„TELL ME WHAT YOU USE AND I WILL TELL YOU WHO YOU ARE“ (Ch.II)

Swiss Cyber Storm 2023

InfoGuard
SWISS CYBER SECURITY

InfoGuard
SWISS CYBER SECURITY

INTRODUCTION

PROFILE

HUMAN-IST
HUMAN CENTERED
INTERACTION
SCIENCE
& TECHNOLOGY

- Penetration Tester@Infoguard AG
- External PhD Student@Human-IST Institute, University of Fribourg
- Applied science research collaboration

WHAT IS THIS TALK ABOUT?

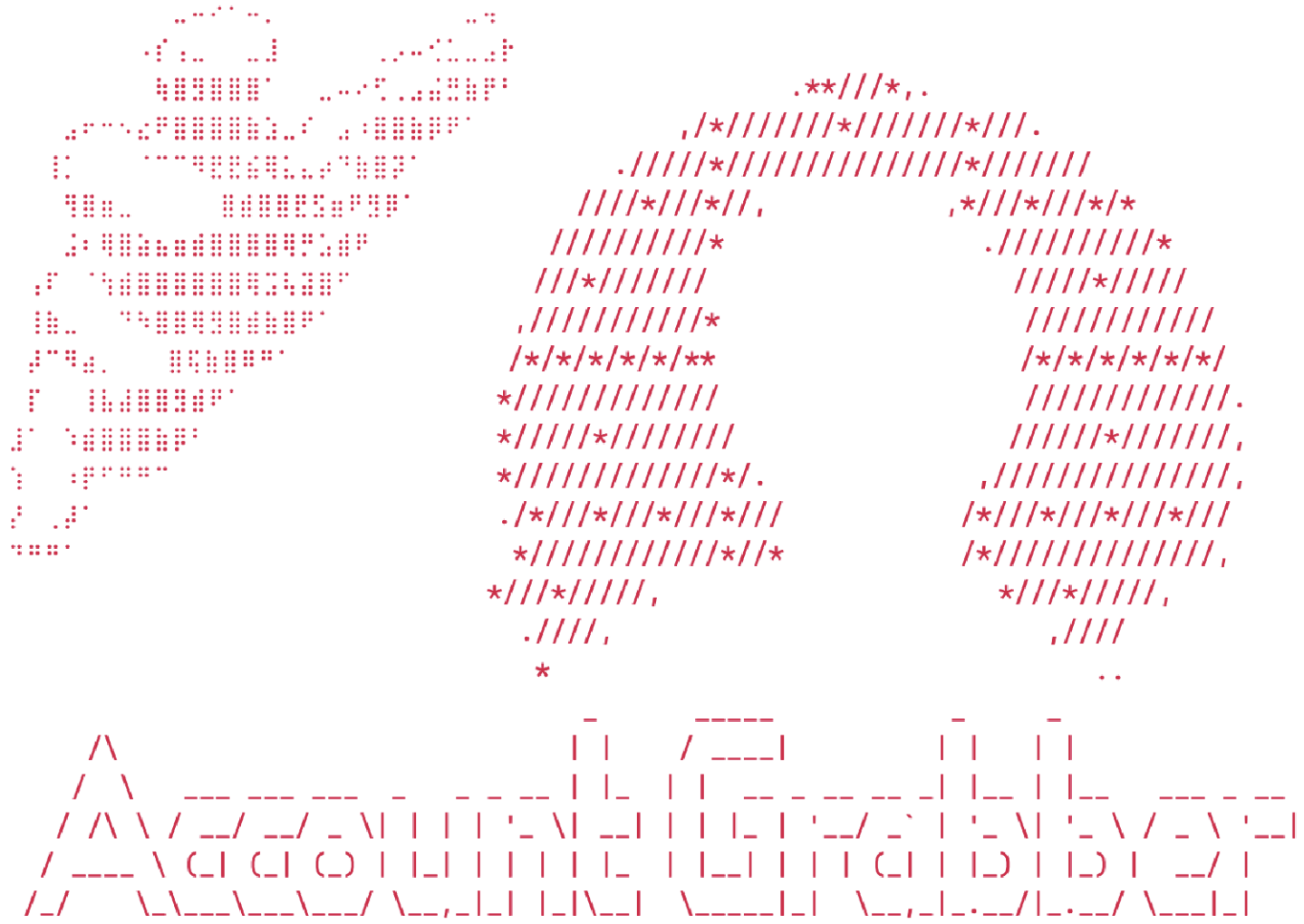
InfoGuard
SWISS CYBER SECURITY

- interdisciplinary security research
- work in progress
- no final conclusions (yet)
- interesting observations, hypotheses
- outlook

InfoGuard
SWISS CYBER SECURITY

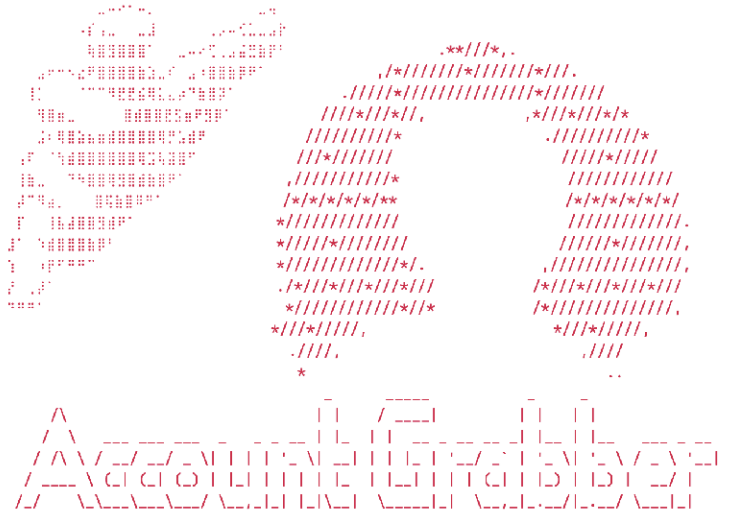
TOOLING

TOOL



WHAT IS THE TOOL?

- Open source intelligence (OSINT) project
- Does user enumeration on webapps
- Analyzed 300-400 webapps
- Tool currently supports >110
- Enhancements w/ IPA

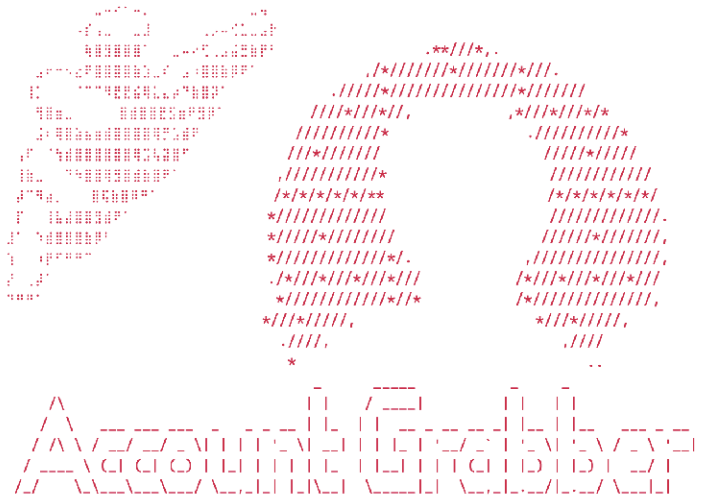


WHAT DOES THE OUTPUT LOOK LIKE?

e-mail	first name	last name	gender	byear	[redacted]					
[redacted]	[redacted]	[redacted]	male	1973	1	0	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1962	1	1	1	1	0	0
[redacted]	[redacted]	[redacted]	male	1975	1	0	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1971	1	0	1	1	0	0
[redacted]	[redacted]	[redacted]	male	1961	0	0	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1985	1	0	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1972	0	0	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1979	1	0	0	0	0	0
[redacted]	[redacted]	[redacted]	male	1955	0	0	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1972	0	0	0	0	0	0
[redacted]	[redacted]	[redacted]	male	1962	1	0	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1984	1	0	1	0	0	0
[redacted]	[redacted]	[redacted]	female	1991	0	0	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1976	1	0	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1987	1	0	1	0	0	0
[redacted]	[redacted]	[redacted]	male	1983	1	1	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1991	0	0	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1987	1	0	0	0	0	0
[redacted]	[redacted]	[redacted]	male	1965	0	0	0	0	0	0
[redacted]	[redacted]	[redacted]	male	1960	0	1	0	0	0	0
[redacted]	[redacted]	[redacted]	female	1970	1	0	0	0	0	0

WHAT IS THE STATUS?

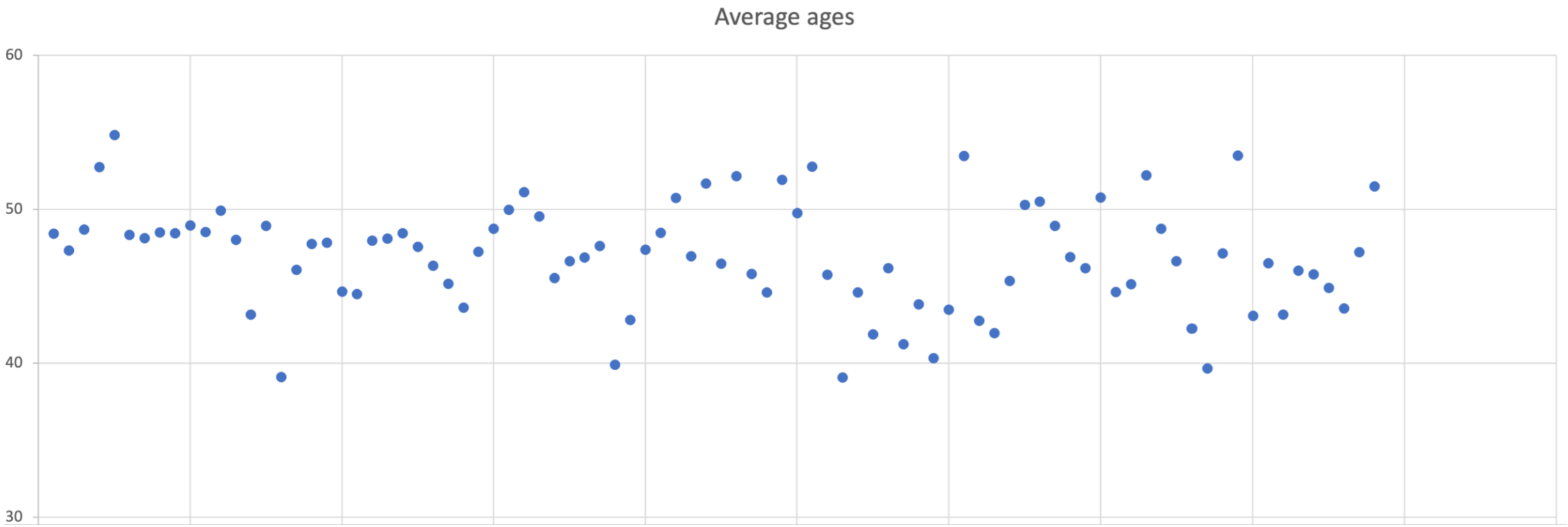
- Still in data aquisition phase
- Current dataset w/ 400 random emails, «training data»
- gender/age are known (self-declared), singup often only 2 options, whole spectrum uncovered, not an ideal repr. of community



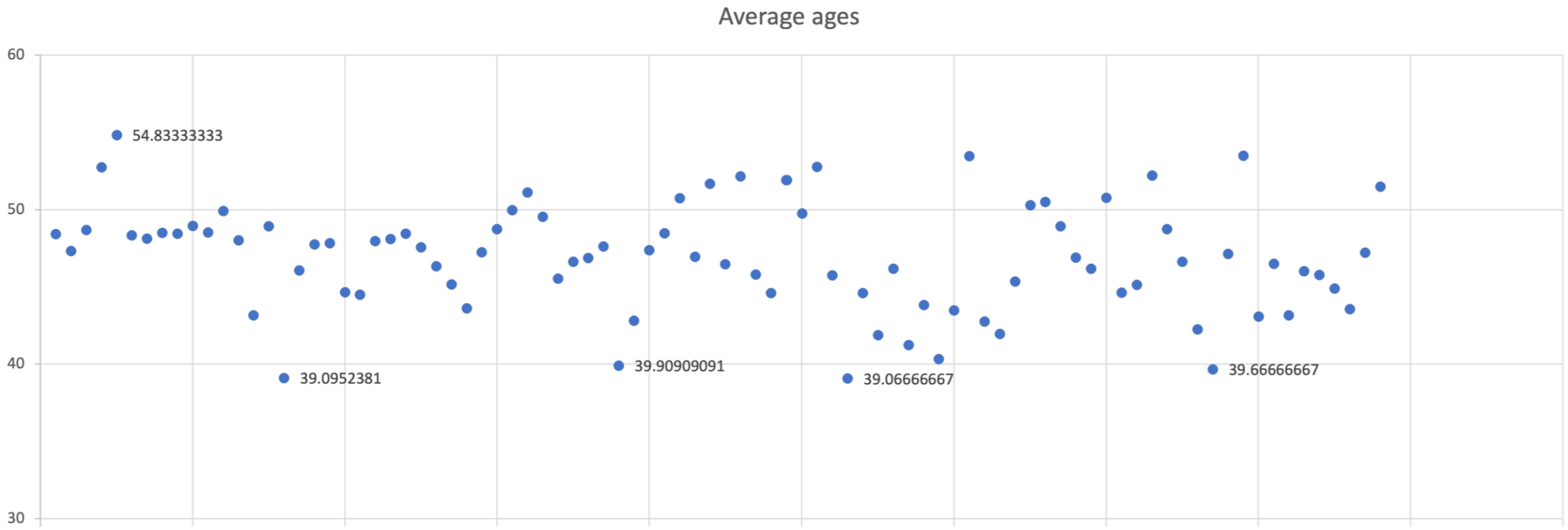
InfoGuard
SWISS CYBER SECURITY

OBSERVATIONS

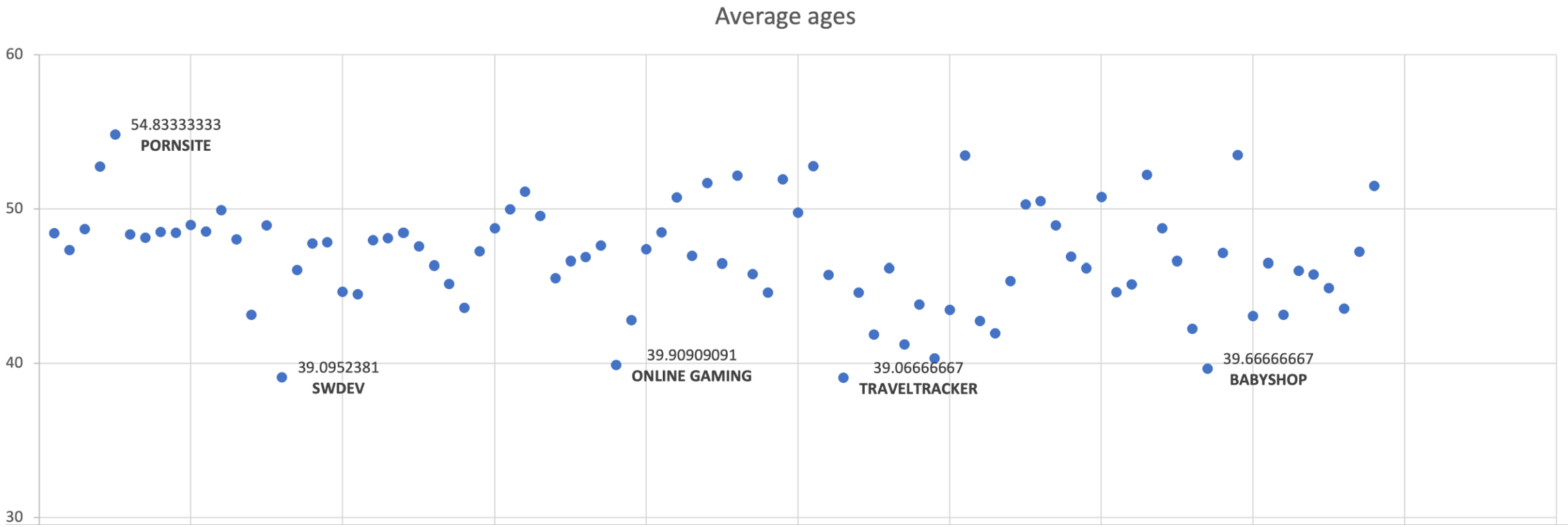
OBSERVATIONS I – AVERAGE AGES



OBSERVATIONS I – AVERAGE AGES MINIMA/MAXIMA

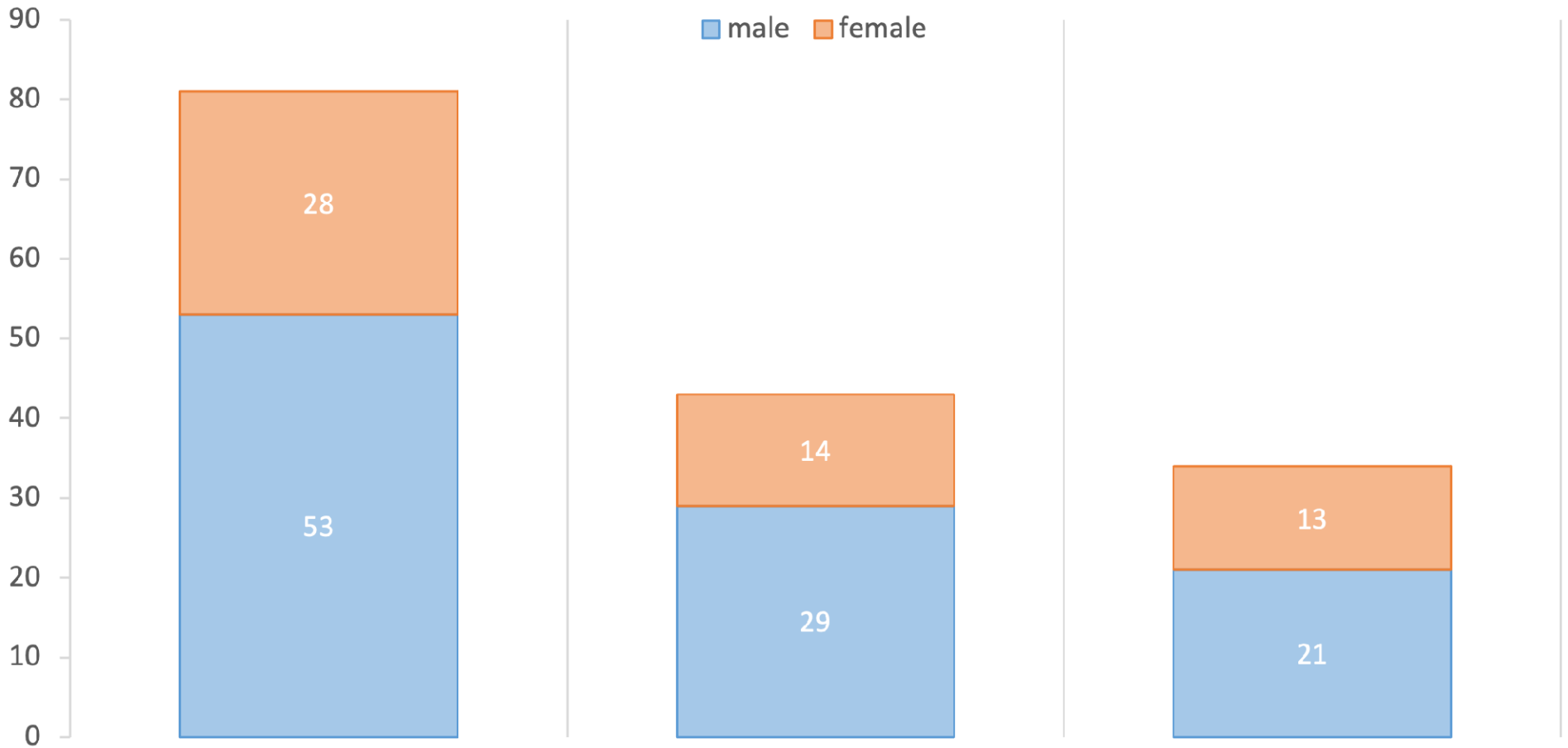


OBSERVATIONS I – AVERAGE AGES MINIMA/MAXIMA

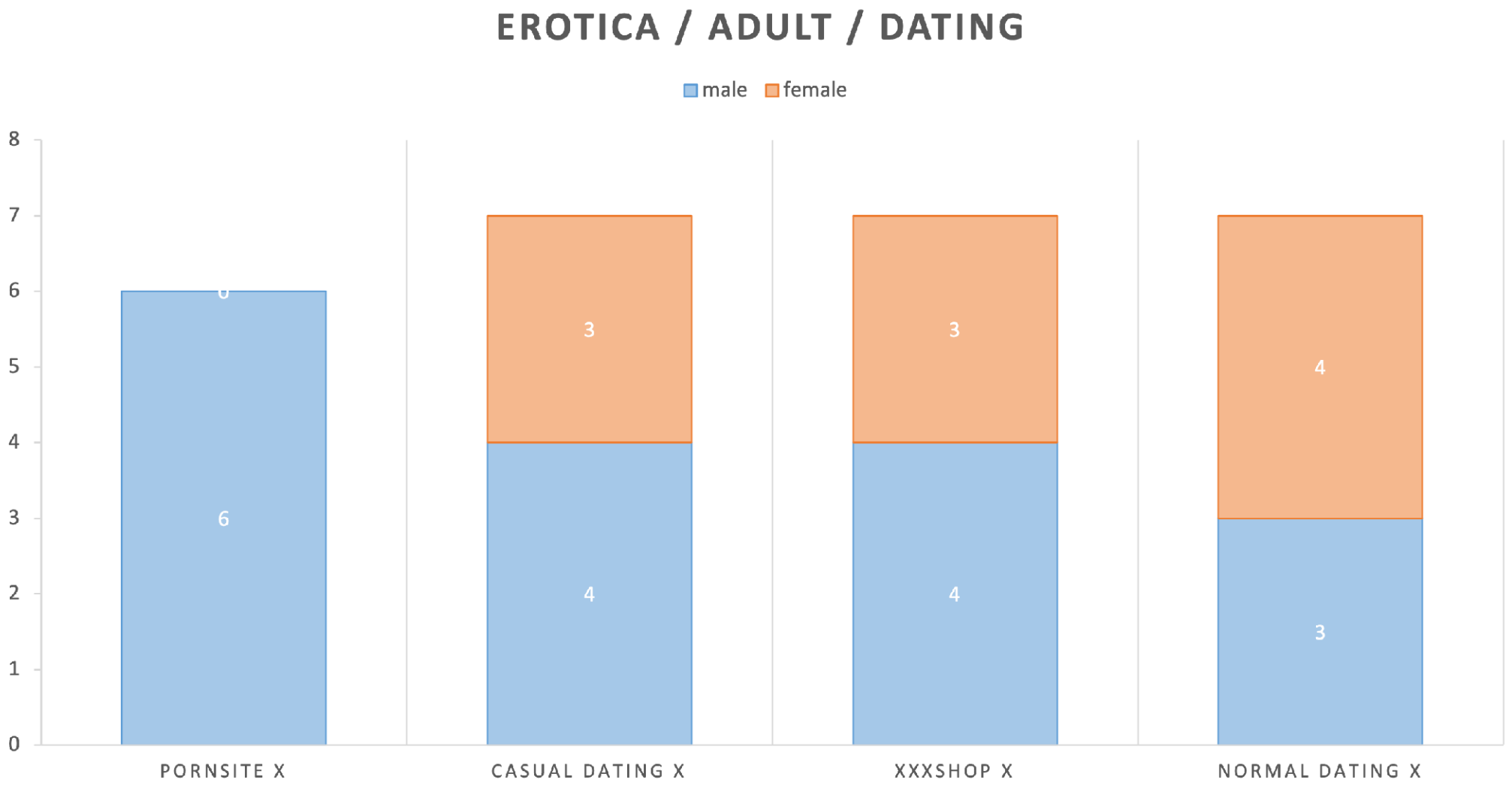


OBSERVATIONS II – GENDER ID DISTRIBUTION TECHSHOPS

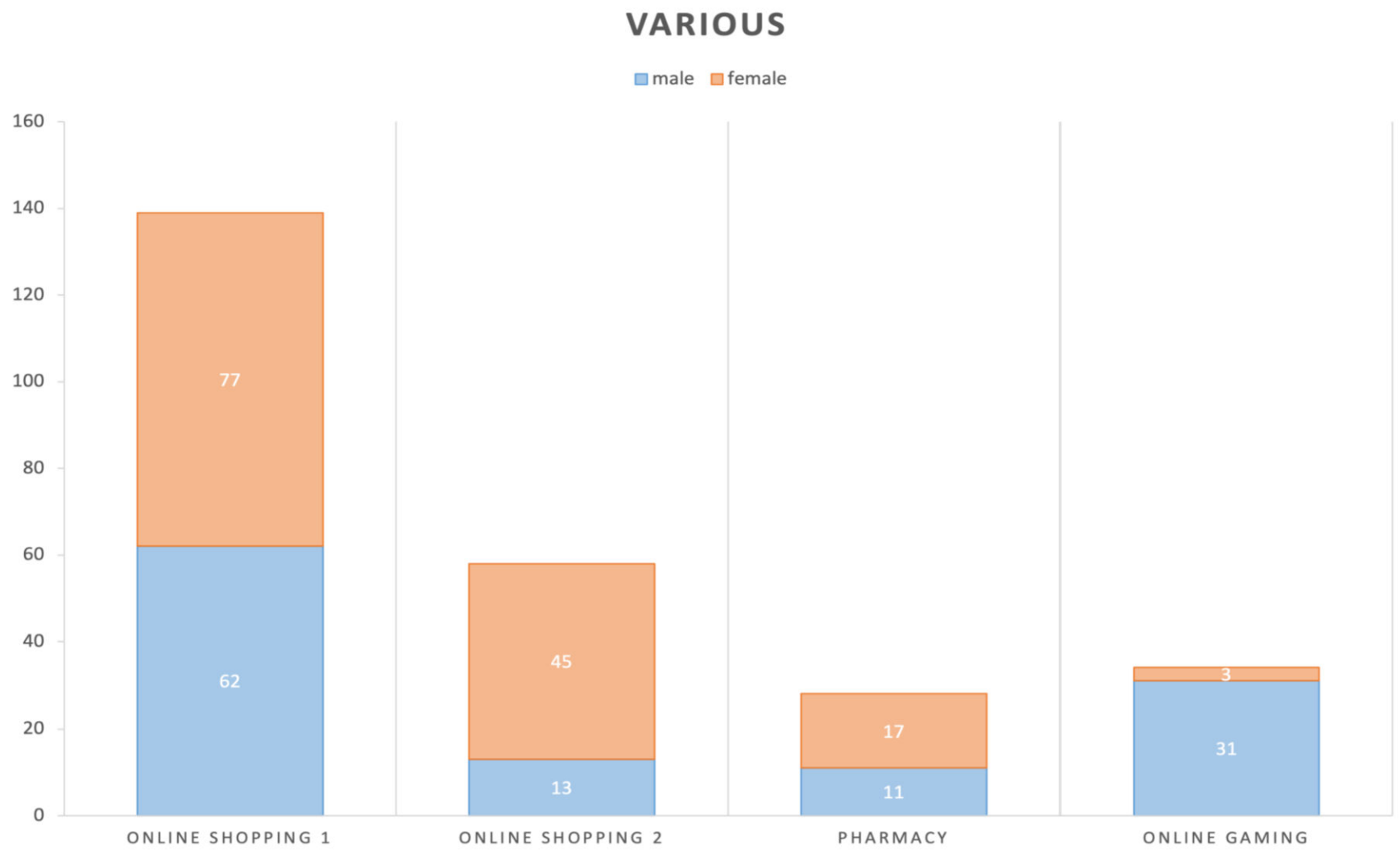
ONLINE SHOPS TECHNOLOGY



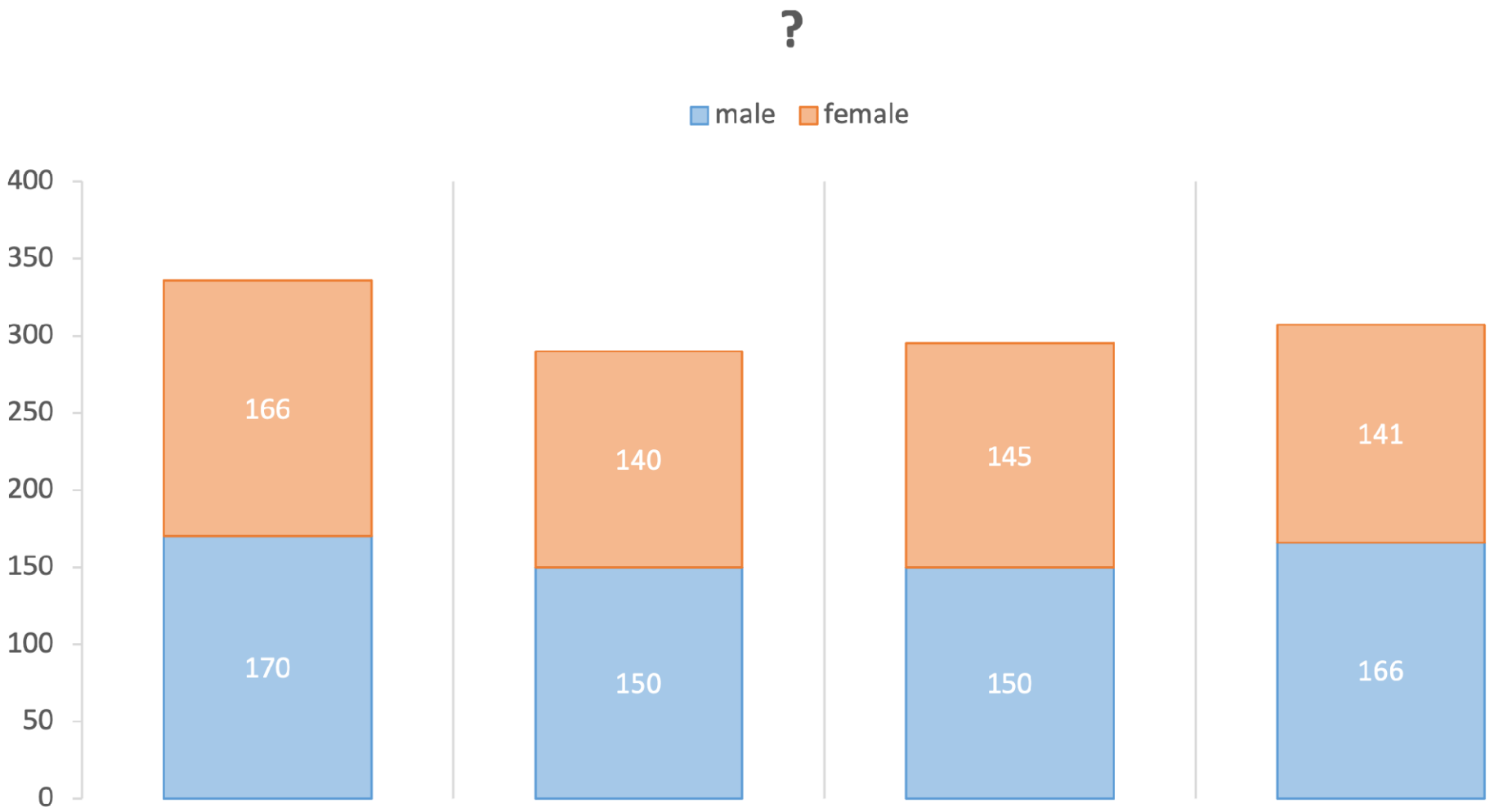
OBSERVATIONS II – GENDER ID DISTRIBUTION ADULT



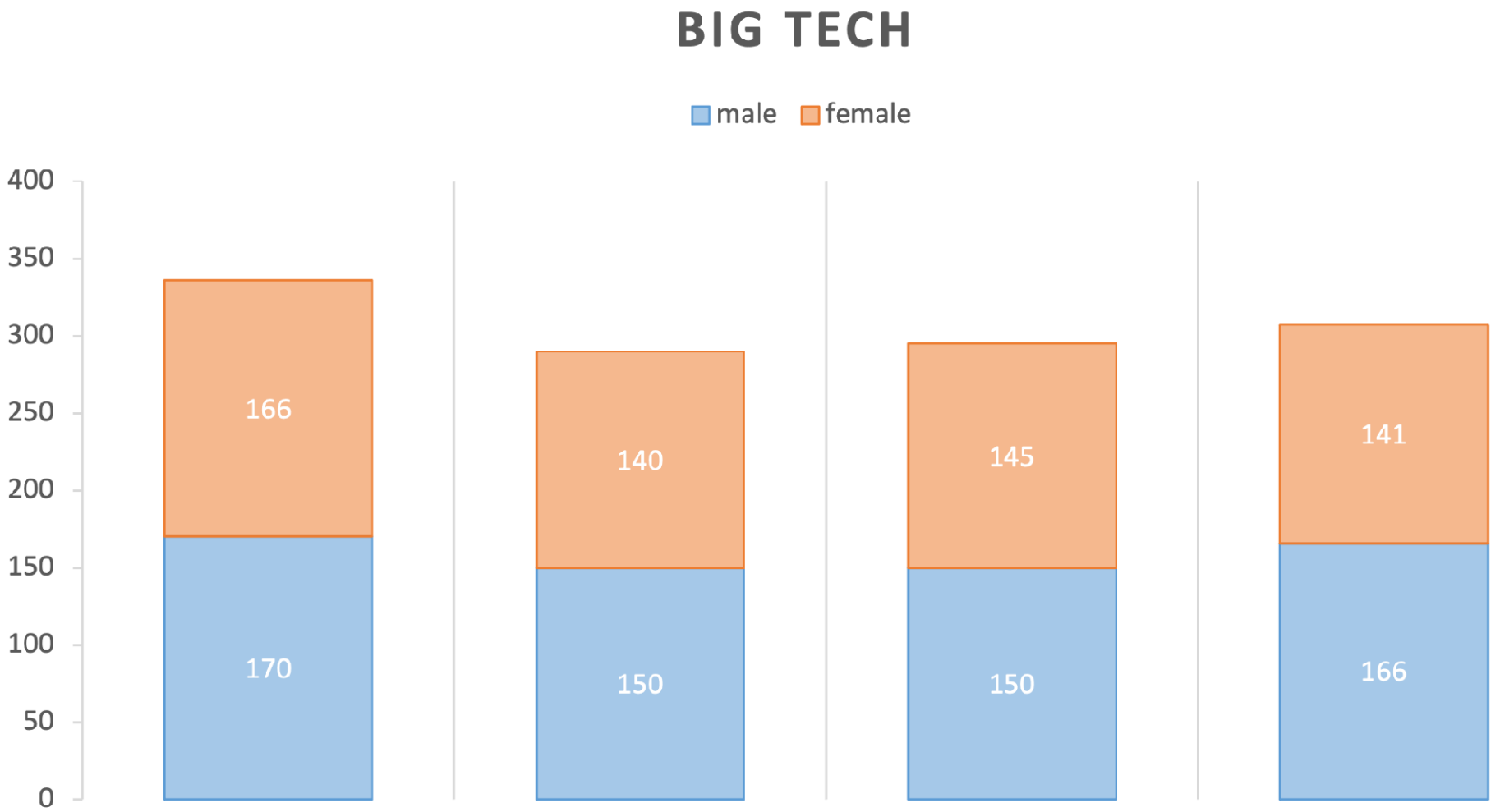
OBSERVATIONS II – GENDER ID DISTRIBUTION MISC



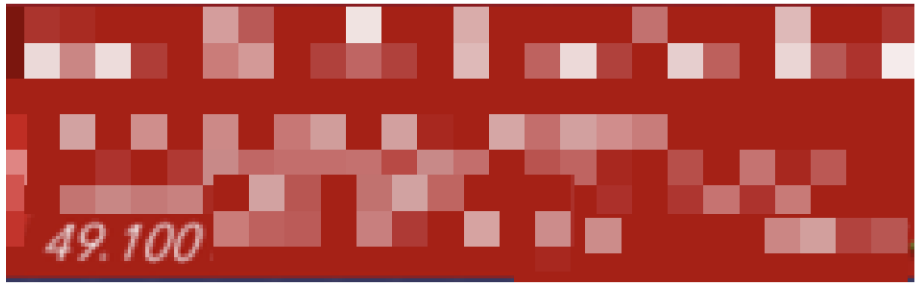
OBSERVATIONS II – ?



OBSERVATIONS II



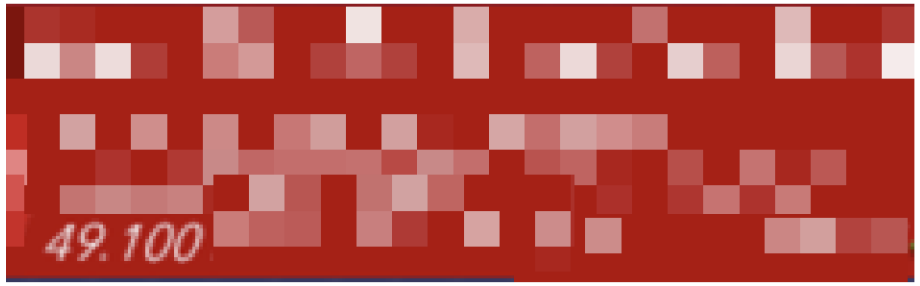
OBSERVATIONS III – XXX FORUM (STATISTICS)



Assumption:

- 100/8700000*49100 ~ 0.5% of population registered
- 8700000/49100 ~ 177, every 177th citizen registered

OBSERVATIONS III – XXX FORUM (STATISTICS)



Assumption:

- 100/8700000*49100 ~ 0.5% of population registered
- 8700000/49100 ~ 177, every 177th citizen registered

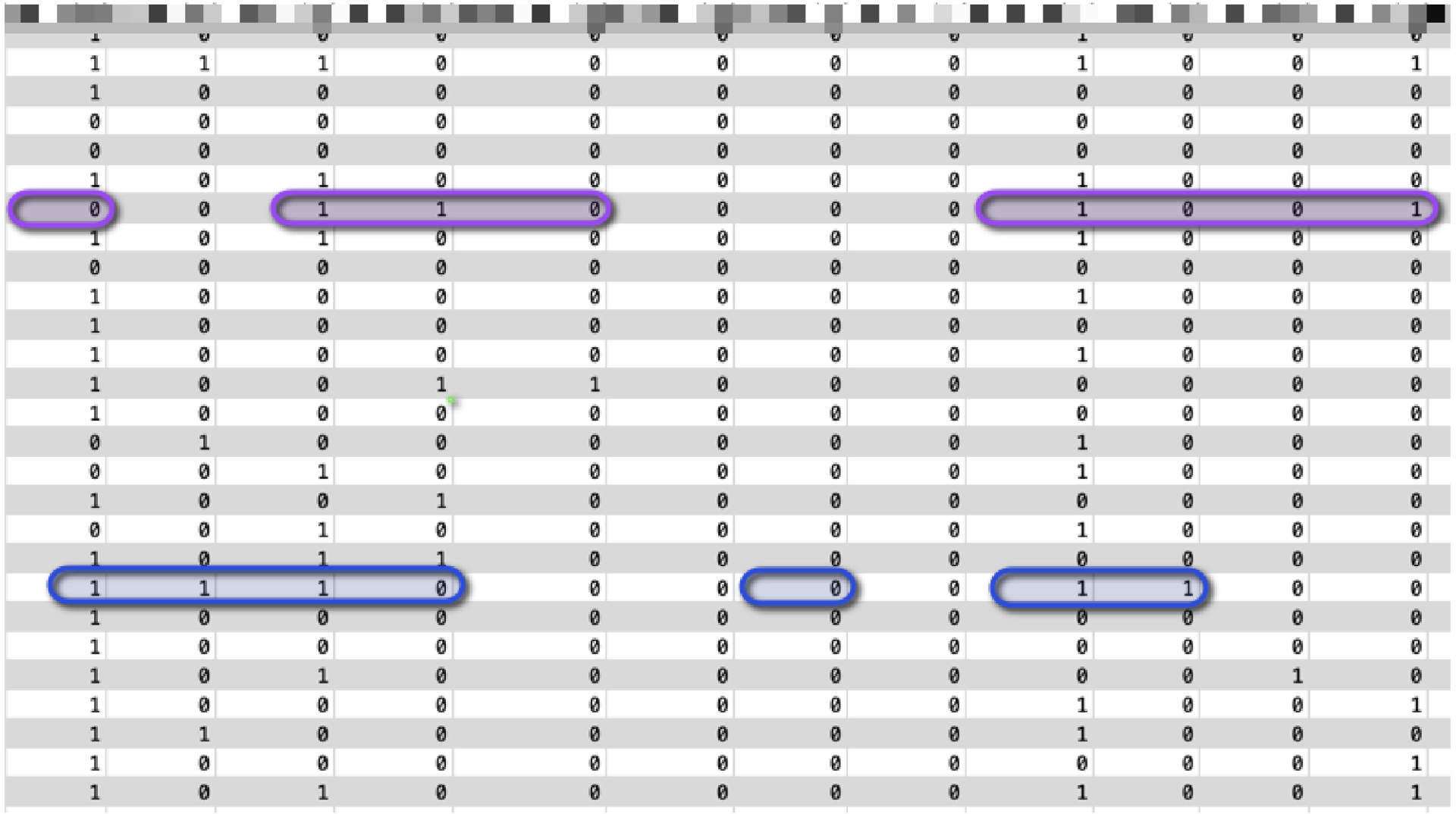
400 dataset:

- #144 and #353 registered
- blackmailable? → use throwaway e-mail address

InfoGuard
SWISS CYBER SECURITY

PROFILING

PROFILING I - FINDING RELATIONSHIPS: PROTOTYPES

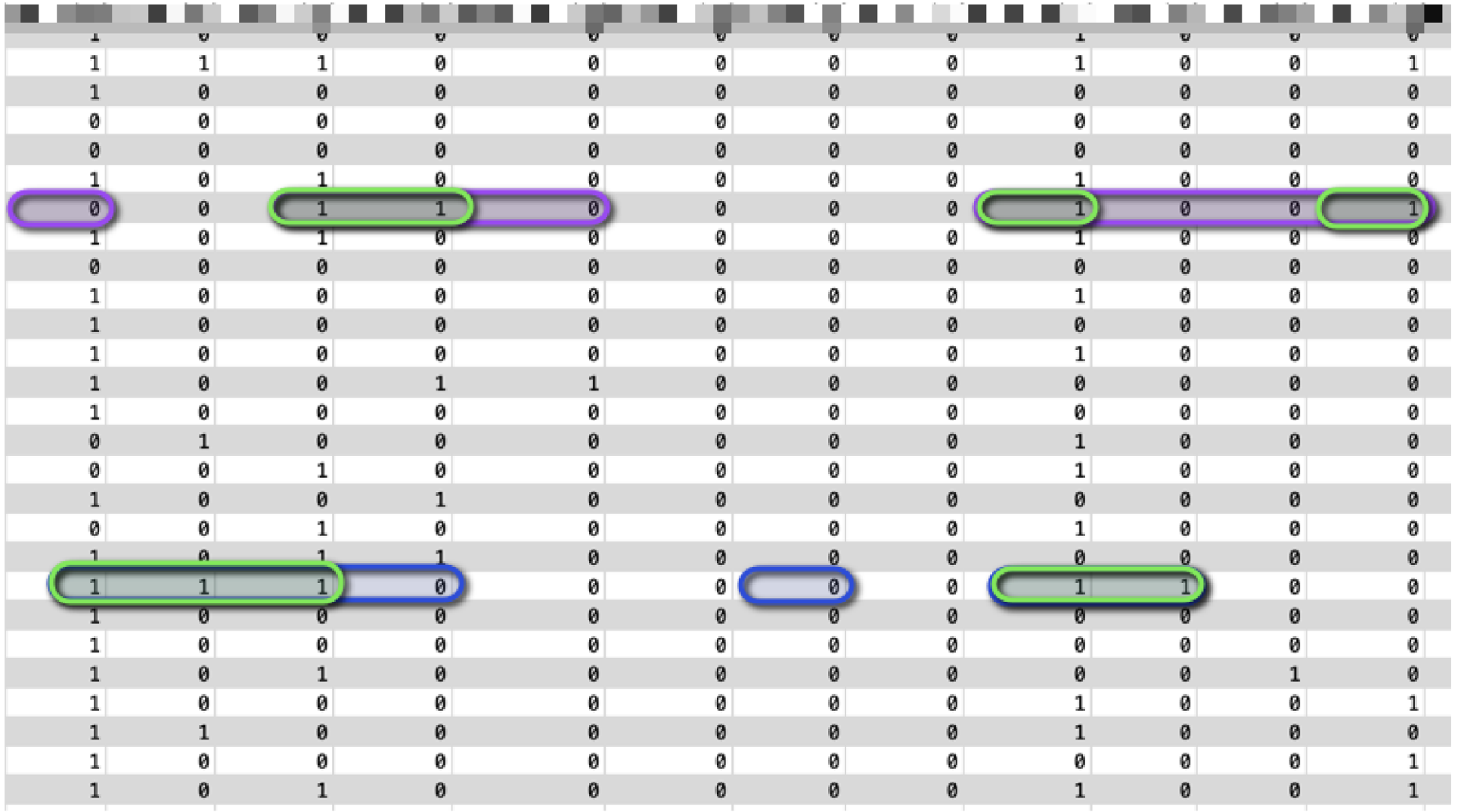


1	1	1	0	0	0	0	0	1	0	0	1
1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	1	0	0	0
0	0	1	1	0	0	0	0	1	0	0	1
1	0	1	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	1	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	1	0	0	0
1	0	0	1	1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	1	0	0	0
0	0	1	0	0	0	0	0	1	0	0	0
1	0	0	1	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	1	0	0	0
1	0	1	1	0	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	1	1	0	0
1	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	1	0
1	0	0	0	0	0	0	0	1	0	0	1
1	1	0	0	0	0	0	0	1	0	0	0
1	0	0	0	0	0	0	0	0	0	0	1
1	0	1	0	0	0	0	0	1	0	0	1

Prototype 1
„Outdoor enthusiast“

Prototype 2
„Hipster Coder“

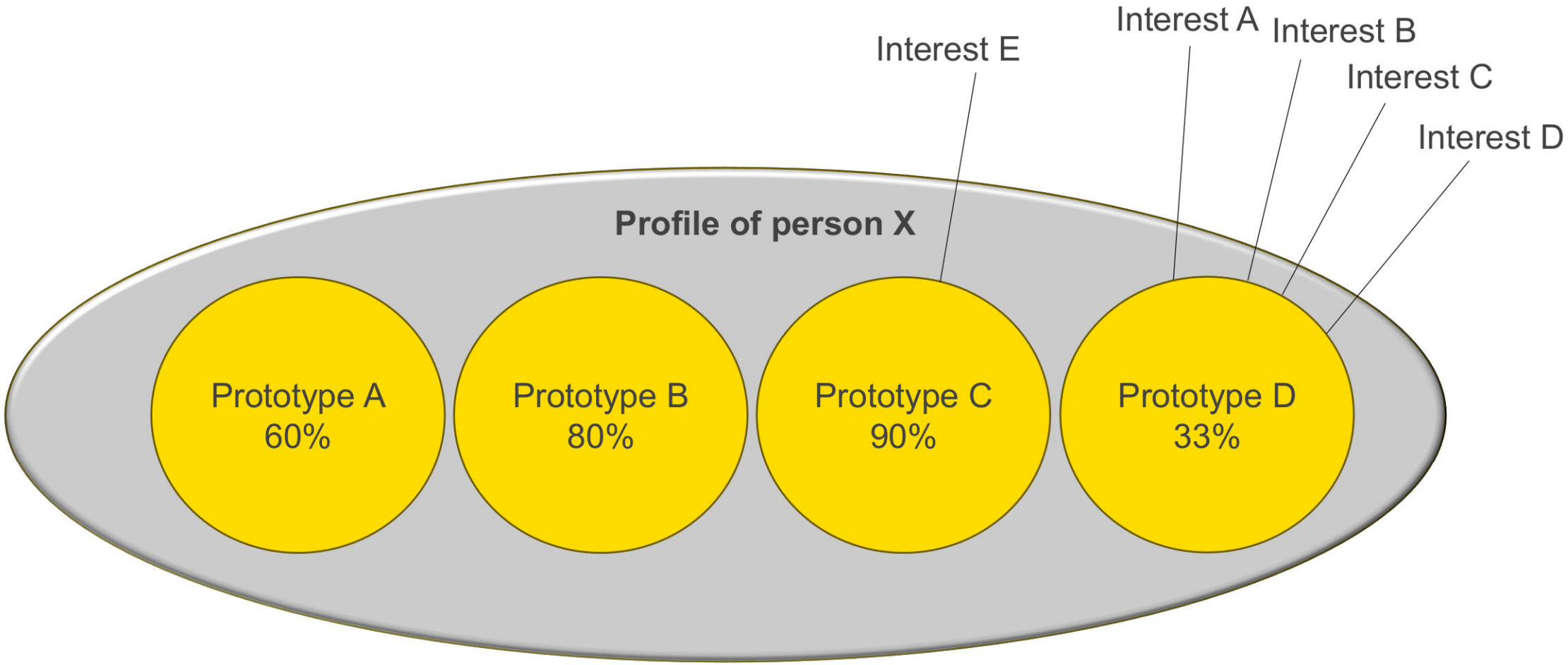
PROFILING I - FINDING RELATIONSHIPS: PROTOTYPES



Prototype 1
„Outdoor enthusiast“
50% match

Prototype 2
„Hipster programer“
71.4% match

PROFILING I - FINDING RELATIONSHIPS: PROTOTYPES



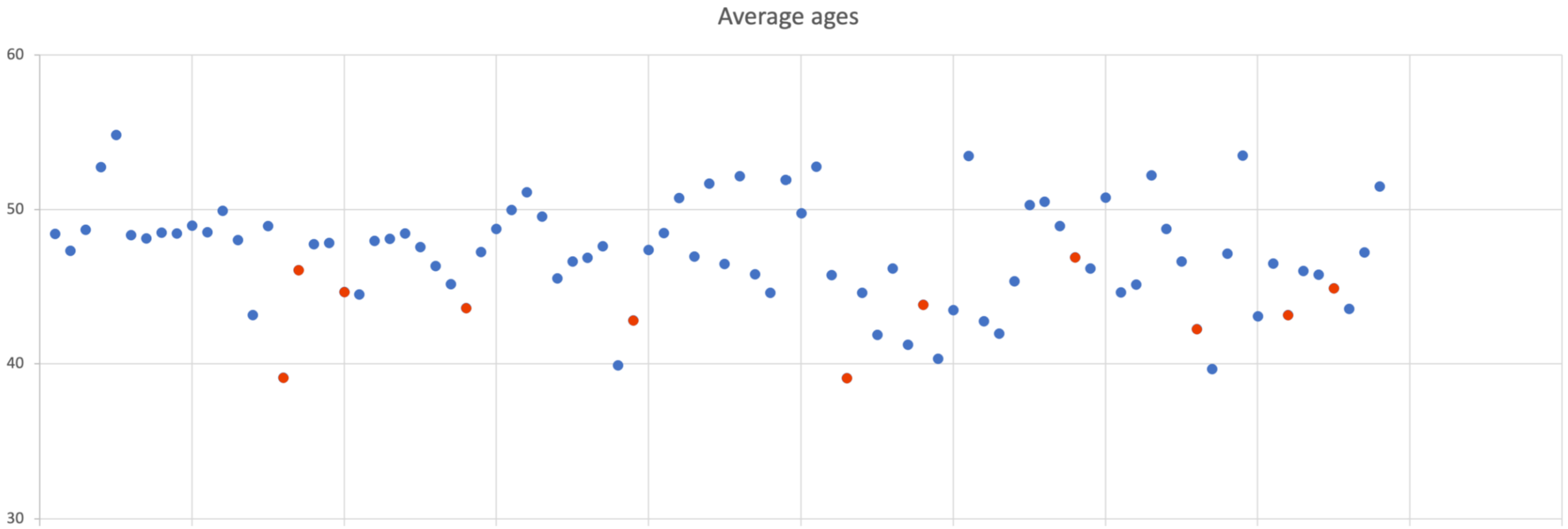
Fact that s.o. is NOT registered at plattform X may also be valuable (meta-)information (or # of regs)

PROFILING II – PREDICTING PROPERTIES

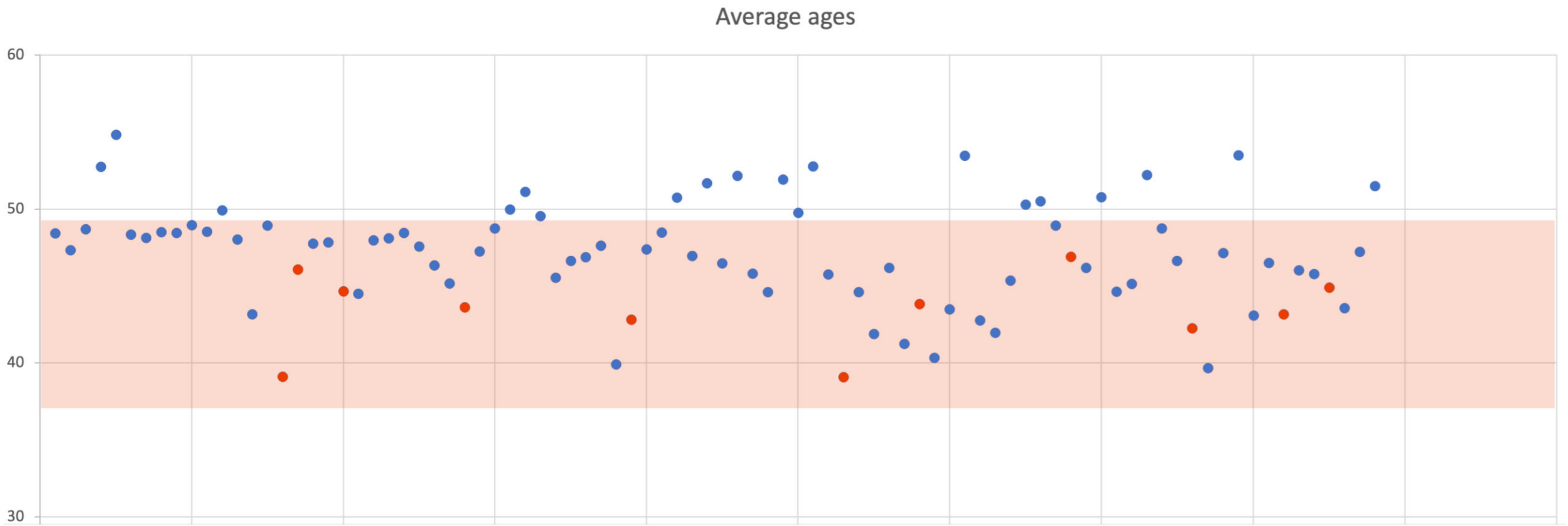
InfoGuard
SWISS CYBER SECURITY

- What if we do not know age, gender id, (profession)?
- Using the data and tool, can we predict these for an unknown person behind an arbitrary email?
- (estimate userbase?)

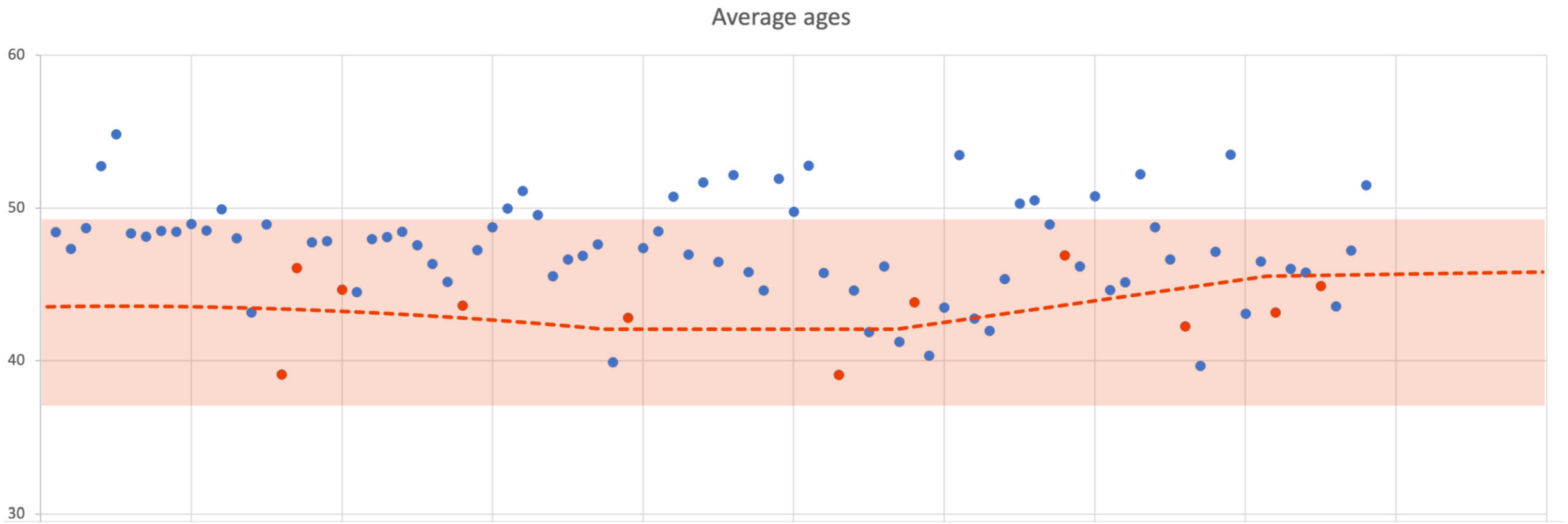
PROFILING II – PREDICTING AGE



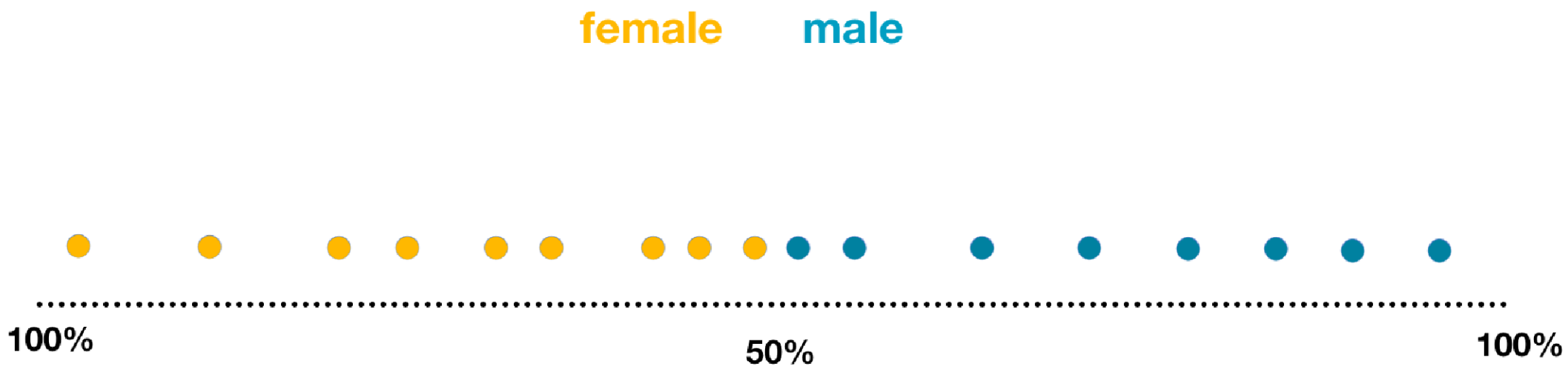
PROFILING II – PREDICTING AGE



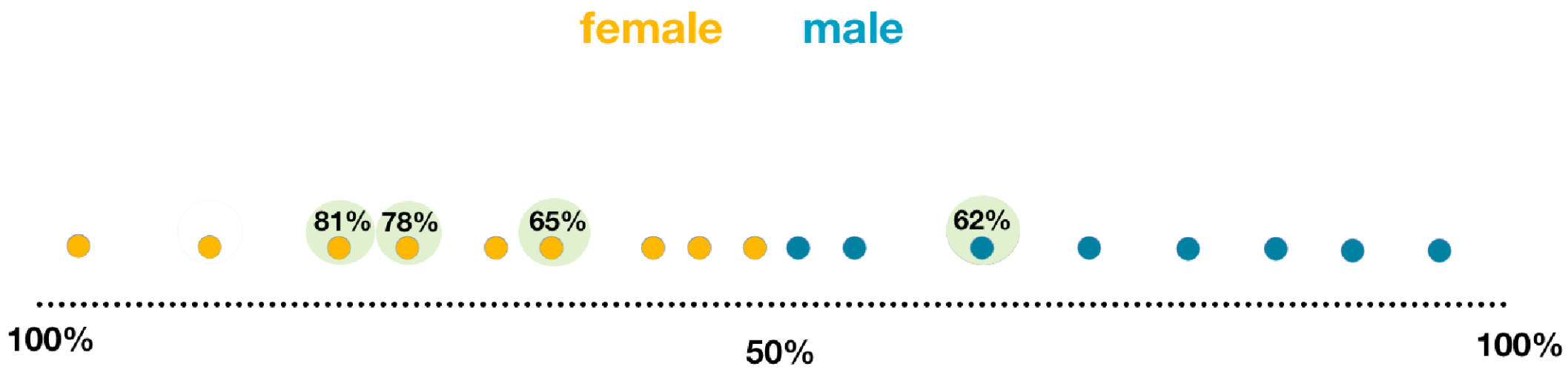
PROFILING II – PREDICTING AGE



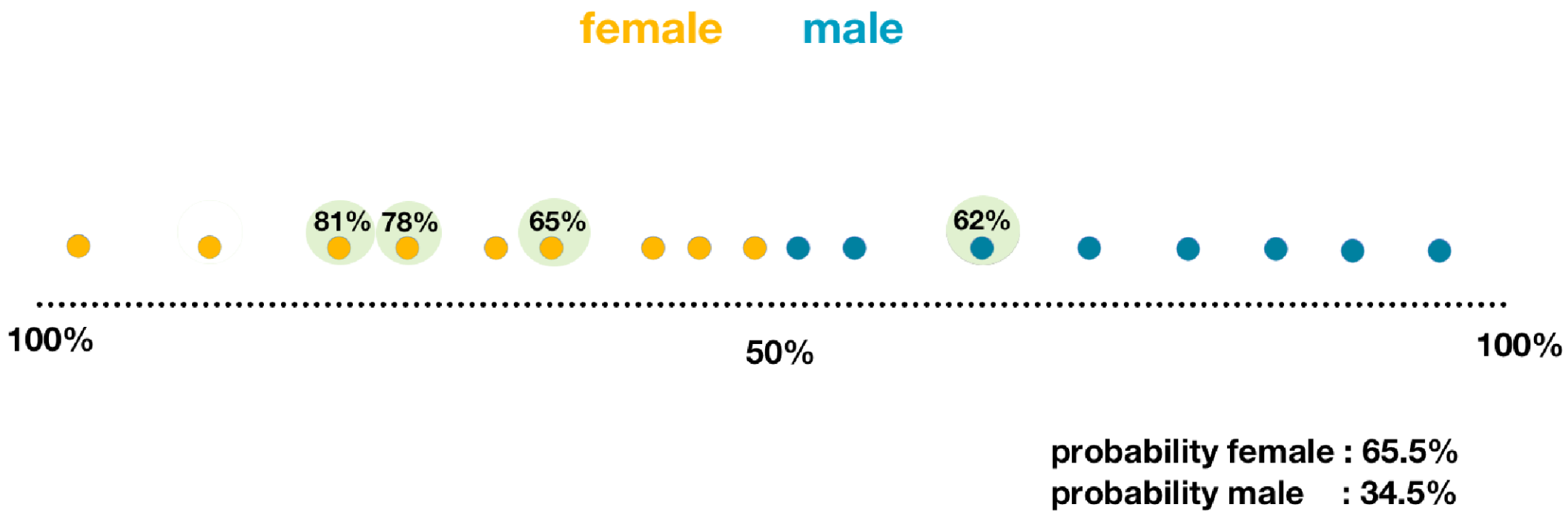
PROFILING II – PREDICTING GENDER ID



PROFILING II – PREDICTING GENDER ID



PROFILING II – PREDICTING GENDER ID

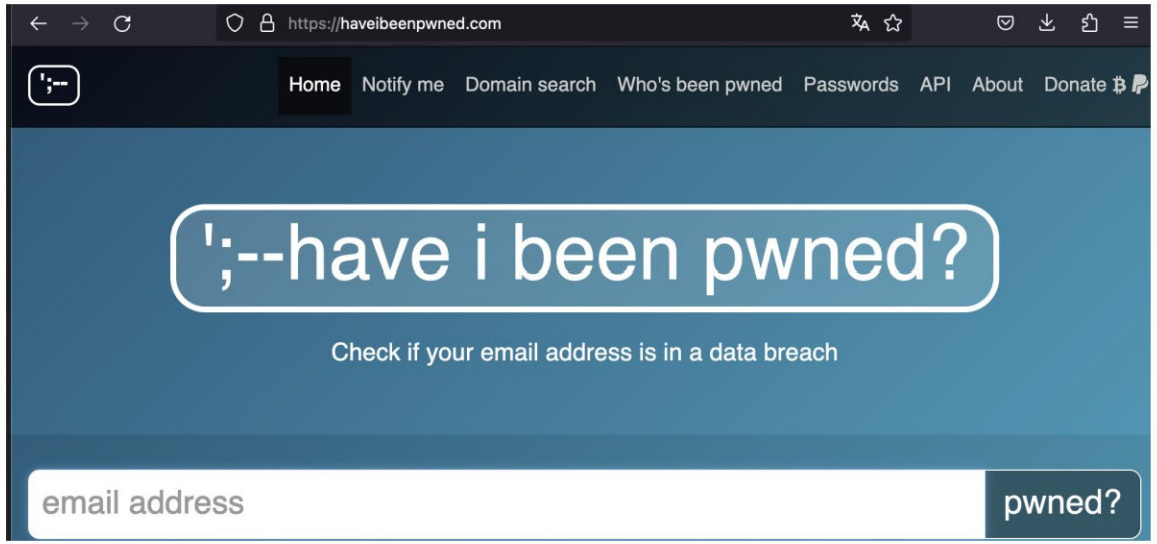


InfoGuard
SWISS CYBER SECURITY

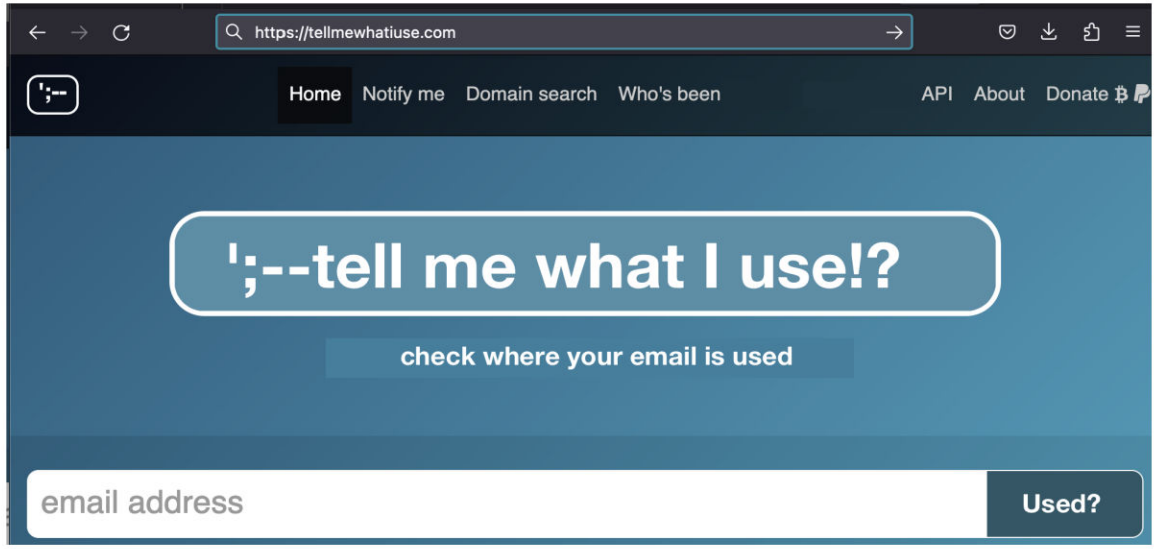
OUTLOOK

SERVICE?

haveibeenpwned.com



tellmewhatiuse.com



FINAL REMARKS

InfoGuard
SWISS CYBER SECURITY

- Still in data aquisition phase
- Analysis of results incomplete
- Trends already visible with few data (400)
- Does not scale well, a study would need more data to be meaningful
- Evaluating ML methods, model training, tbd
- More advanced predictions: political orientation? religion?
- If its possible, it will (or is) be done


InfoGuard
SWISS CYBER SECURITY
QUESTIONS/DISCUSSION

THANK YOUR FOR YOUR ATTENTION!

Lindenstrasse 10
6340 Baar / Schweiz
T +41 41 749 19 00

info@infoguard.ch
www.infoguard.ch

InfoGuard
SWISS CYBER SECURITY