Ethical Dilemmas in Take-down Research

Tyler Moore
tmoore@seas.harvard.edu

joint work with

Richard Clayton
richard.clayton@cl.cam.ac.uk

WELLESLEY

WECSR 2011
1st March 2011

HARVARD
School of Engineering and Applied Sciences

Phishing research

- Phishing is the theft of credentials by the use of fake websites
  - though "phish by email reply" is also relevant these days
- We have a series of papers measuring this from 2007 onwards
  - all are branded "T. Moore & R. Clayton"
- So we are studying the actions of criminals – but we mainly studied (and carefully measured) the main countermeasure which is the "take-down" (the removal of) the fake websites
- This "take-down" was first done by the banks themselves, but various "brand-protection" companies now do most of the work
- In today's talk, I will describe nine ethical issues that we have run up against during the course of our research. It's "war stories" rather than philosophy!

Dilemma 1: Should researchers notify affected parties in order to expedite take-down?

- We were measuring take-down and didn't want to interfere
  - found log-normal distributions (long tails) and that lack of information sharing was damaging effectiveness
- Who could we tell anyway?
  - no organized way to report data to banks
- Our NDAs forbade this!
  - take-down companies make money by selling data feeds
- c.f. clinical trials: These trials can and should be stopped prematurely once the results become statistically significant and the divergence in treatment outcome is substantial
- We recommend that researchers avoid direct interference during data collection, but once the conclusions have been drawn, assistance to relevant stakeholders should be encouraged.

Dilemma 2: Should researchers intervene to assist victims?

- We reported on 414 compromised users whose details we recovered from phishing sites (and have found more since)
- We repatriated these to banks where we could (and there is since 2010 a formal scheme for this run by the NCFTA)
- Common issue: Torpig takeover, 180000 infections, 70G data, 1 million Windows passwords, 100000 SMTP logins, 12000 FTP credentials. Disrupted research for 6+ months
- But may be hard to locate victims, BBC's "Click" changed "wallpaper" on botted machines; apparently without having considered that this is clearly a s1 offence under the UK's Computer Misuse Act 1990.

Rod Rasmussen (Internet Identity)

The normal admin for the machine had been deployed to Iraq as part of his National Guard unit, and his backup was busy and hundreds of miles away that weekend because of his father's funeral. There were plenty of people looking at the machine (as in had their physical eyeballs on it) including the local sheriff, but no one was touching it since it ran the 911 Dispatch system and no one had the knowledge (as in passwords and expertise) to fix it.

We've also had take-downs on machines that were in hospitals, railroad stations, airports, and government facilities. While those could be just public access terminals, there's no way we can tell from the outside if that is the case or they are running life-saving equipment, switching operations, air-traffic control systems, or have sensitive data on them respectively. That's why we have a very bright line barring any sort of “write access”, resetting or otherwise monkeying with content on compromised servers. Not only is it usually illegal in the US, someone's life can literally be on the line!

Dilemma 3: Should researchers fabricate plausible content to conduct "pure" experiments of take-down?

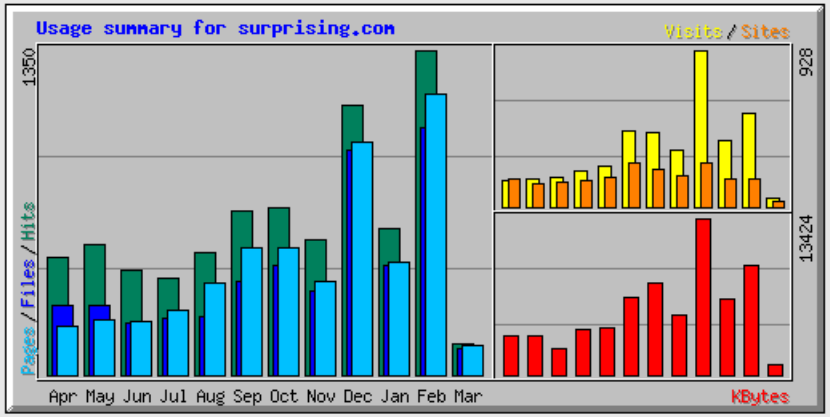
- Most empirical research in computer security is "observational"
- Some attempts at experiments, eg copyright issues, but considerable flaws since didn't investigate whole process (especially US DMCA "put-back")
- Risk of wasting the time and energy of frontline responders on fabricated requests suggests real harm is caused by the experiments. In particular, the responders typically have substantial resource constraints and already find it difficult to keep up with the number of legitimate take-down requests
- We believe the fabrication of reports to study take-down is usually unethical

Dilemma 4: Should researchers collect world-readable data from "private" locations?

www.surprising.com/webalizer/

Usage Statistics for surprising.com

Summary Period: Last 12 Months
Generated 03-Mar-2011 02:11 EST



Summary by Month

Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Mar 2011	66	54	63	28	30	932	56	126	109	133
Feb 2011	48	36	41	19	165	9411	558	1166	1027	1350
Jan 2011	20	15	15	13	163	6491	393	468	454	607
Dec 2010	36	30	31	29	262	13424	928	966	936	1121
Nov 2010	18	11	12	11	180	5202	333	389	352	561
Oct 2010	22	14	17	14	225	7891	439	527	453	694
Sep 2010	22	12	17	14	257	6673	446	527	389	681
Aug 2010	16	7	12	7	174	4013	242	382	245	512
Jul 2010	13	7	8	7	160	3858	217	272	235	405
Jun 2010	14	7	7	5	148	2203	174	226	218	434
May 2010	17	9	7	5	137	3343	162	228	292	542

Dilemma 4: Should researchers collect world-readable data from "private" locations?

- We used "Webalizer" data from compromised websites to study the number of victims and how sites were located by criminals
- Owners of these sites may not have intended to publish their visitor data – but they did; so there's an ethics question
- Our view was that the data enabled us to answer questions that we could not have done otherwise; and it was not personally identifiable data, just summary data for website visits
- On balance, we feel the opportunities for scientific advancement outweigh the risks to an individual website operator in collecting the data. However, it is a judgment call, and one that should be weighed on a case-by-case basis.

Dilemma 5: What if our analysis will assist criminals?

- This is always an issue; the issue is sometimes described as the suitability of "full disclosure"
- We observed that fast-flux systems used multiple servers but this is unnecessarily cautious.
- We also made some observations about DNS time-to-live values
- We tried not to emphasize these issues and the criminals have not changed how they operate
- Long history of ethical analysis of this issue, for example Hobbs (1853) re locksmiths, and Wilkins (1641) re crypto
- General view is that the benefits of explaining how the criminals operate benefits the good guys far more than the bad guys, who already know how to do their crimes and already understand what works and what doesn't

Dilemma 6: Should investigatory techniques be revealed?

NETCRAFT

Rackspace Managed Hosting - Web Hosting - Hosting

What's that site running?

netcraft.com

Subscribe to our RSS feed
Get News updates by email
Follow Us on Twitter

Netcraft Services

Phishing & Security

- Anti-Phishing Toolbar
- Phishing Site Feed
- Hosting Phishing Alerts
- Bank Fraud Detection
- Phishing Site Countermeasures
- Audited by Netcraft
- Open Redirect Detection
- Web Application Security Testing
- Web Application Security Course

Internet Data Mining

- Hosting Provider Analysis
- Million Busiest Websites
- Busiest Sites Switching Analysis
- Hosting Provider Switching Analysis
- Hosting Provider Server Count
- Hosting Reseller Survey
- SSL Survey

Internet Exploration

- Whats that site running?

Mr-Brain: Stealing Phish from Fraudsters

A recurrent group of Moroccan fraudsters calling themselves Mr-Brain has launched a website dedicated to offering easy-to-use phishing site code, email templates and other hacking tools. The website offers phishing kits for many of the most common targets, such as Bank of America, eBay, PayPal and HSBC.

The tools and code provided by Mr-Brain are designed to make it extremely easy for other fraudsters to deploy realistic phishing sites. Only a very basic knowledge of programming is required to configure the PHP scripts to send victims' details to the fraudsters' chosen electronic mail address. Deploying one of these fully working kits can be done in as little as one minute – another factor that adds to their appeal.

Tricking the Fraudsters

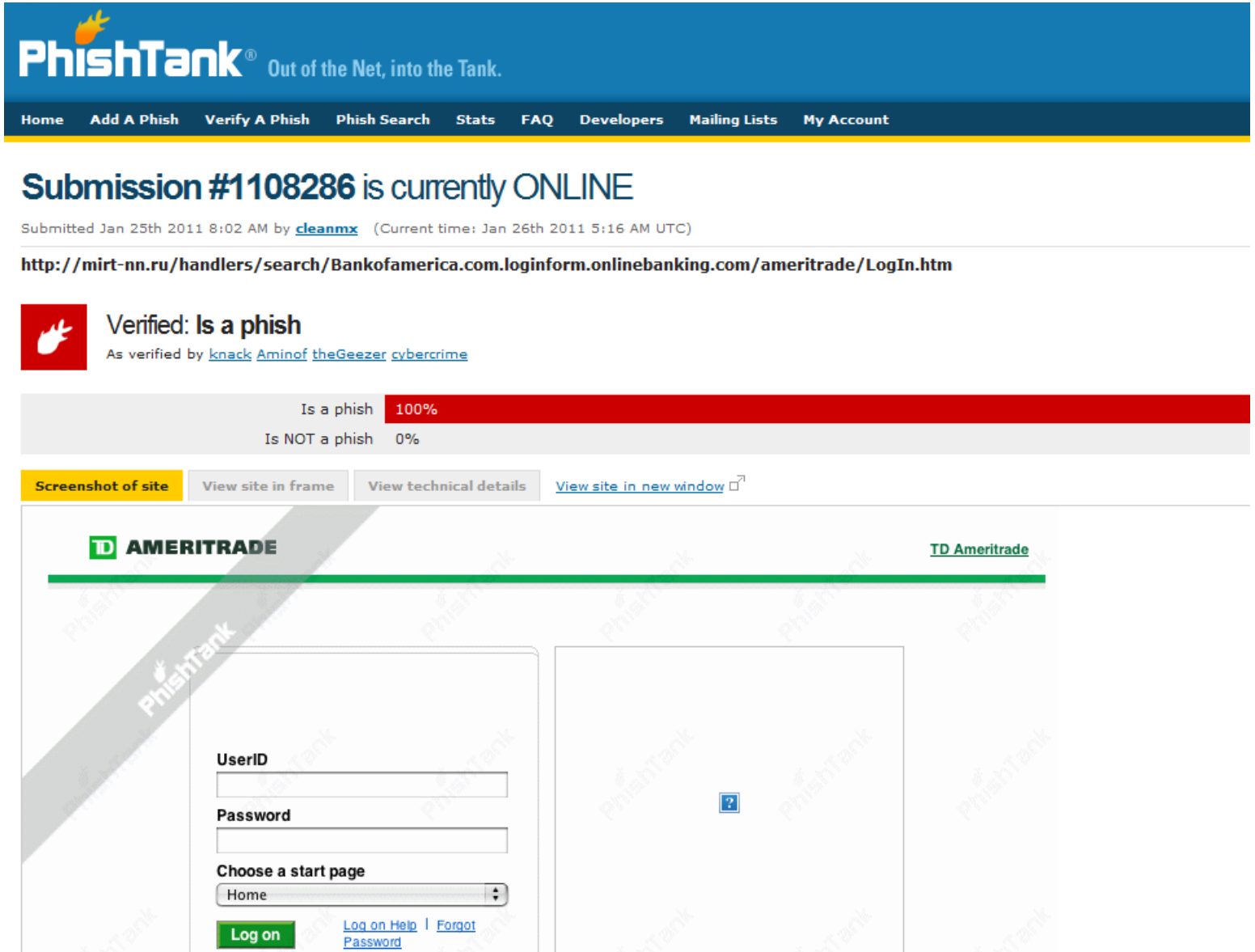
Mr-Brain's intentions are to encourage as many people as possible to use their phishing kits, for all is not what it seems at first glance. Careful inspection of the configuration script reveals deceptive code that hides the true set of electronic mail addresses that are contacted by the kit – every fraudster who uses these kits will unwittingly send a copy of each victim's details back to the Mr-Brain group.

The configuration script exploits the case-sensitivity in PHP variable names to disguise Mr-Brain's electronic mail address as an unrelated but seemingly essential part of the script, encouraging fraudsters not to alter it. The injected electronic mail address is actually contained in a completely separate PHP file, where it is

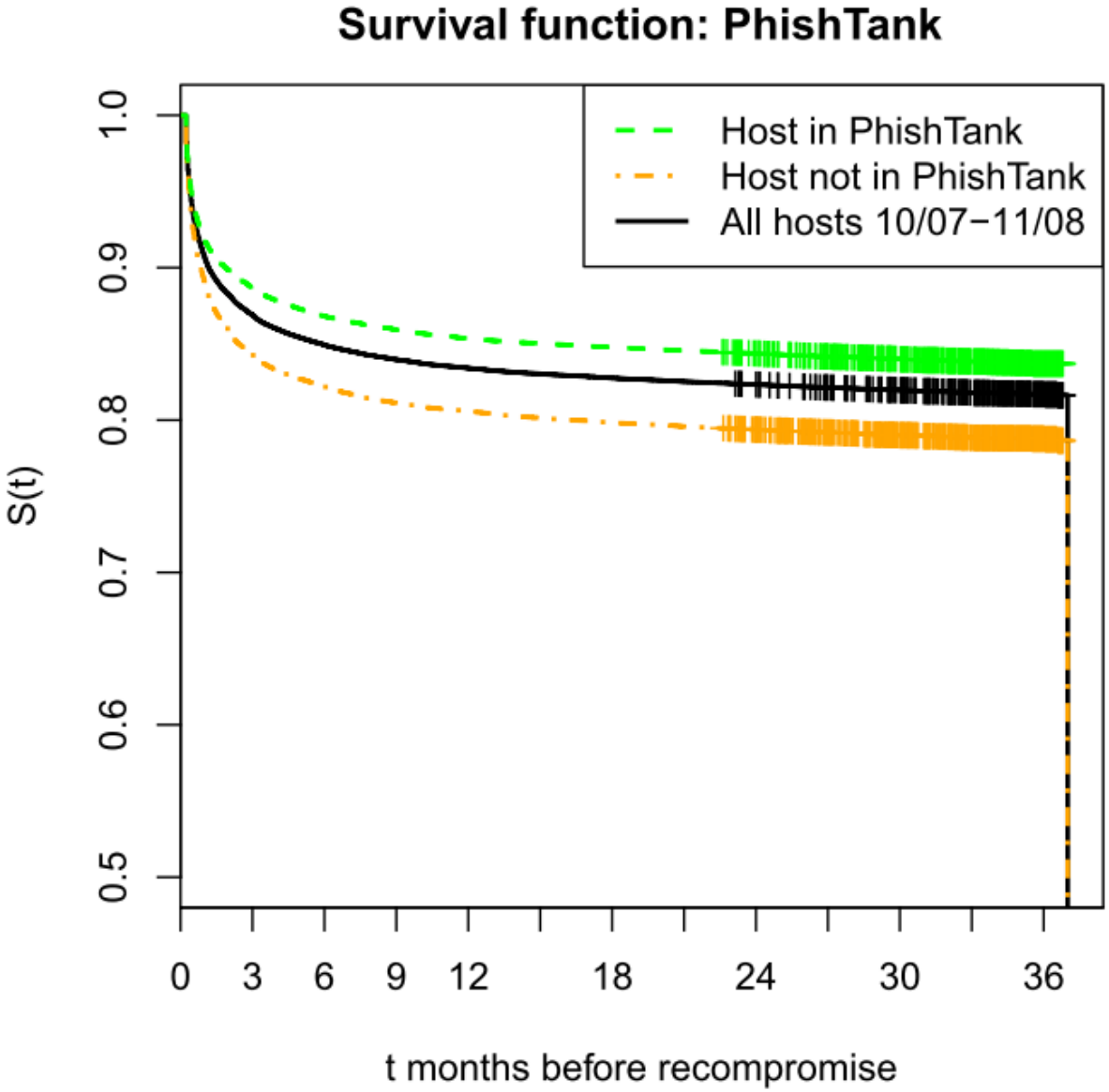
Dilemma 6: Should investigatory techniques be revealed?

- It’s not science if an academic paper does not explain the methodology of an investigation
- The main effect of this has been for us to suppress more minor bits of research
- Others take a different line, such as Netcraft explaining about phishing kit “back doors” and Billy Rios discussing a file injection vulnerability in Zeus
- We don’t agree that “full disclosure” is always the overriding principle and so we choose not to publish when the details of our paper would disrupt investigations by the authorities

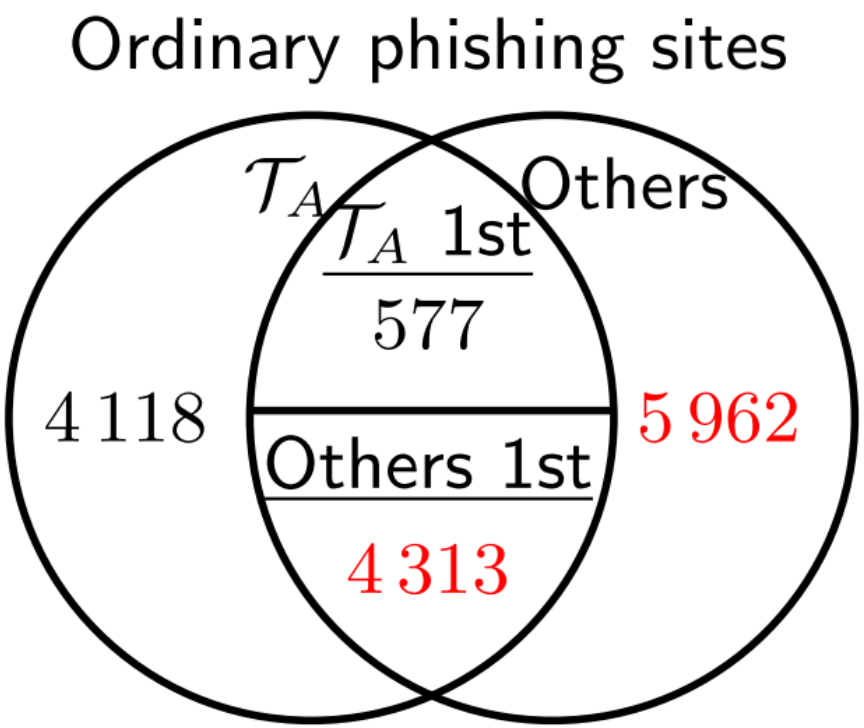
Dilemma 7: When should datasets be made public or kept secret?

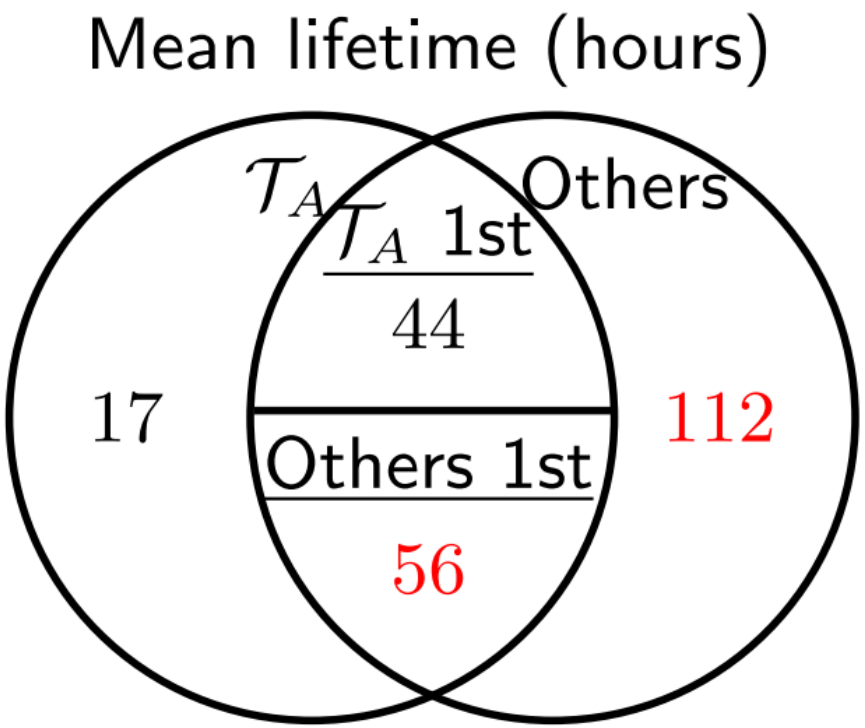


Websites in PhishTank's public list are less likely to be recompromised



Sharing data between take-down companies would reduce phishing website lifetimes





Dilemma 7: When should datasets be made public or kept secret?

- There are downsides to public sharing: it shows what the defenders know, it "names & shames", and criminals can collect caches of credentials (as we did)
- But there are less obvious downsides to *not* sharing
  - Reduced effectiveness of countermeasures
  - Many systems hash the URLs so they can be compared, but the actual values are not revealed – this prevents research and prevents defenders from being proactive.
- The ethical choice is to minimize overall harm, which sometimes means publishing and other times not

Dilemma 8: Is the fix realistic, and does it consider the incentives of all the participants?

- We believe in security economics, and so one of the ways we look to solve security problems is to align incentives
- However, when we proposed list sharing we did not take full account of the incentives of the take-down companies; though we fixed this after they complained of our naivety
- We have proposed improvements in the take-down of child sexual abuse, unfortunately they are not realistic whilst INHOPE prevents cross-border notifications. We still think this would be the right thing to do.
- It is unethical to propose fixes to security problems that cannot be made to work in the real world

Dilemma 9: What if the fix is worse than the problem?

- Restrictions on registering domain names would help the problem of misleading names but would not be proportionate, which is why we never suggest that type of solution
- Paul Vixie has proposed an efficient way (RPZ) of publishing lists of domains that are to be suppressed by Domain Name Servers (similar to the way that spam senders are currently handled). However, it seems unwise to institutionalize this type of suppression in a world where many groups see removal of domain names as a way to impose their world view
- We think that it is unethical to propose fixes without considering their impact, and seeking to minimize the side-effects

Conclusion

http://people.seas.harvard.edu/~tmoore/wecsr11.pdf

- We're not proposing ethical principles but telling "war stories"
- But ethicists might usefully take our stories into account
- Also, if you're going to work in this area it would be worthwhile to learn from our mistakes!