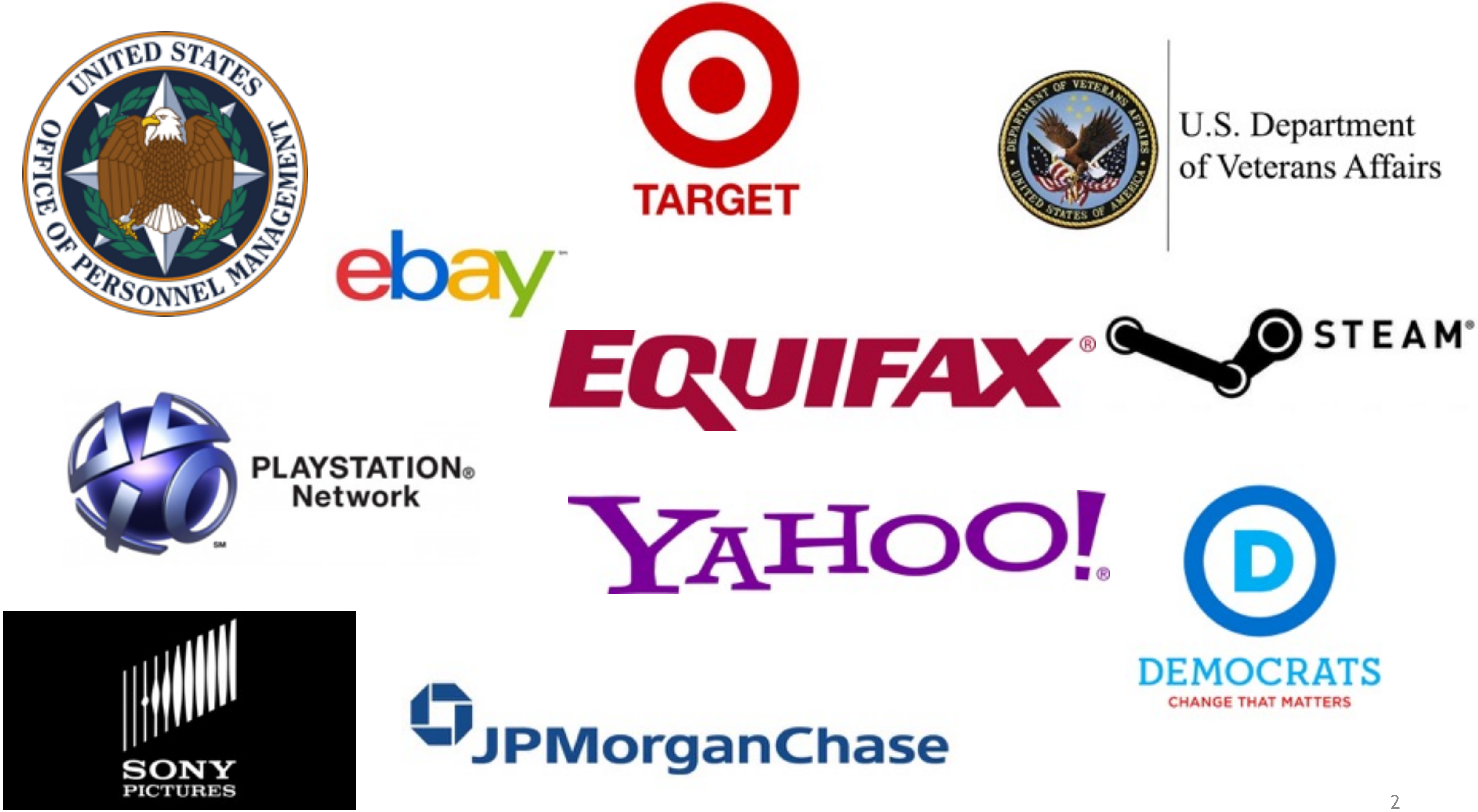
Pixek

Seny Kamara,Tarik Moataz, Martin Zhu


UNITED STATES OFFICE OF PERSONNEL MANAGEMENT
TARGET
DEPARTMENT OF VETERANS AFFAIRS UNITED STATES OF AMERICA
U.S. Department of Veterans Affairs
ebay
EQUIFAX
STEAM
PLAYSTATION Network
YAHOO!
DEMOCRATS
CHANGE THAT MATTERS
SONY PICTURES
JPMorganChase

9,198,580,293*

4%

* since 2013

Why so Few?

Incompetence?

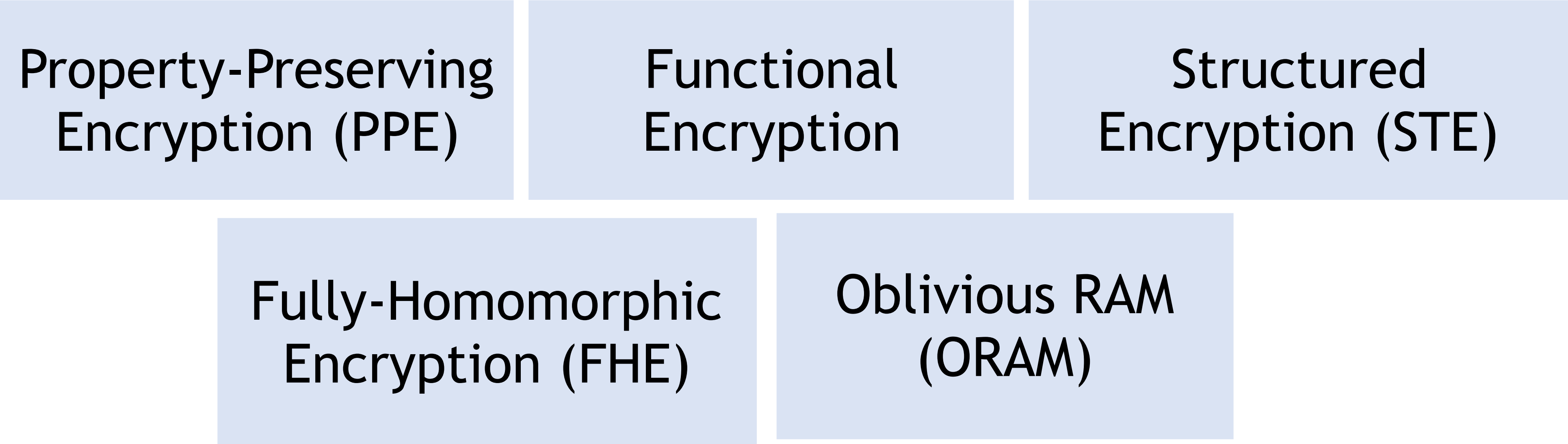
Lazyness?

Cost?

> “…because it would have hurt Yahoo’s ability to index and search message data…”
>
> — J. Bonforte in NY Times

Q: can we search on encrypted data?

Encrypted Search (Building Blocks)



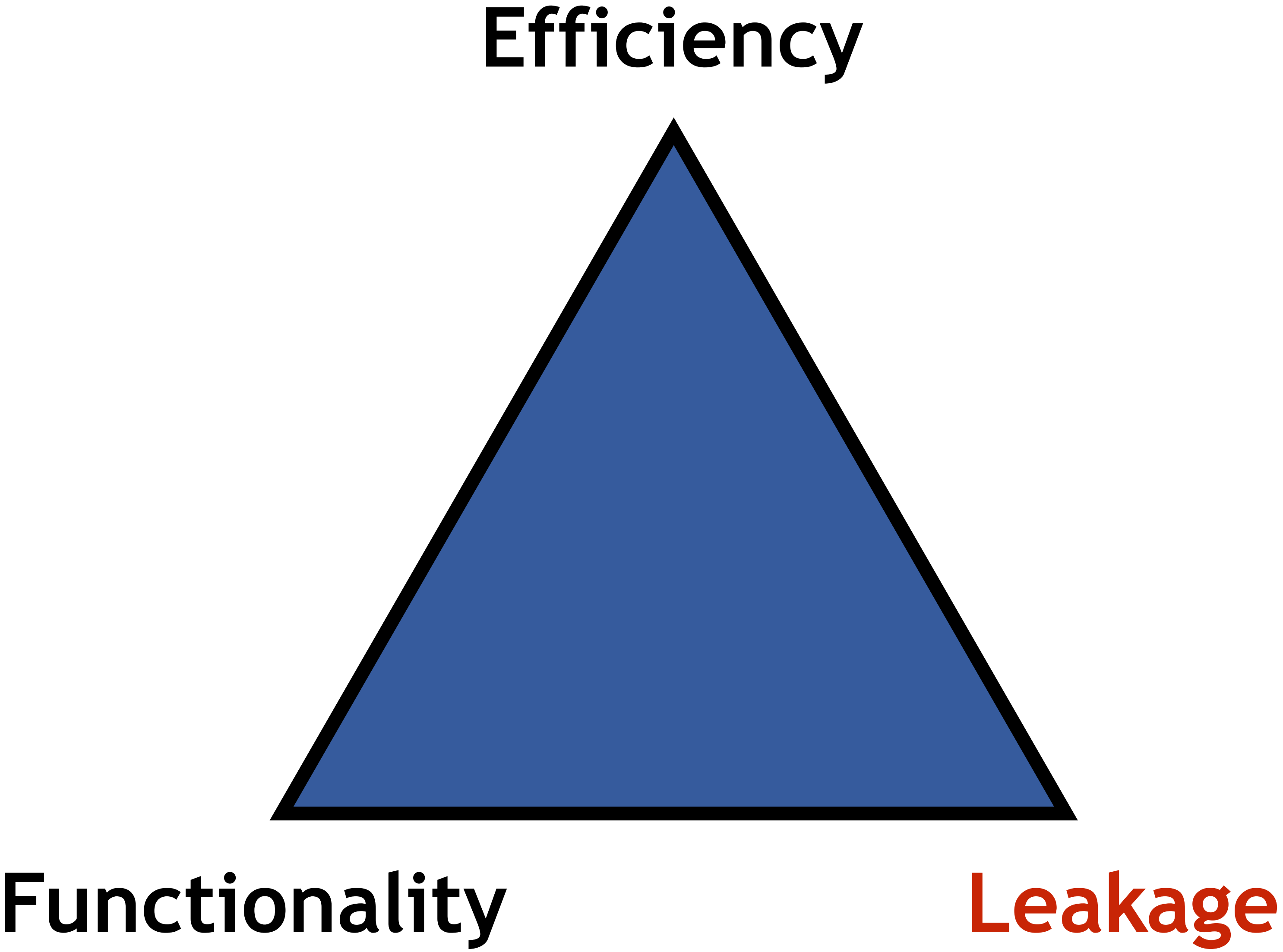
Property-Preserving Encryption (PPE)

Functional Encryption

Structured Encryption (STE)

Fully-Homomorphic Encryption (FHE)

Oblivious RAM (ORAM)

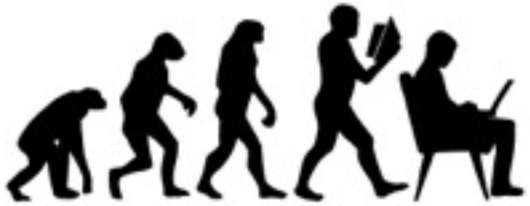


Efficiency

Functionality

Leakage

Evolution from 2001-2018

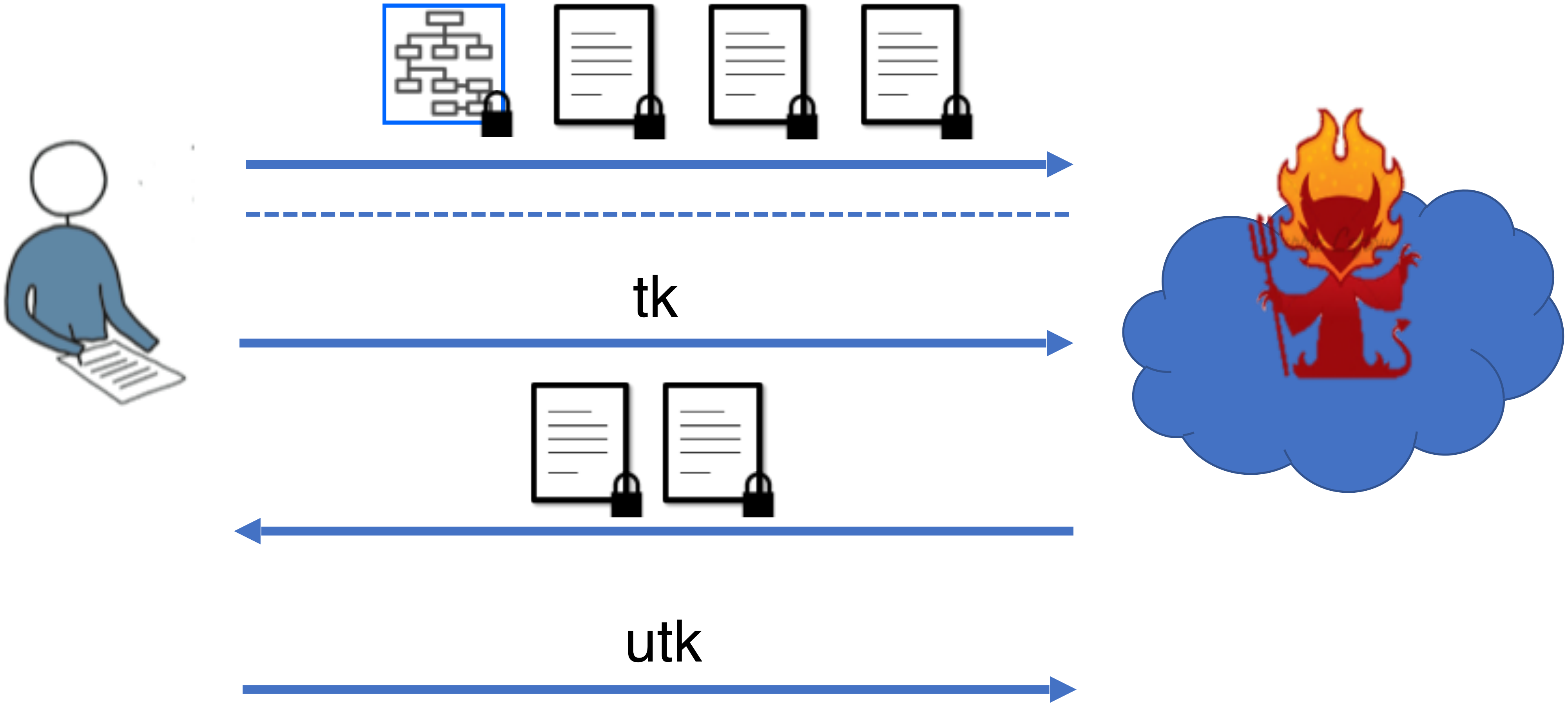


Property-Preserving Encryption (PPE)	
‘06	DET
‘09	OPE
‘11	OPE proofs
‘12	CryptDB
‘15	MS Always Enc
‘15	NKW attacks
‘16	Snapshot PPE

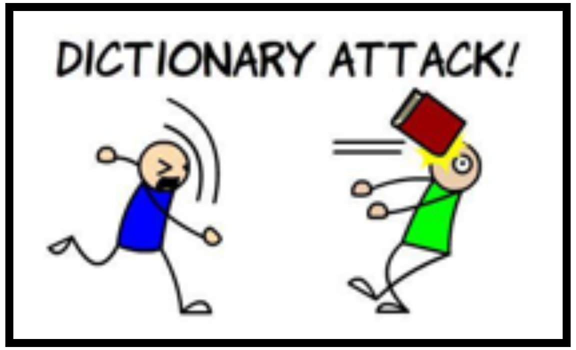
Oblivious RAM (ORAM)	
‘96	ORAM
‘12	Tree-based ORAM
‘13	Path ORAM; ObliviStore
‘16	Obliv P2P; TaoStore
‘16	KKNO attacks

Structured Encryption (STE)	
‘01	SSE
‘06	Efficient SSE
‘10	STE
’12	IKK attacks
’12	CS2
’13	Boolean SSE
’14	OSPIR; BlindSeer
’16	Clusion; OpenSSE
’17	SQL

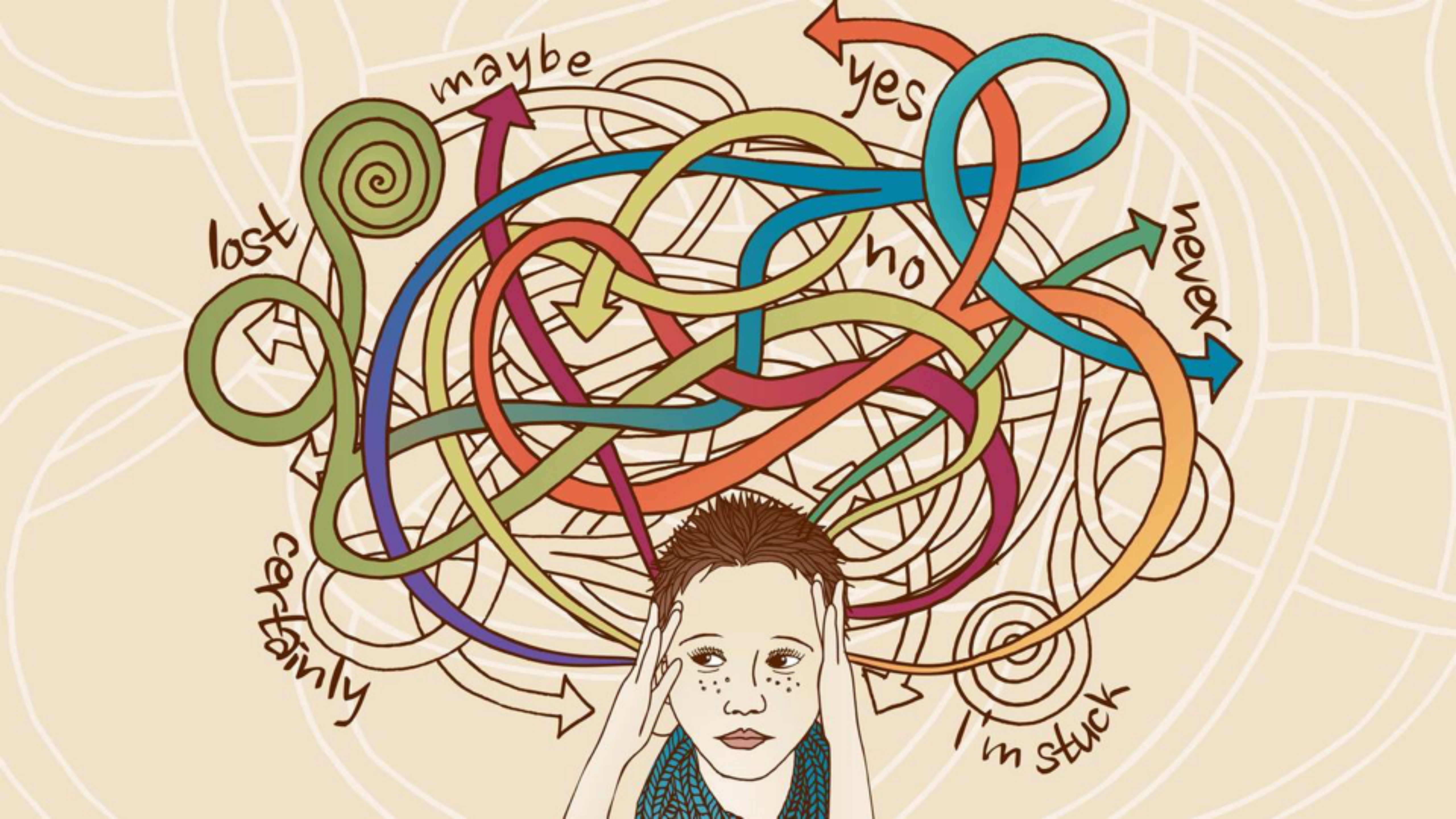
Structured Encryption



Would Encryption Even Prevent Breaches?



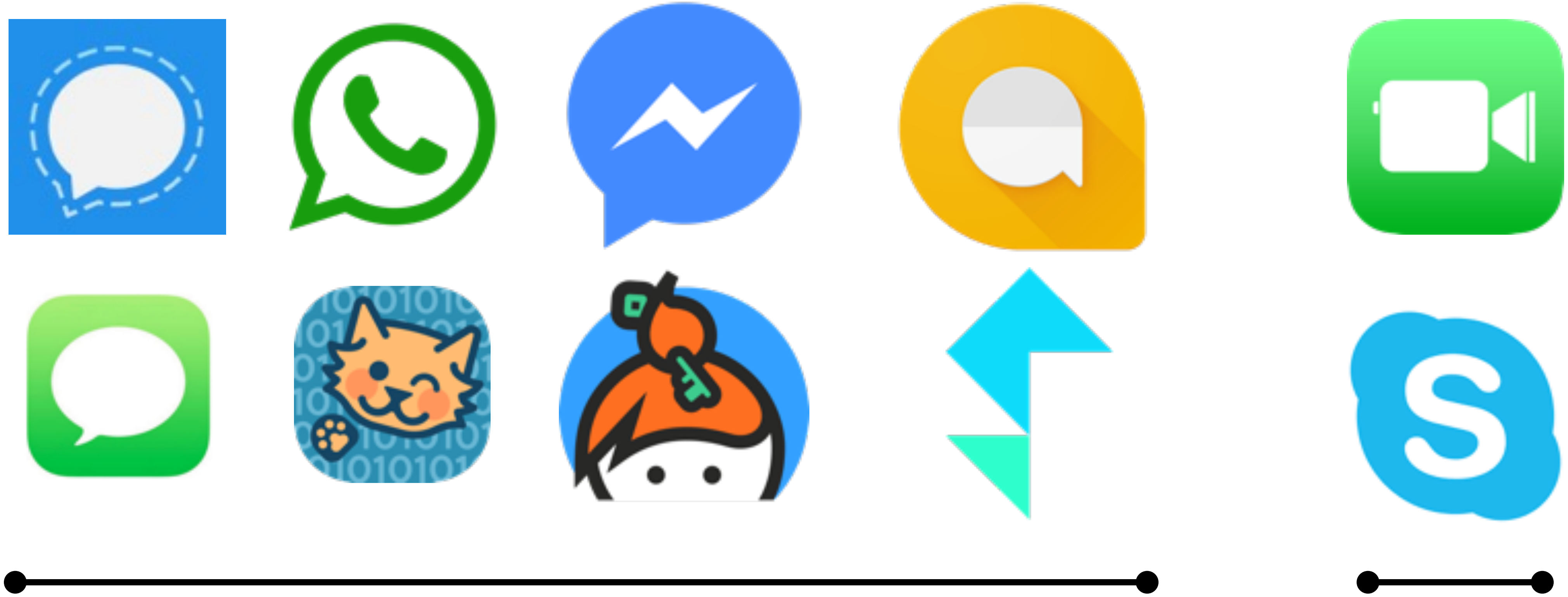
Q: can encrypted search be deployed?


maybe
yes
lost
no
never
certainly
I'm stuck

Tarik

Martin

End-to-End Encryption



messaging

video

Digital Photos - 1.2 Trillion (2017)

85%

4.7%

10.3%

Photo Collections

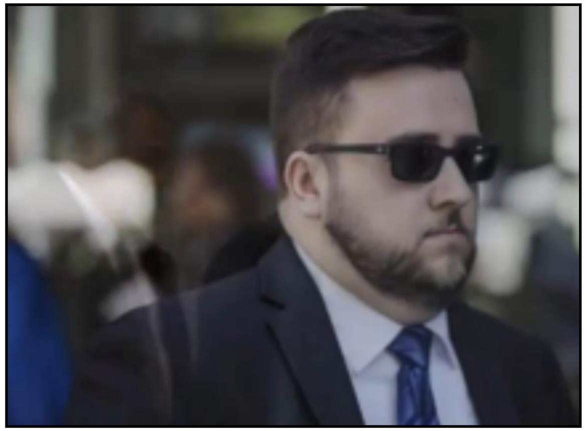
Large

Sentimental value

Private

Cloud
Encryption

Celebgate (2014)



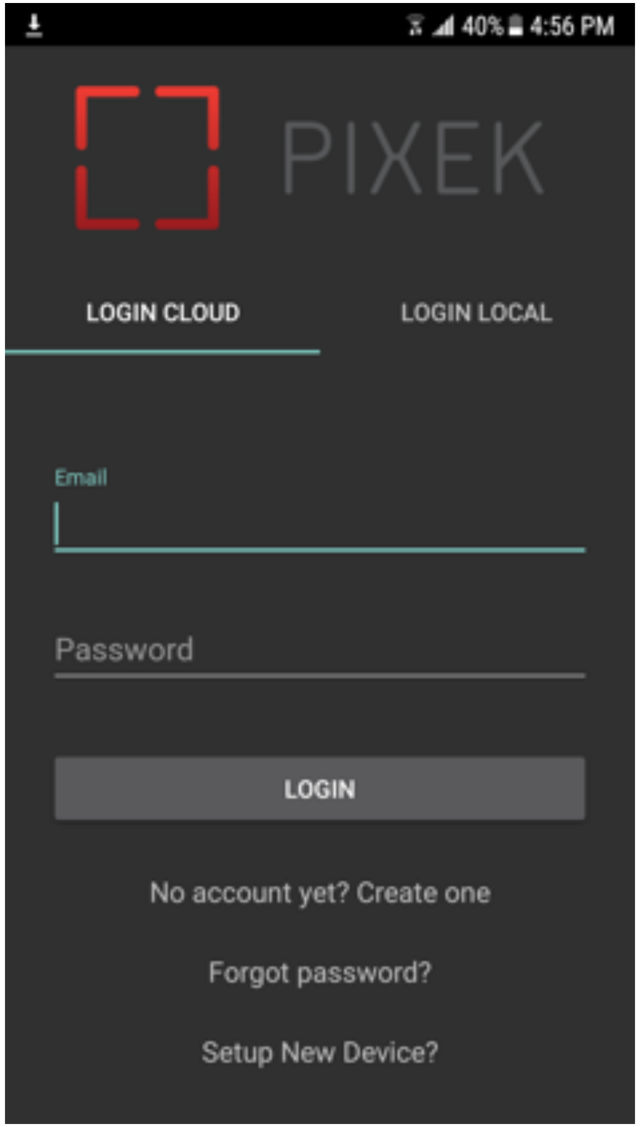
- Edward Majerczyk
  - hacked 30 Gmail & iCloud accounts
  - 500 private photos leaked including of many celebrities

BELL
AP

AM I NEXT?
CUANDO LA TIRANIA ES LEY LA REVOLUCION ES OR
FERGUSON IS E
ASS
WHITE P
UP

Pixek

End-to-end encrypted camera app



LOGIN CLOUD

LOGIN LOCAL

Email

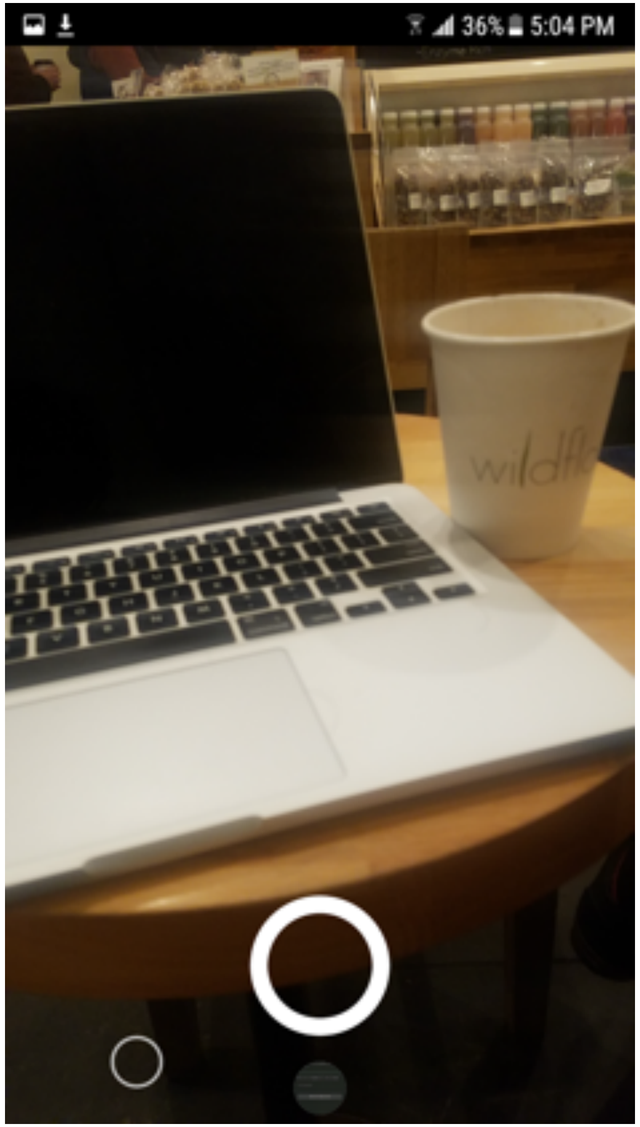
Password

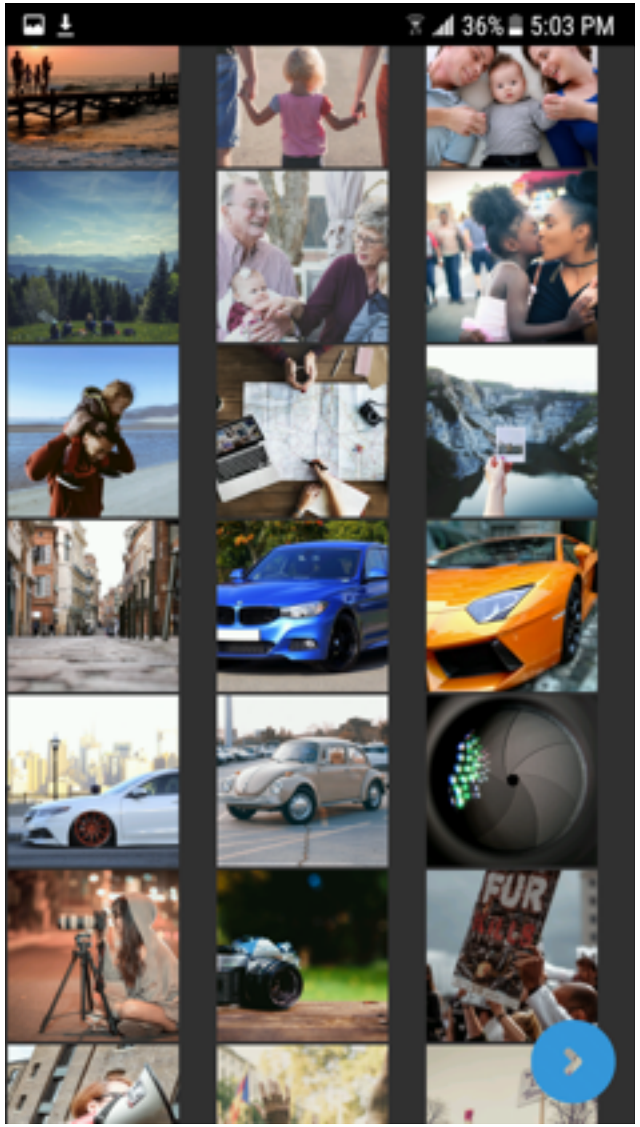
LOGIN

No account yet? Create one

Forgot password?

Setup New Device?





Building Blocks

Clusion

open source (GPLv3) encrypted search library from Brown ESL

pibase, pidyn, 2Lev, ZMF, IEX-2Lev, IEX-ZMF

coming: DLS, SPX, REX, PBS

TensorFlow Mobile

open source machine learning from Google

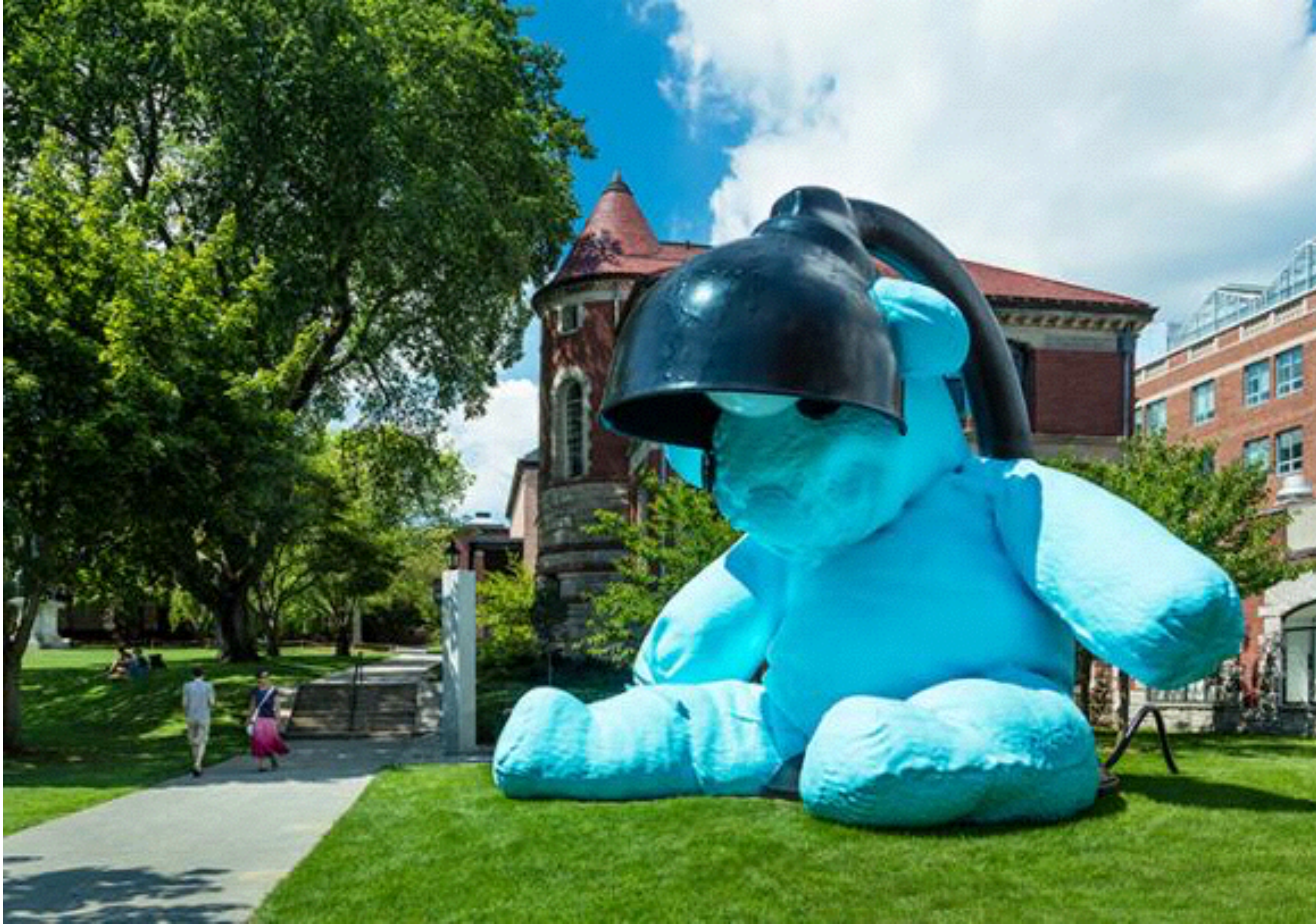
pre-trained model

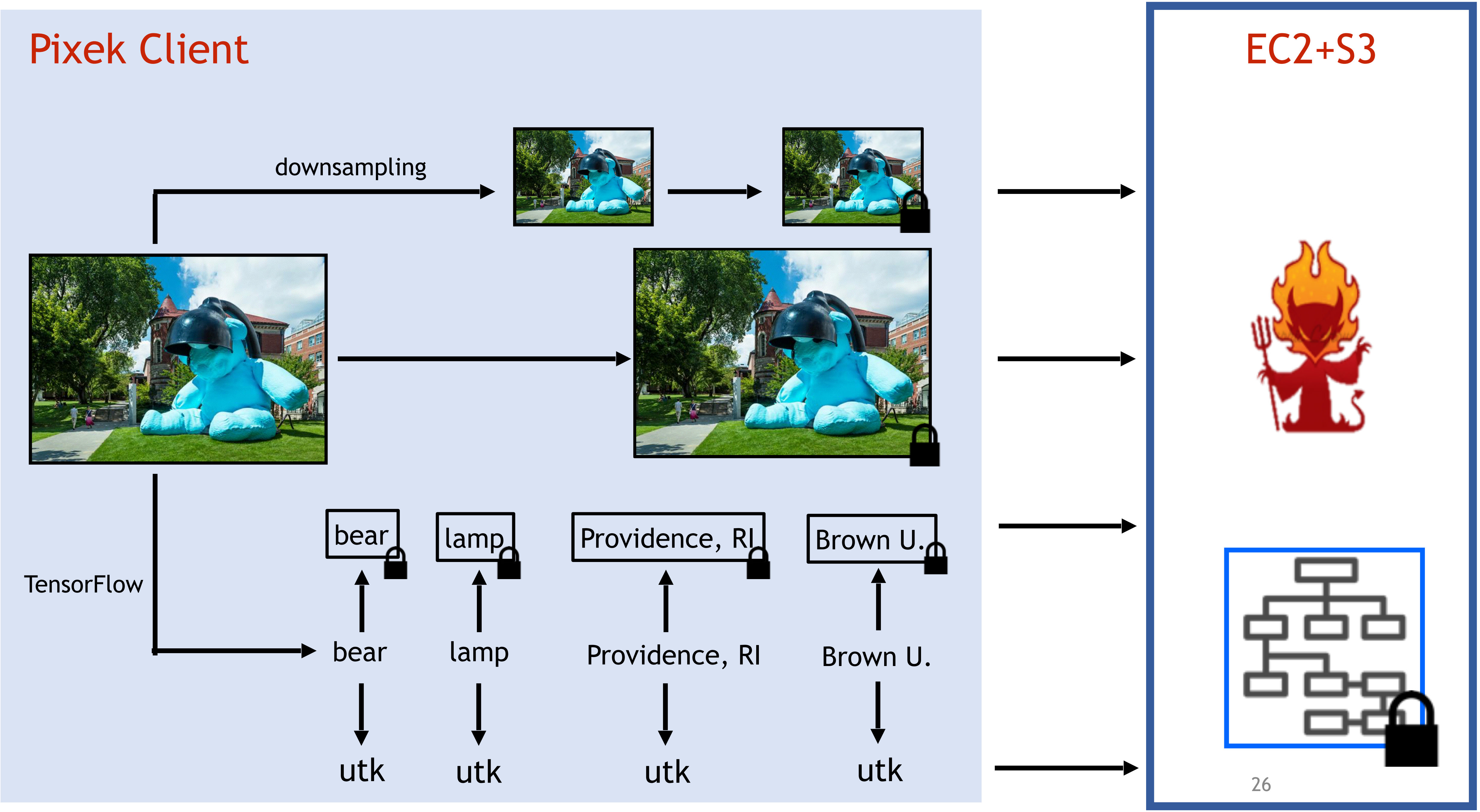
Geomobile

open source geolocation

Lamp/Bear

23’x21’x24’





Pixek Client

EC2+S3

downsampling

TensorFlow

bear

lamp

Providence, RI

Brown U.

bear

lamp

Providence, RI

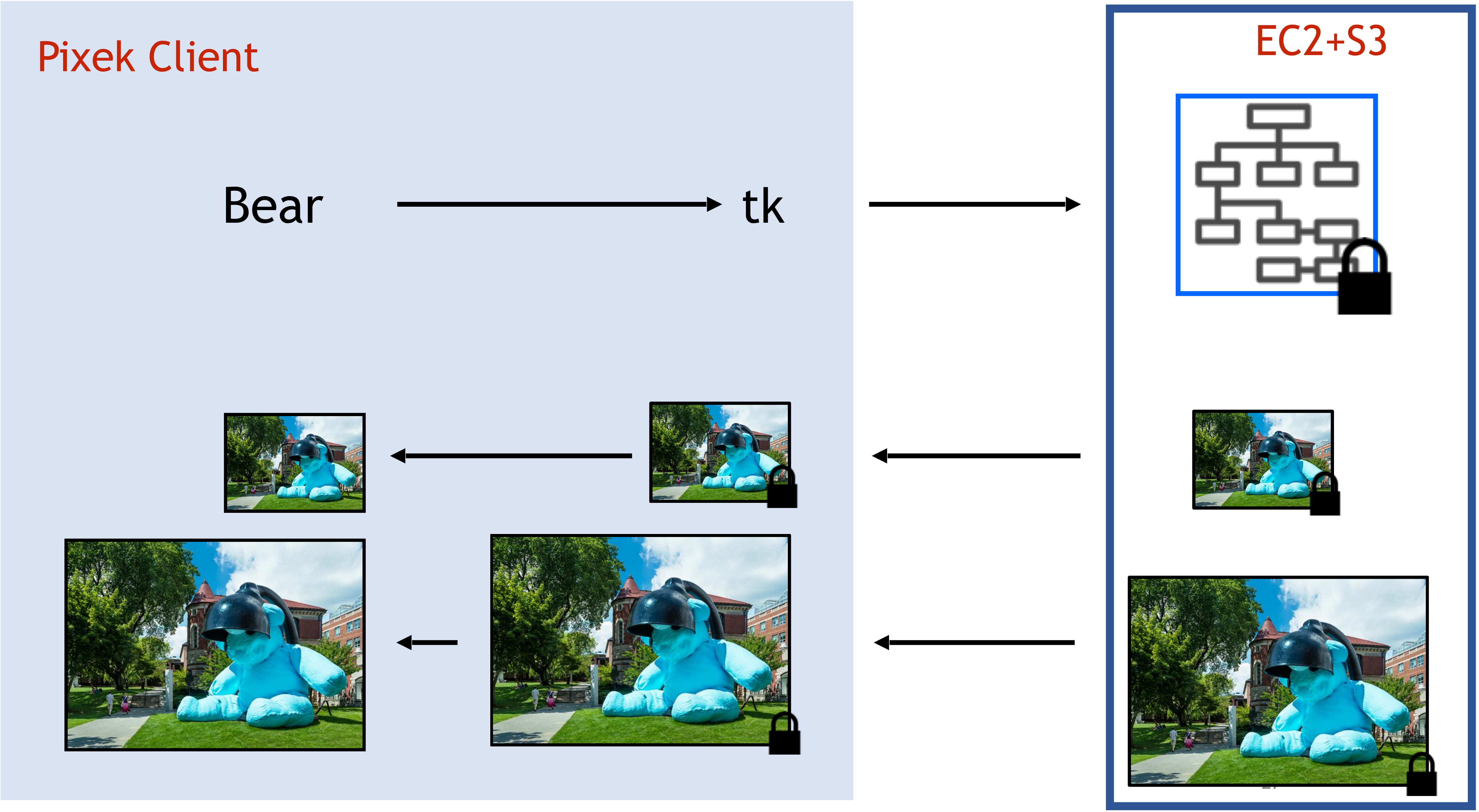
Brown U.

utk

utk

utk

utk



Pixek Client

Bear

tk

EC2+S3

What I Didn’t Cover

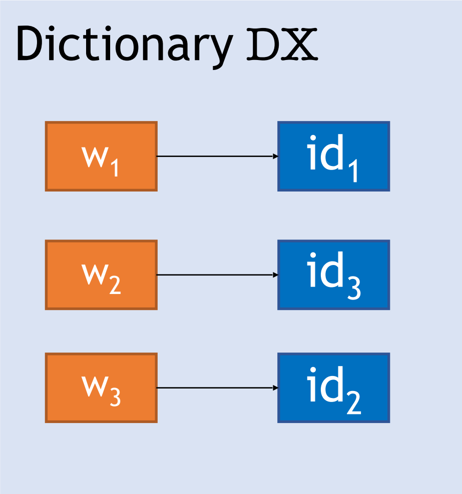
- Caching
- Crash recovery
- Password recovery
- Multi-device
- Local mode

Pixek v0.1.0 (Current)

- Tags & photos are streamed
    - Encrypted structure needs forward-privacy
- Published state-of-the-art
    - Sophos [Bost16]
    - Diana [Bost-Minaud-Ohrimenko17]
- New scheme
    - pidyn [Cash-Jaeger-Jarecki-Jutla-Krawczyk-Rosu-Steiner14]
    - no public-key operations
    - no constrained PRFs

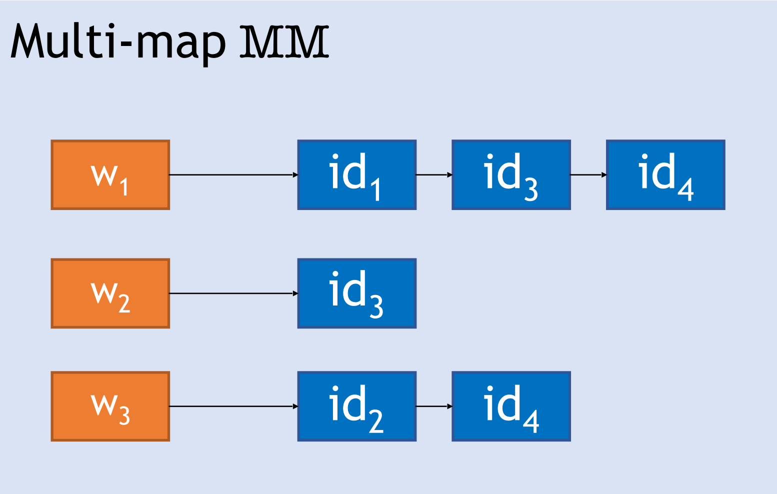
Background: Data Structures

- DXs map labels to values



- Get: $DX[w_3]$ returns id_2

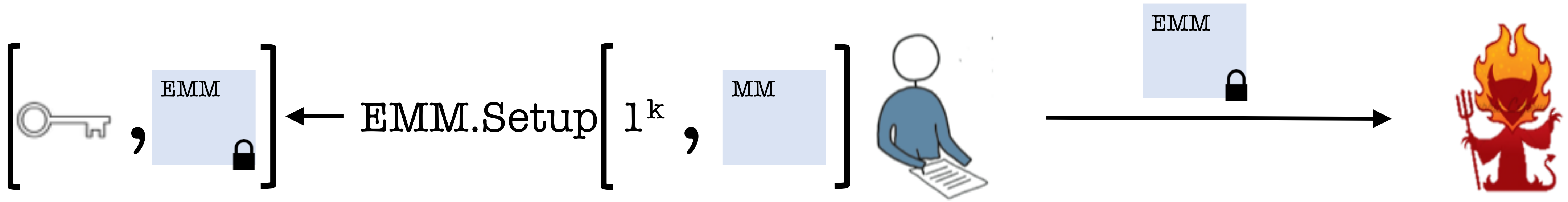
- MMs map labels to tuples



- Get: $MM[w_3]$ returns (id_2 , id_4)

π_{dyn} [CJJJKRS'14]

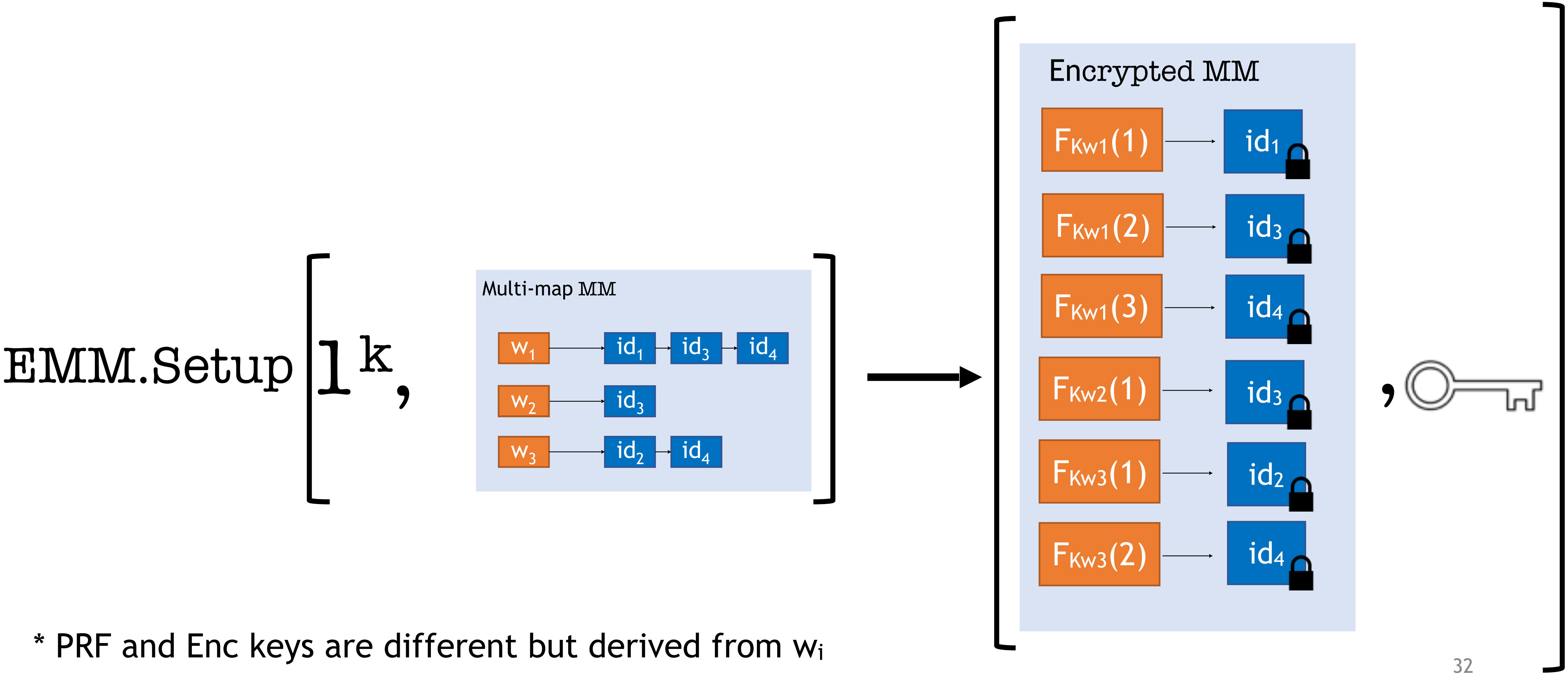
Setup



$\left[\ \text{🔑}\ ,\ \text{EMM}\ \right] \leftarrow \text{EMM.Setup}\left[\ 1^k\ ,\ \text{MM}\ \right]$

π_{dyn} [CJJJKRS'14]

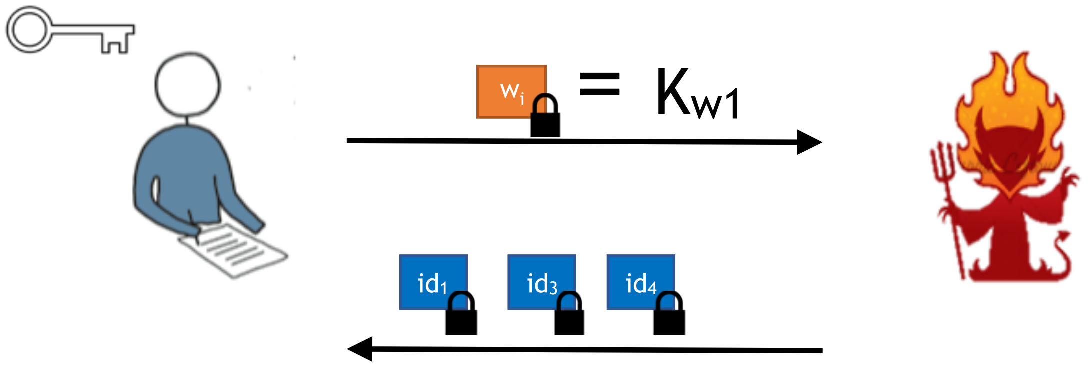
Setup



* PRF and Enc keys are different but derived from w_i

π_{dyn} [CJJJKRS'14]

Get

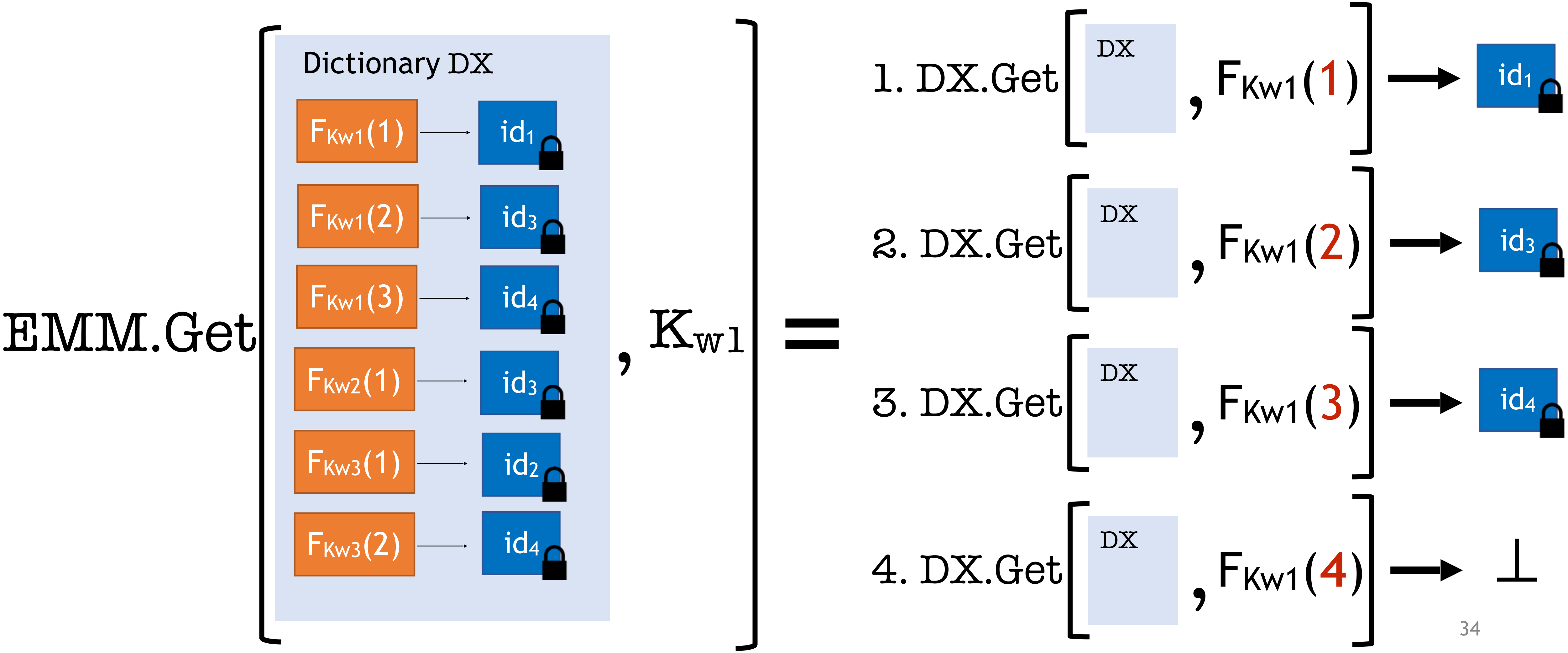


EMM.Get[EMM , K_{w1}]

1. DX.Get[DX , $F_{K_{w1}}(1)$] → id₁
2. DX.Get[DX , $F_{K_{w1}}(2)$] → id₃
3. DX.Get[DX , $F_{K_{w1}}(3)$] → id₄
4. DX.Get[DX , $F_{K_{w1}}(4)$] → $\perp$

π_{dyn} [CJJJKRS'14]

Get



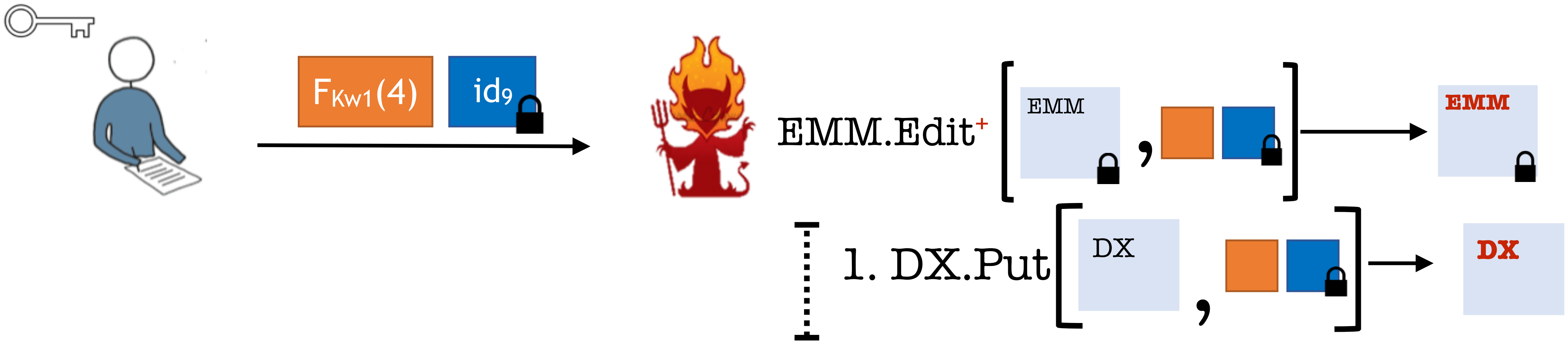
EMM.Get[Dictionary DX , K_{w1}] =

Dictionary DX	
$F_{Kw1}(1)$	id_1
$F_{Kw1}(2)$	id_3
$F_{Kw1}(3)$	id_4
$F_{Kw2}(1)$	id_3
$F_{Kw3}(1)$	id_2
$F_{Kw3}(2)$	id_4

1. DX.Get[DX , $F_{Kw1}(1)$] → id_1
2. DX.Get[DX , $F_{Kw1}(2)$] → id_3
3. DX.Get[DX , $F_{Kw1}(3)$] → id_4
4. DX.Get[DX , $F_{Kw1}(4)$] → $\perp$

π_{dyn} [CJJJKRS'14]

Edit$^+$



$F_{K_{w1}}(4)$ id$_9$

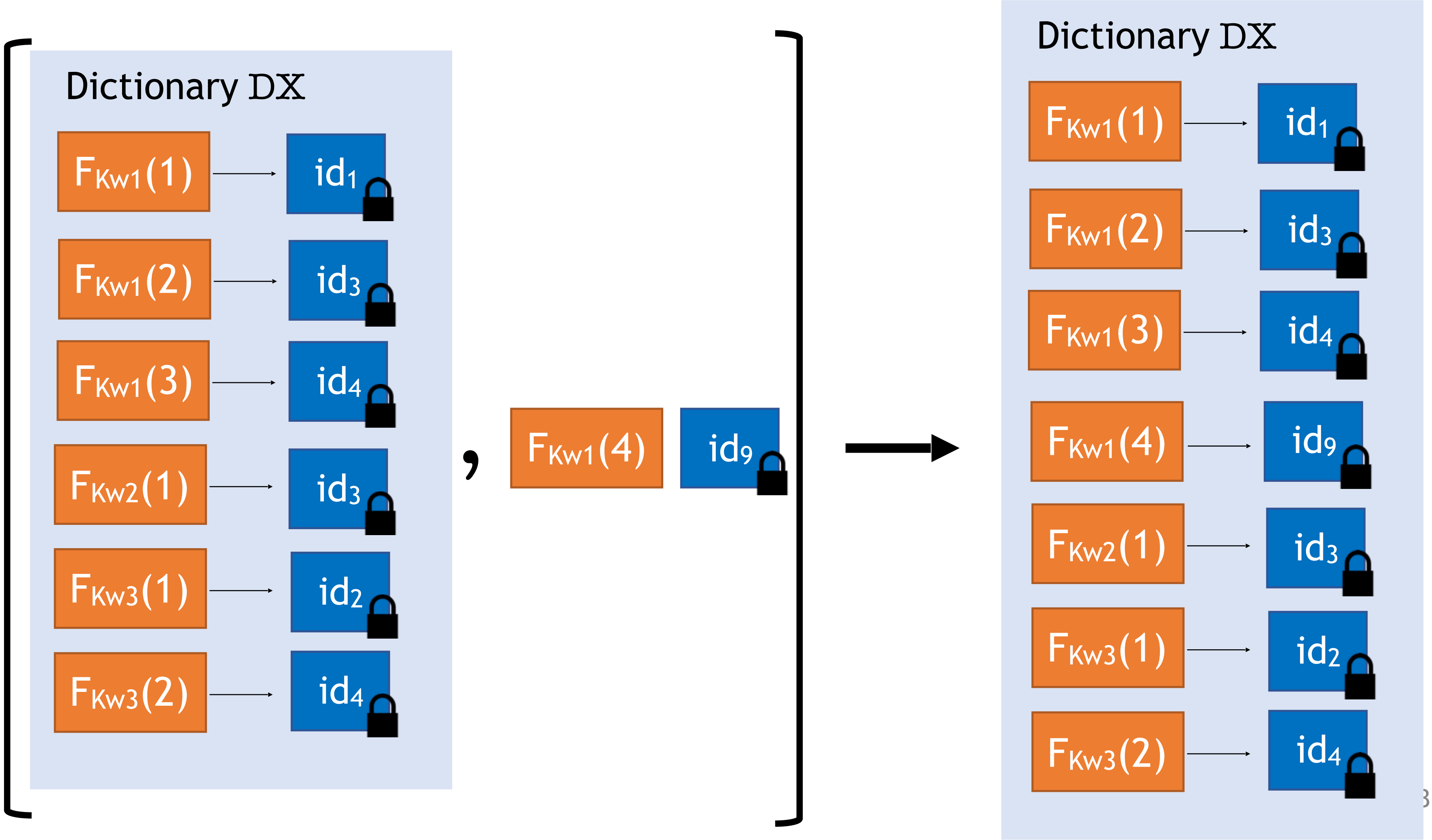
EMM.Edit$^+$ [EMM ,] → EMM

1. DX.Put [DX ,] → DX

π_{dyn} [CJJJKRS'14]

Edit^{+}

EMM.Edit^{+}



Dictionary DX

$F_{K_{w1}}(1)$ → id_1

$F_{K_{w1}}(2)$ → id_3

$F_{K_{w1}}(3)$ → id_4

$F_{K_{w2}}(1)$ → id_3

$F_{K_{w3}}(1)$ → id_2

$F_{K_{w3}}(2)$ → id_4

, $F_{K_{w1}}(4)$ id_9

→

Dictionary DX

$F_{K_{w1}}(1)$ → id_1

$F_{K_{w1}}(2)$ → id_3

$F_{K_{w1}}(3)$ → id_4

$F_{K_{w1}}(4)$ → id_9

$F_{K_{w2}}(1)$ → id_3

$F_{K_{w3}}(1)$ → id_2

$F_{K_{w3}}(2)$ → id_4

Forward-Private $\pi_{\mathbf{dyn}}$

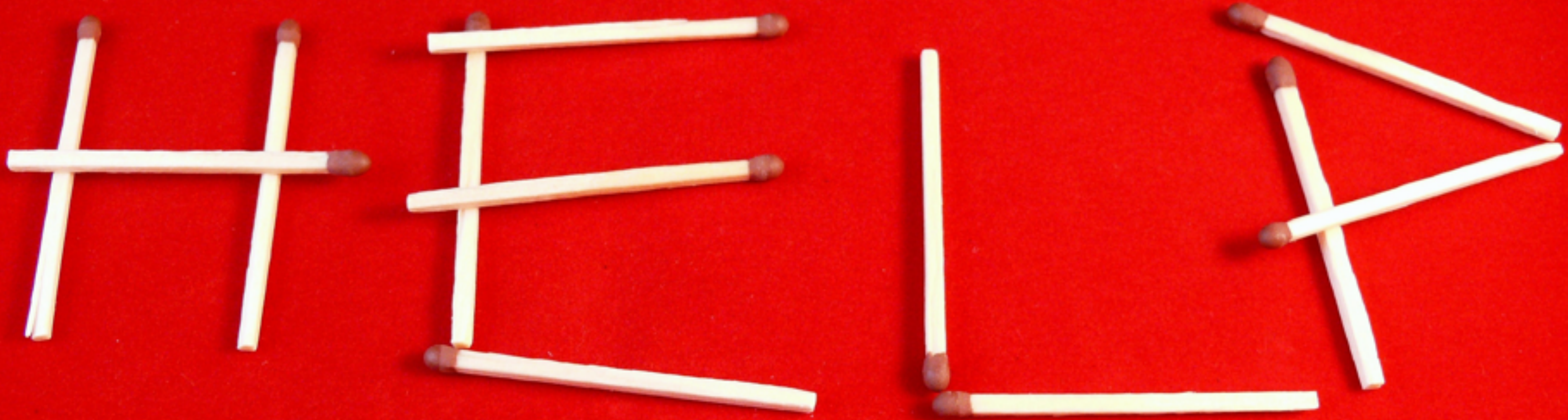
- Why is $\pi_{\mathbf{dyn}}$ not forward-private?
    - new pairs encrypted under same key used for search,
        - K_{wi} := $F_K(w_i||1)$
    - so previously searched w's can be linked to new pairs
- Making $\pi_{\mathbf{dyn}}$ forward-private
    - use keys with version number that rotates at each update
        - K_{wi} := $F_K(w_i||\text{version}||1)$
    - To search send keys for all versions
        - $F_K(w_i||\text{version1}||1)$, …, $F_K(w_i||\text{version8}||1)$

Forward-Private π_{dyn}

- Search complexity
  - optimal O(#MM[w])
- Token size
  - non-optimal O(#MM[w])
  - new technique makes it O(1) (not implemented yet)

Leakage

- Search pattern
  - ***we see if a query is repeated***
  - ex: if you search for “bear” 3x, we see you searched for **?** 3x
- Access pattern
  - ***we see which encrypted photo matched your query***
  - ex: if you search for “bear”, we see which encrypted photos match query
- What are the consequences of this leakage?
  - To see your photos we have to break AES
  - To learn about your queries we have to know/guess > 90% of your tags & know the occurrence of each tag


HELP

Testers & Feedback

Google Play

- Only available on Android
- Let us know @pixekapp if you want access

pixek.io

🐦 @pixekapp