Ethical NREN Data Sharing
Going beyond the bare legal requirements

Roland van Rijswijk-Deij
roland.vanrijswijk@surfnet.nl
Outline

• Why is SURFnet doing this?

• State-of-the-art

• Background: Ethics 101

• Our approach

• Next steps
Why is SURFnet doing this? (1)

• NRENs are (usually) publicly funded bodies
  - Implies a responsibility to society
  - Our work should benefit our constituency, but also society at large

• Researchers need real data from operational networks to establish ground truth
  - Researchers know we have data
  - We fund Research on Networks (RoN)
  - RoN benefits us, and thus our constituencies
What data are we talking about?

Some examples of requests:

• Network flow data
• Packet captures
• DNS data
• Topology information
• Network link saturation
• Power consumption
• … (and much more)
• Most of the data mentioned on the previous slide is *highly privacy sensitive*

• Sometimes the privacy of *individual users* is at risk, sometimes that of *connected institutions*

• We have an obligation to protect the privacy of our users and of our connected institutions ... but we also want to help researchers
This is NOT what we want!

Betreten den amerikanischen Sektor

Privacy ends here

Vatsphäre endet hier

YES

YES, WE SCAN!

STASI 2.0

US Army

Yahoo!

National Security Agency

SURFnet: we make innovation work
SURFnet currently works with targeted NDAs that specify:

- what the data may be used for;
- who is responsible for safeguarding the data;
- under what conditions results based on the data may be published;
- that data must be stored and transported securely;
- that data must be destroyed upon request *and at the end of the project*;
- what the duration of the agreement is
State-of-the-Art

- Other NRENs have data sharing policy, e.g. Janet part of Jisc

- In the US comp. sci. researchers involving human subjects (vague!) must be approved by IRB
  - There are many examples of how this approach can fail, e.g. Facebook study, depression study

- In the Netherlands some universities are starting ethics committees for CS research
Problems with our current approach

• SURFnet requires data to be destroyed; *clashes with current academic values!*

• We only share data with *trusted researchers* (i.e. people we know personally)

• Decisions about data sharing are often taken by a single person and there is no guiding policy
• In January last year an international group of academics met at Schloss Dagstuhl to discuss “Ethics in Data Sharing”
Ethics 101 (intro)

• Ethics is about good and bad, right and wrong

• Ethics deals with core values: privacy, safety, fairness, transparency, ...

• It’s related to, but separate from law, religion and etiquette
• The relation between ethics and law explained to computer scientists :-)

Ethics

Law
Ethics 101 (when?)

- When values, beliefs or intuitions clash we have an ethical dilemma.
- Ethics helps examine such dilemmas but does not provide a single good thing to do, nor does analysis by an ethicist sanction certain behaviour.
Ethics 101 (theories)

- **Consequentialism** - the consequences of an action determine if it is the right thing to do.

- **Deontology** - adhering to a certain duty determines if it is the right thing to do.

- **Virtue Ethics** - whether or not a virtuous person would do it determines if it is the right thing to do.
Our approach: goals

• We want to improve on our current situation
• Share data more widely (i.e. not only with trusted researcher that we know in person)
• Strict adherence to privacy law (EU + national)
• Go beyond what the law requires and explicitly consider ethics
### Our approach: risk levels

<table>
<thead>
<tr>
<th>Risk level</th>
<th>Privacy impact</th>
<th>Ethics self assessment</th>
<th>Ethics review</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>High</td>
<td>Required</td>
<td>Mandatory</td>
<td>Flow data, packet captures</td>
</tr>
<tr>
<td>Medium</td>
<td>Moderate</td>
<td>Required</td>
<td>Some cases*</td>
<td>Data from authoritative DNS</td>
</tr>
<tr>
<td>Low</td>
<td>Low</td>
<td>Required</td>
<td>Not required</td>
<td>Aggregate statistics of network traffic</td>
</tr>
</tbody>
</table>

*at SURFnet’s discretion
Our approach: ethics assessment

• To encourage researchers to think about the ethics of their project

• Short, e.g. 1 or 2 paragraphs

• Often also a requirement for academic conferences

• Not formalised, free form

ACM Code of Ethics
https://www.acm.org/about/code-of-ethics

IEEE Code of Ethics
http://www.ieee.org/about/corporate/governance/p7-8.html
Our approach: ethics review

• Will be a formalised process, not yet fully specified

• Assessment involves: SURFnet employee handling the request, one independent colleague and an external subject matter expert

• Moderated by external ethics advisor
Our approach: teaching Ethics 101

- Ethics booklet created by ethics advisor from the University of Twente, Aimee van Wynsberghe
- Internal SURFnet workshop for employees about ethics
- More comprehensive training for employees who will deal with data sharing requests (?)
Our approach: legal considerations

• Dutch privacy law has two provisions for processing personal data (PII, includes IP addresses in most cases)

• Clause 1: processing is allowed to ensure reliable and secure operation of the [network] service

• Clause 2: processing is allowed [under certain conditions] for scientific research purposes
Our approach: legal considerations

• We choose to extend use of the data under clause 1 (reliable & secure operations)

• We see research as an extension of SURFnet’s obligations to care for the network

• But most importantly: we remain the sole controller of the data (researcher have delegated control); this would not be the case under clause 2 (scientific research)
Next steps

• Approval of policy by management

• Formalise ethics review process

• Start using the policy in practice; review effectiveness after one year

• Public repository of research results based on SURFnet data sets
Draft policy is on the TNC15 site for this session: https://tnc15.terena.org/core/presentation/127

ethics questions? feel free to contact Aimee aimeeovanrobot@gmail.com
• slide 8:
  - "Destroyed Hard Drive“ by IT Liquidators - Own work. Licensed under CC BY-SA 3.0 via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Destroyed_Hard_Drive.jpg#/media/File:Destroyed_Hard_Drive.jpg
  - “Shaking hands” CC-BY-SA-2.0 https://www.flickr.com/photos/aidan_jones/3575000735
• slide 10: