Recent Advancements in Cloud Security

Matthew Mitchell

“Rather then attempt to restrict usage, the goal should be to enable the freedom employees need to do their jobs better, without compromising company security and liability.”

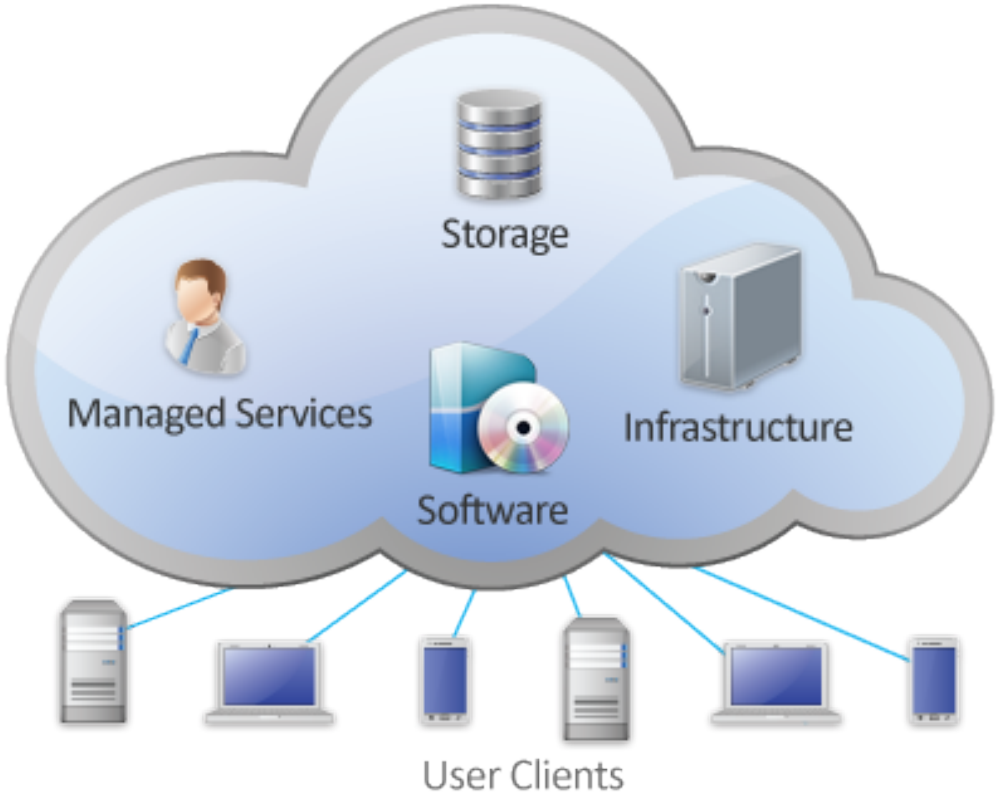
-Lynda Stadtmueller

Outline

- Introduction
  - What is the cloud
  - Different types of cloud services
- Existing approaches to data disclosure
  - Security circle
- Browserflow
  - Data flow control
- Conclusion

The Cloud

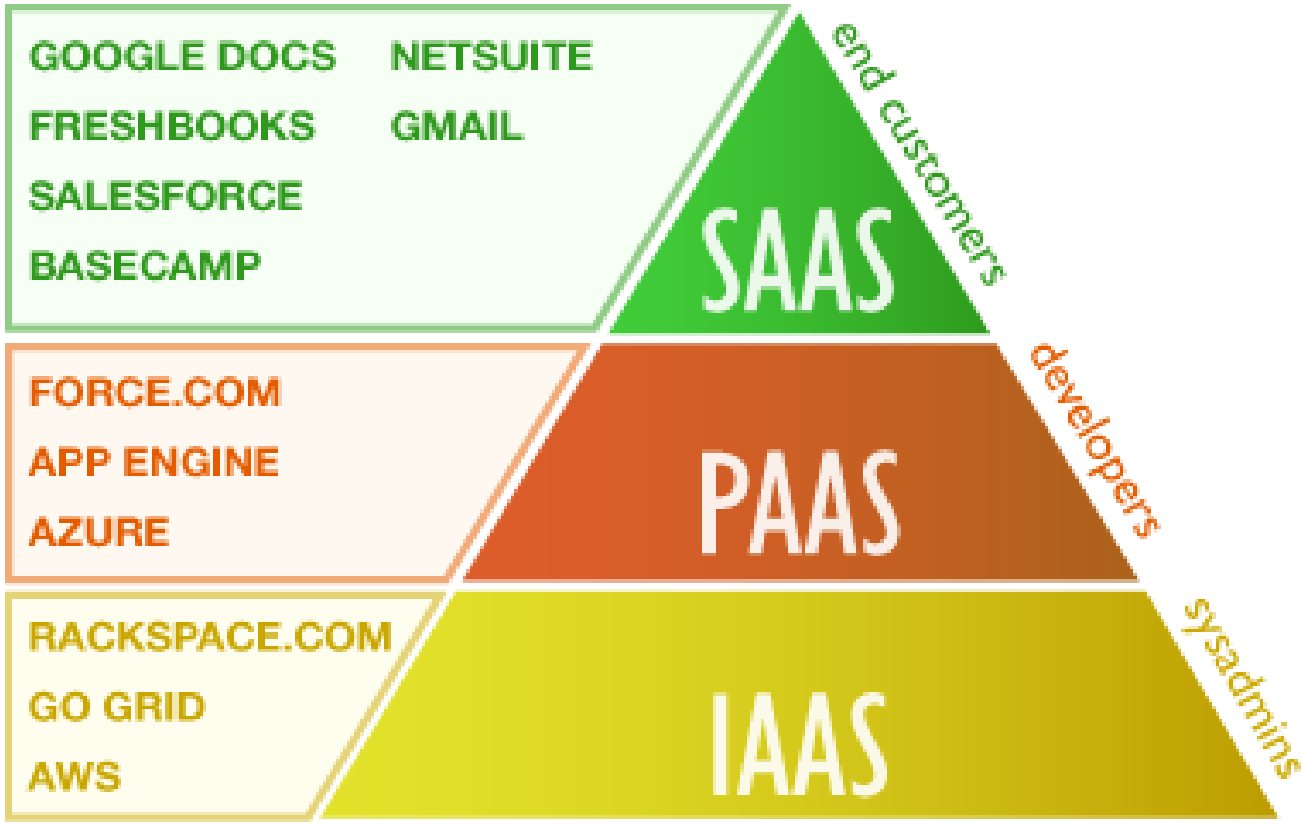
- Access programs over the internet
- Data is stored on vendors servers
- Easily accessible data



https://xrm.com/reference/cloud/

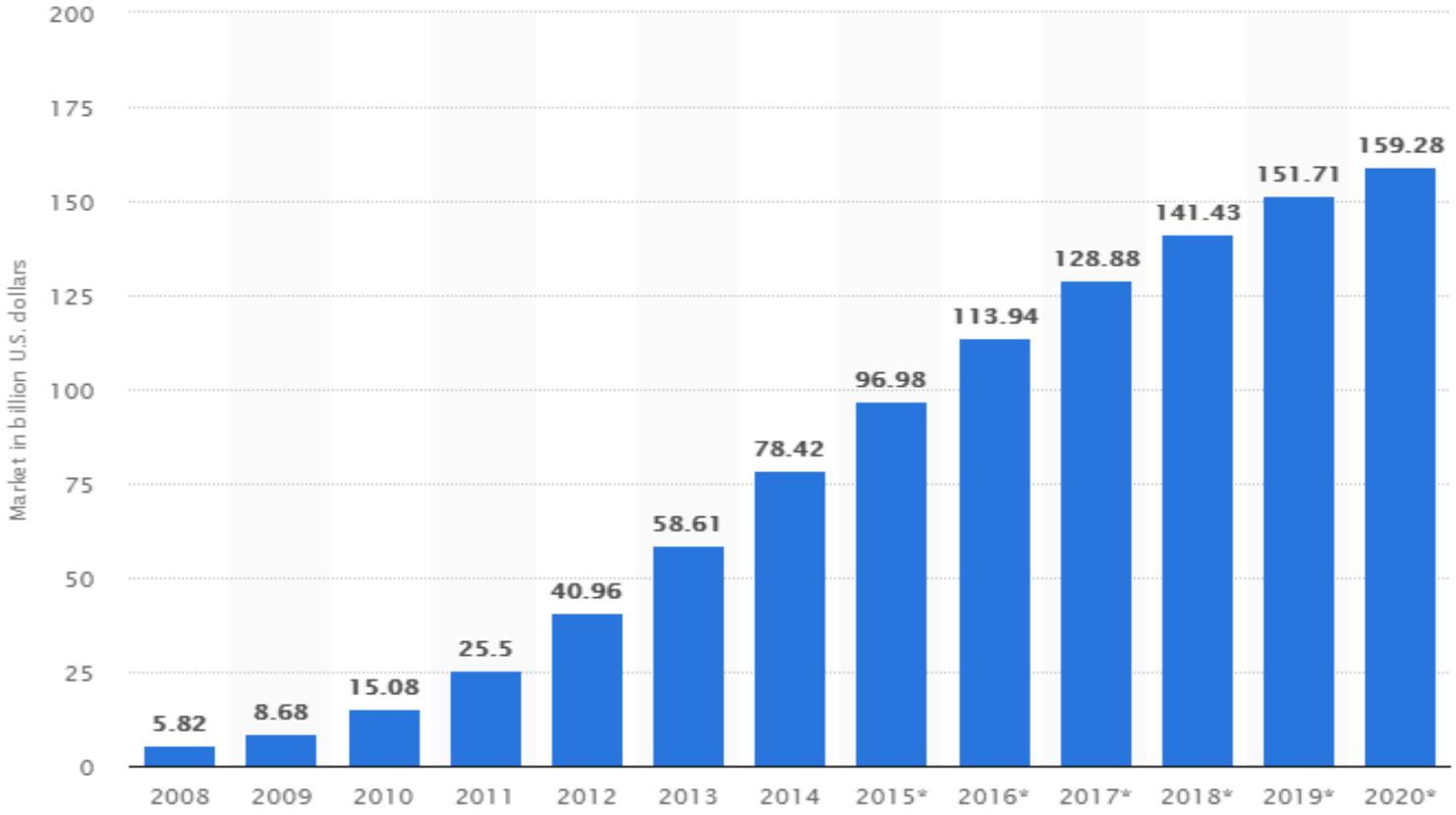
Most Common Cloud Service Provider (CSP)

- Software-as-a-Service (SaaS)
- Platform-as-a-Service (PaaS)
- Infrastructure-as-a-Service (IaaS)



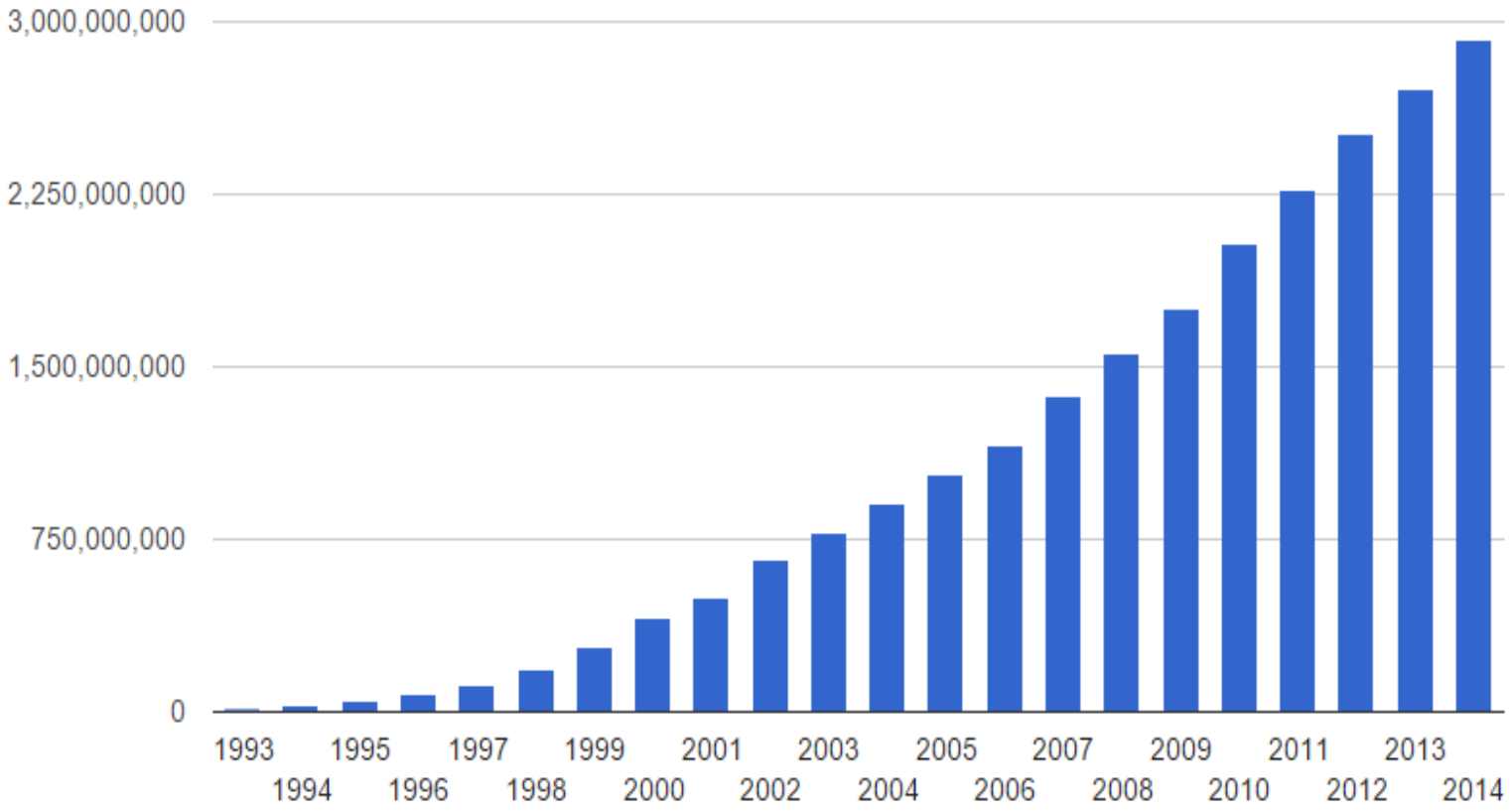
https://www.globaldots.com/cloud-computing-types-of-cloud/

Spending on cloud infrastructure



https://www.statista.com/

Internet Users (World)



https://www.statista.com/

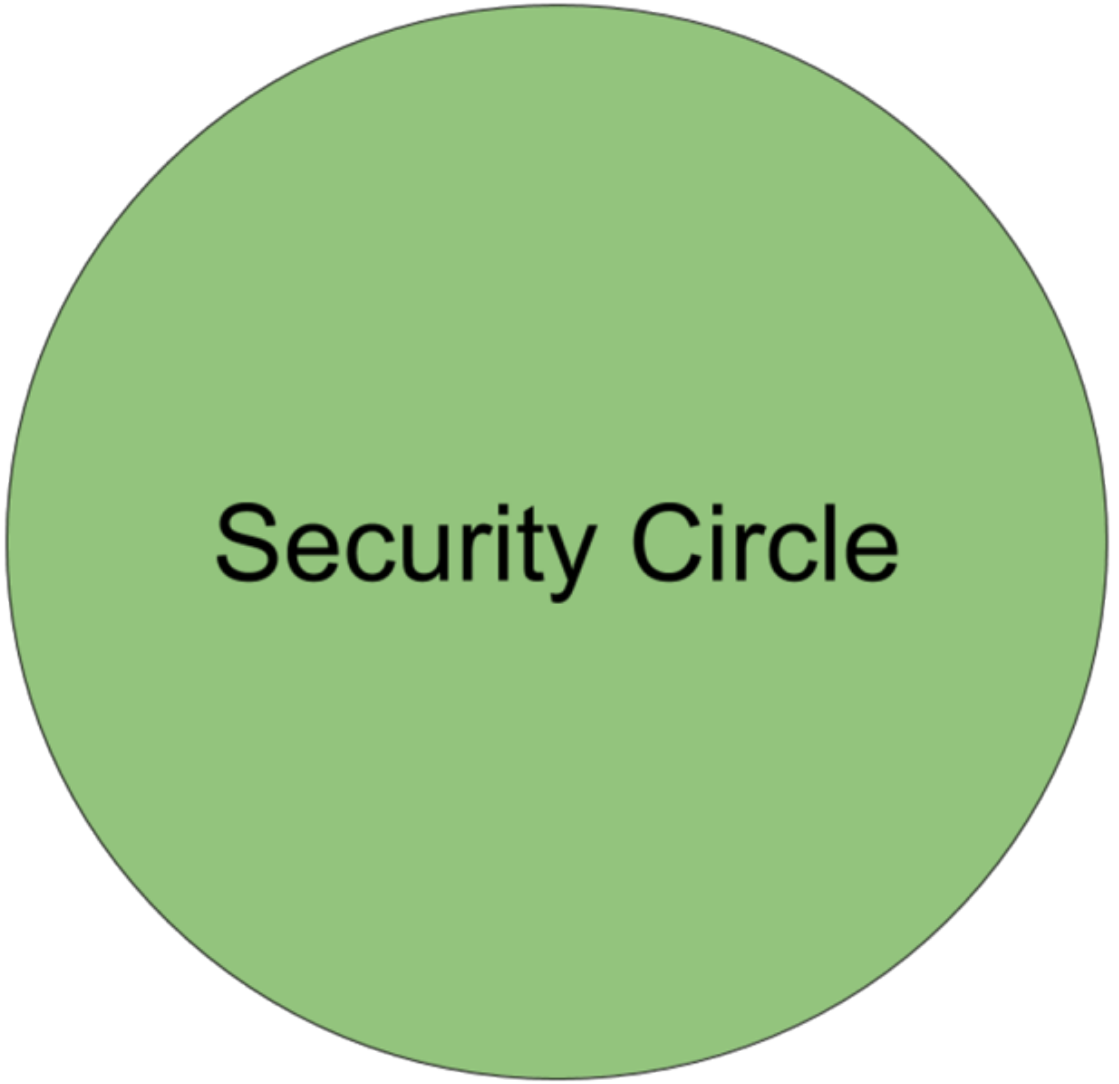
Service level agreement (SLA)

- SLA is a blueprint and warranty for the cloud
- The performance the data center will have and more recently performance of the network
- Cloud service provider(CSP) agrees to what kind of monitoring and reporting

- Users have access to multiple types of cloud services
- How to enforce *data disclose polices*
- Track the users activities with in the browser

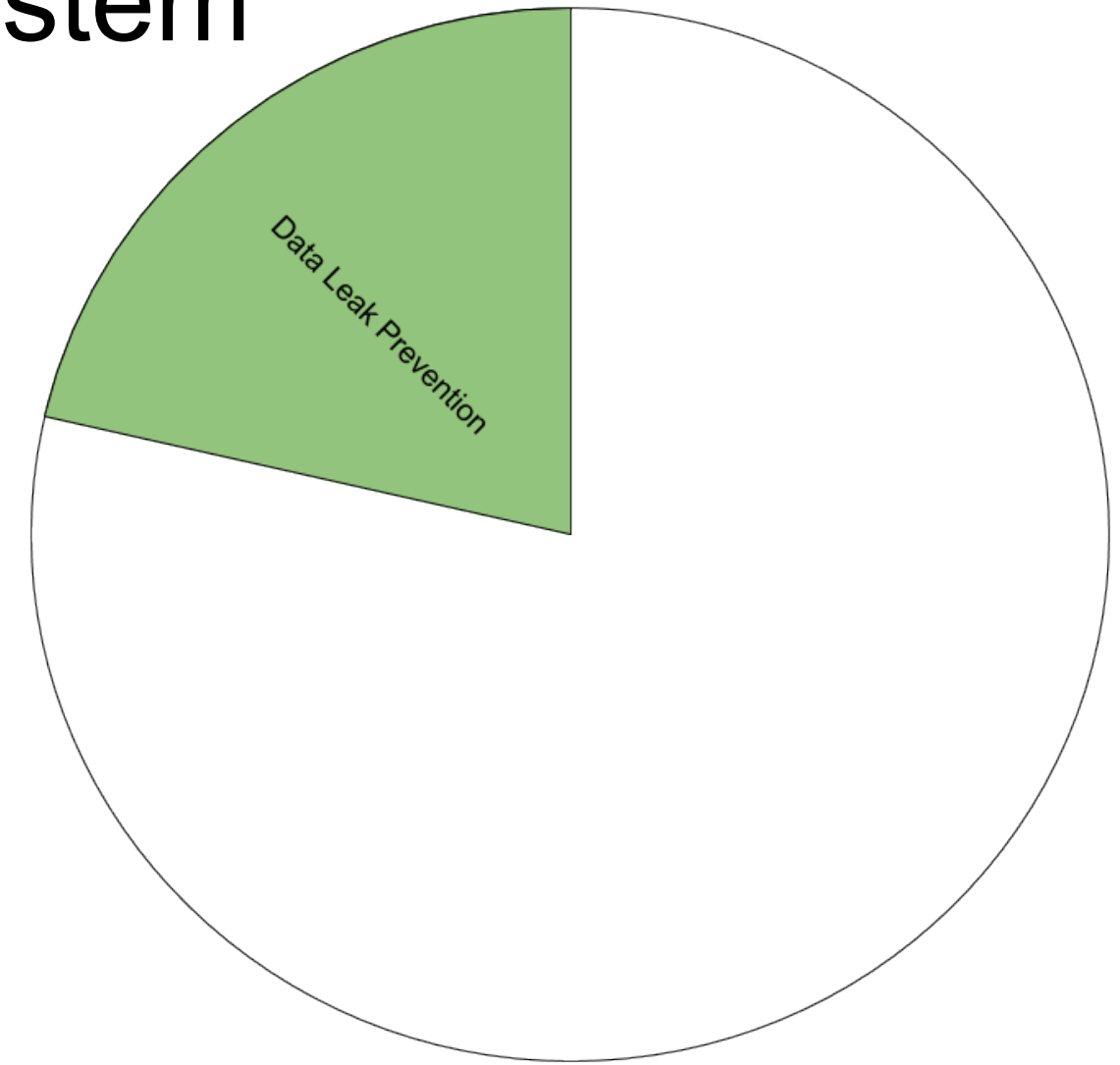
Existing approaches

- Data leak prevention system
- Data flow tracking system
- Static data flow analysis
- Browser-side enforcement
- Client-side middleware



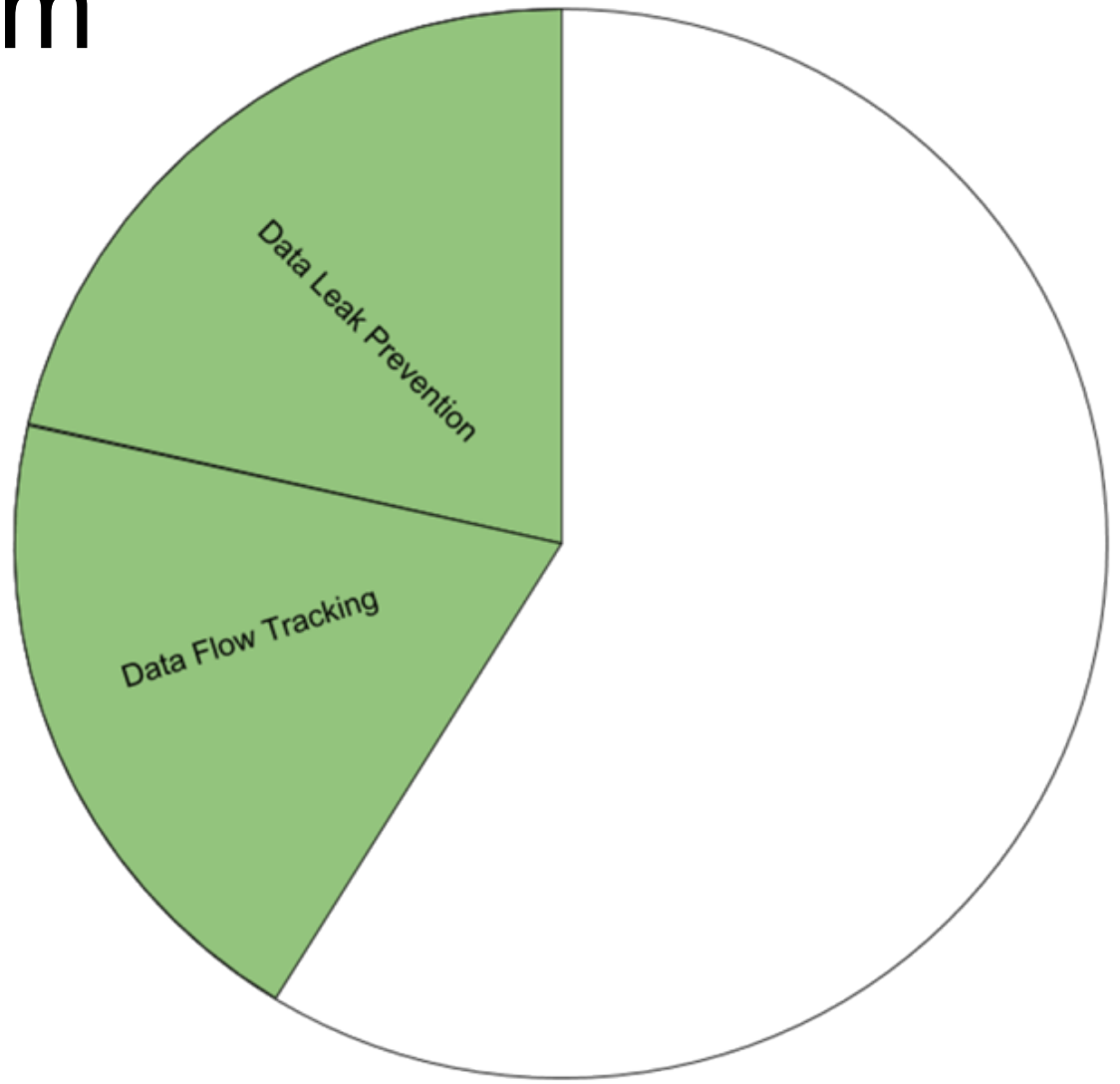
Data leak prevention system

- Protects sensitive data
- Analyzes endpoint
- By inspecting outgoing network traffic



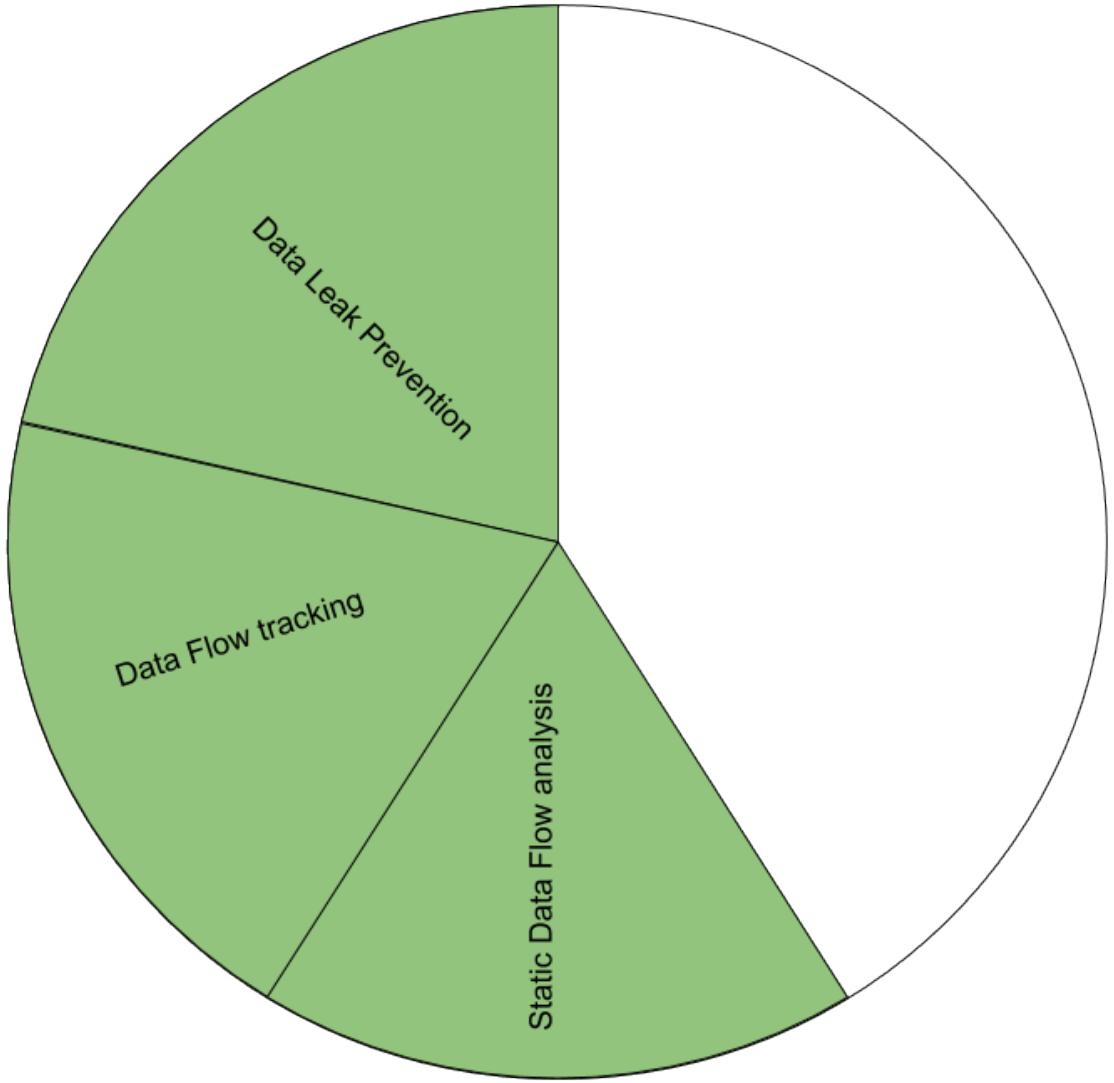
Data flow tracking system

- Tracked by attaching labels to data
- Used for precise data flow, such as tracking passwords
- Computation heavy, thus has a big performance overhead



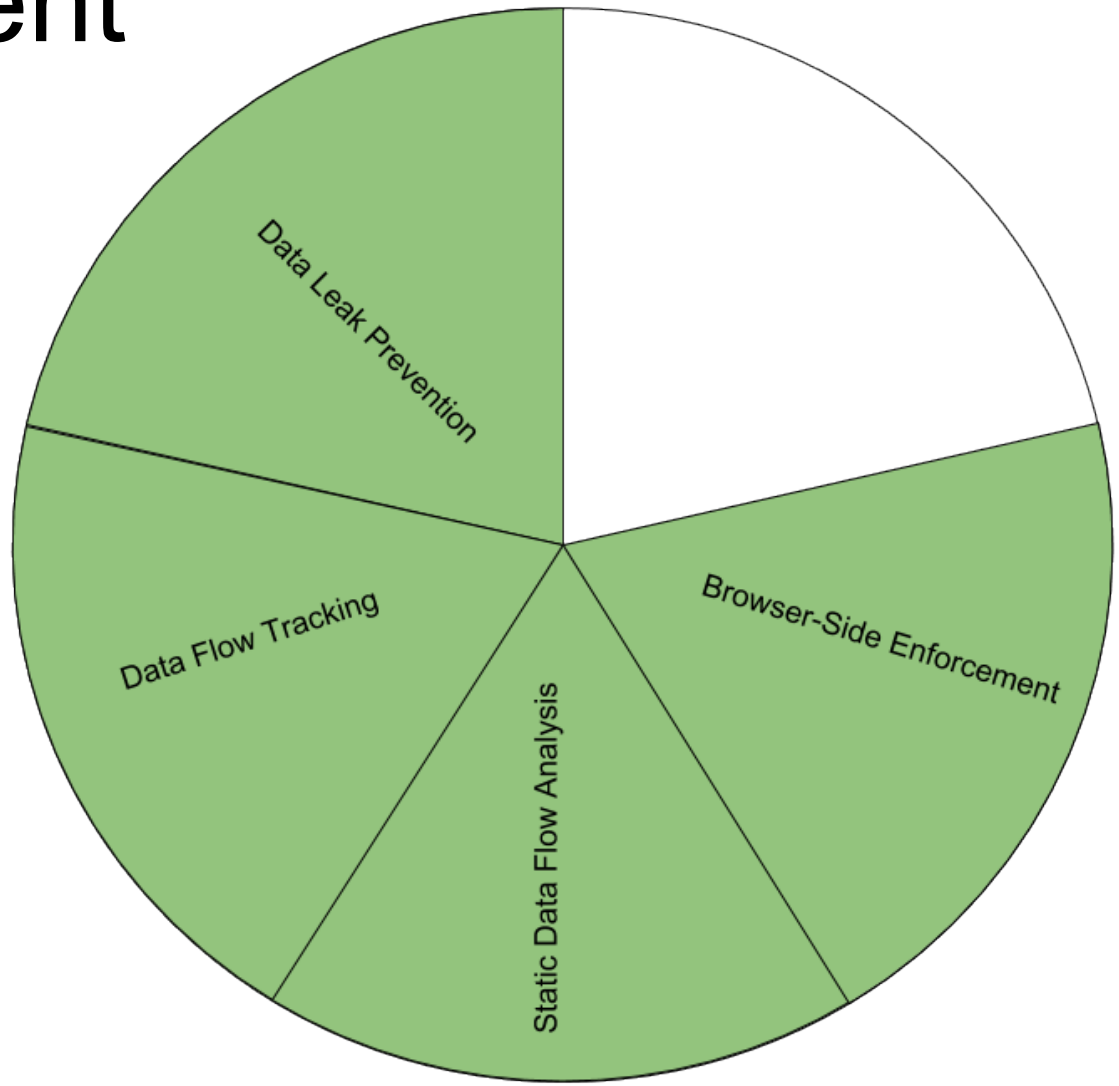
Static data flow analysis

- Tracking data by analyzing source code
- Produces conservative results
- Unusable for legacy programs



Browser-side enforcement

- Data is encrypted before the sensitive data is uploaded
- Not ideal since a CSP may have to index, search and inspect the original data
- Doesn't allow collaborative editing similar to Google Docs

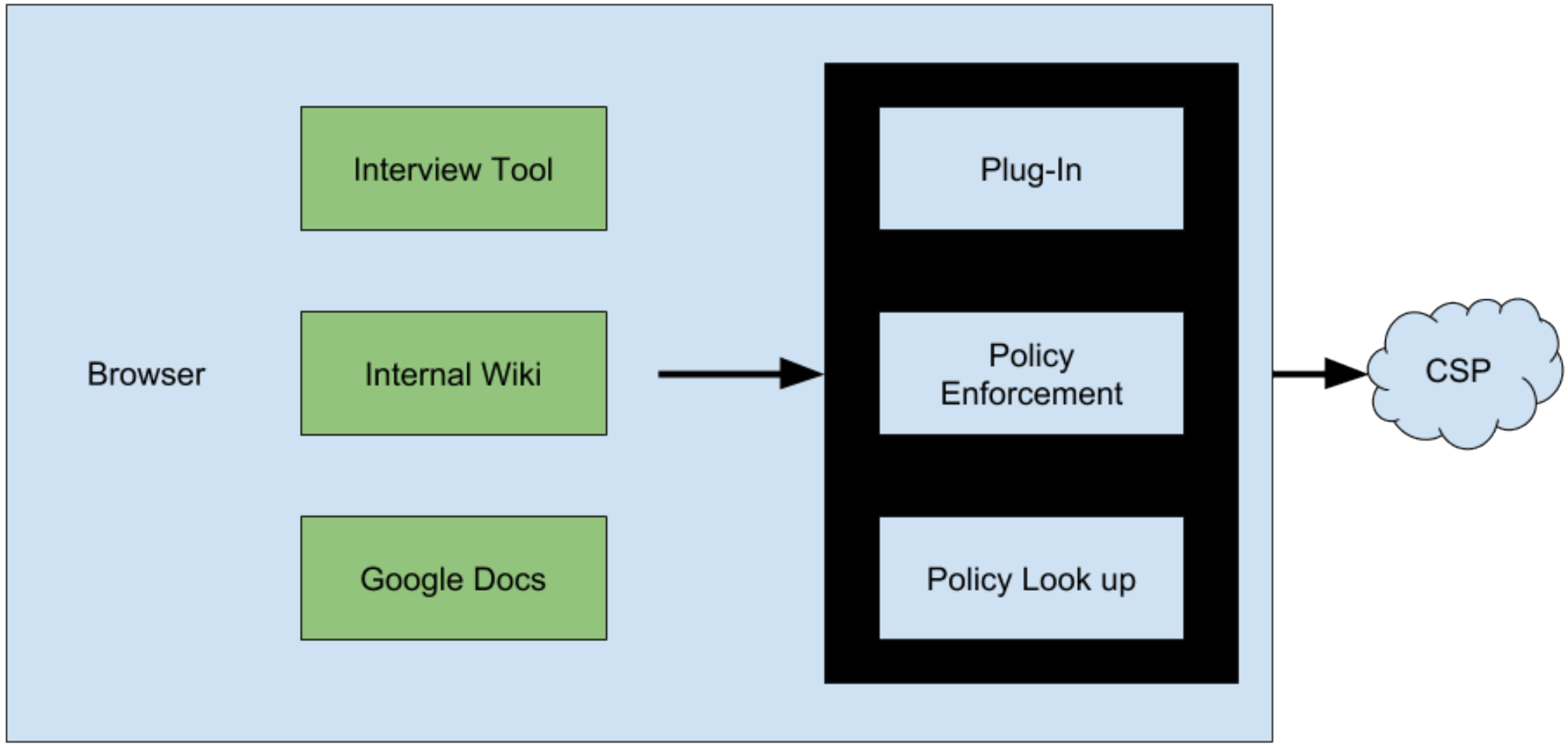


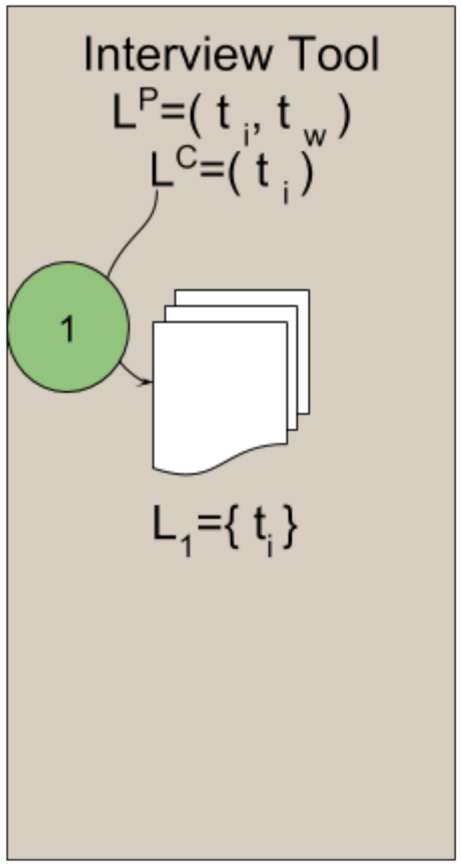
Client-side middleware

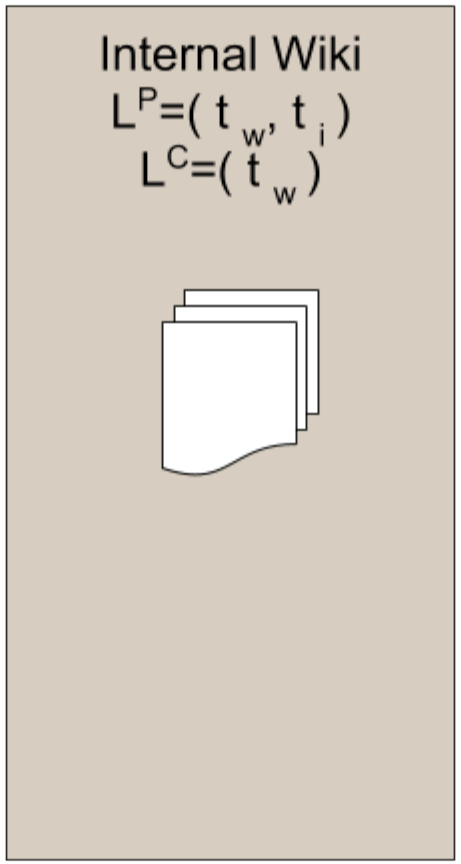
- Protects the confidentiality of users information from an untrusted CSP
- Decoupling the data from application logic.

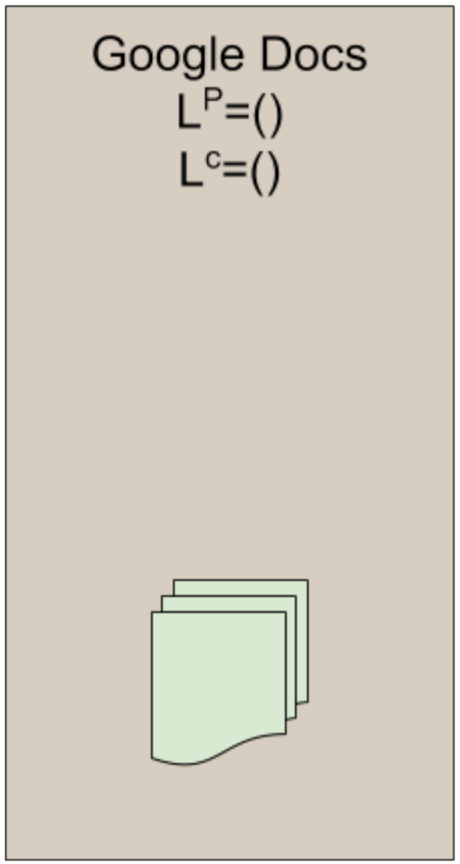
BROWSERFLOW

- Group of researchers from The Imperial College London
- Protects against disclosure by users.
- Actively scans text.
- Enforces data exposure between cloud services.
- Done on the browser


Browser
Interview Tool
Internal Wiki
Google Docs
Plug-In
Policy Enforcement
Policy Look up
CSP



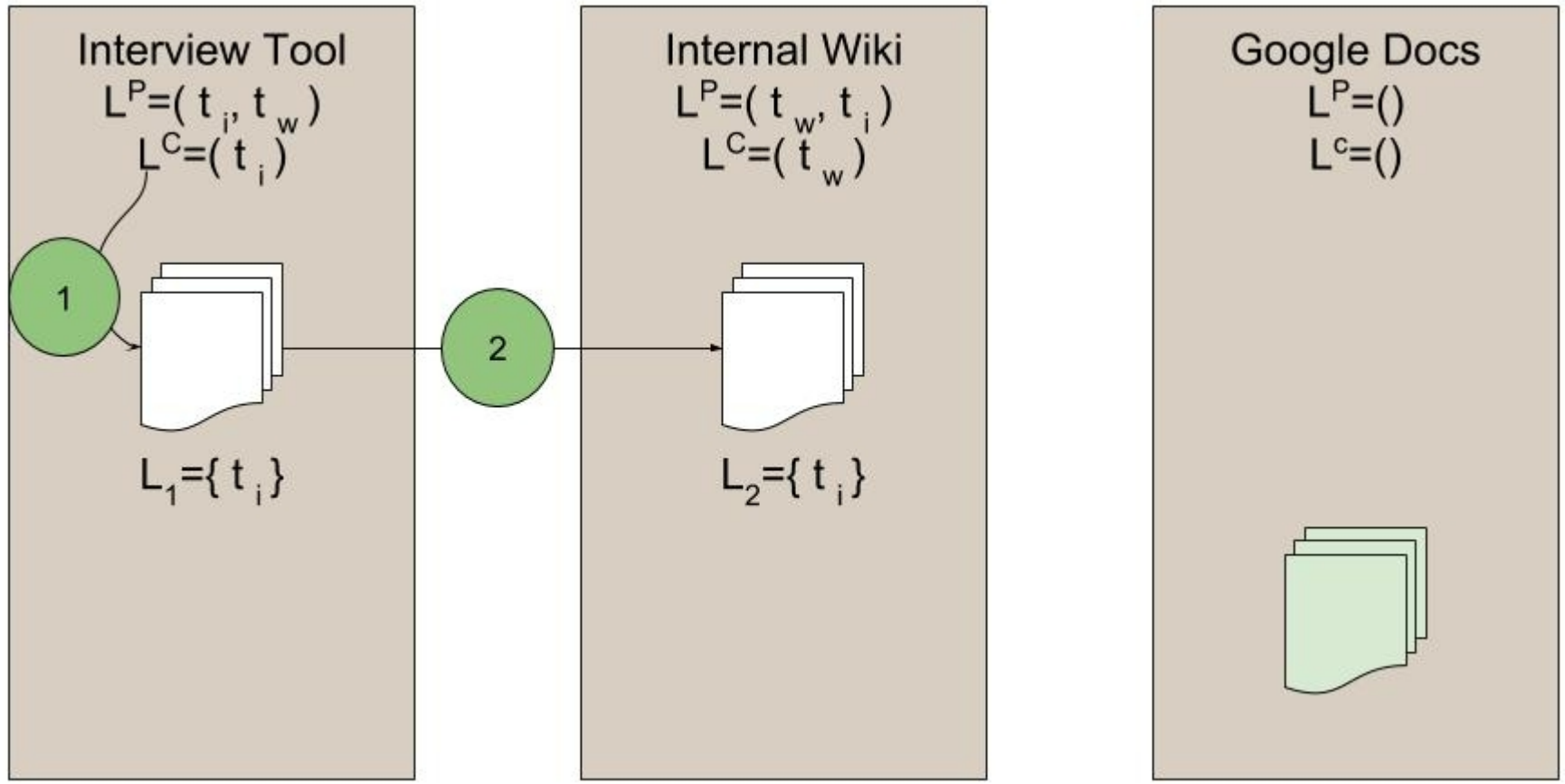




L^p = privilege label
L^c = Confidentiality label

L_i = Text

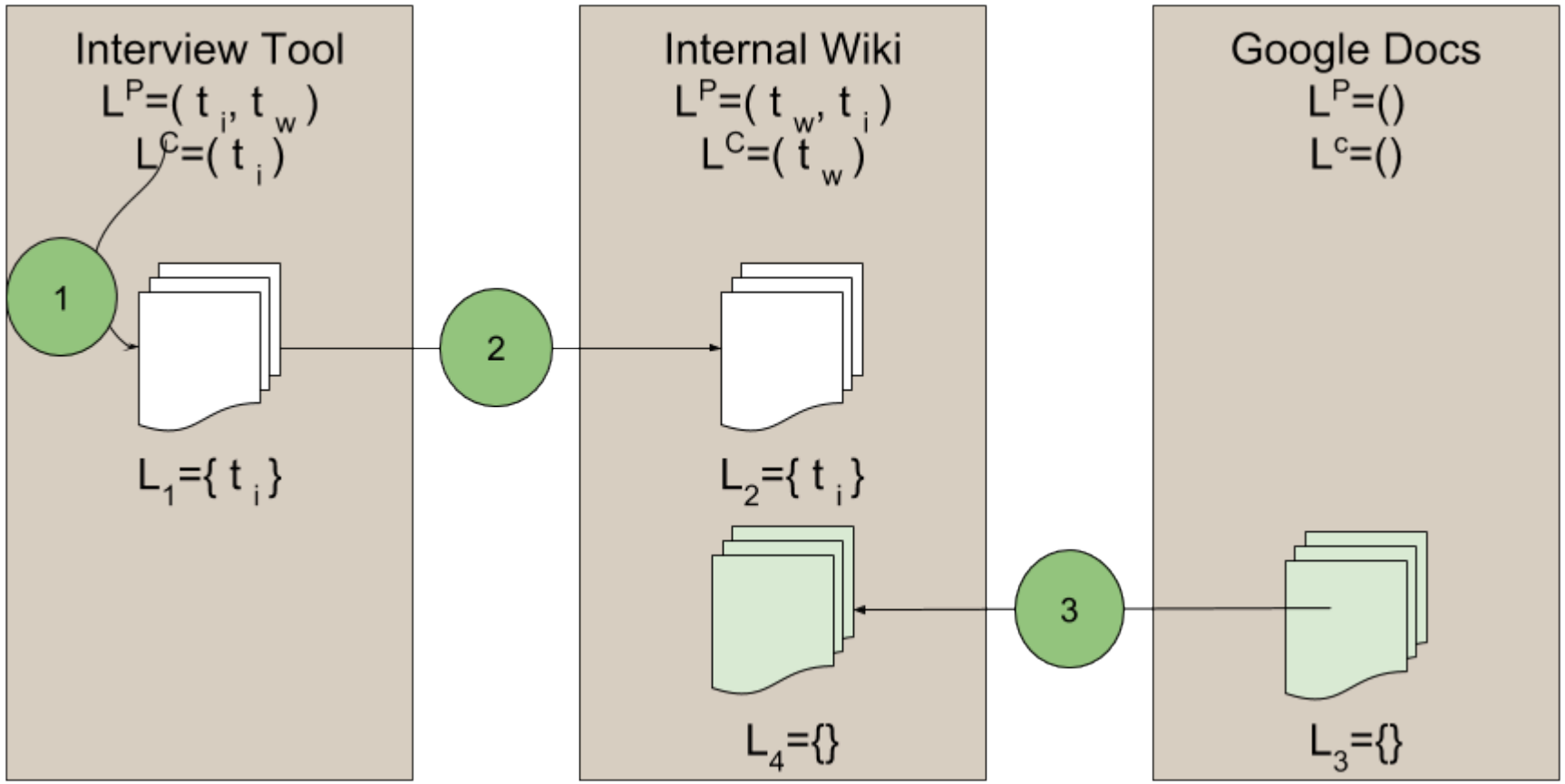
t_i = Tag from Interview



L^p = privilege label
L^c = Confidentiality label

L_i = Text

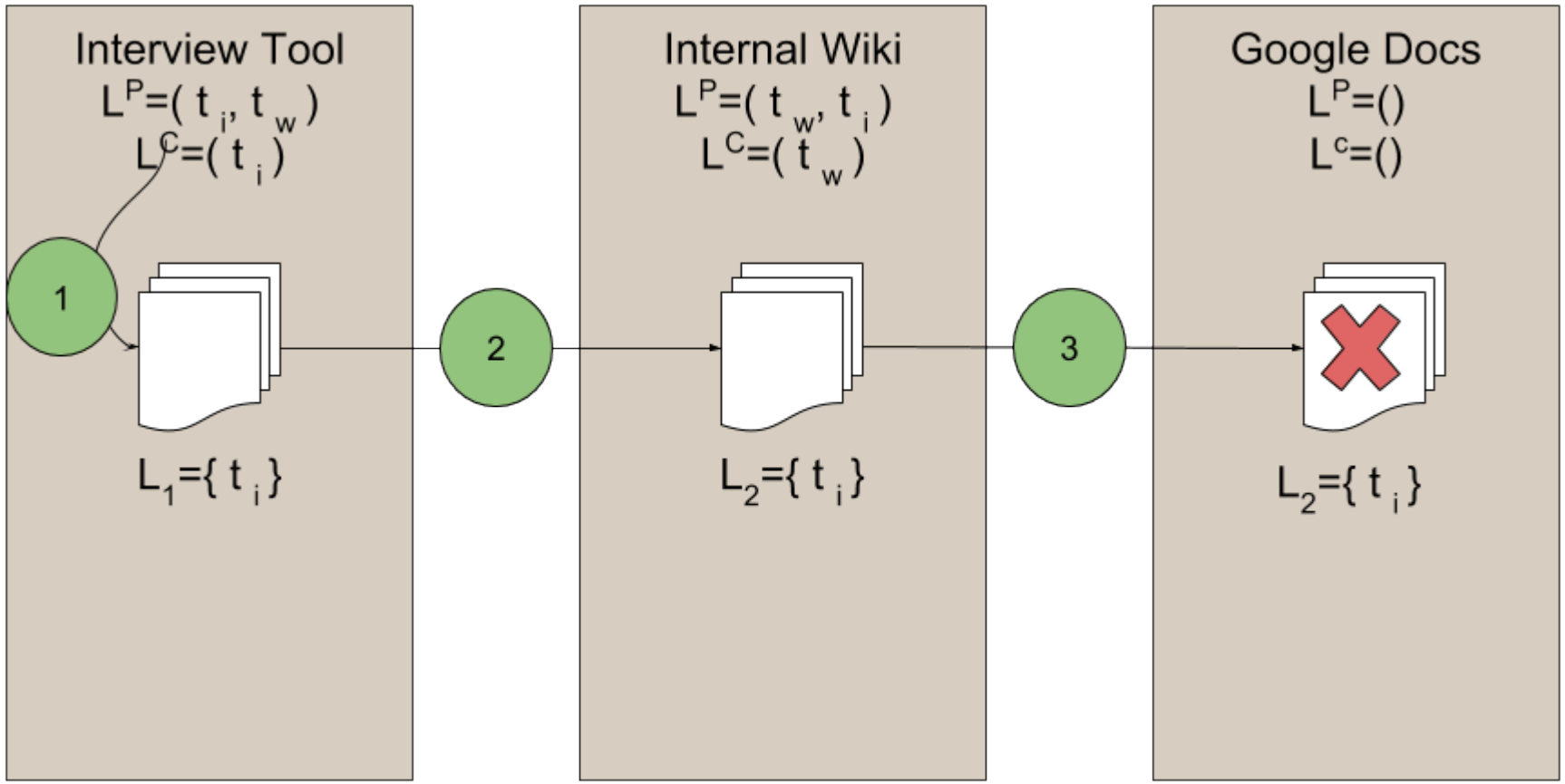
t_i = Tag from Interview



L^p = privilege label
L^c = Confidentiality label

L_i = Text

t_i = Tag from Interview



L^p = privilege label
L^c = Confidentiality label

L_i = Text

t_i = Tag from Interview

- By using an plagiarism detection algorithm to create a digital fingerprint

- There algorithms are well studied problems

- There are four steps that are used to calculate the fingerprint of a text segment

- Normalize the text by removing punctuation, whitespace and character case

“Hello World!”

“helloworld”

- Taking an *n*th-gram
- Example of 6-grams

“hellow”, “ellowo”, “llowor”, “loworl”, “oworld”

- Create a hash values from the grams

{51, 42, 53, 10, 22}

- Creating an overlapping set of hashes

{51, 42, 53}, {42,53,10}, {53,10,22}

- Min value to get the fingerprint

{42,10}

- Moving information from Doc A to B

$$D_{doc}(A, B) = \frac{|F_{authorized}(A) \cap F(B)|}{|F(A)|}$$

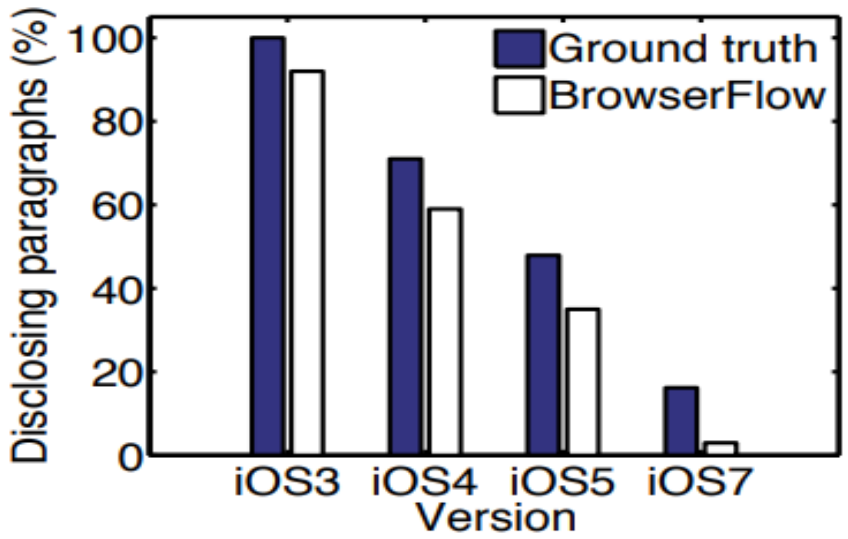
$$D_{par}(A_p, B) = \frac{|F_{authorized}(A_p) \cap F(B)|}{|F(A_p)|}$$

A_p = paragraph

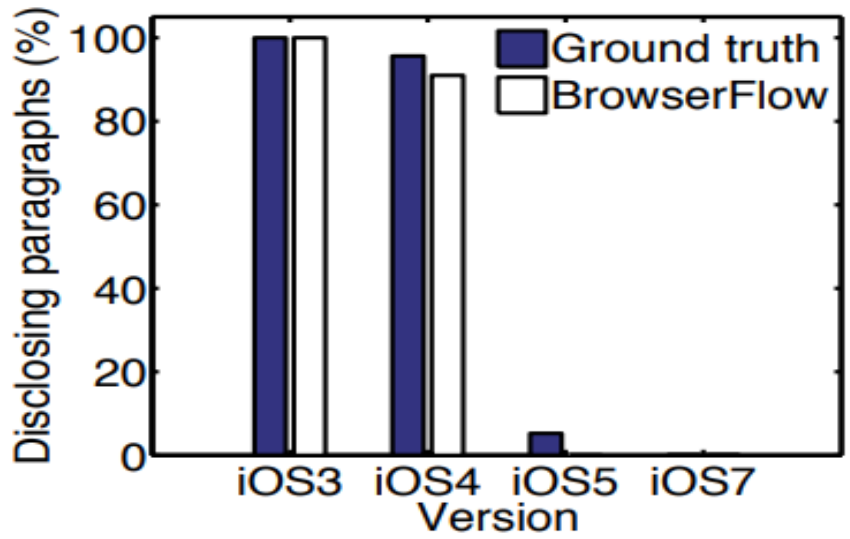
- When the Threshold(T) is $>$ then the amount of disclosure allowed

$$D_{doc}(A,B) \geq T_{doc}(A)\ or\ \exists A_p \in A: D_{par}\left(A_p, B\right) \geq T_{par}(A_p)$$

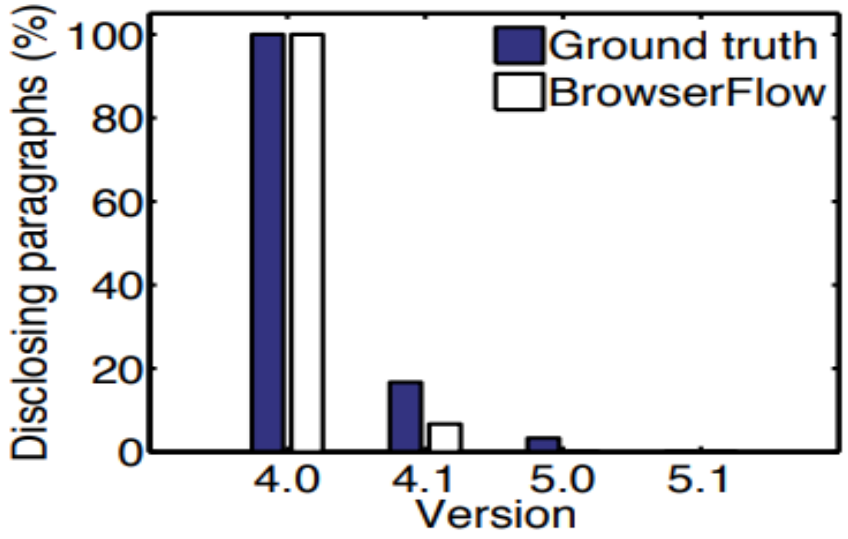
- The document is not allows to upload to an untrusted service.



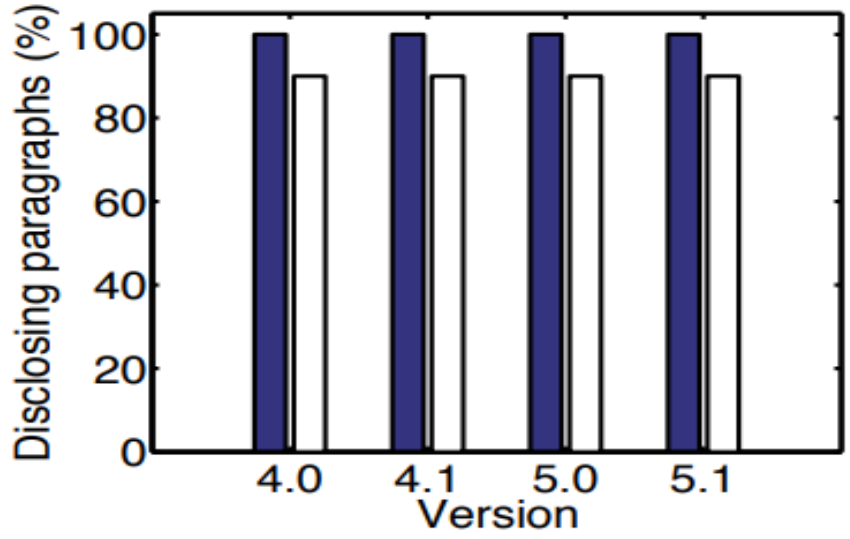
(a) IPhone Camera



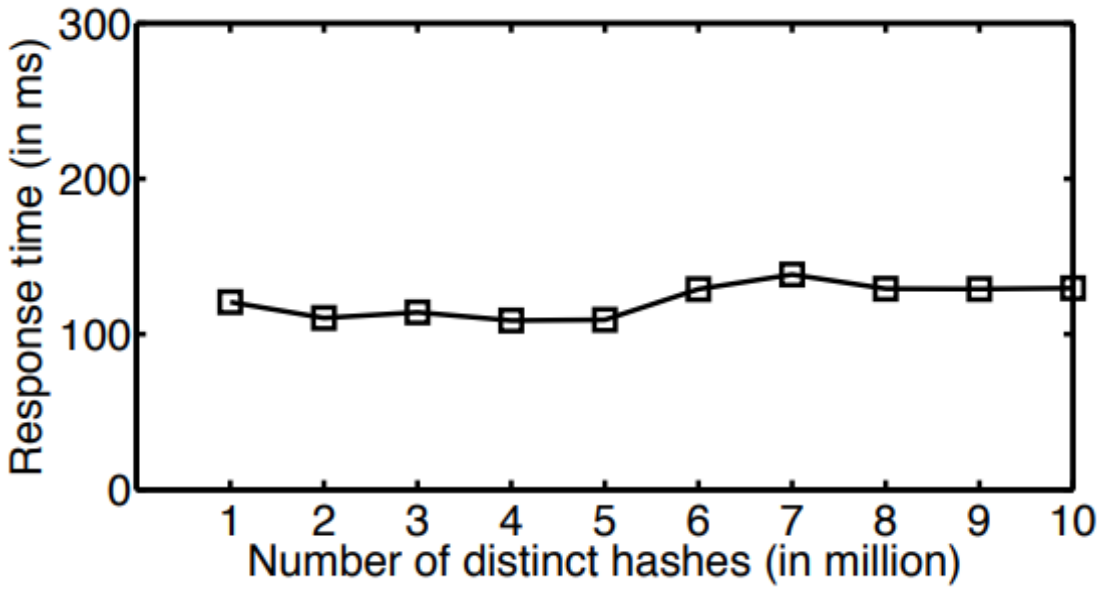
(b) IPhone Message

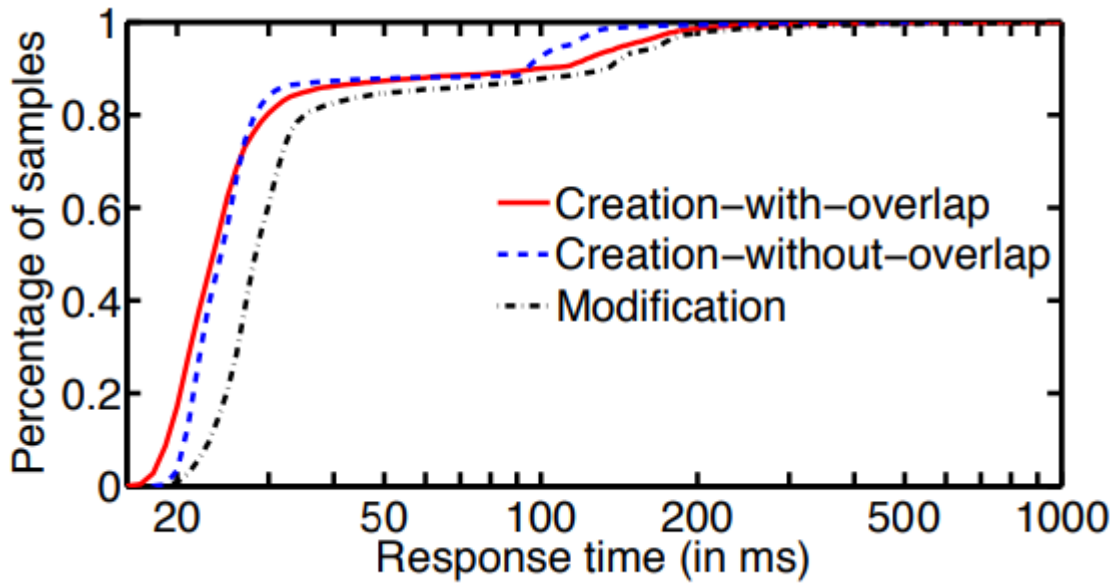


(c) MySQL New Features



(d) MySQL What's MySQL





https://cps-vo.org/

Conclusion

What we covered

- What the Cloud, security that's offered

- Enforce polices by analyzing text on the browser

Questions?

Thanks to Elena Machkasova and Kevin Arhelger for their input