Herbert West – Deanonymizer

Mihir Nanavati, Nathan Taylor, William Aiello and Andrew Warfield
University of British Columbia, Vancouver

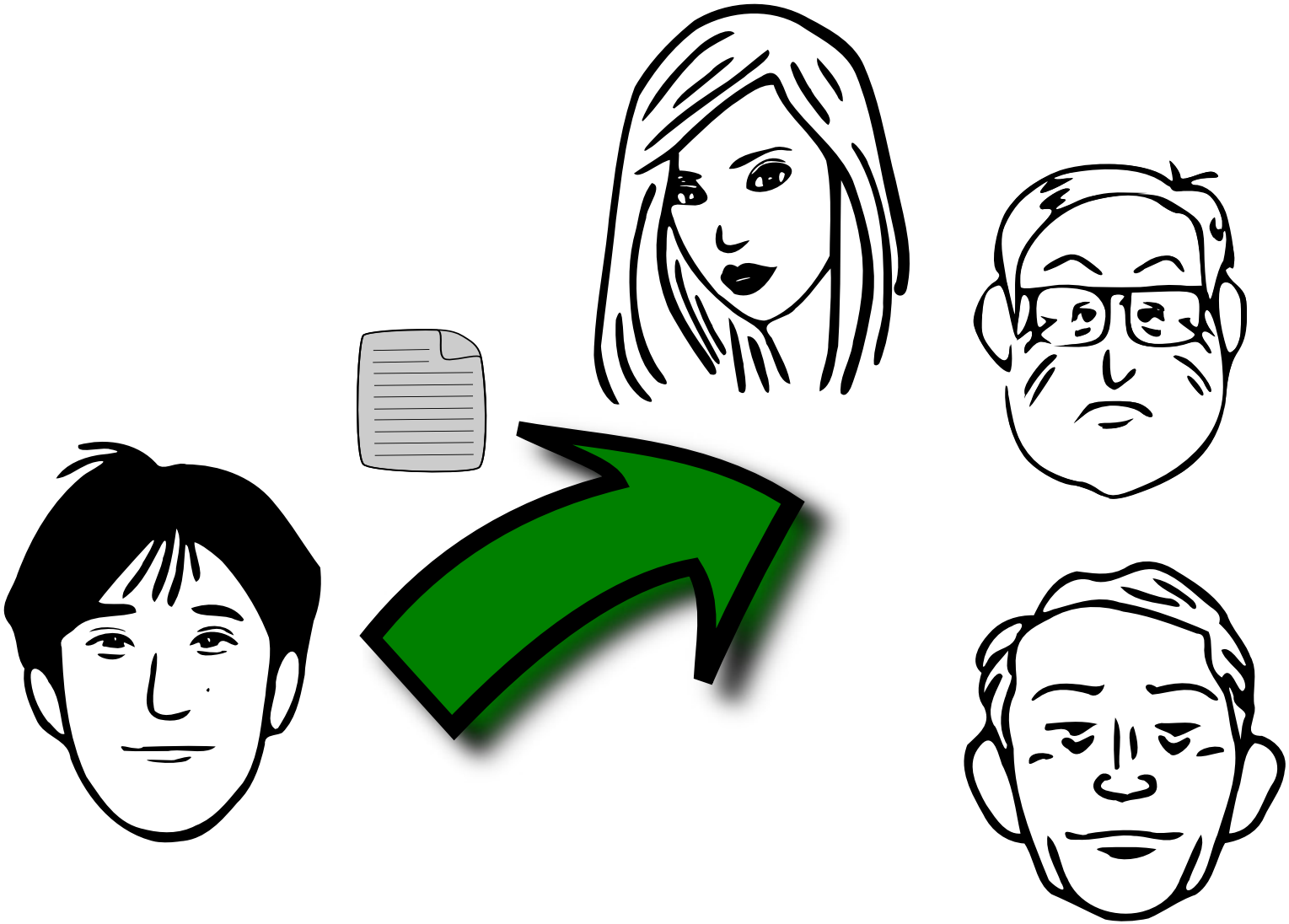
HotSec, August 9, 2011
San Francisco, CA.


UNITED STATES
POSTAL SERVICE

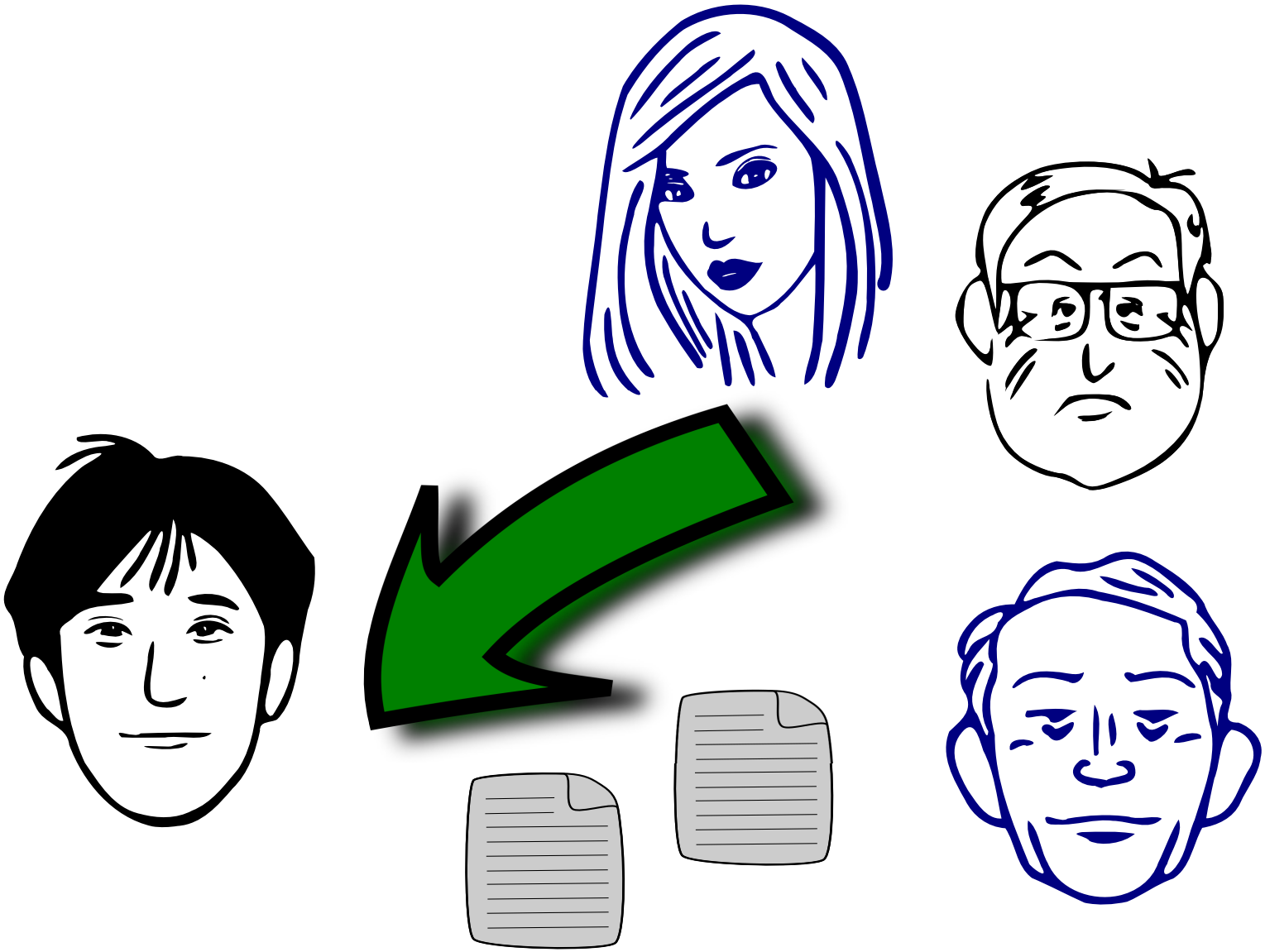
Review Snippets

- *"...I believe the paper in its present form requires substantial modifications before it is fit for publication...."*
- *"The paper doesn't present much that's novel... I don't buy the handwaving."*

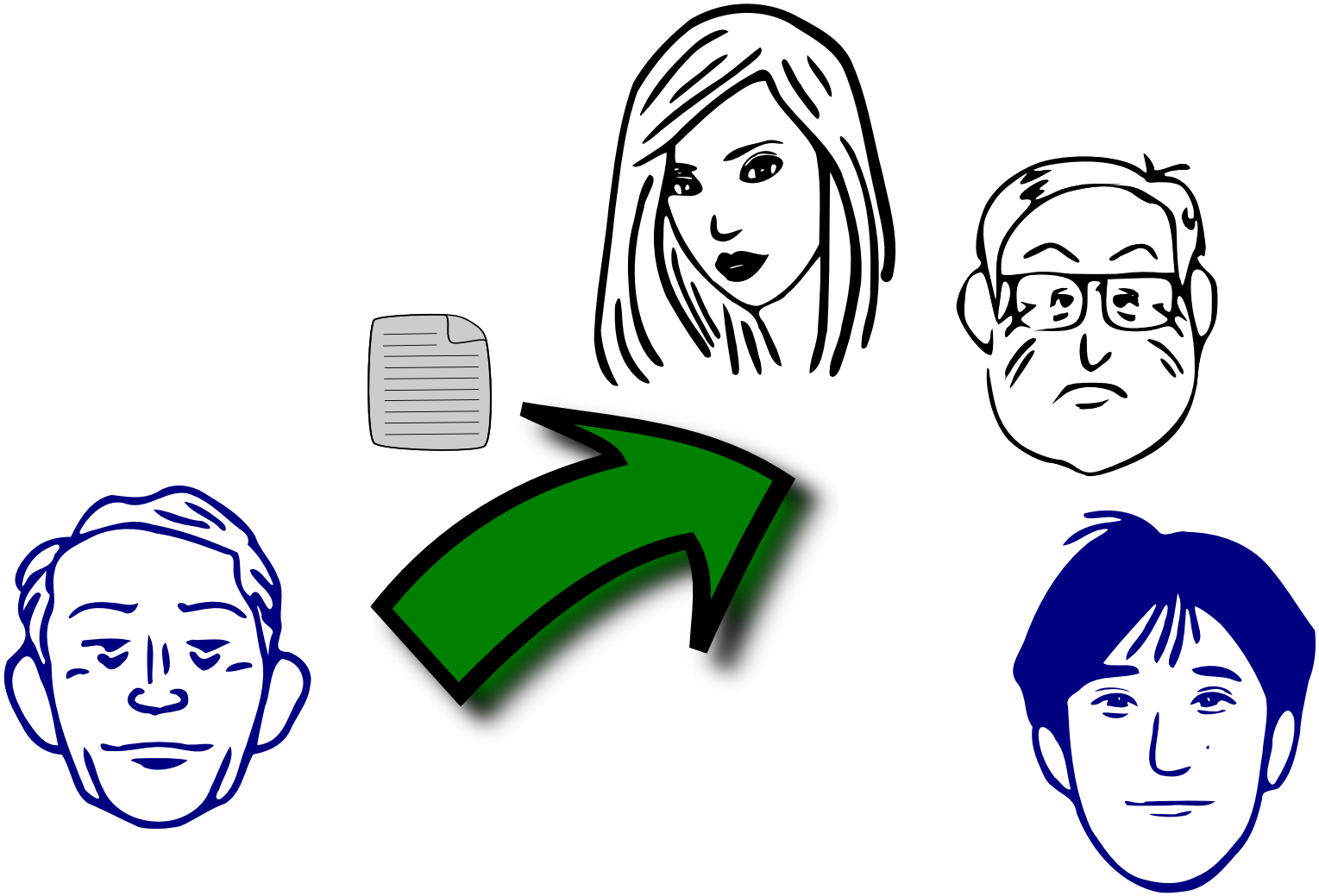
The Peer Review Process



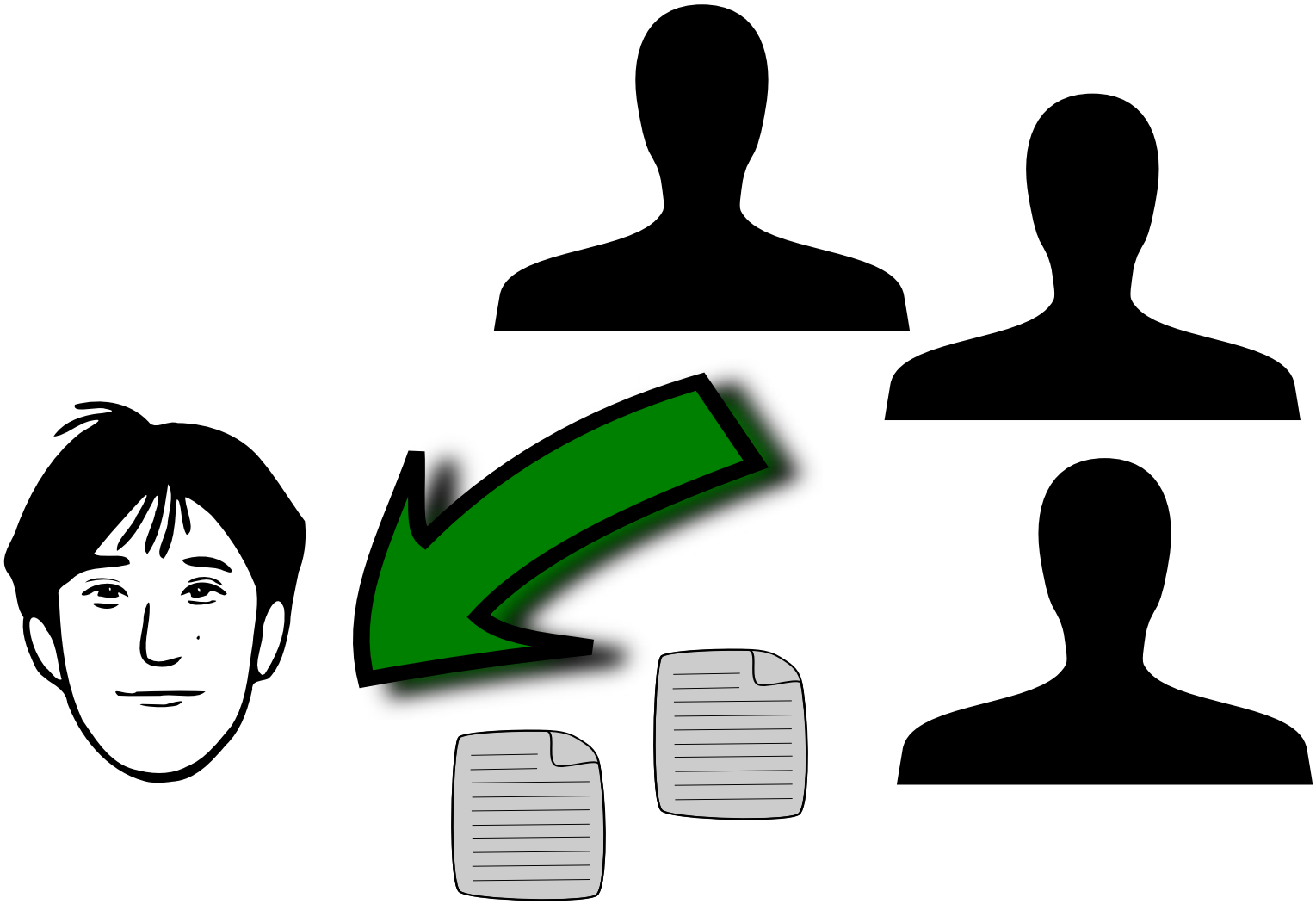
The Peer Review Process



The Peer Review Process



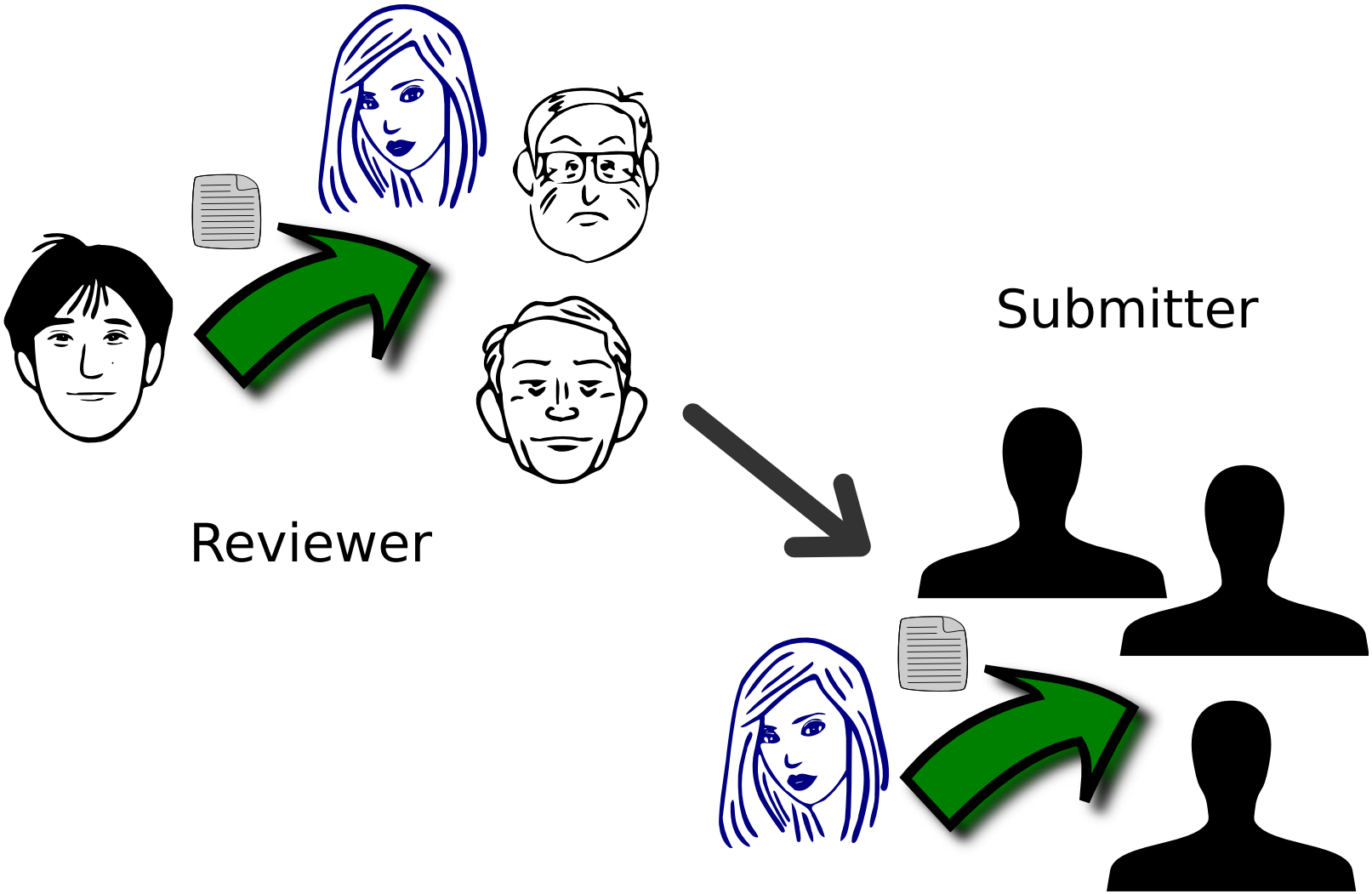
Anonymous Review



Anatomy of an Attack

Stylometric Analysis

Insider Attack



Small Academic Communities

How many people has Patrick McDaniel served with on other program committees in the last five years?

Small Academic Communities

How many people has Patrick McDaniel served with on other program committees in the last five years?

569*

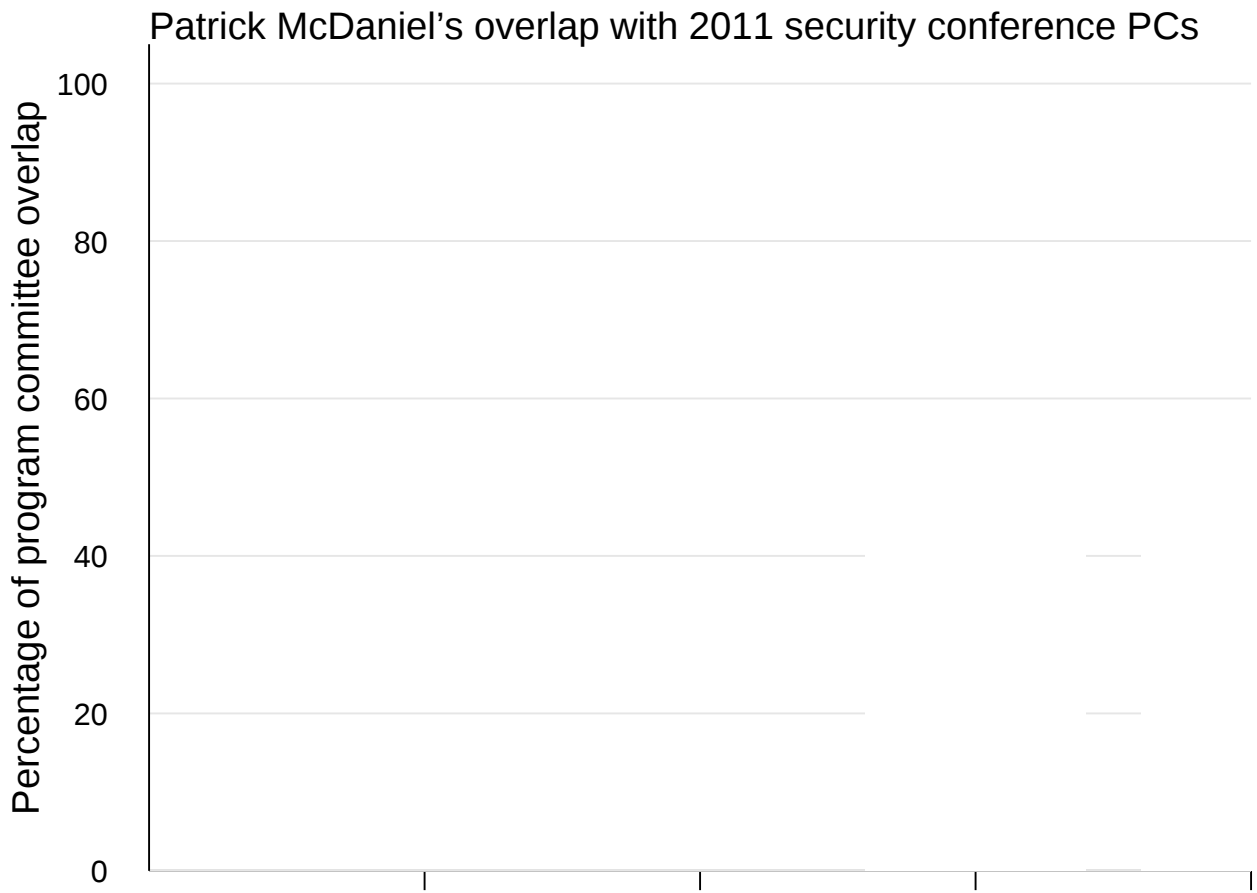
Small Academic Communities

1. Angelos Keromytis
2. Bart Preneel
3. Brian Levine
4. Christoph Schuba
5. David Evans
6. Hao Chen
7. Jelena Mirkovic
8. Jessica Staddon
9. Kevin Butler
10. Konstantin Beznosov
11. Margo Seltzer
12. Matt Blaze
13. Michael Franz
14. Niels Provos
15. Nikita Borisov
16. Patrick Traynor
17. Sam King
18. William Aiello

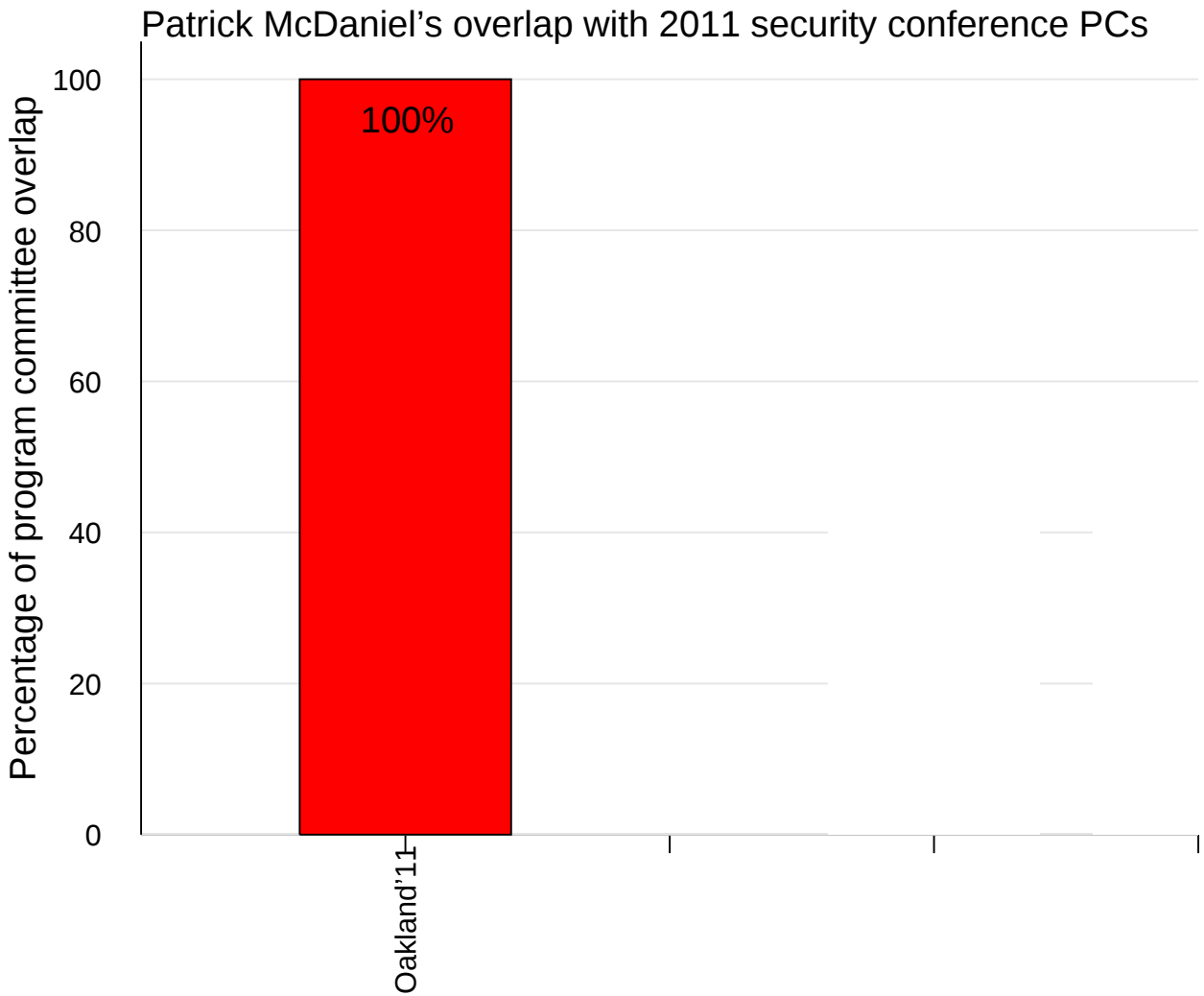
Small Academic Communities

For four top-tier security conferences in 2011, what percentage of the program committee does Patrick potentially have prior reviews for?

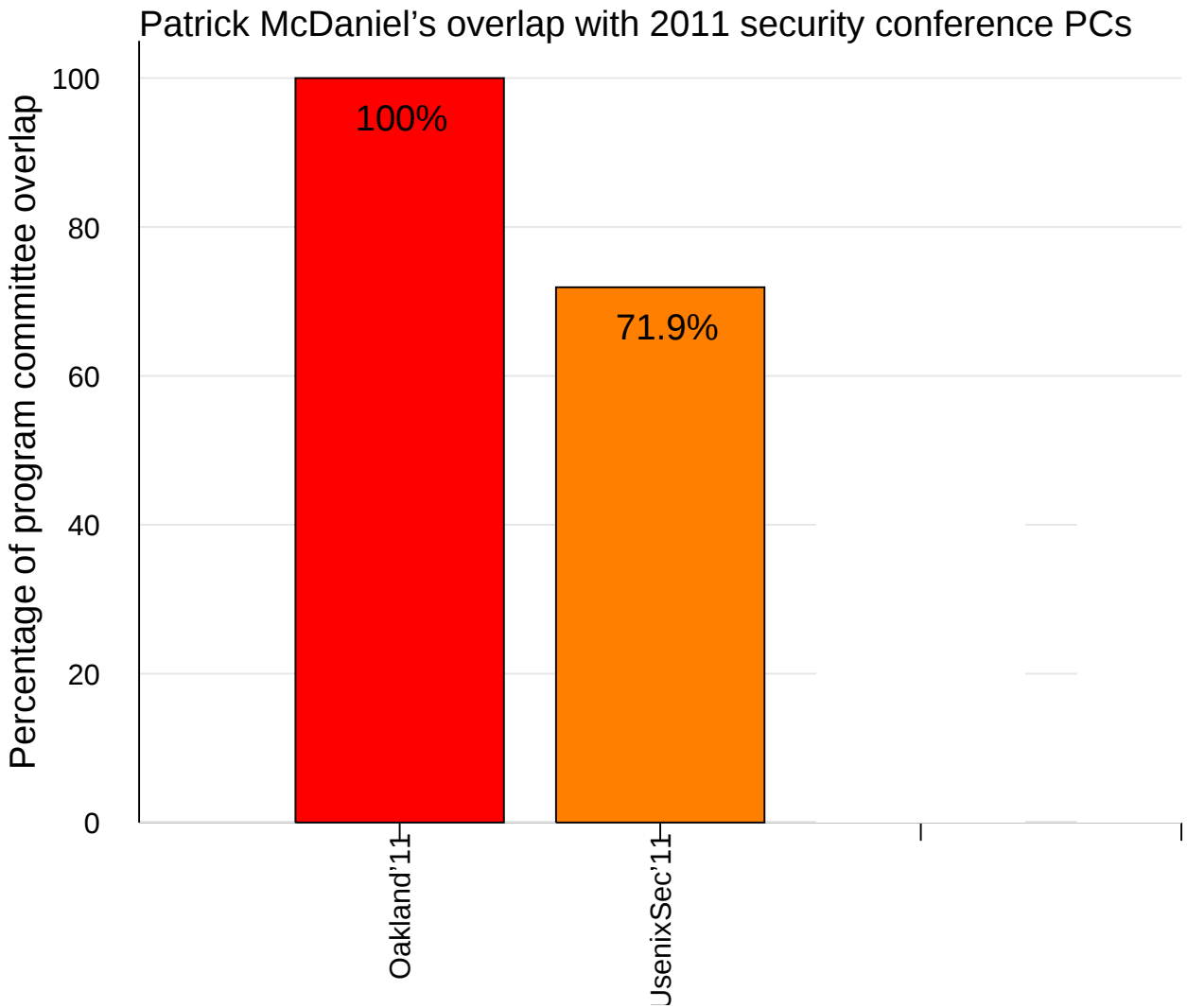
Degree of Overlap



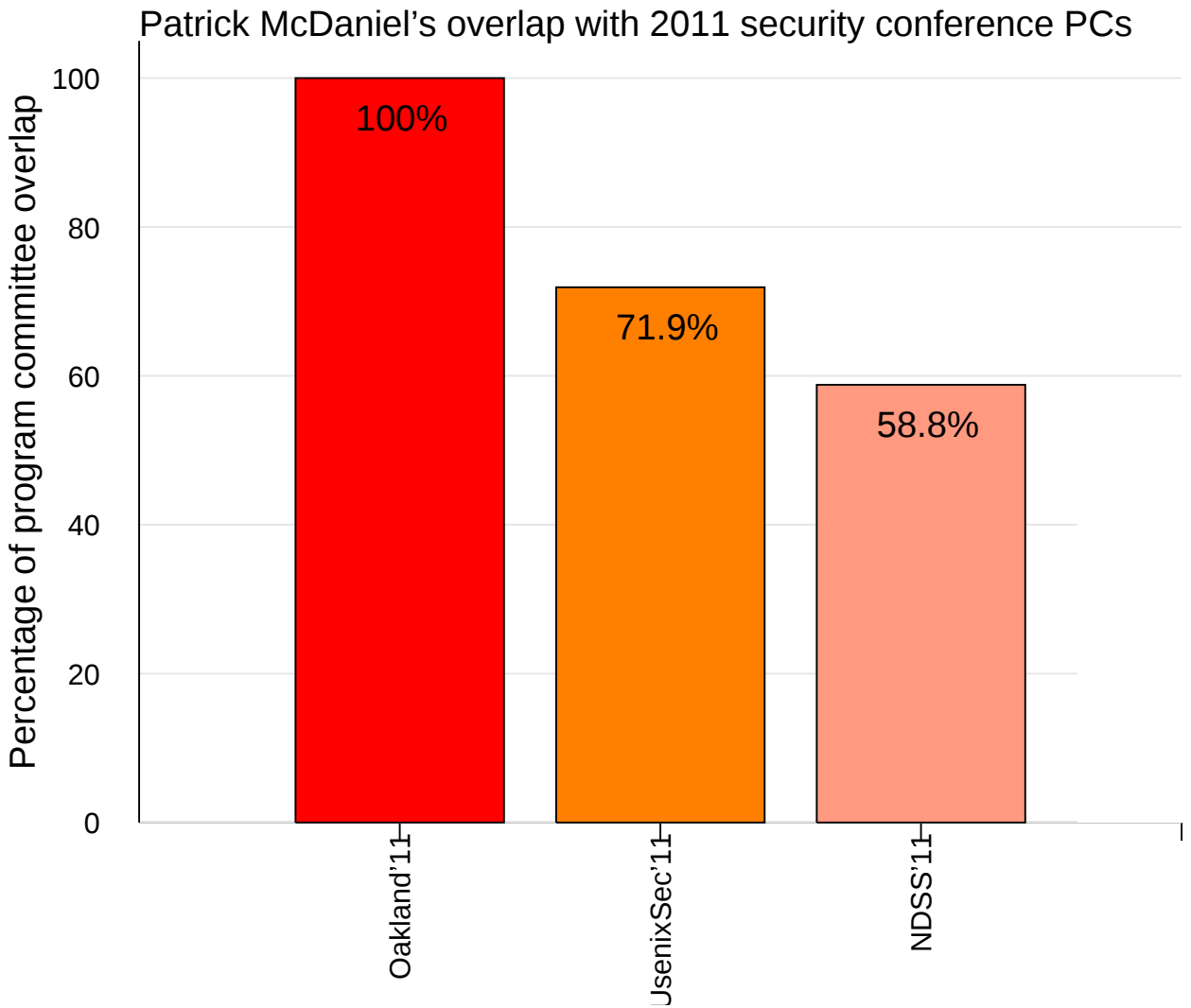
Degree of Overlap



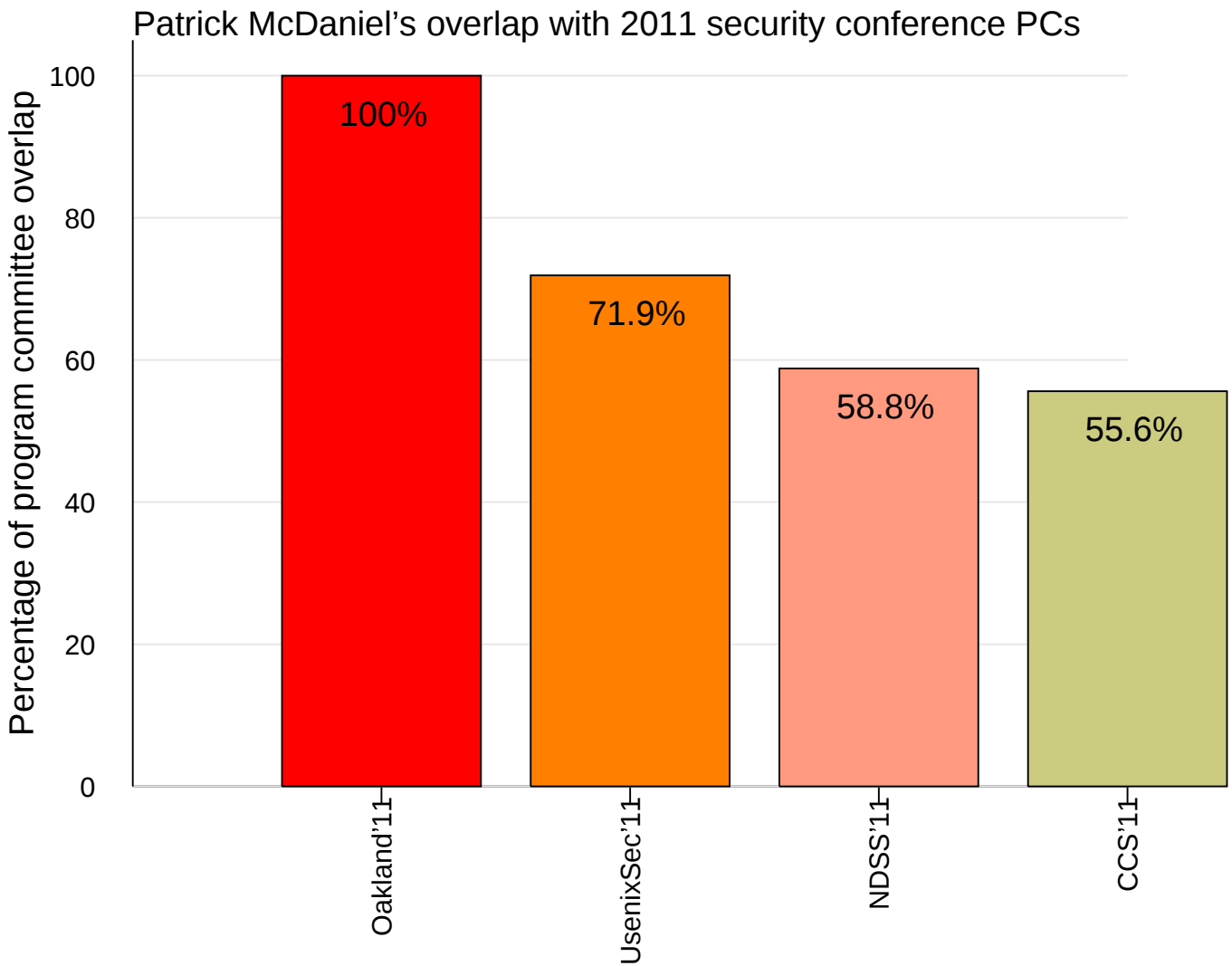
Degree of Overlap



Degree of Overlap



Degree of Overlap



How does it work?

Training: Tiger

Orange, black stripes

Training: Lion

Yellow, a mane

Testing

A tiger!

(Image copyright B_cool, available under Creative Commons Attribution license)

Testing

Lion

(Public domain, original copyright Trisha M Shears)

Testing

Tiger

(Image copyright Monika Betley, available under GFDL)

Testing

Lion

Testing

Tiger

Testing

Lion!

(Public domain, original copyright Jeff Kubina)

Our Classifier

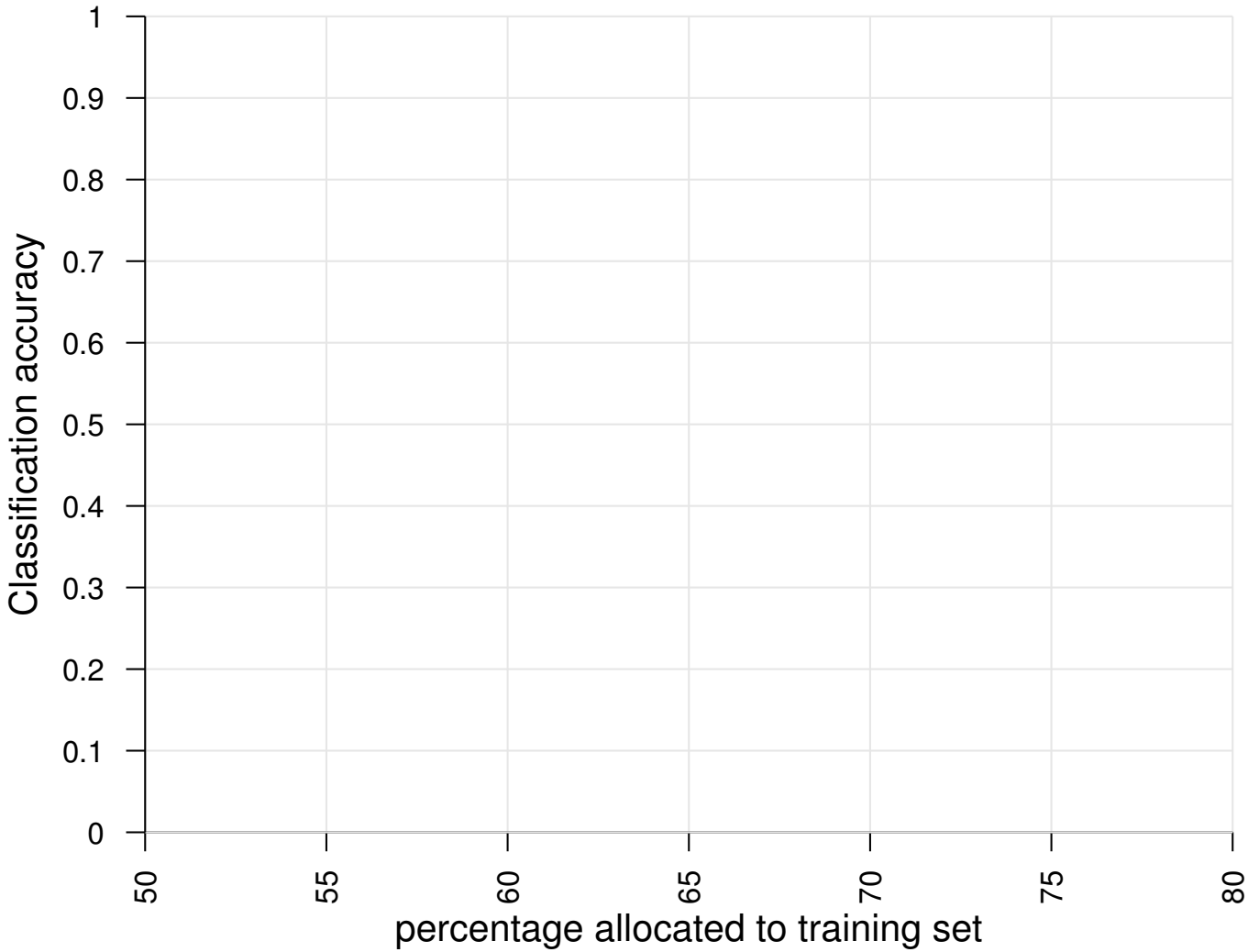
- `NaiveBayesClassifier` class of NLTK in Python
- Features – Unigrams, Bigrams and Trigrams
- Scored on an authorial basis using *tf-idf*

Evaluation Data Sets

Data Set	Reviewers	Corpus Size	Avg. Review Length
class1	9	125,138	424
class2	16	225,067	403
conf1	17	45,619	217
conf2	14	43,922	488

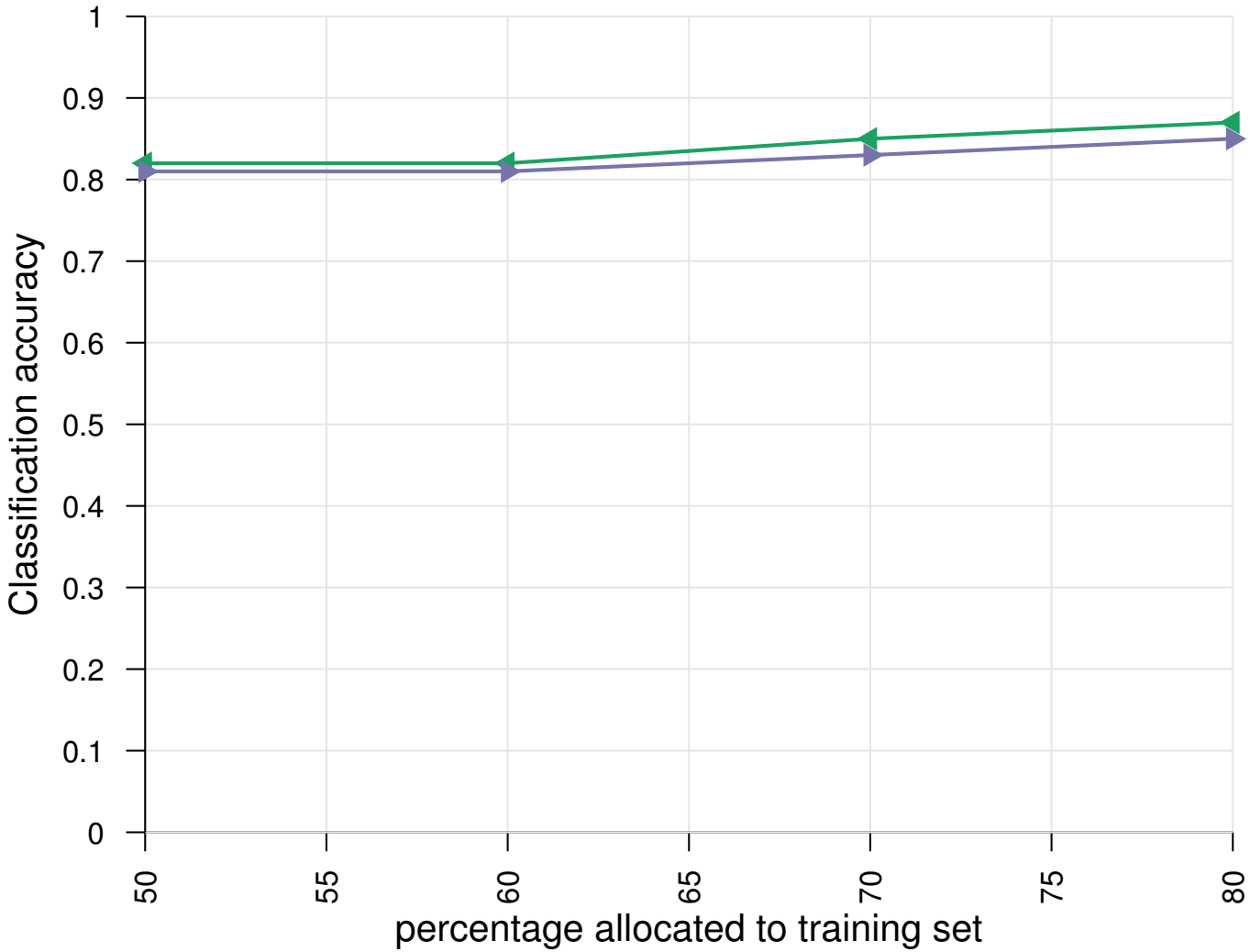
How **well** does it work?

Classifier Accuracy



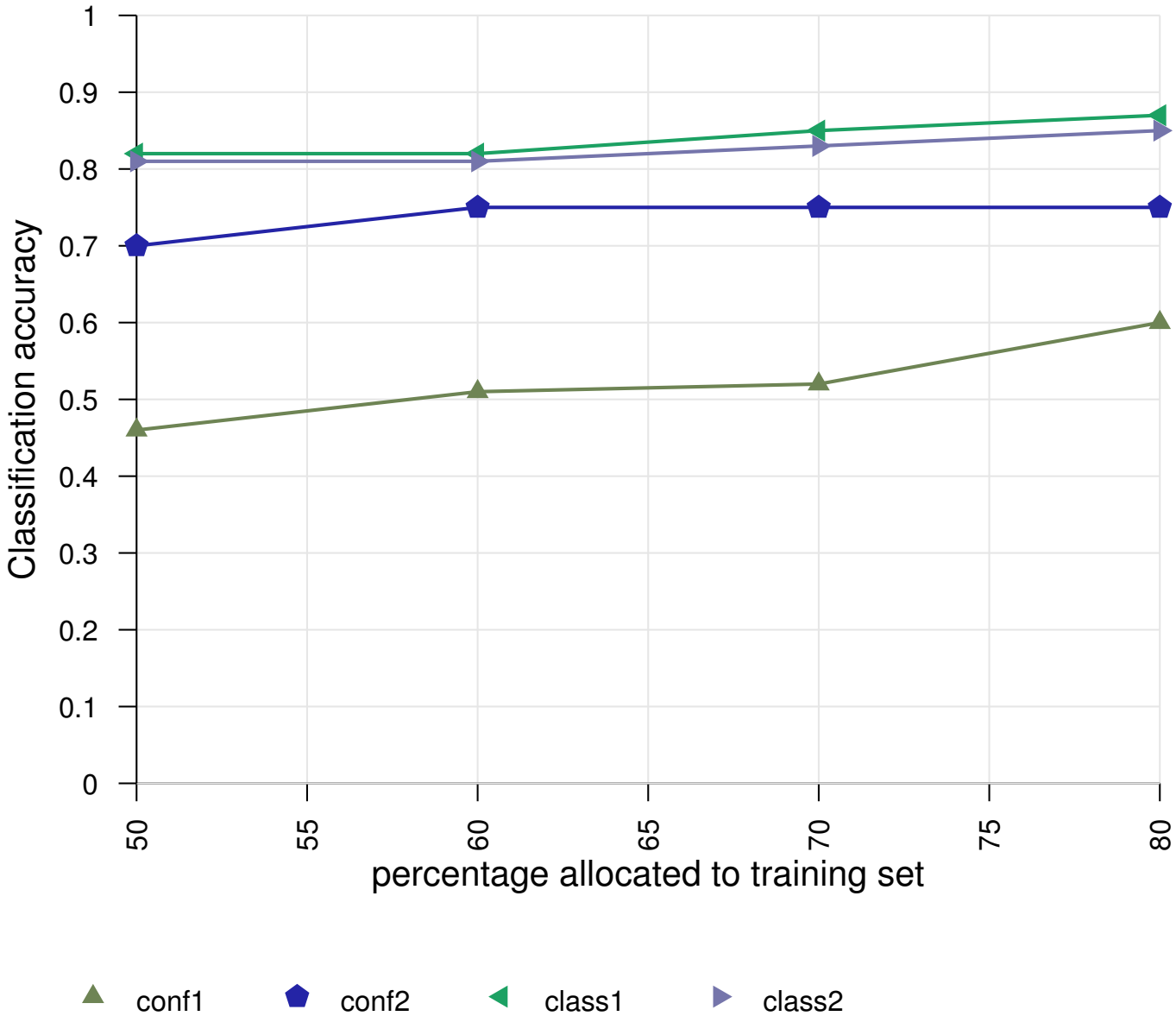
(Median of 15 cross-validated runs)

Classifier Accuracy



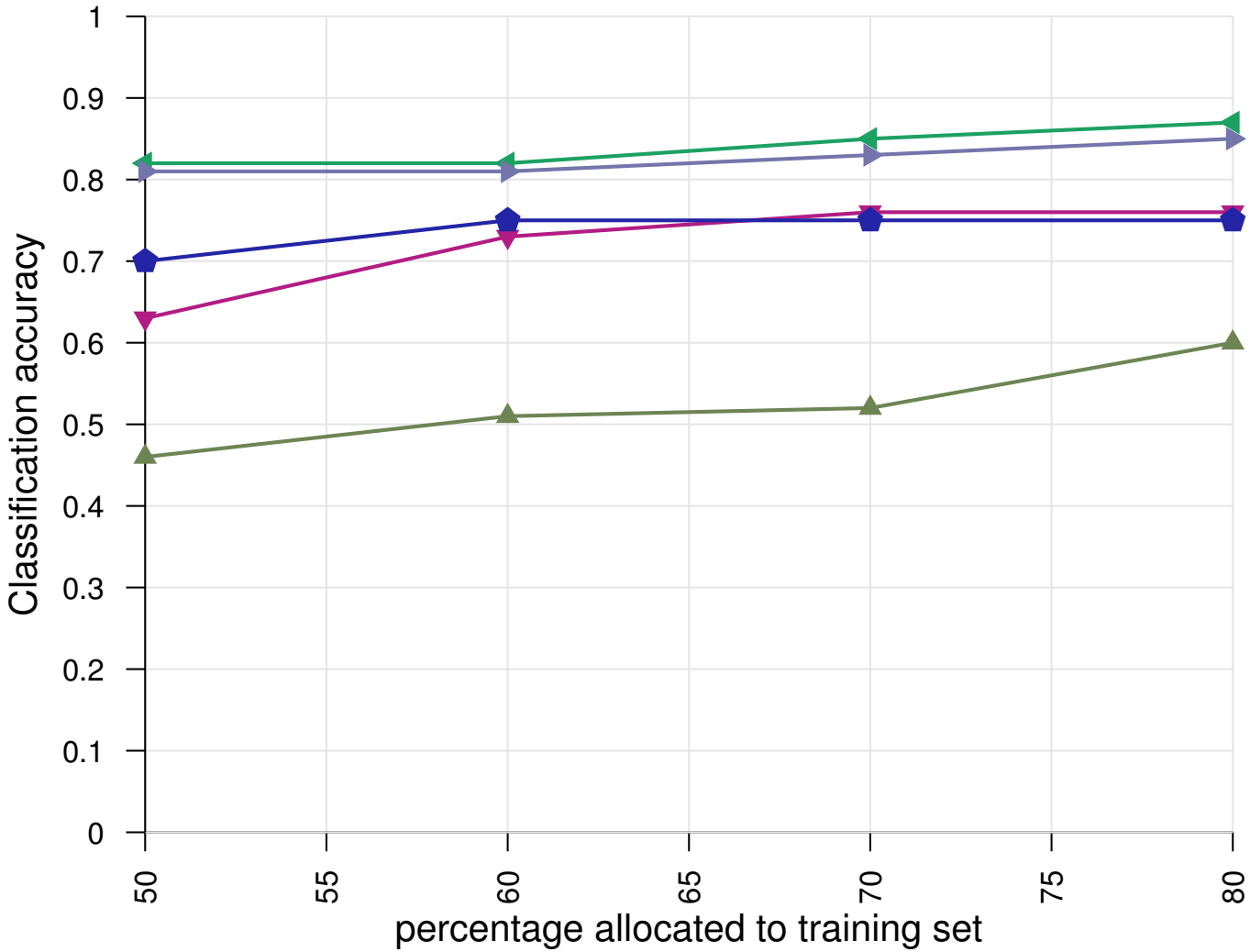
(Median of 15 cross-validated runs)

Classifier Accuracy



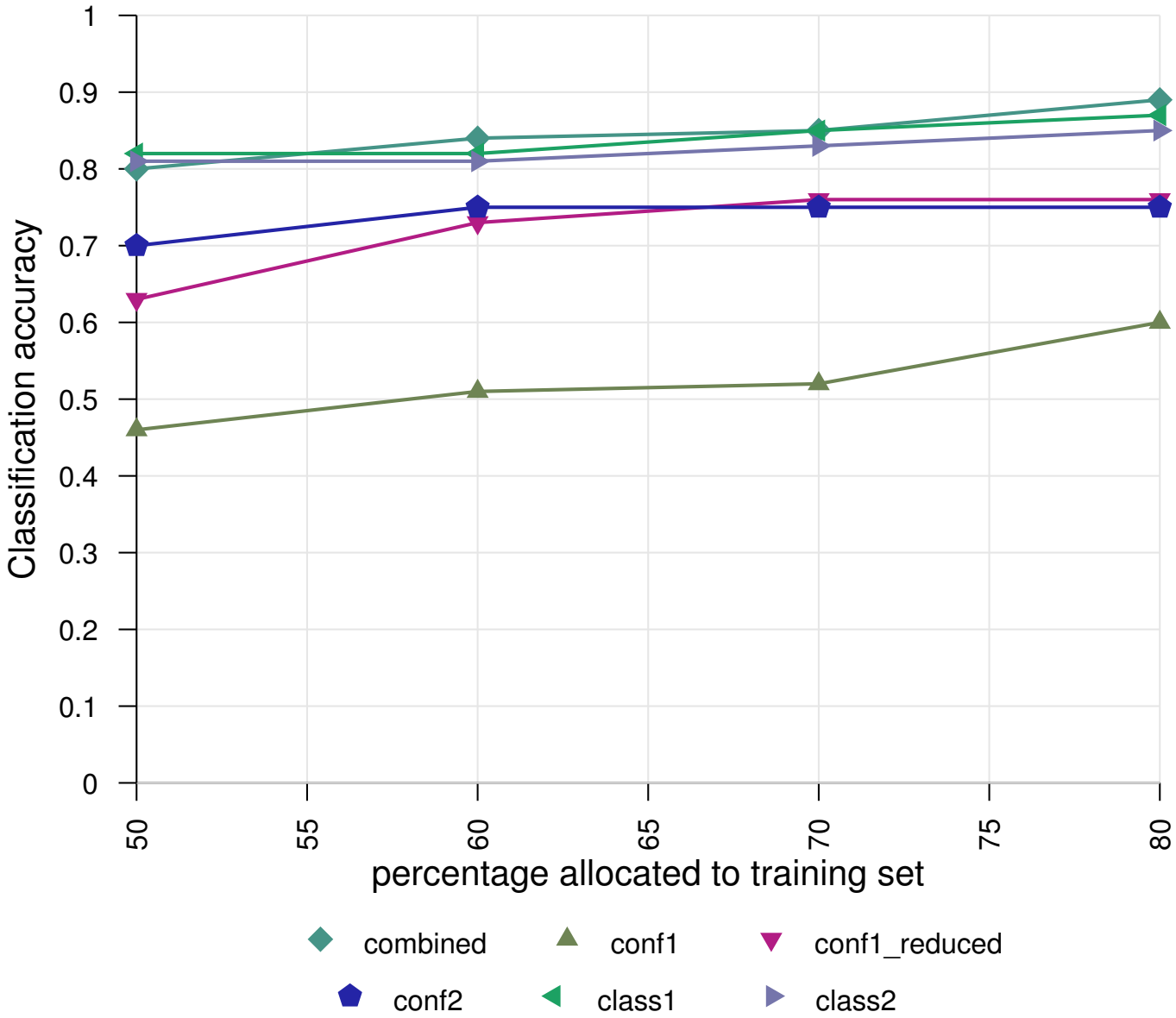
(Median of 15 cross-validated runs)

Classifier Accuracy



(Median of 15 cross-validated runs)

Classifier Accuracy



(Median of 15 cross-validated runs)

What makes **you** identifiable?

Word Features

Author 1	Author 2	Author 3	Author 4
"particularily"	"your system"	"looks at"	"Cons :"
"Lastly,"	**"visualisation"**	"sounds like"	"Pros :"
"Additionally"	"For example"	"what extent"	"awesome"
"so-called"	"is generally well-written"	"paper looks"	"problems existing"
"Indeed,"	"easy to follow"	"anomalous"	"slowdown factor"

BrE vs. AmE, spelling mistakes, connectives, structural demarcations, praise and criticisms

Word Features

Author 1	Author 2	Author 3	Author 4
"particularily"	"your system"	"looks at"	"Cons :"
"Lastly,"	"visualisation"	"sounds like"	"Pros :"
"Additionally"	"For example"	"what extent"	"awesome"
"so-called"	"is generally well-written"	"paper looks"	"problems existing"
"Indeed,"	"easy to follow"	"anomalous"	"slowdown factor"

BrE vs. AmE, **spelling mistakes**, connectives, structural demarcations, praise and criticisms

Word Features

Author 1	Author 2	Author 3	Author 4
"particularily"	"your system"	"looks at"	"Cons :"
"Lastly,"	"visualisation"	"sounds like"	"Pros :"
"Additionally"	"For example"	"what extent"	"awesome"
"so-called"	"is generally well-written"	"paper looks"	"problems existing"
"Indeed,"	"easy to follow"	"anomalous"	"slowdown factor"

BrE vs. AmE, spelling mistakes, **connectives**, structural demarcations, praise and criticisms

Word Features

Author 1	Author 2	Author 3	Author 4
"particularily"	"your system"	"looks at"	**"Cons :"**
"Lastly,"	"visualisation"	"sounds like"	**"Pros :"**
"Additionally"	"For example"	"what extent"	"awesome"
"so-called"	"is generally well-written"	"paper looks"	"problems existing"
"Indeed,"	"easy to follow"	"anomalous"	"slowdown factor"

BrE vs. AmE, spelling mistakes, connectives, **structural demarcations**, praise and criticisms

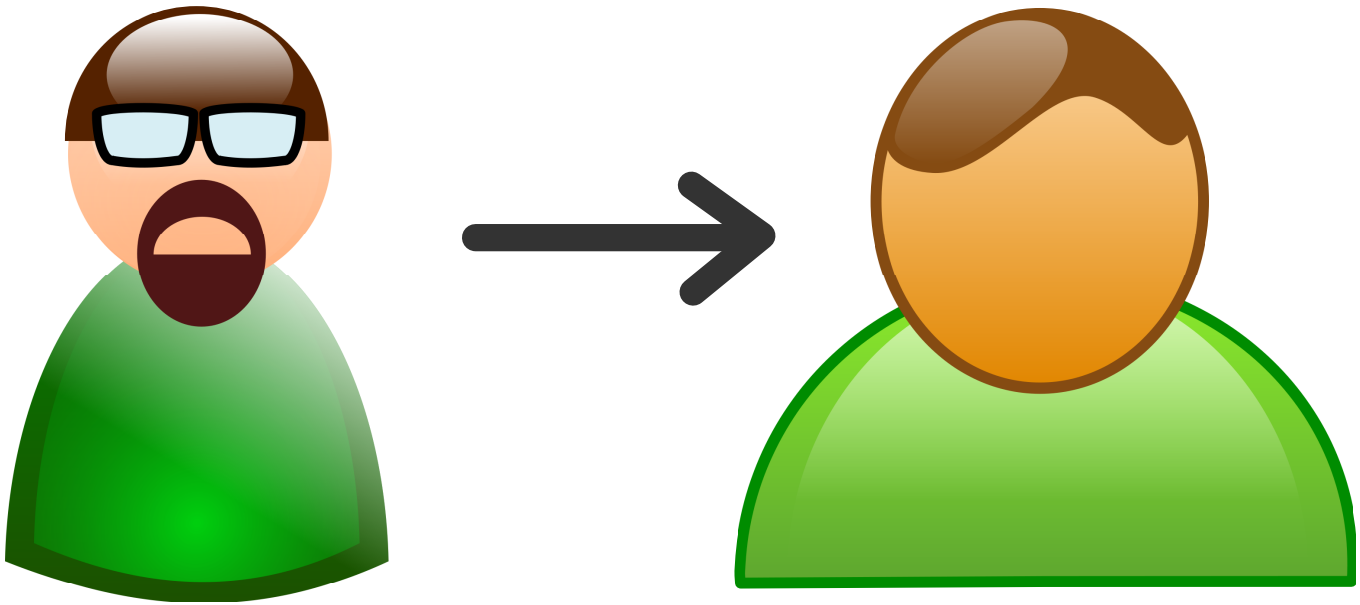
Word Features

Author 1	Author 2	Author 3	Author 4
"particularily"	"your system"	"looks at"	"Cons :"
"Lastly,"	"visualisation"	"sounds like"	"Pros :"
"Additionally"	"For example"	"what extent"	**"awesome"**
"so-called"	**"is generally well-written"**	"paper looks"	"problems existing"
"Indeed,"	**"easy to follow"**	"anomalous"	"slowdown factor"

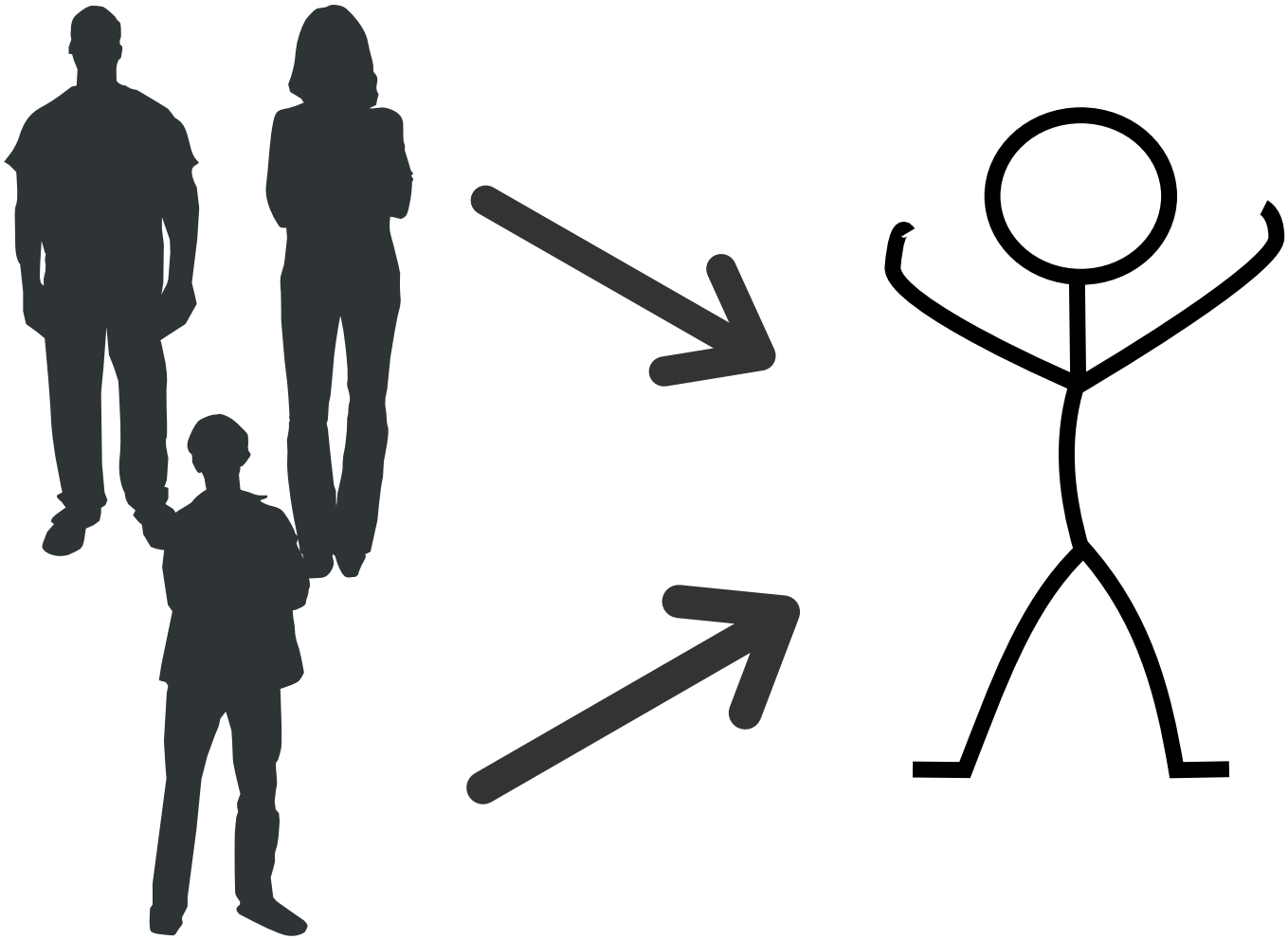
BrE vs. AmE, spelling mistakes, connectives, structural demarcations, **praise and criticisms**

What can we do?

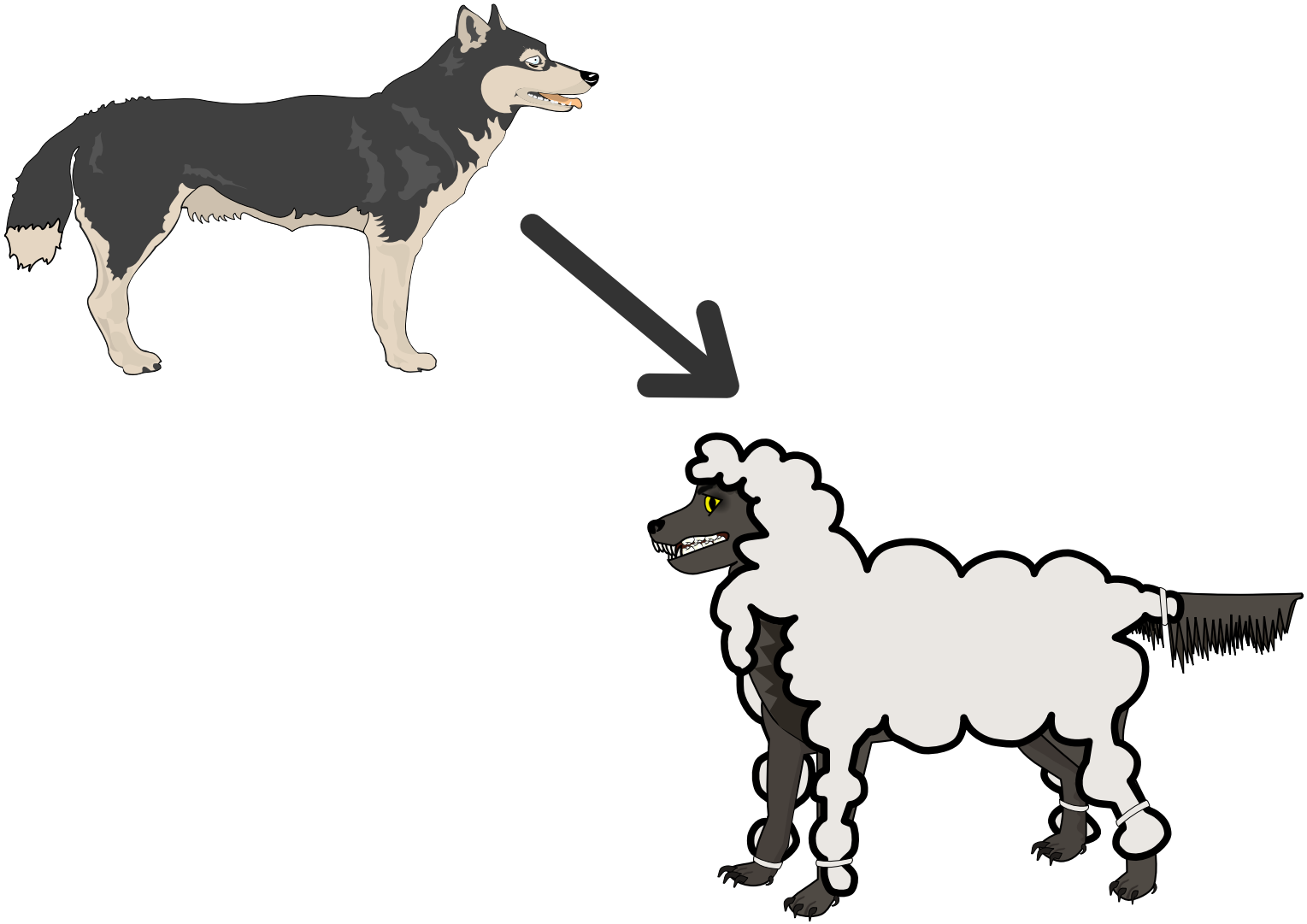
Obfuscation



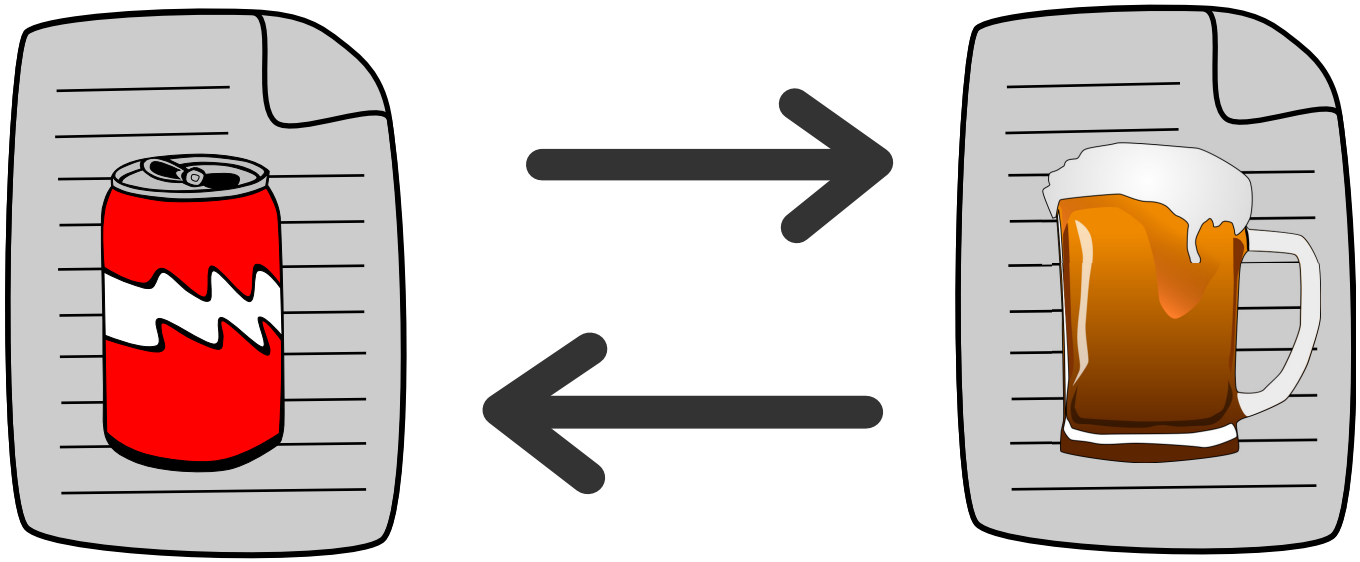
Normalization



Mimicry



Language Translation



English

German

Semantic Accuracy?

Language Translation

1. Sensitive data, present in the memory of the system, is liable to exposure.

Language Translation

1. Sensitive data, present in the memory of the system, is liable to exposure.
2. Sensitive data that resides in the memory of the system is responsible for exposure.

Language Translation

1. Sensitive data, present in the memory of the system, is liable to exposure.
2. Sensitive data that resides in the memory of the system is responsible for exposure.
3. Resident in the memory of the system have a responsibility to expose sensitive data.

Language Translation

1. Sensitive data, present in the memory of the system, is liable to exposure.
2. Sensitive data that resides in the memory of the system is responsible for exposure.
3. Resident in the memory of the system have a responsibility to expose sensitive data.

..

...

..

Language Translation

1. Sensitive data, present in the memory of the system, is liable to exposure.
2. Sensitive data that resides in the memory of the system is responsible for exposure.
3. Resident in the memory of the system have a responsibility to expose sensitive data.

   ..

   ...

   ..
4. Have a responsibility to expose sensitive data and system memory.

Mimicry

Wouldn't it be great if we had a plugin that allowed us to sound $(FAV_PC_MEMBER)-esque?

Mimicry

Wouldn't it be great if we had a plugin that allowed us to sound $(FAV_PC_MEMBER)-esque?

This is an excellent paper.

Chef: *Thees is un ixcellent peper. Hurty flurty schnipp schnipp!*

Valley: *Like, ya know, this is ya know, like, an excellent paper.*

Mimicry

Andy
super, high value, impact, nits, –, intuitive
Bill
policy, audience, magic runes, implications, the core tenets, interesting

Mimicry

Andy
super, **high value**, **impact**, **nits**, **–**, **intuitive**
Bill
policy, audience, magic runes, implications, the core tenets, interesting

Andy: "Despite some **nits**, the paper is **super high value** – the idea is **intuitive** and high **impact**."

Mimicry

Andy
super, **high value**, **impact**, **nits**, **–**, **intuitive**
Bill
policy, **audience**, **magic runes**, **implications**, **the core tenets**, **interesting**

Andy: "Despite some **nits**, the paper is **super high value** – the idea is **intuitive**, but high **impact**."

Bill: "**The core tenets** of this paper, especially the **magic runes** the authors use for their security **policy**, are **interesting** with broad **implications** for the **audience** of this conference."

Conclusions

- Simple machine classifiers are capable of identifying reviewers
- Current anonymization steps are insufficient
- Loss of perceived anonymity is more dangerous than a loss of anonymity *per se*