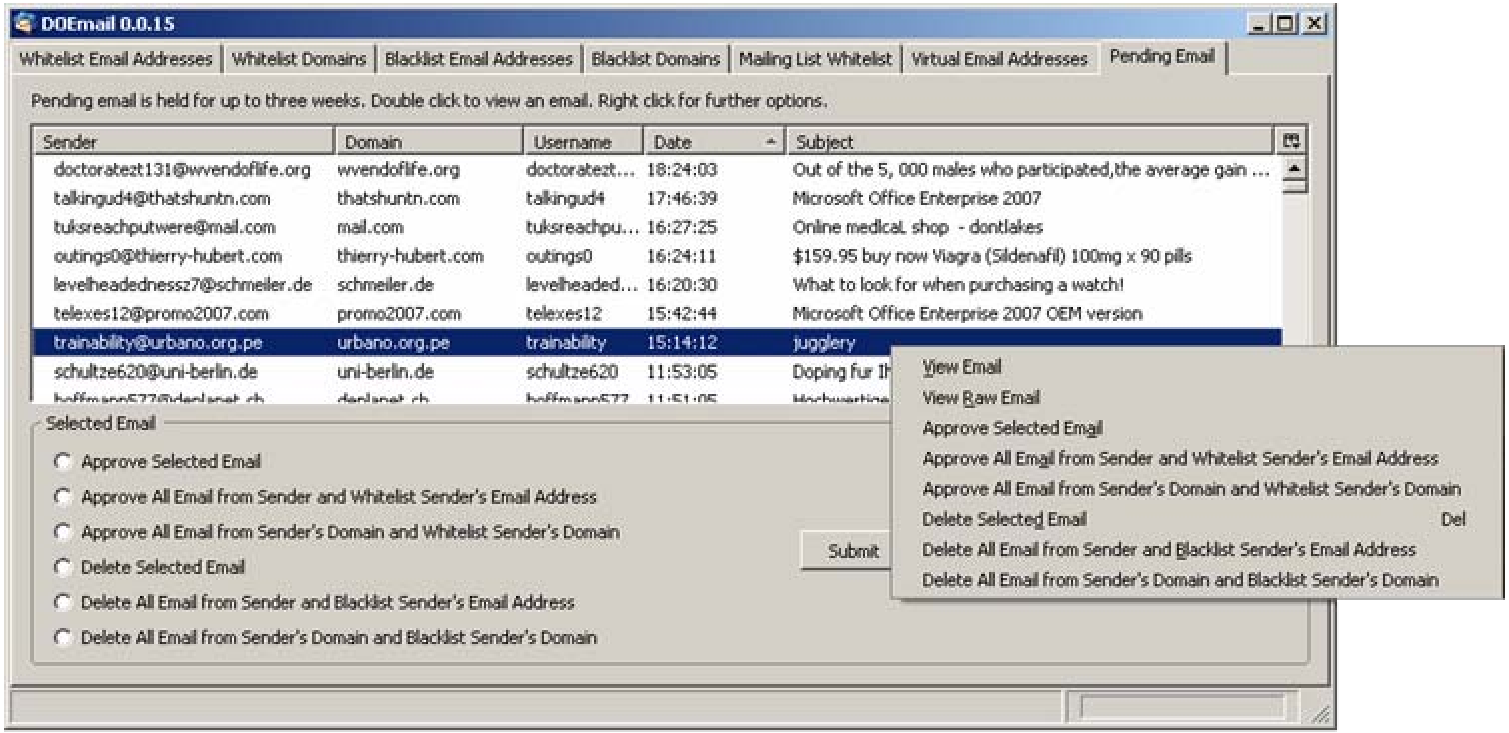
| Domain | Username | Email Address |
|---|---|---|
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] |

Add Whitelist Email Address:

Count: 472

# Thunderbird Add-on

- View and manage pending email


DOEmail 0.0.15
Whitelist Email Addresses
Whitelist Domains
Blacklist Email Addresses
Blacklist Domains
Mailing List Whitelist
Virtual Email Addresses
Pending Email
View Email
View Raw Email
Approve Selected Email
Approve All Email from Sender and Whitelist Sender's Email Address
Approve All Email from Sender's Domain and Whitelist Sender's Domain
Delete Selected Email Del
Delete All Email from Sender and Blacklist Sender's Email Address
Delete All Email from Sender's Domain and Blacklist Sender's Domain
Submit

Pending email is held for up to three weeks. Double click to view an email. Right click for further options.

| Sender | Domain | Username | Date | Subject |
|---|---|---|---|---|
| doctoratezt131@wvendoflife.org | wvendoflife.org | doctoratezt… | 18:24:03 | Out of the 5, 000 males who participated,the average gain … |
| talkingud4@thatshuntn.com | thatshuntn.com | talkingud4 | 17:46:39 | Microsoft Office Enterprise 2007 |
| tuksreachputwere@mail.com | mail.com | tuksreachpu… | 16:27:25 | Online medicaL shop - dontlakes |
| outings0@thierry-hubert.com | thierry-hubert.com | outings0 | 16:24:11 | $159.95 buy now Viagra (Sildenafil) 100mg x 90 pills |
| levelheadednessz7@schmeiler.de | schmeiler.de | levelheaded… | 16:20:30 | What to look for when purchasing a watch! |
| telexes12@promo2007.com | promo2007.com | telexes12 | 15:42:44 | Microsoft Office Enterprise 2007 OEM version |
| trainability@urbano.org.pe | urbano.org.pe | trainability | 15:14:12 | jugglery |
| schultze620@uni-berlin.de | uni-berlin.de | schultze620 | 11:53:05 | Doping fur Ih |
| hoffmann577@deplanet.ch | deplanet.ch | hoffmann577 | 11:51:05 | Hochwertige |

Selected Email

○ Approve Selected Email

○ Approve All Email from Sender and Whitelist Sender's Email Address

○ Approve All Email from Sender's Domain and Whitelist Sender's Domain

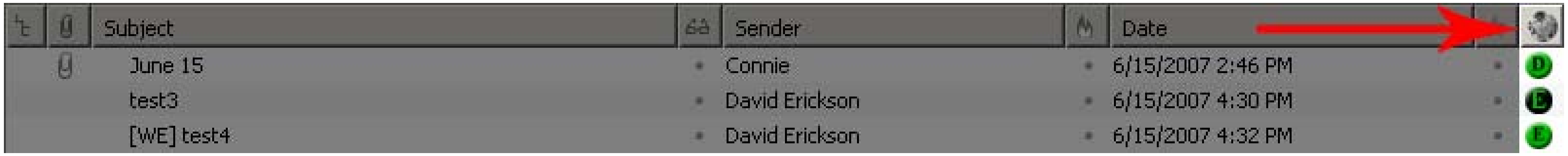
○ Delete Selected Email

○ Delete All Email from Sender and Blacklist Sender's Email Address

○ Delete All Email from Sender's Domain and Blacklist Sender's Domain

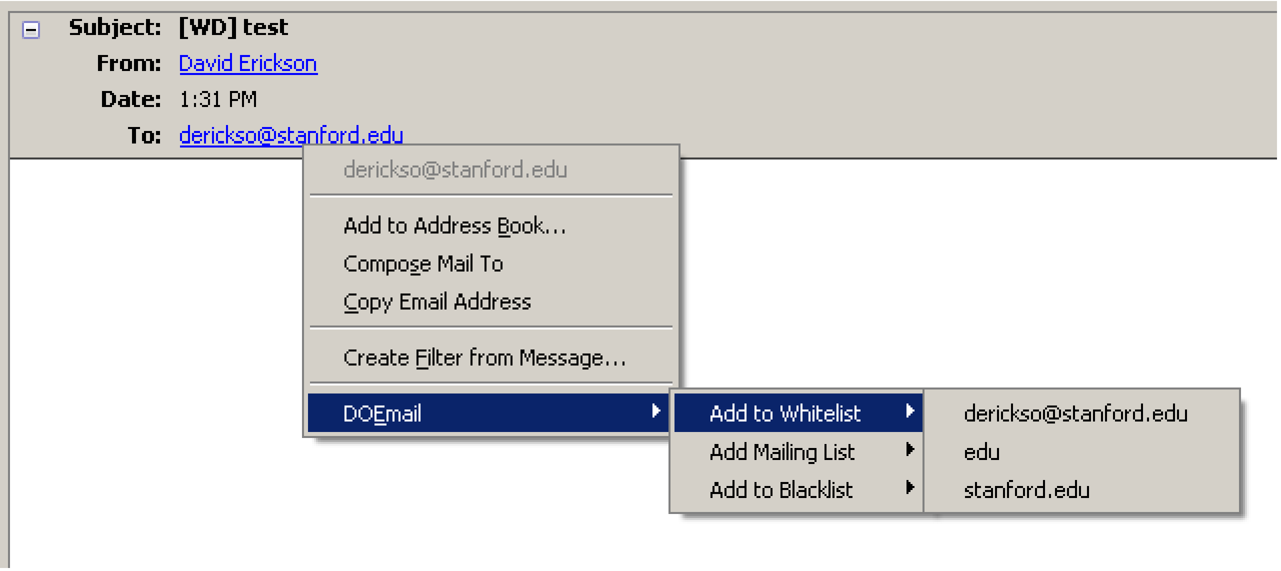
# Thunderbird Add-on

- View the type of rule the email matched



| Subject | Sender | Date | |
|---|---|---|---|
| June 15 | Connie | 6/15/2007 2:46 PM | D |
| test3 | David Erickson | 6/15/2007 4:30 PM | E |
| [WE] test4 | David Erickson | 6/15/2007 4:32 PM | E |

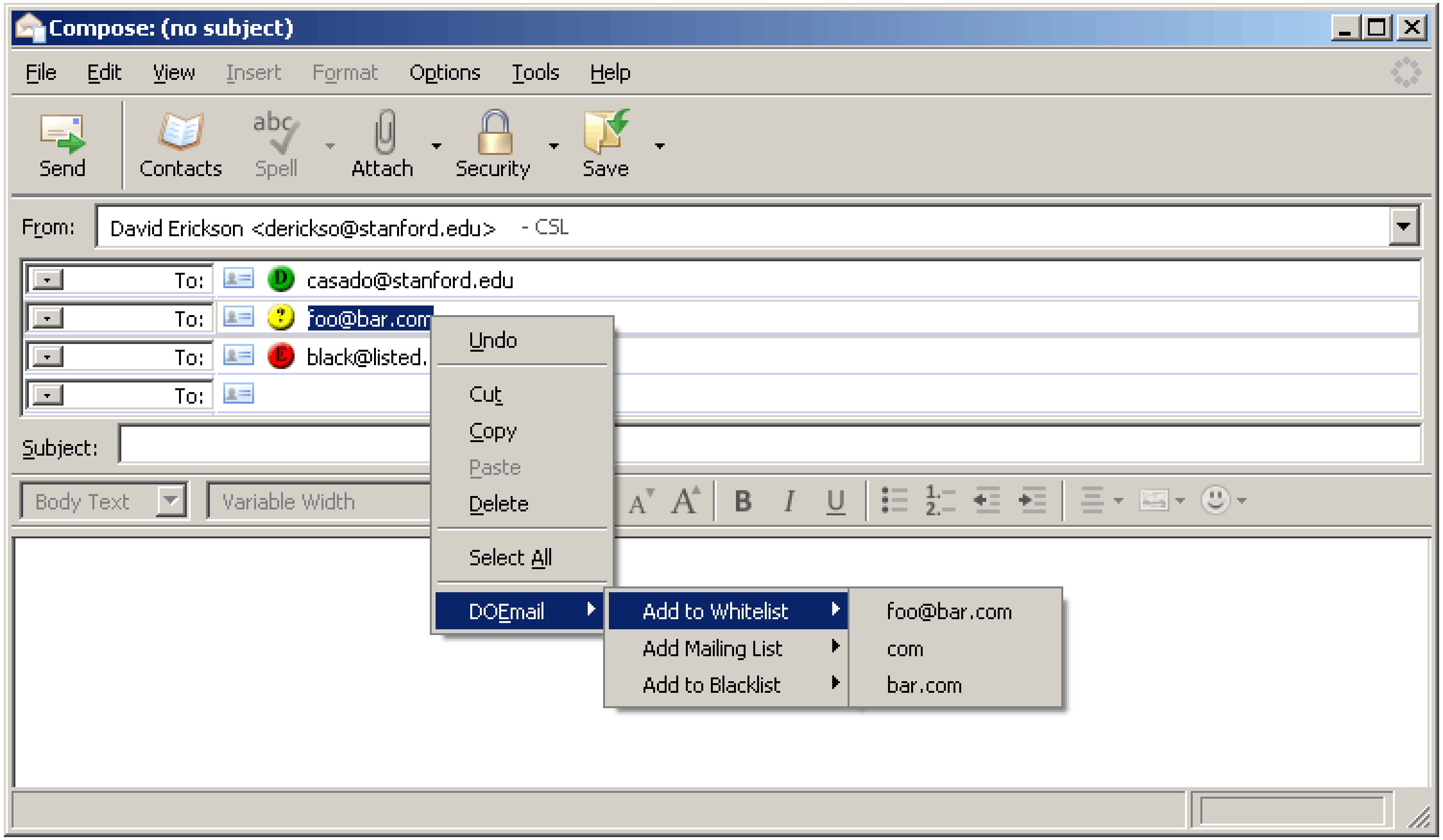
- Add/remove entries by right clicking addresses



**Subject:** **[WD] test**
**From:** David Erickson
**Date:** 1:31 PM
**To:** dericksо@stanford.edu

# Thunderbird Add-on

- See if recipients are on your lists, if not, add them!



Compose: (no subject)

From: David Erickson <dericksо@stanford.edu> - CSL

To: casado@stanford.edu

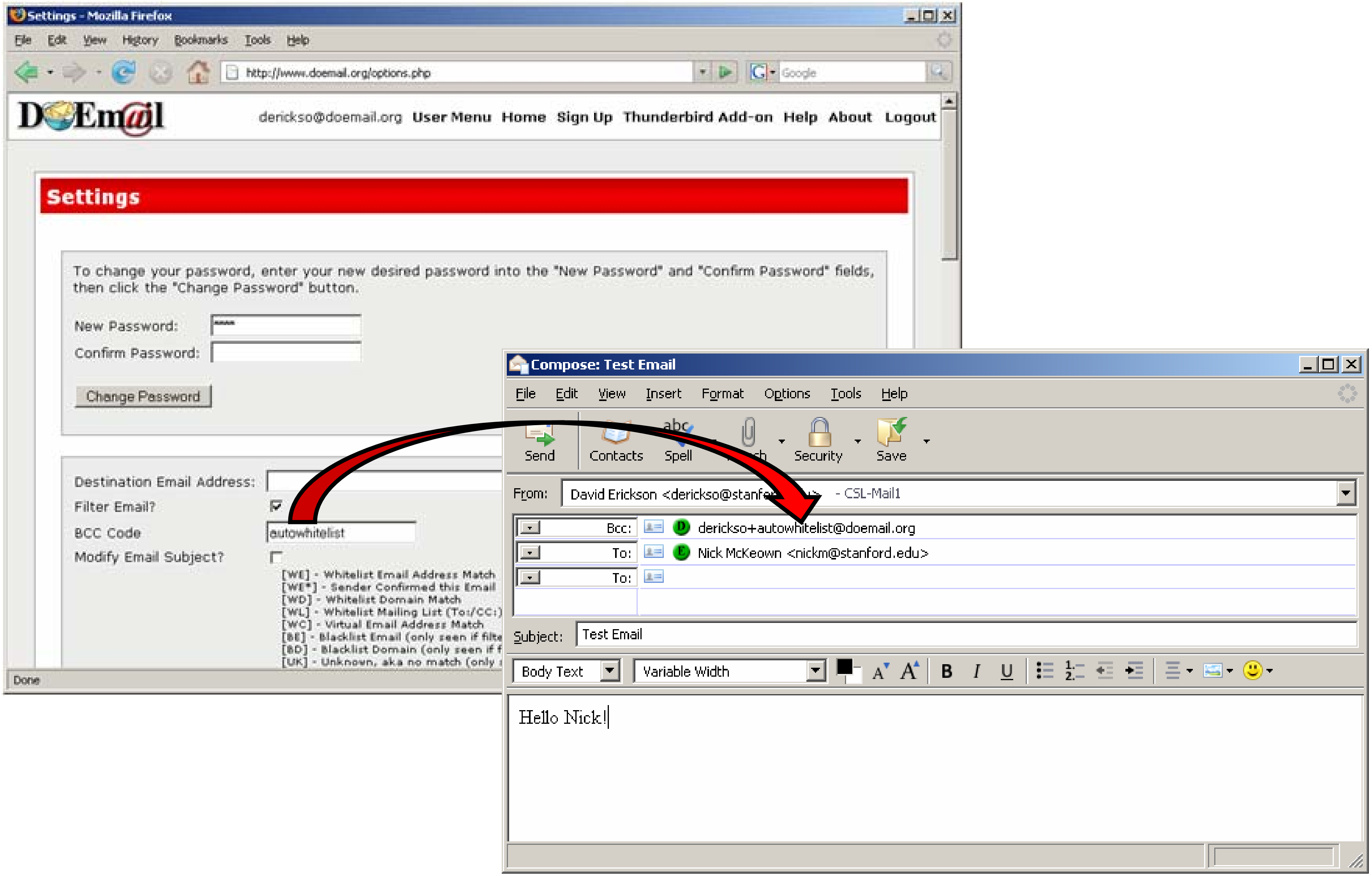
To: foo@bar.com

To: black@listed.

To:

Subject:

# Thunderbird Add-on


Settings - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://www.doemail.org/options.php
Google
DoEm@il
User Menu Home Sign Up Thunderbird Add-on Help About Logout
Change Password
Done
Compose: Test Email
File Edit View Insert Format Options Tools Help
Send Contacts Spell Attach Security Save
Body Text
Variable Width
B I U

derickso@doemail.org

## Settings

To change your password, enter your new desired password into the "New Password" and "Confirm Password" fields, then click the "Change Password" button.

New Password: ****

Confirm Password:

Destination Email Address:

Filter Email? ☑

BCC Code: autowhitelist

Modify Email Subject? ☐

[WE] - Whitelist Email Address Match
[WE*] - Sender Confirmed this Email
[WD] - Whitelist Domain Match
[WL] - Whitelist Mailing List (To:/CC:)
[WC] - Virtual Email Address Match
[BE] - Blacklist Email (only seen if filte
[BD] - Blacklist Domain (only seen if f
[UK] - Unknown, aka no match (only

From: David Erickson <derickso@stanf - CSL-Mail1

Bcc: derickso+autowhitelist@doemail.org

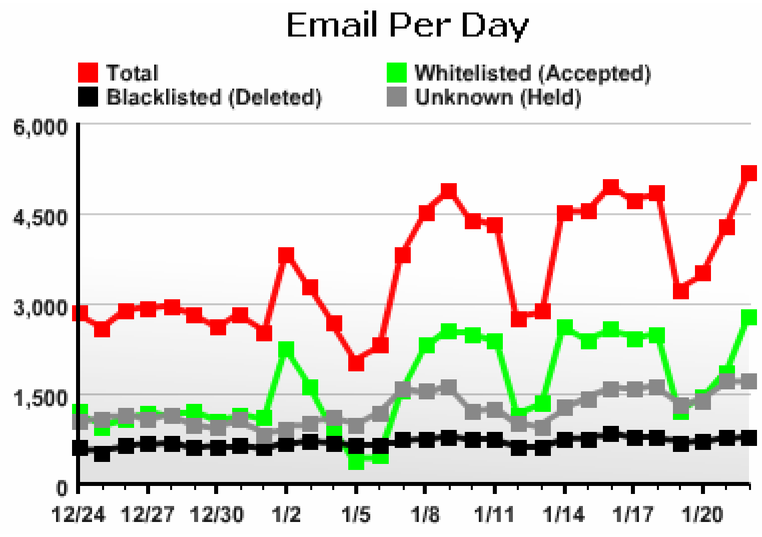
To: Nick McKeown <nickm@stanford.edu>

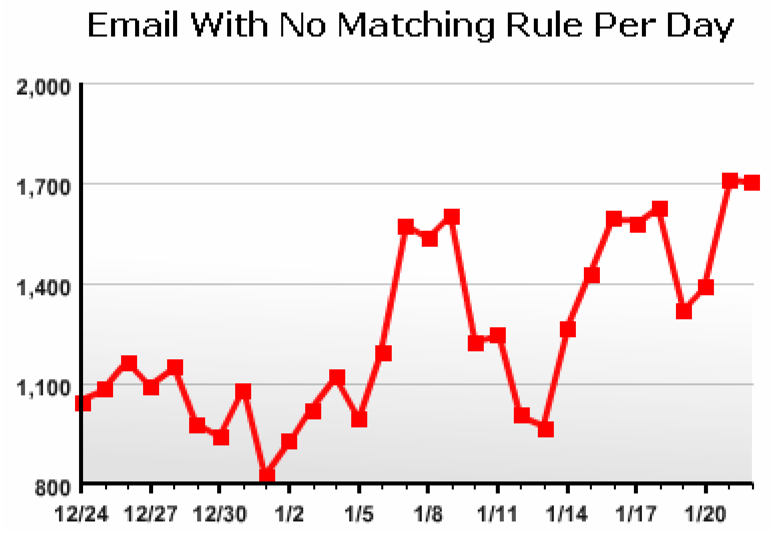
To:

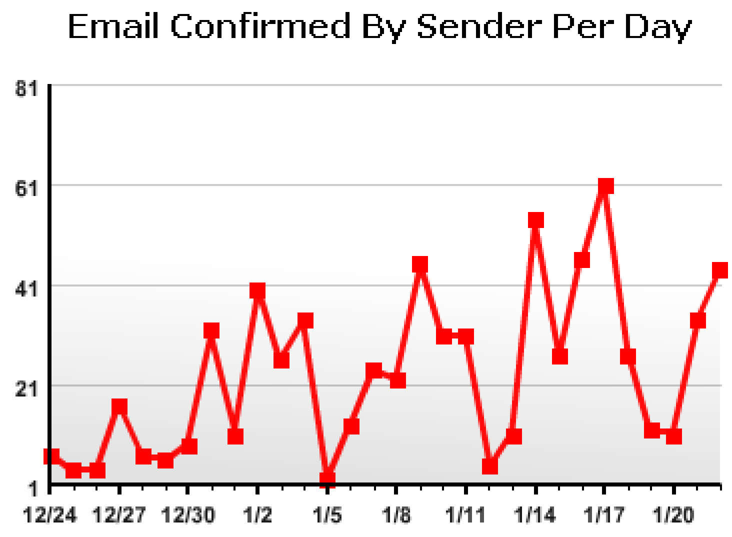
Subject: Test Email

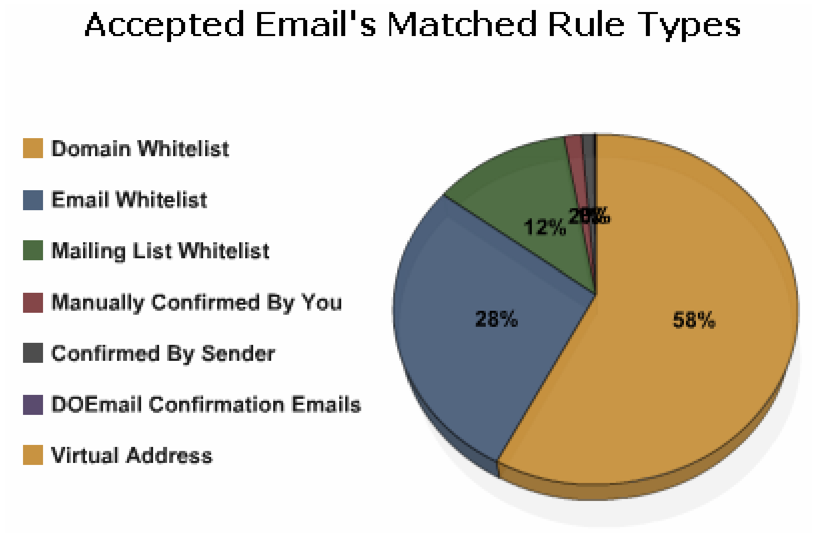
Hello Nick!

# Example Dynamic Graphs


Email Per Day

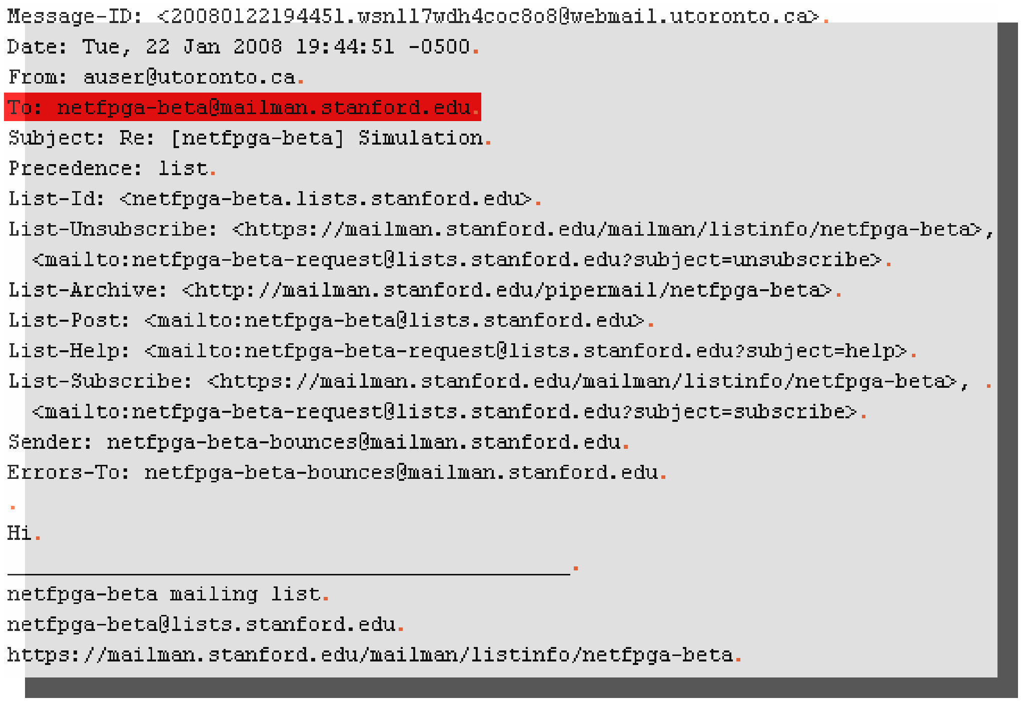

Email With No Matching Rule Per Day


Email Confirmed By Sender Per Day


Accepted Email's Matched Rule Types

# Lists

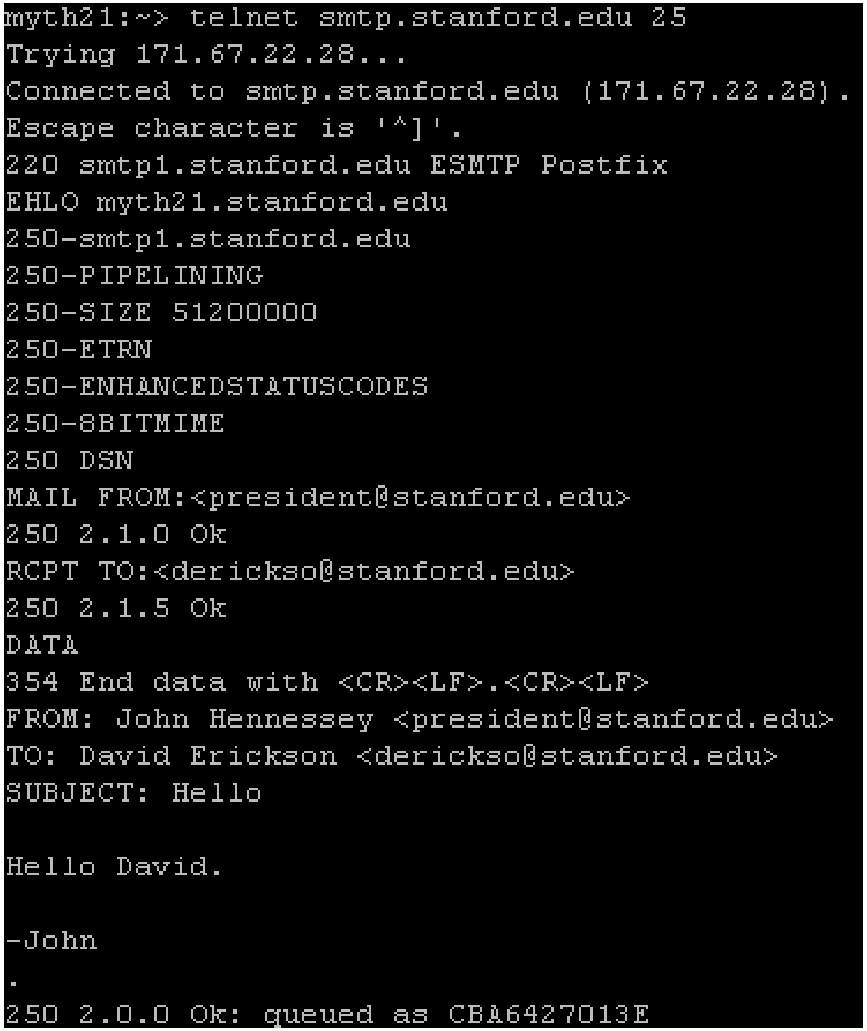
- To: / CC: Whitelist
- Auto Detection



```
Message-ID: <20080122194451.wsnl17wdh4coc8o8@webmail.utoronto.ca>.
Date: Tue, 22 Jan 2008 19:44:51 -0500.
From: auser@utoronto.ca.
To: netfpga-beta@mailman.stanford.edu.
Subject: Re: [netfpga-beta] Simulation.
Precedence: list.
List-Id: <netfpga-beta.lists.stanford.edu>.
List-Unsubscribe: <https://mailman.stanford.edu/mailman/listinfo/netfpga-beta>,
  <mailto:netfpga-beta-request@lists.stanford.edu?subject=unsubscribe>.
List-Archive: <http://mailman.stanford.edu/pipermail/netfpga-beta>.
List-Post: <mailto:netfpga-beta@lists.stanford.edu>.
List-Help: <mailto:netfpga-beta-request@lists.stanford.edu?subject=help>.
List-Subscribe: <https://mailman.stanford.edu/mailman/listinfo/netfpga-beta>, .
  <mailto:netfpga-beta-request@lists.stanford.edu?subject=subscribe>.
Sender: netfpga-beta-bounces@mailman.stanford.edu.
Errors-To: netfpga-beta-bounces@mailman.stanford.edu.
.
Hi.
_____________________________________________.
netfpga-beta mailing list.
netfpga-beta@lists.stanford.edu.
https://mailman.stanford.edu/mailman/listinfo/netfpga-beta.
```

# Limitations

- Backscatter
- Header spoofing
  - DomainKeys/DKIM
    - Hash/Sign Email
- Mailing list detection
  - Poor standardization
- Challenge Emails
  - Filtered

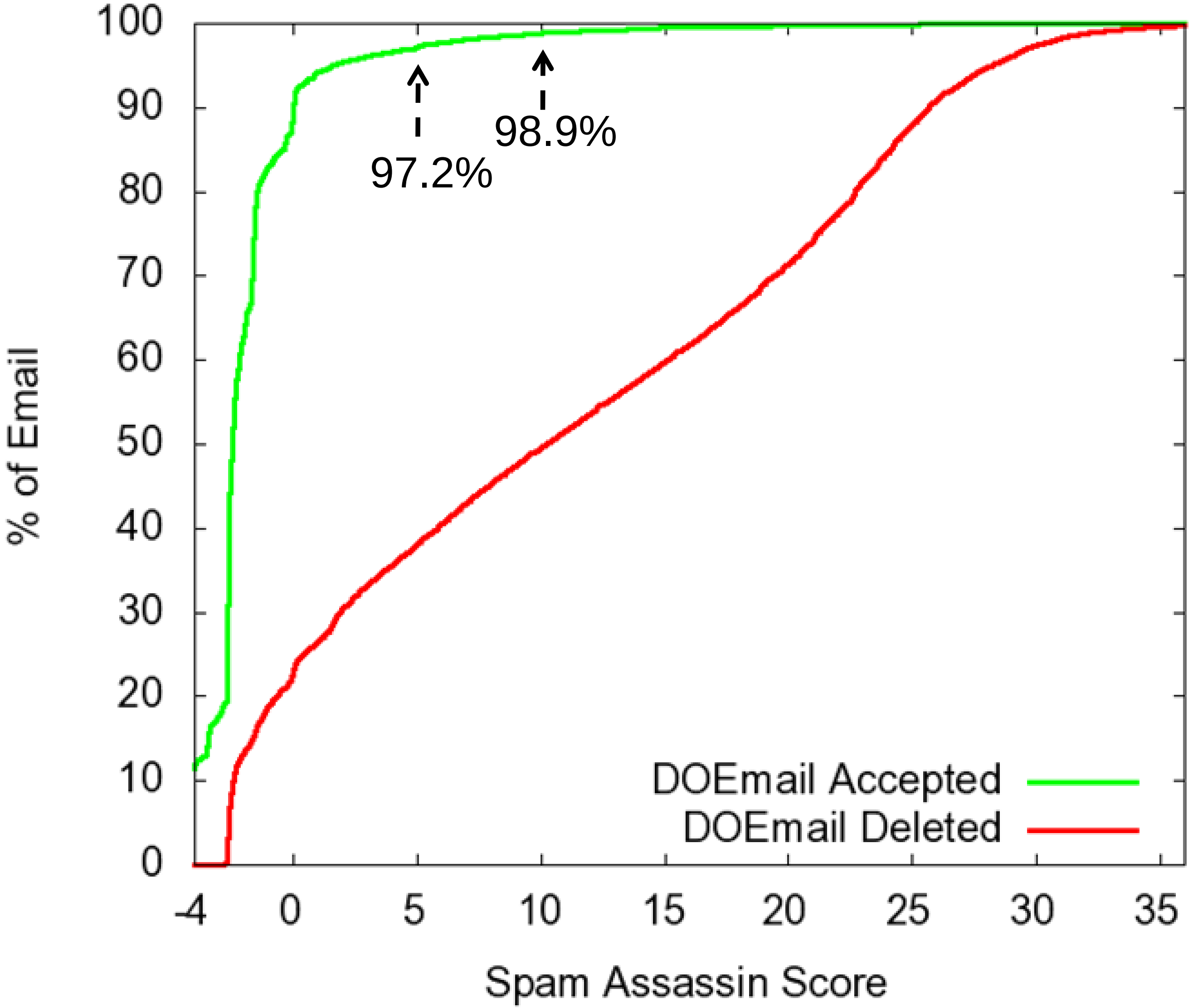


```
myth21:~> telnet smtp.stanford.edu 25
Trying 171.67.22.28...
Connected to smtp.stanford.edu (171.67.22.28).
Escape character is '^]'.
220 smtp1.stanford.edu ESMTP Postfix
EHLO myth21.stanford.edu
250-smtp1.stanford.edu
250-PIPELINING
250-SIZE 51200000
250-ETRN
250-ENHANCEDSTATUSCODES
250-8BITMIME
250 DSN
MAIL FROM:<president@stanford.edu>
250 2.1.0 Ok
RCPT TO:<dericksо@stanford.edu>
250 2.1.5 Ok
DATA
354 End data with <CR><LF>.<CR><LF>
FROM: John Hennessey <president@stanford.edu>
TO: David Erickson <derickso@stanford.edu>
SUBJECT: Hello

Hello David.

-John
.
250 2.0.0 Ok: queued as CBA6427013E
```

# Results

- Measured from 7/13/07 – 2/29/08
- 112 user accounts received email
- 592,794 emails processed
- Two main questions:
  - What are DOEmail’s detection rates?
    - Compare with Spam Assassin
  - How much effort is required?
    - Track user behavior

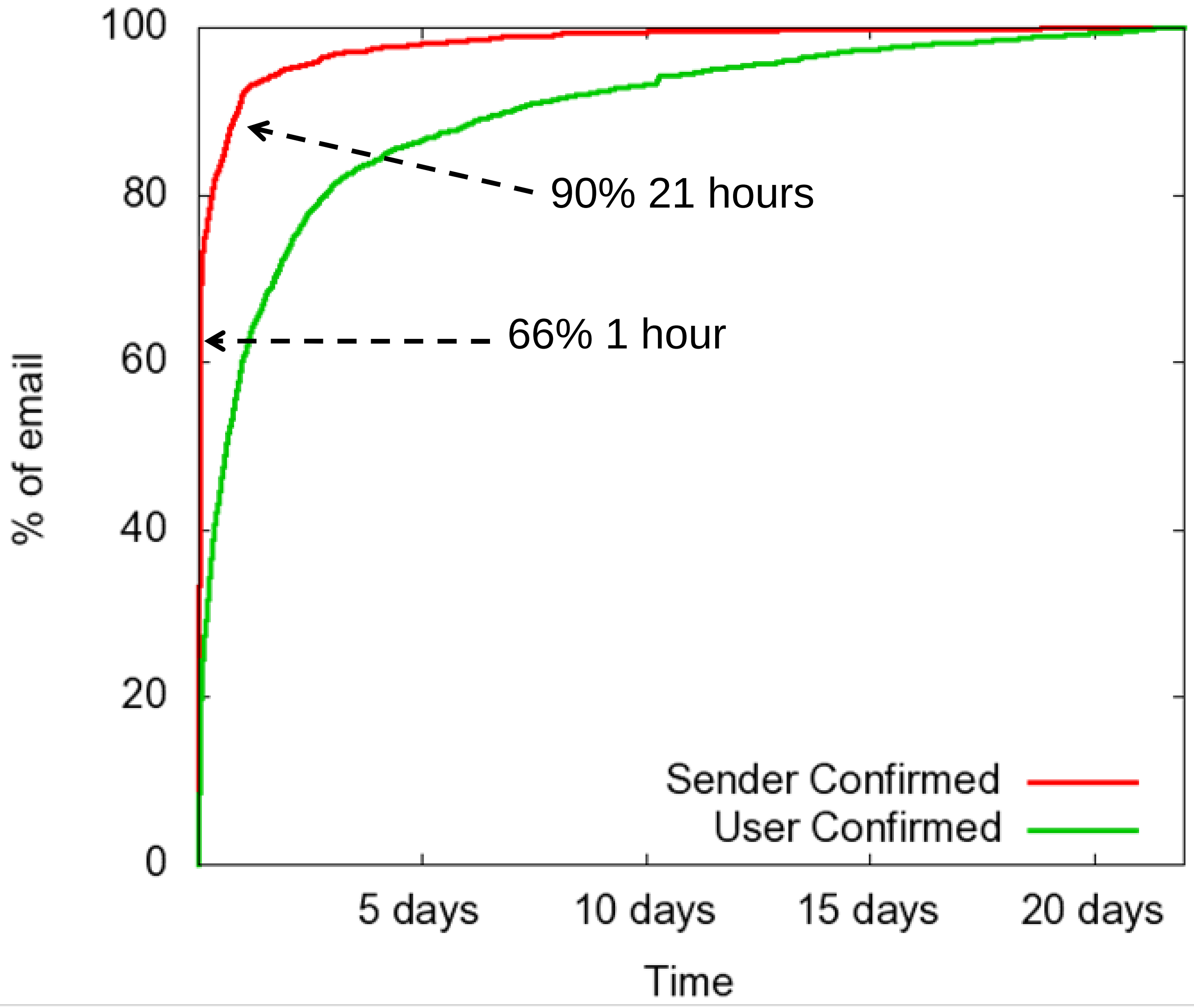
# Spam Assassin Comparison CDF



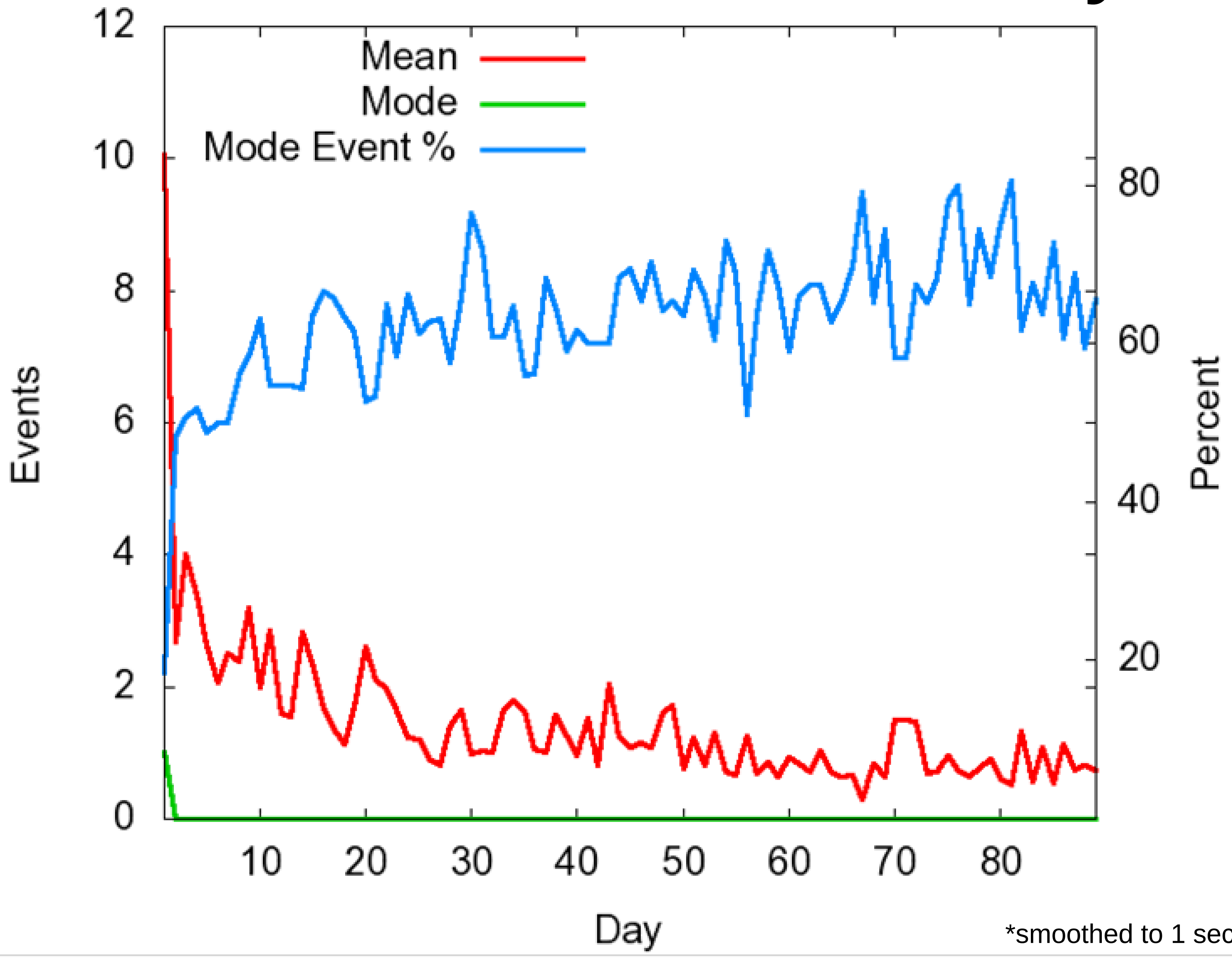
# Pending Email

- 9180 (1.55%) pending emails confirmed
  - 4382 (0.74%) by sender
  - 4798 (0.81%) by user (False Positive Rate)
    - 3864 (0.65%) sent challenges
    - 934 (0.16%) not sent challenges
- 58+% sender confirmation rate

# Pending Email Delay



# User Events First 90 days



*smoothed to 1 sec granularity

# Conclusions

- Whitelisting enables powerful filtering
  - Can achieve high degrees of accuracy
    - Based on user's rule preferences
  - Low rate of false positives
  - Content filtering limitations
    - Fundamental tradeoff between FPs and FNs
- Negligible email delay
  - Applies only to first email from a new sender
- Low user overhead