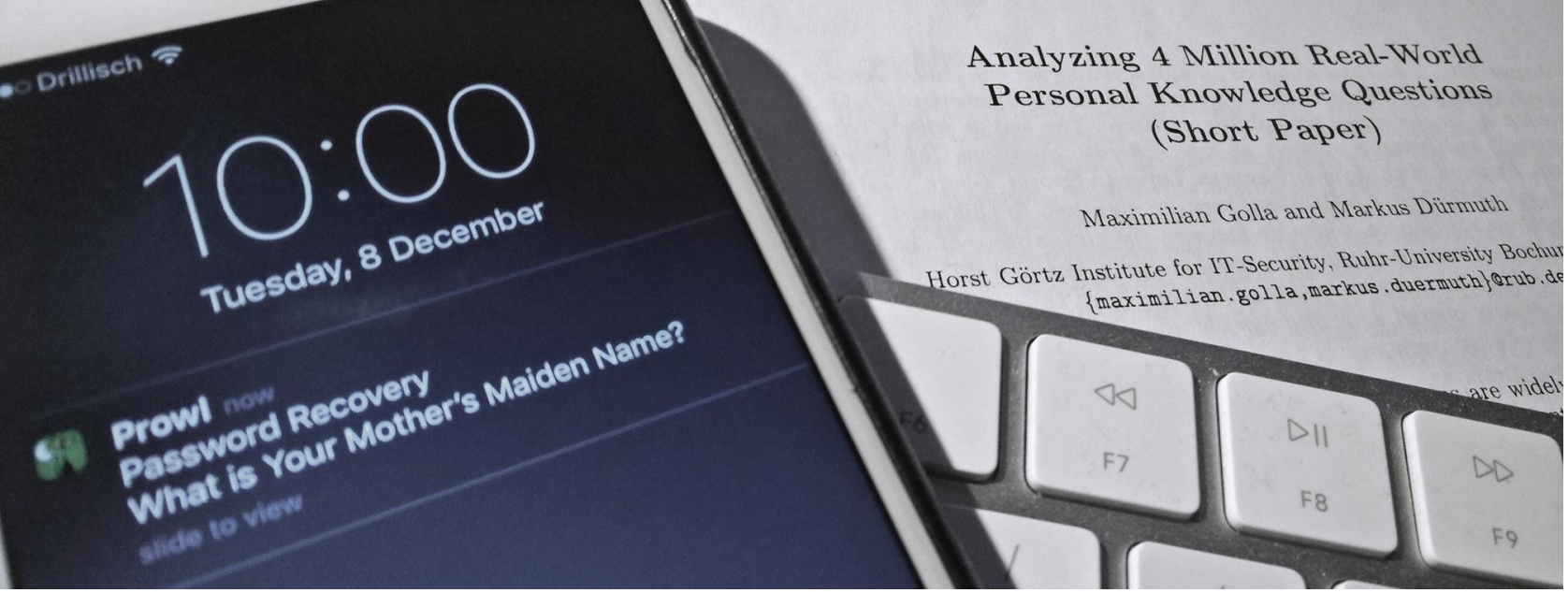


RUHR-UNIVERSITÄT BOCHUM

Analyzing 4 Million Real-World Personal Knowledge Questions

Maximilian Golla and Markus Dürmuth

Fallback Authentication?

Used to regain access if the primary means of authentication is lost!

Out-of-band Communication

Email, SMS, not always applicable (security, privacy, transport)

Personal Knowledge Questions

Rather insecure, answers can be guessed

Alternatives

Vouching, Preference Questions, Implicit Memory, …

Different to Primary Authentication

Memorability

Rate limiting

Time required to authenticate

[Image 1] "Retouched Kitty" by Ozan Kilic

PKQs Attack Scenarios

Targeted attacker: Specific user

Social networks are a good source for personal information

- Politician Sarah Palin, 2008
- WIRED author Mat Honan, 2012

Trawling attacker: Any user

Guesses answers based on population-wide statistics

Simultaneously attacks many accounts

In depth analysis of real-world data is given in [3]

[Ref. 3] Joseph Bonneau and Elie Bursztein and Ilan Caron and Rob Jackson and Mike Williamson. Secrets, Lies, and Account Recovery: Lessons from the Use of Personal Knowledge Questions at Google. (WWW'15)

[Image 1] "Isolated cat on white background" by Sergiu Bacioiu

Our Attacker Model

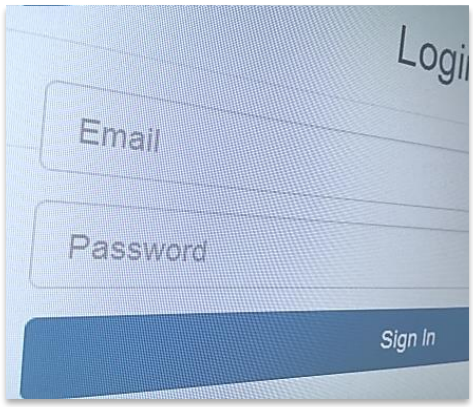
Statistical Guessing Attacks:

No knowledge about a specific user

Ideal attacker, knows approx. distribution of the answers

Resistance against …

Lower-bound security



Very limited no. of guesses allowed

Online Guessing Attacks

e.g., $\lambda_1 = 3.15\%$

Lower → more secure

Partial Guessing Entropy

Offline Guessing Attacks

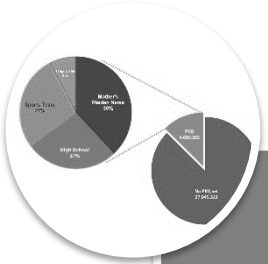
e.g., $G_{0.25} = 5.82$ bit

Higher → more secure

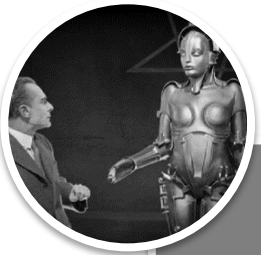
[Ref. 2] Joseph Bonneau. The Science of Guessing: Analyzing an Anonymized Corpus of 70 Million Passwords. (SP'12)

[Image 1] "Wordfence: Announces Password Auditing" by Mark Maunder

Outline



Results



Robots

Dataset

Introduction

[Image 1] "Metropolis" by "Fritz Lang" on gizmodo.com

Dataset

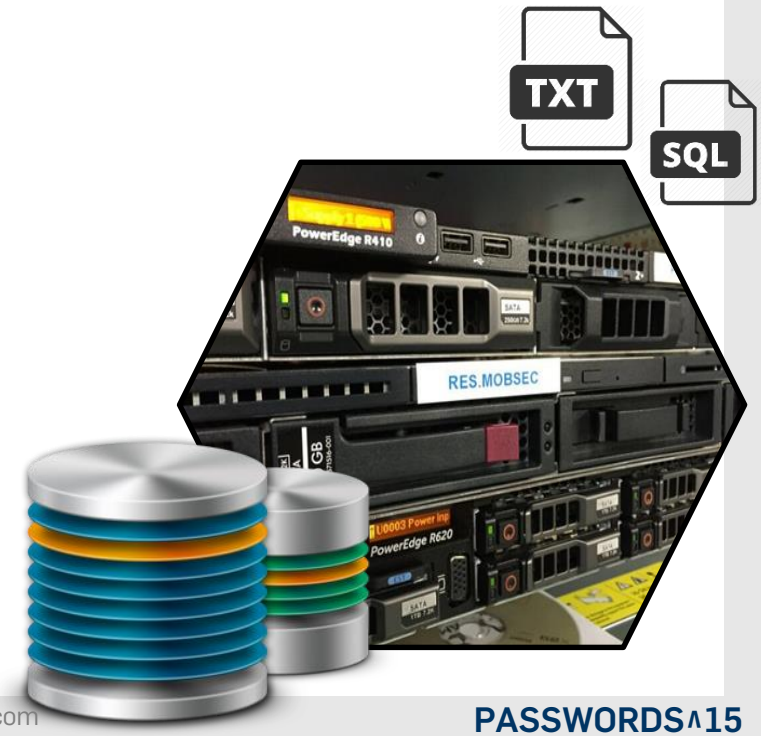
Dating Website

"Ashley Madison" opened in 2002, currently ~37 million "users"

Server Breach on 11. July 2015

Data was leaked via BitTorrent in August including

- Databases
- Source code
- Credit card transaction logs
- CEO emails



[Image 1] Icon by Freepik on flaticon.com, [Image 2] Icon by Barry Mieny on findicons.com

Database

Downloaded, Extracted, and Imported into DB

am_am.dump.gz (2.59GB)

~~CreditCardTransactions.7z~~, ~~member_login.dump.gz~~, ~~aminno_member_email.dump.gz~~, ...

Not used

am_am_member (Table)

id	...	security_question	security_answer
101234	...	4	1234
101235	...	3	MAPLE LEAFS
101236	...	0	**NULL**
101237	...	1	SMITH
101238	...	2	LINCOLN
...	...	...	...

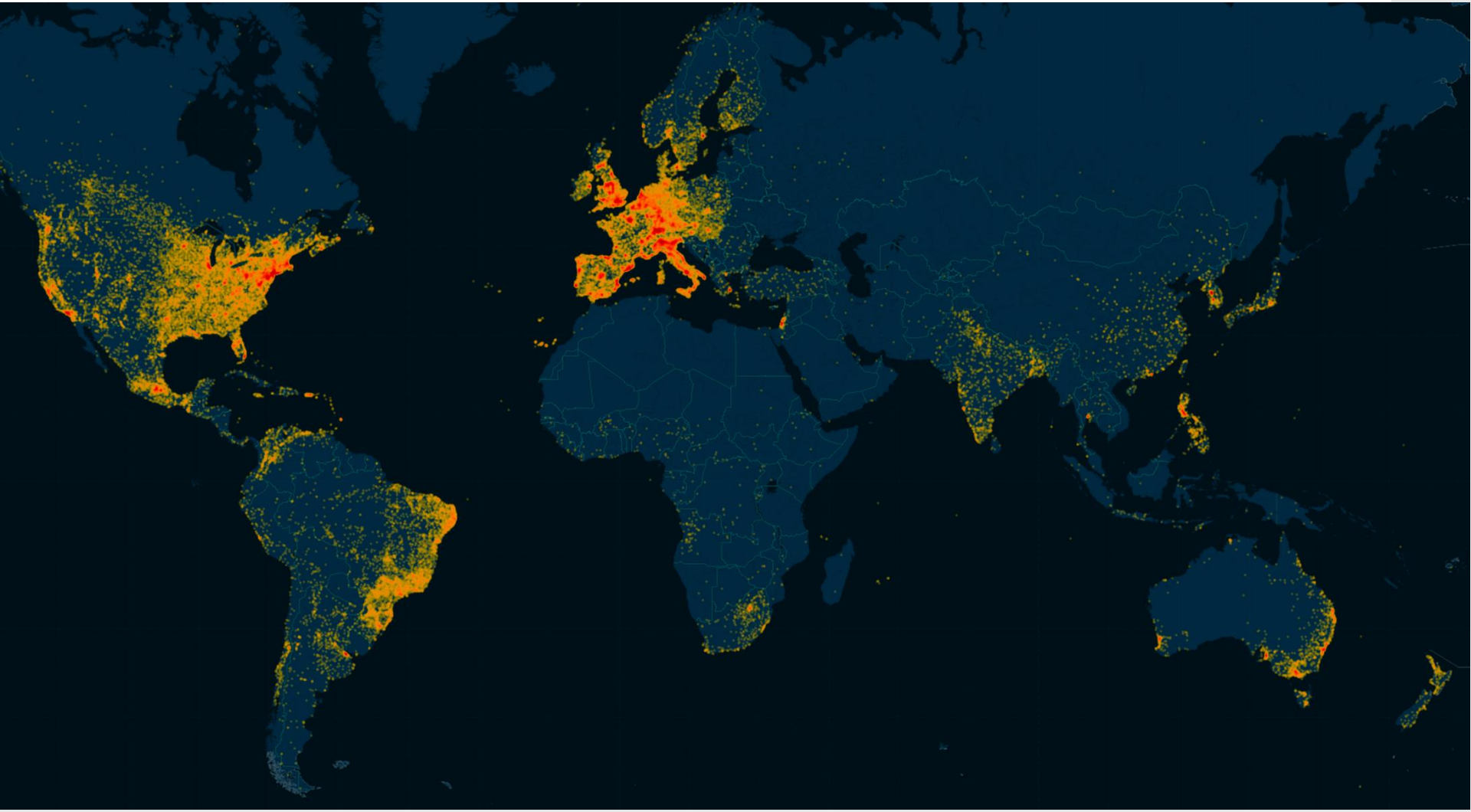
Bot filtering and aggregation

Used columns

Affected “Users”?

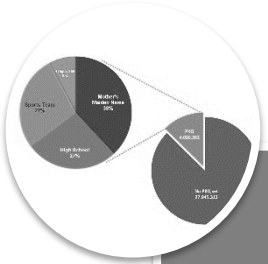
The “Malfideleco” map by Tecnilógica

Juan Alonso visualized the locations of the Ashley Madison users

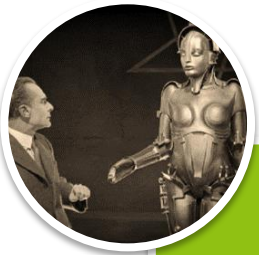


[Image 1] “Malfideleco" by Juan Alonso - blog.tecnilogica.com

Outline



Results



Robots

Dataset

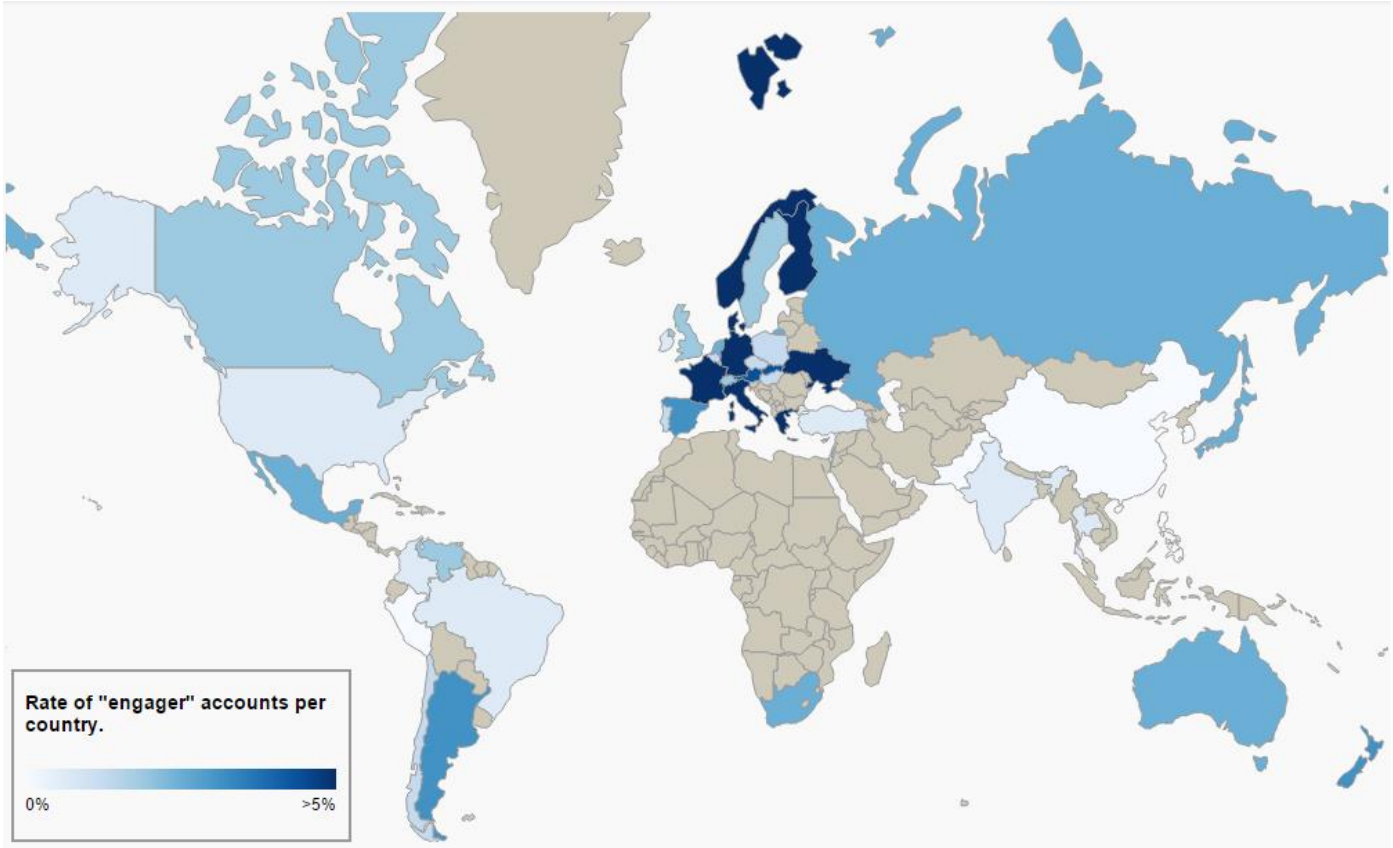
Introduction

Bots Called “Engagers”

Business Model:

Contacting new members requires “credits” (money)

Ashley Madison used bots to send millions of fake messages



[Image 1] "Map of female engagers" by Jake Perkowski

Bot Filter Criteria

Bot		Not a bot	
email	*@ashleymadison.com, …	**password** (bcrypt hash)	<paid_delete>, 111111Iwillneverdoitagain
bc_mail/chat_last_time	0000-00-00 00:00:00		
ishost	1		
ip address	“home addresses” 127.0.0.1/192.168.*/10.*		

Only available in payment information

Not used

[Ref. 1] “Ashley Madison Code Shows More Women, and More Bots” by Annalee Newitz

Ethical Considerations

Well, it is difficult …

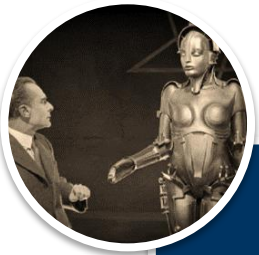
1. Data is public domain at this point
2. Aggregated data analysis, while ensuring total anonymity
3. Outcomes outweigh any potential harm?

Visit the tutorial "**To whom it's not concern**" (*Alexandra Strigunkova)* here at Passwords^15, today at 05:00 pm

[Image 1] Icon by Freepik on flaticon.com

Outline

Results



Robots

Dataset

Introduction

The Questions …

Security Question	Please Select ▼ We will ask you this security question if you forget your login details
Security Answer	 Your answer to the above security question will identify you as the owner of this account

- **What is Your Mother's Maiden Name?**
  SMITH, JONES, BROWN

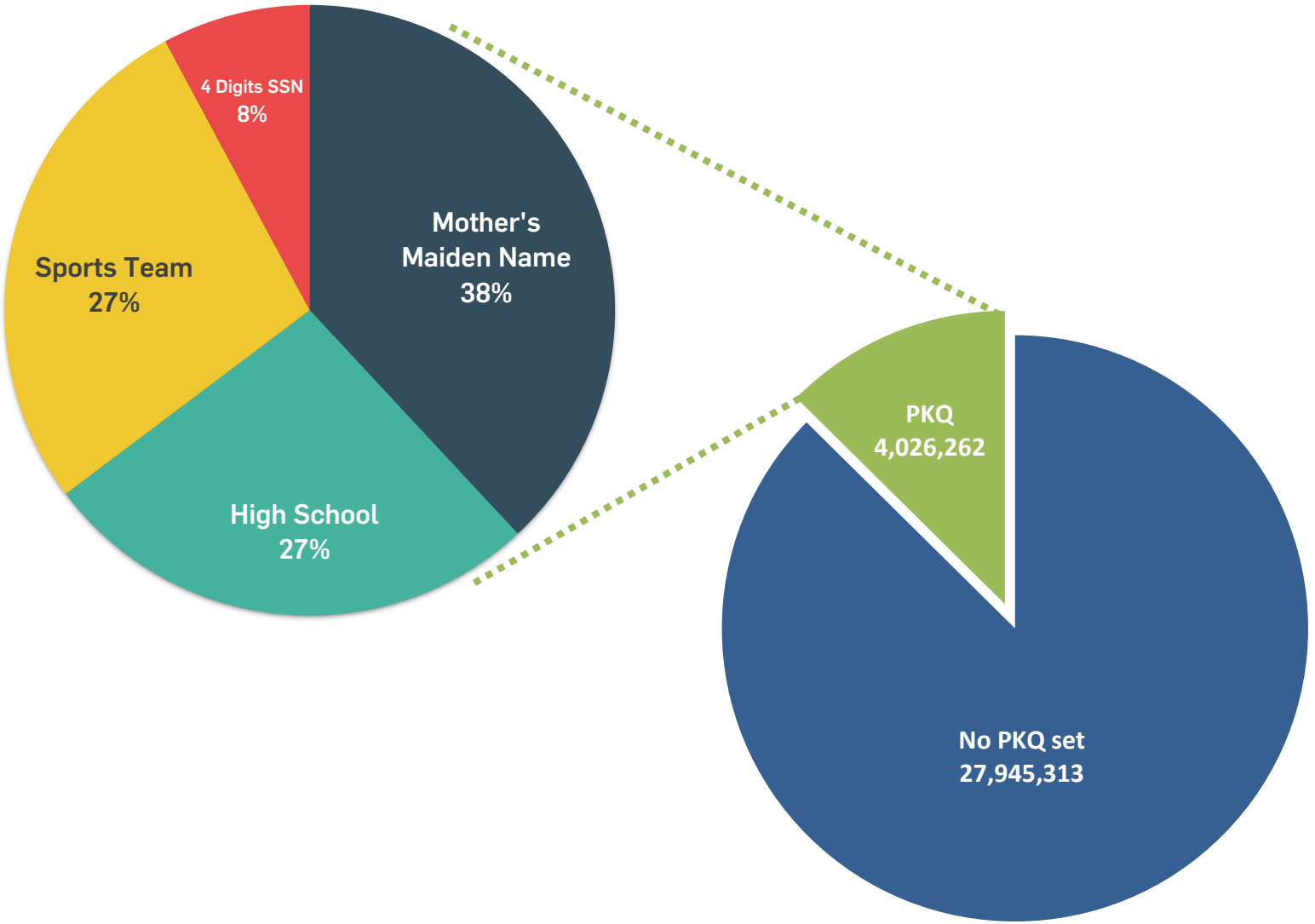
- **What is the Name of Your High School?**
  CENTRAL, LINCOLN, EAST

- **What is Your Favorite Sports Team?**
  YANKEES, COWBOYS, LEAFS

- **What are the last 4 Digits of Your SSN?**
  1234, 1111, 0000

[Image 1] Screenshot PKQs from ashleymadison.com (© Avid Life Media, Inc.), [Image 2] Icons by Freepik on flaticon.com

Basic Statistics



Significance

Re-sampling Approach:

1. Sample 10% of the original size
2. Test if the value falls into a 5% or a 10% error band or not, with probability p ≥ 0.98

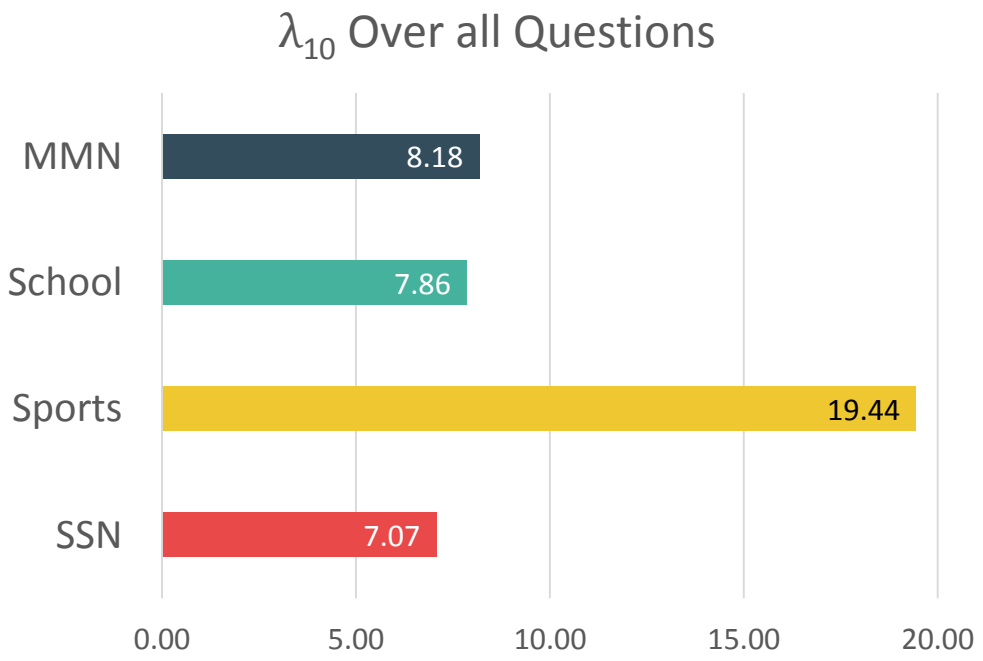
	size	online guessing (success in %) λ_1	λ_3	λ_{10}	λ_{100}	λ_{1000}
mother's maiden name						
all	903 255	3.21	5.12	8.18	22.41	48.10
country USA	612 890	*3.15*	5.31	8.89	24.41	51.74
Canada	125 101	-	*4.09*	*6.91*	*21.64*	*48.49*

Comparable to: [Ref. 3] Joseph Bonneau and Elie Bursztein and Ilan Caron and Rob Jackson and Mike Williamson. Secrets, Lies, and Account Recovery: Lessons from the Use of Personal Knowledge Questions at Google. (WWW'15)

MMN	Size	λ_1	λ_3	λ_{10}	λ_{100}
All	903 255	**3.21**	**5.12**	**8.18**	**22.41**
USA	612 890	3.15	**5.31**	**8.89**	**24.41**
Canada	125 101	-	4.09	6.91	21.64
UK	26 912	-	-	-	-
<= 35	169 571	2.65	4.06	6.79	**20.73**
36 - 45	293 868	3.08	5.12	**8.08**	**22.63**
46 - 55	284 594	3.48	5.55	**8.82**	**23.38**
> 55	155 222	3.58	5.50	8.91	**23.74**
School					
All	632 484	**2.63**	**5.60**	**7.86**	**16.91**
USA	461 582	2.68	**6.03**	**8.55**	**18.25**
Canada	86 465	-	-	9.83	26.79
UK	8 740	-	-	-	-
<= 35	127 707	2.56	5.76	7.92	16.45
36 - 45	217 157	2.57	**5.51**	**7.72**	**16.88**
46 - 55	194 608	2.80	5.53	**7.87**	**17.27**
> 55	93 012	-	5.86	8.39	18.60
Sports					
All	650 680	**2.83**	**7.84**	**19.44**	**62.94**
USA	432 129	**4.11**	**10.75**	**26.51**	**74.15**
Canada	85 520	11.16	**18.61**	**36.03**	**74.61**
UK	12 011	-	23.23	41.93	73.49
<= 35	128 520	-	6.03	**16.38**	**58.78**
36 - 45	250 901	2.73	**7.69**	**19.73**	**63.39**
46 - 55	199 321	3.07	**8.67**	**21.94**	**65.83**
> 55	71 938	-	9.48	**23.49**	**66.95**
SSN					
All	186 134	-	5.32	7.07	-
USA	128 611	-	5.15	6.73	-
Baseline					
PIN (Rnd.)	4-digit	0.01	0.03	0.1	1.0
PW	RockYou	0.9	1.4	2.1	4.6
PIN (Real)	4-digit	4.3	9.2	14.4	29.3

Diff. in Questions

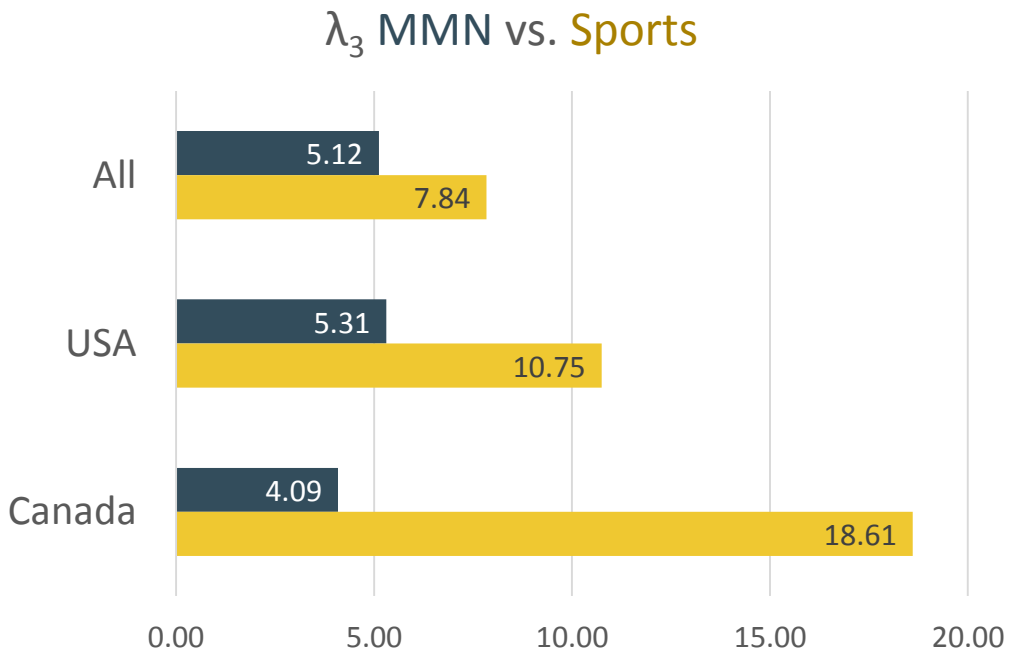
- λ_1 is consistent over all questions
- Sports Team is very weak!



MMN	Size	λ_1	λ_3	λ_{10}	λ_{100}
All	903 255	**3.21**	**5.12**	**8.18**	**22.41**
USA	612 890	3.15	**5.31**	**8.89**	**24.41**
Canada	125 101	-	4.09	6.91	21.64
UK	26 912	-	-	-	-
<= 35	169 571	2.65	4.06	6.79	**20.73**
36 - 45	293 868	3.08	5.12	**8.08**	**22.63**
46 - 55	284 594	3.48	5.55	**8.82**	**23.38**
> 55	155 222	3.58	5.50	8.91	**23.74**
School					
All	632 484	**2.63**	**5.60**	**7.86**	**16.91**
USA	461 582	2.68	**6.03**	**8.55**	**18.25**
Canada	86 465	-	-	9.83	26.79
UK	8 740	-	-	-	-
<= 35	127 707	2.56	5.76	7.92	16.45
36 - 45	217 157	2.57	**5.51**	**7.72**	**16.88**
46 - 55	194 608	2.80	5.53	**7.87**	**17.27**
> 55	93 012	-	5.86	8.39	18.60
Sports					
All	650 680	**2.83**	**7.84**	**19.44**	**62.94**
USA	432 129	**4.11**	**10.75**	**26.51**	**74.15**
Canada	85 520	11.16	**18.61**	**36.03**	**74.61**
UK	12 011	-	23.23	41.93	73.49
<= 35	128 520	-	6.03	**16.38**	**58.78**
36 - 45	250 901	2.73	**7.69**	**19.73**	**63.39**
46 - 55	199 321	3.07	**8.67**	**21.94**	**65.83**
> 55	71 938	-	9.48	**23.49**	**66.95**
SSN					
All	186 134	-	5.32	7.07	-
USA	128 611	-	5.15	6.73	-
Baseline					
PIN (Rnd.)	4-digit	0.01	0.03	0.1	1.0
PW	RockYou	0.9	1.4	2.1	4.6
PIN (Real)	4-digit	4.3	9.2	14.4	29.3

Diff. in Nationality

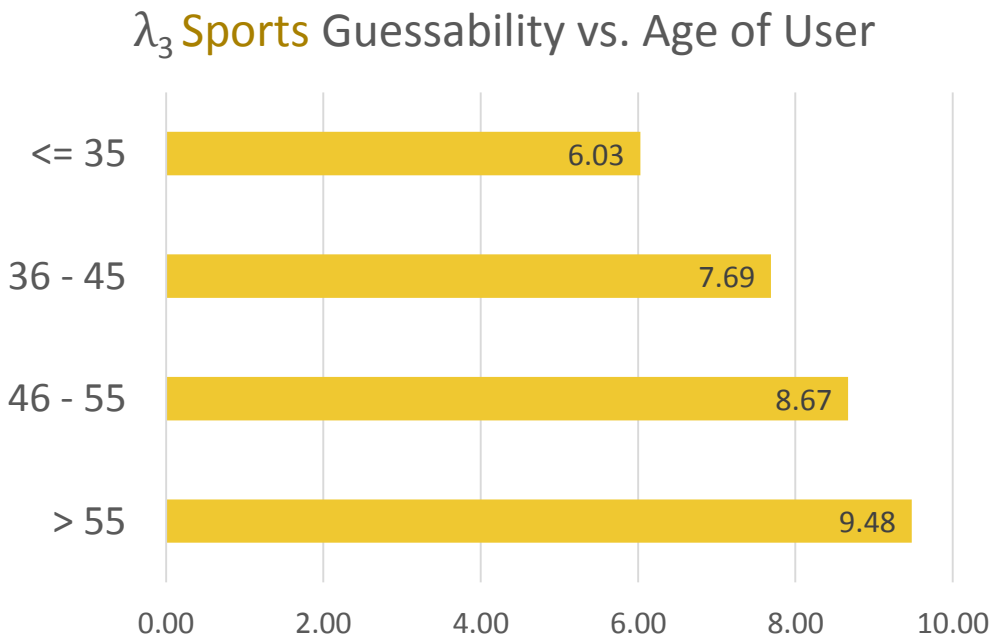
- Origin of user influences Sports
- Canada's teams are easier to guess
- Fake SSNs are popular: "1234" (2.23%)



MMN	Size	λ_1	λ_3	λ_{10}	λ_{100}
All	903 255	**3.21**	**5.12**	**8.18**	**22.41**
USA	612 890	3.15	**5.31**	**8.89**	**24.41**
Canada	125 101	-	4.09	6.91	21.64
UK	26 912	-	-	-	-
<= 35	169 571	2.65	4.06	6.79	**20.73**
36 - 45	293 868	3.08	5.12	**8.08**	**22.63**
46 - 55	284 594	3.48	5.55	**8.82**	**23.38**
> 55	155 222	3.58	5.50	8.91	**23.74**
School					
All	632 484	**2.63**	**5.60**	**7.86**	**16.91**
USA	461 582	2.68	**6.03**	**8.55**	**18.25**
Canada	86 465	-	-	9.83	26.79
UK	8 740	-	-	-	-
<= 35	127 707	2.56	5.76	7.92	16.45
36 - 45	217 157	2.57	**5.51**	**7.72**	**16.88**
46 - 55	194 608	2.80	5.53	**7.87**	**17.27**
> 55	93 012	-	5.86	8.39	18.60
Sports					
All	650 680	**2.83**	**7.84**	**19.44**	**62.94**
USA	432 129	**4.11**	**10.75**	**26.51**	**74.15**
Canada	85 520	11.16	**18.61**	**36.03**	**74.61**
UK	12 011	-	23.23	41.93	73.49
<= 35	128 520	-	6.03	**16.38**	**58.78**
36 - 45	250 901	2.73	**7.69**	**19.73**	**63.39**
46 - 55	199 321	3.07	**8.67**	**21.94**	**65.83**
> 55	71 938	-	9.48	**23.49**	**66.95**
SSN					
All	186 134	-	5.32	7.07	-
USA	128 611	-	5.15	6.73	-
Baseline					
PIN (Rnd.)	4-digit	0.01	0.03	0.1	1.0
PW	RockYou	0.9	1.4	2.1	4.6
PIN (Real)	4-digit	4.3	9.2	14.4	29.3

Diff. in Age

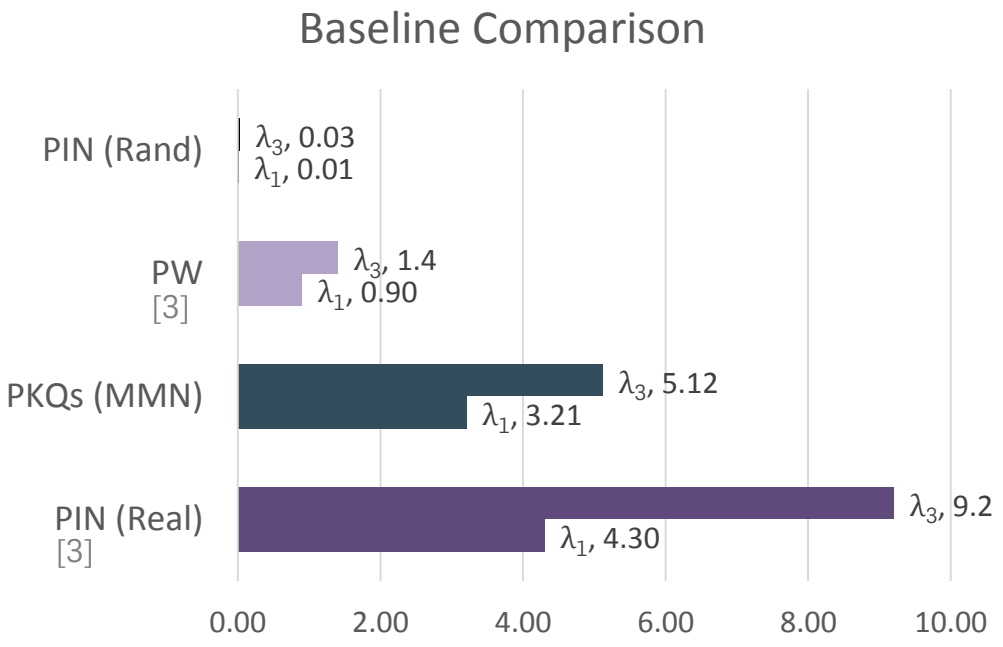
- Age has a little influence on guessability
- With increasing age the answers are slightly easier to guess



MMN	Size	λ_1	λ_3	λ_{10}	λ_{100}
All	903 255	**3.21**	**5.12**	**8.18**	**22.41**
USA	612 890	3.15	**5.31**	**8.89**	**24.41**
Canada	125 101	-	4.09	6.91	21.64
UK	26 912	-	-	-	-
<= 35	169 571	2.65	4.06	6.79	**20.73**
36 - 45	293 868	3.08	5.12	**8.08**	**22.63**
46 - 55	284 594	3.48	5.55	**8.82**	**23.38**
> 55	155 222	3.58	5.50	8.91	**23.74**
School					
All	632 484	**2.63**	**5.60**	**7.86**	**16.91**
USA	461 582	2.68	**6.03**	**8.55**	**18.25**
Canada	86 465	-	-	9.83	26.79
UK	8 740	-	-	-	-
<= 35	127 707	2.56	5.76	7.92	16.45
36 - 45	217 157	2.57	**5.51**	**7.72**	**16.88**
46 - 55	194 608	2.80	5.53	**7.87**	**17.27**
> 55	93 012	-	5.86	8.39	18.60
Sports					
All	650 680	**2.83**	**7.84**	**19.44**	**62.94**
USA	432 129	**4.11**	**10.75**	**26.51**	**74.15**
Canada	85 520	11.16	**18.61**	**36.03**	**74.61**
UK	12 011	-	23.23	41.93	73.49
<= 35	128 520	-	6.03	**16.38**	**58.78**
36 - 45	250 901	2.73	**7.69**	**19.73**	**63.39**
46 - 55	199 321	3.07	**8.67**	**21.94**	**65.83**
> 55	71 938	-	9.48	**23.49**	**66.95**
SSN					
All	186 134	-	5.32	7.07	-
USA	128 611	-	5.15	6.73	-
Baseline					
PIN (Rnd.)	4-digit	0.01	0.03	0.1	1.0
PW	RockYou	0.9	1.4	2.1	4.6
PIN (Real)	4-digit	4.3	9.2	14.4	29.3

Comparison

- All questions are slightly more secure than a **real-world** 4 digit PIN
- All questions are less secure than a **random** 4 digit PIN



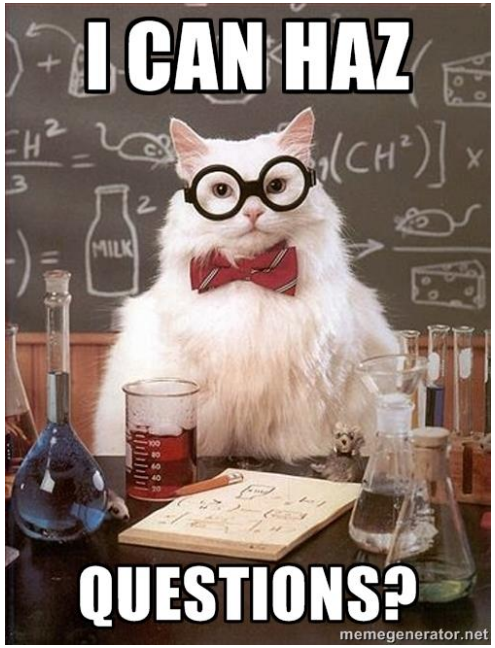
[Ref. 3] Joseph Bonneau and Elie Bursztein and Ilan Caron and Rob Jackson and Mike Williamson. Secrets, Lies, and Account Recovery: Lessons from the Use of Personal Knowledge Questions at Google. (WWW'15)

Takeaway

- Personal knowledge questions offer only a low level of security
- Security depends on age and origin of the user
- Sports Team question is particular easy to guess

[Image 1] “Broken Rusty Lock: Security" by Nick Carter

Questions?



- Personal knowledge questions offer only a low level of security
- Security depends on age and origin of the user
- Sports Team question is particular easy to guess

[Image 1] "Chemistry Cat" by memegenerator.net