IRBs and Security Research: Myths, Facts and Mission Creep

Simson L. Garfinkel
- Center for Research on Computation and Society
- Naval Postgraduate School
Since the late 1990s, security researchers have increasingly focused on "the weakest link."

As computers became more connected, they became less secure. This, despite:

- Revolution in cryptography (RSA & faster CPUs)
- Java (no buffer overflows)
- 20+ years of secure operating system research.

Why? Most operational security problems result from *human factors*:

- User error (failure to use cryptography; improper use)
- Configuration error
- Programmer error
- Specification error
- Poorly understood problem
Human factors dominate today's security landscape.

Phishing, Wireless Security, Sanitization Failures

If we want to make real improvements, we need to see where and why people are making errors, and then either:

- train the people so they don't make errors
- fix the software so that training is not required.

We can't do this without working with human subjects or data from humans.

This brings us under Federal Regulations and the IRB structure.
Why do we have IRBs?  
(Institutional Review Boards)

A lot of scientists did a lot of bad things in the 1960s.

- "Tuskegee Study of Untreated Syphilis in the Negro Male" (1932-1972)
- Stanley Milgram shock psychology experiments (1961; 1974)
- Timothy Leary LSD experiments at Harvard (1961)
- Stanford Prison Experiment (1971)

Results:

- National Research Act (PL 93-348) signed into law July 12, 1974
- The Belmont Report ("Ethical Principles and Guidelines for the Protection of Human Subjects of Research," April 18, 1979)
1979: The Belmont Report's key findings

1. Respect for Persons
   • "Individuals should be treated as autonomous agents"
   • "Persons with diminished autonomy are entitled to protection."

2. Beneficence
   • "Persons are treated in an ethical manner not only by respecting their decisions and protecting them from harm, but also by making efforts to secure their well-being"
   • "Do not harm"
   • "Maximize possible benefits and minimize possible harms."

3. Justice
   • Fairness in distribution of the results of the research.

http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.htm

Originally adopted by HHS to govern use of humans in research. Adopted by other federal agencies in 1991 (EPA's is 40 CFR 26)

Applies to:

Agency for International Development
Consumer Product Safety Commission
Department of Agriculture
Department of Commerce
Department of Defense
Department of Education
Department of Energy
Department of Health and Human Services
Centers for Disease Control
Food and Drug Administration
National Institutes of Health
Department of Housing and Urban Development
Department of Justice
Department of Veterans Affairs
Department of Transportation
Environmental Protection Agency
National Aeronautics and Space Administration
National Science Foundation

In addition, the Central Intelligence Agency and Social Security Administration are required by Executive Order and statute, respectively, to follow the DHHS regulations (including all subparts).
45 CFR 46 has very broad definitions for "Research" and "Human Subjects"

Research:

• "systematic investigation including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge."

• "whether or not they are conducted or supported under a program which is considered research for other purposes."

Human subject:

• "A living individual about whom an investigator (whether a professional or student) conducting research obtains"

  (1) Data through intervention or interaction with the individual, or

  (2) Identifiable private information
Enforcement through the Institutional Review Board.

Each organization receiving Federal research funds must designate an Institutional Review Board (IRB)

At least five members:

- "Varying backgrounds" to promote complete and adequate review of research activities commonly conducted by the institution."
- "Diversity:" race, gender, cultural backgrounds
- Assures compliance with "institutional commitments and regulations, applicable law, and standards of professional conduct and practice"
- Both genders
- At least one scientist
- At least one person not otherwise affiliated with the institution
The IRB has very broad powers.

- IRB approval is required before work involving human subjects can commence.
- IRB decides if application can be "expedited."

The IRB has no jurisdiction over research that is exempt or not federally funded:
- Research on educational practices or with educational tests
- Research involving "existing data, documents, [and] records" (provided data is "publicly available" or subjects "cannot be identified".)
- Research involving surveys or interviews, unless results could identify the humans and place subjects at risk of "criminal or civil liability."

Nevertheless, most organizations require that all work involving human subjects go through IRB review.
What is "IRB Approval"?

IRBs have several ways of "approving" research.

The IRB can:

- **EXEMPT** — Declare research does not require IRB approval.
- **EXPEDITE** — Approve as "minimal risk" without a review by the full IRB.
- **APPROVE WITH FULL REVIEW**

From the point of view of a Computer Security researcher, all of these require:

- Notifying the IRB
- Submitting *something* (email, application, etc)
- Getting a response.

Even "EXEMPT" research require some kind of involvement and approval.
Myth or Fact?

Because the Common Rule exempts research involving subjects that cannot be identified, IRB approval is not required when using anonymized data.
Myth or Fact?

Because the Common Rule exempts research involving subjects that cannot be identified, IRB approval is not required when using anonymized data.

Myth
This would be convenient, but most institutions require the determination to be made by the IRB.
Myth or Fact?

"Pilot studies" do not require IRB approval.
Myth or Fact?

"Pilot studies" do not require IRB approval.

**Myth**
The common rule makes no reference to "pilot" or "preliminary" studies.

Most policies I reviewed have require IRB approval for all research.
Myth or Fact?

IRB approval is not required if you are working with data that you already have.
Myth or Fact?

IRB approval is not required if you are working with data that you already have.

**Myth**
IRB approval is for a specific experimental protocol.

Minor changes in protocol may be granted "expedited" review.
Myth or Fact?

IRB approval is not required when using publicly available data.
Myth or Fact?

IRB approval is not required when using publicly available data.

Fact!
The Common Rule exempts research with "publicly available" records.
Myth or Fact?

IRB approval is not required when using publicly available data.

Fact!
The Common Rule exempts research with "publicly available" records.

But most institutions (Harvard, NPS, UC) still require IRB review!
What does IRB approval require?

Administrative overhead for the application:
- What is the protocol?
- What human subjects are involved?

Respect for the human subjects:
- Will the subjects be informed? If so, how? If not, why not?
- What specifically will the subjects be told?
- How will their information be protected?

Social Justice:
- How are the subjects recruited?
- Who will benefit from the research?
For many computer [security] researchers, IRB regulations are an unexpected complication.

Much of today's research involves use of computers by people:
  - User interface work.
  - Applications (email, web)
  - Operating systems (file systems)
  - Programming languages

Much of the data on computers was generated by people:
  - email messages
  - Program samples

A surprising number of experiments that you can imagine doing with data you already have is probably covered by IRB regulations.
Scenario 1: Security toolbar with anonymized summary statistics.

Alice has developed an anti-phishing toolbar. To assist in development and research, the toolbar sends a small anonymized report to the experimenter once a day. Because each toolbar reports only once every 24 hours, it is easy for the experimenter to measure adoption and use of the toolbar.
Alice needs IRB approval

Alice is:

- Recruiting subjects.
- Interacting with her subjects.
- Collecting information from her subjects.

Furthermore:

- Alice's users reveal their IP address when the toolbar reports its statistics.
- IP addresses do not necessarily reveal personal information, but they frequently do.
- The European Union considers IP addresses to be PII.
- At Harvard, IP addresses are frequently assigned to a specific person.
Scenario 2: Web server logfile analysis.

Bob’s research group operates a popular web-based discussion forum. Bob:

- Analyzes the server’s log file to report # of password resets each day.
- Records # of new passwords that do not pass password quality rules.
- Web server does not collect IP address.

Research question: how do restrictive rules affect password resets?

At least 3 letters and 2 numbers
UPPERCASE and lowercase
2 digits and 3 symbols (*)&^%$#)
big letters and small letters
At least 3 different colors
Bob needs IRB approval.

Bob is not collecting IP addresses!
But Bob needs IRB approval because the information in the webserver logs was generated by human subjects and is not publicly available.
Scenario 3: Popular security search terms.

Christine is a graduate student who also writes articles for a major security-related website.

Christine is working on a project that correlates search terms on the website with news stories.

The security-related website prepares a report which shows, for each hour, the number of times each term is searched.

The report is sent as a PGP-encrypted file to Christine’s Gmail account.
Christine needs IRB approval!

The data is generated by human beings and is not publicly available.

Christine could avoid IRB involvement if:

- The website published the search results on a public web page (rather than protecting the information and controlling its release.)

  and

- If Christine is at an institution that does not require IRB approval for exempt work.

Alternatively, Christine can avoid the IRB if she does the study for the website without using Federal research funds.
Scenario 4: Building better spam filters.

Don is creating a better spam filter. He wants to test it on his inbox.
Don needs IRB approval!

Common Rule does not exempt information already in Don's possession.

**RESPECT:** The people who sent mail to Don did not consent for their email to be used in the experiment.

**PROTECTION OF SUBJECTS:**

- Who sent Don email? Why?
- Is there confidential information in Don's inbox?
- What procedures did Don follow to make sure that there will be **minimal risk** for the people who sent him mail?
Scenario 5: Wi-Fi Security Survey

Elaine installs NetStumber on a laptop and drives around the neighborhood with a GPS.

Elaine compares names & locations of Wi-Fi sites she finds with an online database.

Research results:

- Older Wi-Fi access points that were open are now closed or have been removed from service.
- Newer Wi-Fi access points are closed.
Elaine might not require IRB approval, but she might.

Elaine is not observing people, she is observing APs
- But they were configured by people.
- The names might be identifiable.
- The GPS coordinates might be identifiable.
Frank downloads 100,000 Microsoft Word files from public websites. 15% of the files contain significant amounts of hidden information. Guy randomly chooses 100 of the 15,000 files and confirms the findings.
Frank doesn't need IRB approval under the Common Rule!

The research is exempt!

- The word files are "existing ... documents" that are "publicly available."
- Even though the information may be confidential!
- Even though the information may be leaked without the author's knowledge!

(Of course, if Frank is at Harvard he still needs IRB approval.)
Scenario #7: Online EXIFs

Gail downloads 10,000 JPEGs from a social network website. By examining the camera serial numbers in the images she is able to determine which images were shot by the same camera.

Felicity shows:

- She can reconstruct "friends" networks.
- She can find pseudonyms sharing the same camera.

Are these the same people?
Or people living together?
Gail probably doesn't need IRB approval either.

The documents (photographs) are publicly available.

- But what if the website requires registration?
Can we trust these researchers to do the right thing?

- Alice: anti-phishing toolbar
- Bob: Web server logfile analysis (bad passwords)
- Christine: popular security search terms
- Don: better spam filters
- Elaine: Wi-Fi Security
- Frank: hidden data survey
- Gail: EXIF analysis

We got the National Research Act because researchers in the 1960s said "trust me" and they were wrong.

*Research can blind the researcher to the needs of the research subjects.*
Each researcher has access to sensitive data that could be misused.

The IRB process forces the researchers to:

- Document what they are going to do.
- Think about how human subject data will be protected.
- Treat the human subjects with respect.

These scenarios involve no more than "minimal risk" and should be approved under the Common Rule's "expedited review" procedure.
The Human Test

"Would the experiment be useful if the data were generated by simulation or random processes and not by a human?"

- If YES, then **don't use humans!** (Respect for persons)

- If NO, then **get IRB approval!** (Follow the law.)
Advice for working with IRBs

Be intimately familiar with the Common Rule and your local regulations.

Make clear arguments that research should be approved under "expedited review procedures."

Ask your IRB to waive informed consent requirements (§46.116(c,d)).

Be familiar with protocols that other IRBs have approved.

Security researchers should volunteer to serve on their organization's IRBs.
IRB Mission Creep

IRBs are being applied to more areas of research:

- Oral histories.
- Ethnography
- Journalism

Some IRBs are quite conservative:

- Some IRBs are protecting the institution, rather than enforcing the rules.
- Some IRB members refuse to approve studies that they think "aren't science."

Retroactive approval question:

- How do you "wash" data that was collected w/o IRB approval?
IRB Resources

NSF FAQ — advocates consent forms in plain language, not legalese
  • http://www.nsf.gov/bfa/dias/policy/hsfqas.jsp


"Ethical Escape Routes for Underground Ethnographers," Jack Katz, UCLA, working paper