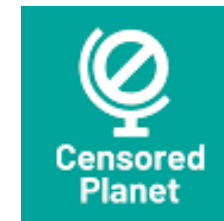


# Modeling and Detecting Internet Censorship Events

Elisa Tsai, Ram Sundara Raman, Atul Prakash, Roya Ensafi  
University of Michigan



# Internet censorship is pervasive

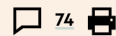
## Turkey tightens internet censorship ahead of elections

New controls on VPNs underline Erdoğan's moves to curb freedom of information



The internet curbs, which curtail access to independent sources of news and information, come ahead of local elections in March © Kerem Uzel/Bloomberg

Adam Samson in Ankara JANUARY 14 2024



Financial Times, Jan 2024

## Russia's Online Censorship Has Soared 30-Fold During Ukraine War

A report from Citizen Lab laid out how much online censorship has increased on one of Russia's biggest social media sites.



The New York Time, Jul 2023

# Internet censorship is pervasive

Numerous peer-reviewed publications on Internet censorship:

**A Study of China's Censorship and Its Evasion Through the Lens of Online Gaming**

**How Great is the Great Firewall? Measuring China's DNS Censorship**

**Network Responses to Russia's Invasion of Ukraine in 2022:  
A Cautionary Tale for Internet Freedom**

**TSPU: Russia's Decentralized Censorship System**

**Measuring and Evading Turkmenistan's Internet Censorship**

A Case Study in Large-Scale Measurements of a Low-Penetration Country

**CensorWatch: On the Implementation of Online Censorship in India**

**A multi-perspective view of Internet censorship in Myanmar**

# Internet censorship is pervasive

Numerous peer-reviewed publications on Internet censorship studies:

**A Study of China's Censorship and Its Evasion Through the Lens of Online Gaming**

**How Great is the Great Firewall? Measuring China's DNS Censorship**

**Network Responses to Russia's Invasion of Ukraine in 2022:  
A Cautionary Tale for Internet Freedom**

**TSPU: Russia's Decentralized Censorship System**

**Measuring and Evading Turkmenistan's Internet Censorship**

A Case Study in Large-Scale Measurements of a Low-Penetration Country

**CensorWatch: On the Implementation of Online Censorship in India**

**A multi-perspective view of Internet censorship in Myanmar**



# Internet censorship is pervasive

Numerous peer-reviewed publications on Internet censorship studies:

**A Study of China's Censorship and Its Evasion Through the Lens of Online Gaming**

**How Great is the Great Firewall? Measuring China's DNS Censorship**

**Network Responses to Russia's Invasion of Ukraine in 2022:  
A Cautionary Tale for Internet Freedom**

**TSPU: Russia's Decentralized Censorship System**

**Measuring and Evading Turkmenistan's Internet Censorship**

A Case Study in Large-Scale Measurements of a Low-Penetration Country

**CensorWatch: On the Implementation of Online Censorship in India**

**A multi-perspective view of Internet censorship in Myanmar**

# Internet censorship is pervasive

Numerous peer-reviewed publications on Internet censorship studies:

**A Study of China's Censorship and Its Evasion Through the Lens of Online Gaming**

**How Great is the Great Firewall? Measuring China's DNS Censorship**

**Network Responses to Russia's Invasion of Ukraine in 2022:  
A Cautionary Tale for Internet Freedom**

**TSPU: Russia's Decentralized Censorship System**

**Measuring and Evading Turkmenistan's Internet Censorship**

A Case Study in Large-Scale Measurements of a Low-Penetration Country

**CensorWatch: On the Implementation of Online Censorship in India**

**A multi-perspective view of Internet censorship in Myanmar**

# Internet censorship is pervasive

Numerous peer-reviewed publications on Internet censorship studies:

**A Study of China's Censorship and Its Evasion Through the Lens of Online Gaming**

**How Great is the Great Firewall? Measuring China's DNS Censorship**

**Network Responses to Russia's Invasion of Ukraine in 2022:  
A Cautionary Tale for Internet Freedom**

**TSPU: Russia's Decentralized Censorship System**

**Measuring and Evading Turkmenistan's Internet Censorship**

A Case Study in Large-Scale Measurements of a Low-Penetration Country

**CensorWatch: On the Implementation of Online Censorship in India**

**A multi-perspective view of Internet censorship in Myanmar**

# Internet censorship is pervasive

Numerous peer-reviewed publications on Internet censorship studies:

**A Study of China's Censorship and Its Evasion Through the Lens of Online Gaming**

**How Great is the Great Firewall? Measuring China's DNS Censorship**

**Network Responses to Russia's Invasion of Ukraine in 2022:  
A Cautionary Tale for Internet Freedom**

**TSPU: Russia's Decentralized Censorship System**

**Measuring and Evading Turkmenistan's Internet Censorship**

A Case Study in Large-Scale Measurements of a Low-Penetration Country

**CensorWatch: On the Implementation of Online Censorship in India**

**A multi-perspective view of Internet censorship in Myanmar**

# Internet censorship

## 1. Domain blocking



**This site can't be reached**

Check if there is a typo in thepiratebays.com

DNS\_PROBE\_FINISHED\_NXDOMAIN

Reload

**Данный ресурс  
заблокирован по решению  
органов государственной власти.**

**Access to the requested resource has been blocked by decision of public authorities.**



# Internet censorship

1. Domain blocking

**2. Content takedowns**

1. Domain admins

2. Social media platforms

...



The Atlantic, Mar 2019

# Internet censorship

1. Domain blocking
2. Content takedowns
3. Circumvention tools blocking

Great Firewall Report Home Publications ▾ Projects

## How China Detects and Blocks Shadowsocks

**Authors:** Anonymous, Anonymous, Anonymous, David Fifield, Amir Houmansadr

Date: Sunday, December 29, 2019

[中文版: Shadowsocks是如何被检测和封锁的](#)

REPORT

## OpenVPN is Open to VPN Fingerprinting

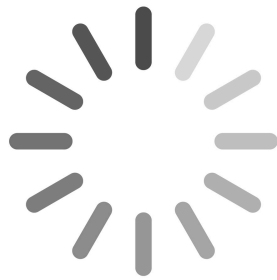
**AUTHORS:** Diwen Xue, Reethika Ramesh, Arham Jain, Michalis Kallitsis, J. Alex Halderman, Jedidiah R. Crandall, Roya Ensafi

**AFFILIATIONS:** Censored Planet at the University of Michigan

**DATE:** November 2, 2022

# Internet censorship

1. Domain blocking
2. Content takedowns
3. Circumvention tools blocking
- 4. Throttling**



Loading...

## REPORT

# Throttling of Twitter in Russia

**AUTHORS:** Diwen Xue, Reethika Ramesh, ValdikSS, Leonid Evdokimov, Andrey Viktorov, Arham Jain, Eric Wustrow, Simone Basso, Roya Ensafi (ensafi@umich.edu)

**AFFILIATIONS:** Censored Planet at the University of Michigan

**DATE:** April 6, 2021

## Elon Musk's X is throttling traffic to websites he dislikes

The site formerly known as Twitter added a five-second delay when a user clicked on a shortened link to the New York Times, Facebook and other sites Musk commonly attacks, a Washington Post analysis found

By [Jeremy B. Merrill](#) and [Drew Harwell](#)

Updated August 16, 2023 at 11:53 a.m. EDT | Published August 15, 2023 at 1:27 p.m. EDT

The New York Time, Aug 2023

# Internet censorship

1. Domain blocking
2. Content takedowns
3. Circumvention tools blocking
4. Throttling
5. Internet shutdowns



**Access Now**  
@accessnow



Election-related shutdowns are much more than a ‘mere inconvenience’ — they take away people’s right to participate in democracy!

If you were affected by the suspension of mobile internet in Pakistan on February 8, share your story with us. [#KeptItOn](#)



[docs.google.com](https://docs.google.com)

Share Your Internet Shutdown Story

Welcome to the Shutdown Stories project by Access Now. If you have ever been a victim of an internet shutdown, we ...

Share Your Internet Shutdown Story

Join the Shutdown Stories project by Access Now. If you have ever been a victim of an internet shutdown, we want to hear from you. This includes disruptions of mobile internet, fixed-line internet, telephone services, or specific apps like WhatsApp, Twitter, or Facebook.

If you write here will guide Access Now's work fighting internet shutdowns through the KeptItOn campaign (<https://www.accessnow.org/keptit-on>). We may publish it on our website through a media outlet.

We respect your privacy and are happy to accept your story under an alias instead of your name. You are also welcome to submit on behalf of someone else. Please use the contact form ([feedback@accessnow.org](mailto:feedback@accessnow.org)) for questions and for email story submissions if you prefer to use this form. (PGP: [info@accessnow.org](mailto:info@accessnow.org))

12:30 AM · Feb 17, 2024 · **1,243** Views



# Internet censorship

## 1. Domain blocking

2. Content takedowns

3. Circumvention tools blocking

4. Throttling

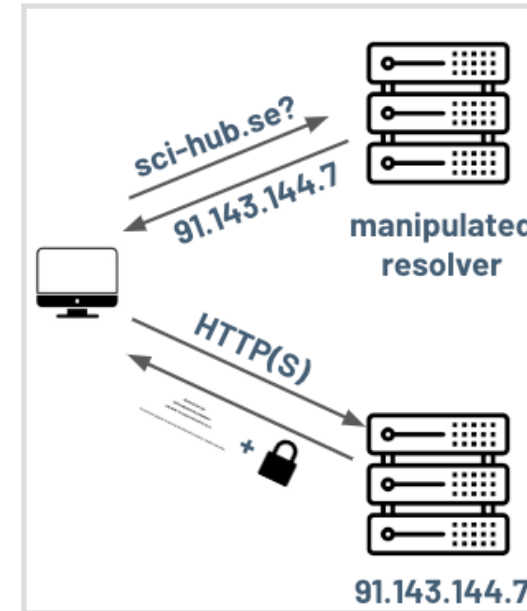
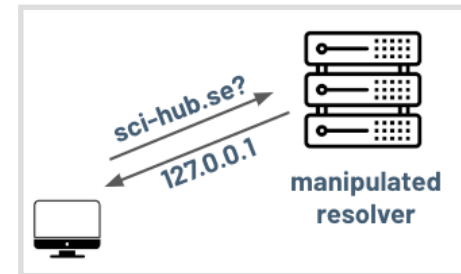
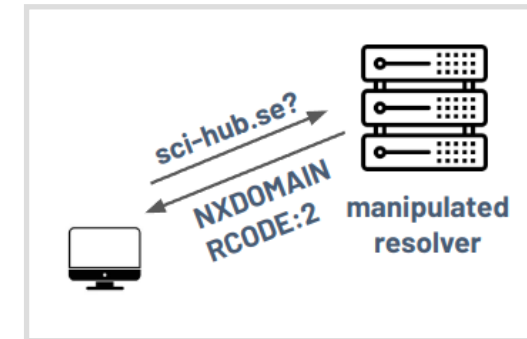
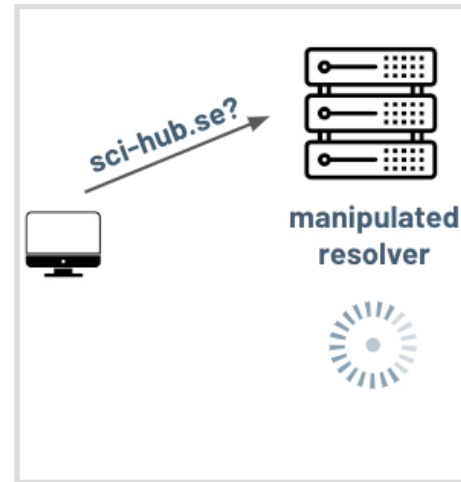
5. Internet shutdowns



# Internet censorship

## 1. Domain blocking

### 1. DNS manipulation

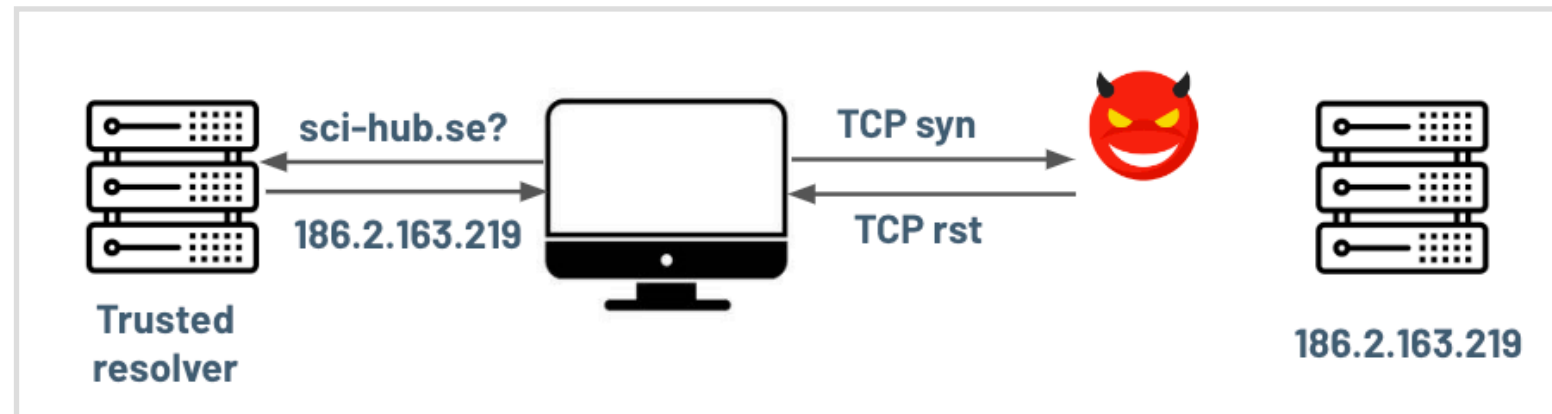


# Internet censorship

## 1. Domain blocking

### 1. DNS manipulation

### 2. IP Blocking



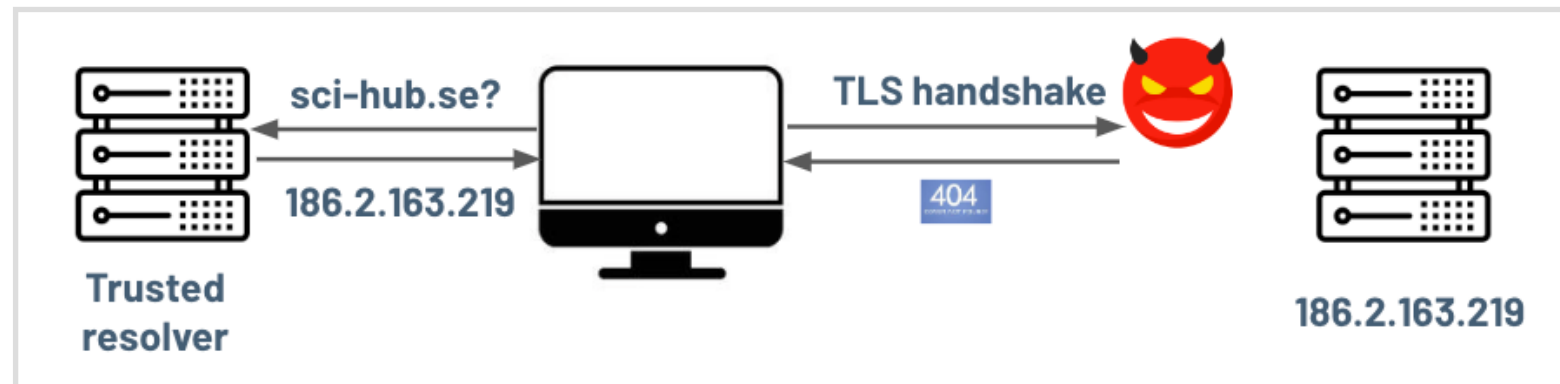
# Internet censorship

## 1. Domain blocking

1. DNS manipulation

2. IP Blocking

**3. HTTP(S) blocking**



# Global Internet censorship observatories



Censored Planet  
Remote Measurement



OONI  
In Situ Measurement

# Global Internet censorship observatories



Censored Planet

**Remote Measurement**



OONI

**In Situ Measurement**



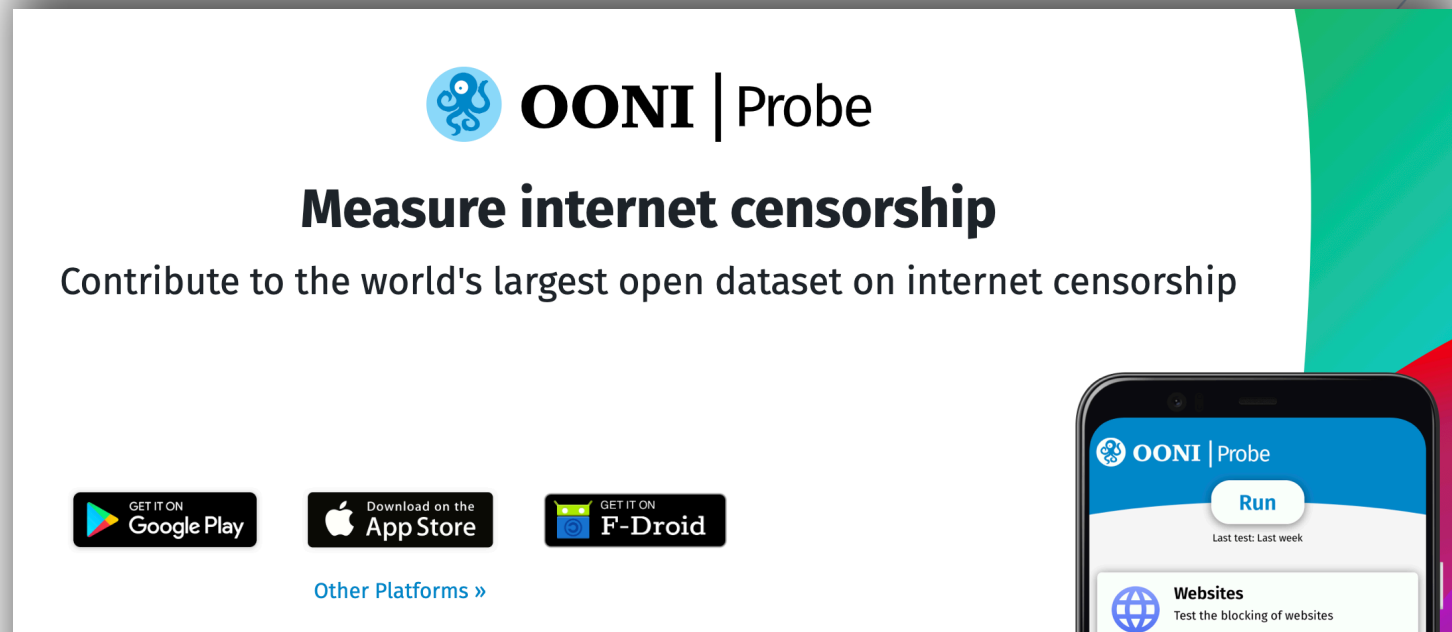
# Global Internet censorship observatories




Censored Planet  
Remote Measurement



OONI  
In Situ Measurement



 **OONI | Probe**

**Measure internet censorship**

Contribute to the world's largest open dataset on internet censorship

GET IT ON  
Google Play

Download on the  
App Store

GET IT ON  
F-Droid

Other Platforms »

**Run**

Last test: Last week

**Websites**  
Test the blocking of websites

# Global Internet censorship observatories



Censored Planet

Remote Measurement

78 B Measurements  
Accumulated (Apr 2023)

241 Countries and Regions



OONI

In Situ Measurement

1.38 B Measurements  
Accumulated (Apr 2023)

223 Countries and Regions

# Global Internet censorship observatories



Censored Planet

Remote Measurement

**78 B Measurements  
Accumulated (Apr 2023)**

241 Countries and Regions



OONI

In Situ Measurement

**1.38 B Measurements  
Accumulated (Apr 2023)**

223 Countries and Regions

**Massive data has been collected for inspection!**

# Global Internet censorship observatories



Censored Planet

Remote Measurement

**78 B Measurements  
Accumulated (Apr 2023)**

241 Countries and Regions



OONI

In Situ Measurement

**1.38 B Measurements  
Accumulated (Apr 2023)**

223 Countries and Regions

**Massive data has been collected for inspection!**

# Internet censorship event detection

Classical Internet censorship event detection pipeline



# Internet censorship event detection

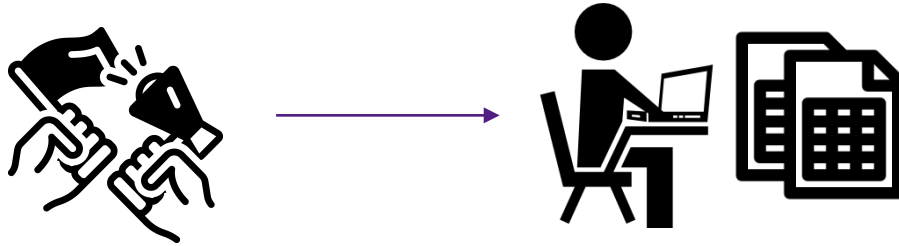
## Classical Internet censorship event detection pipeline



Journalists or on-the-ground activists reach out

# Internet censorship event detection

## Classical Internet censorship event detection pipeline

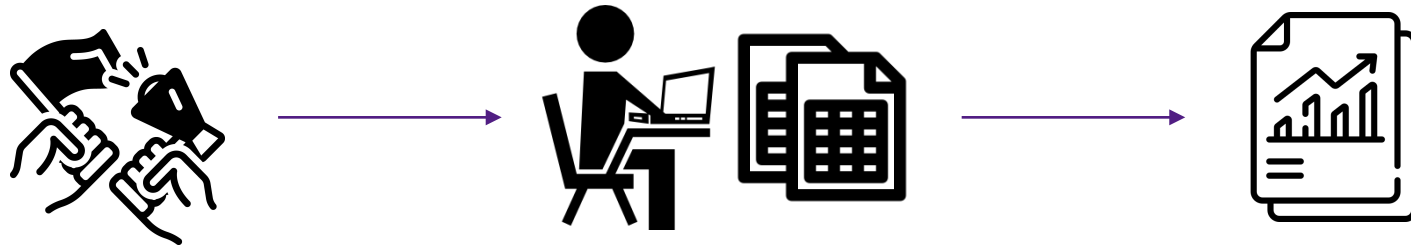


Journalists or on-the-ground activists reach out

Researchers investigate data from recent measurements in the country of interest

# Internet censorship event detection

## Classical Internet censorship event detection pipeline



Journalists or on-the-ground activists reach out

Researchers investigate data from recent measurements in the country of interest

Generate reports

# Internet censorship event detection



OONI

[About](#)

[Tests](#)

[Data](#)

[Get Involved](#)

[Reports](#)

[Blog](#)

[Donate](#)

## Research reports

Grindr blocked in Jordan: Shrinking LGBTQ spaces

20 Sep 2023

Senegal: Social media blocks and network outages amid political unrest

1 Aug 2023

China is blocking OONI

28 Jul 2023

Brazil: OONI data on the blocking of Telegram

28 Apr 2023

Throttling of news media amid Kazakhstan's 2022 presidential election

28 Apr 2023

OONI measurements show ongoing internet censorship in Azerbaijan

28 Feb 2023

How Internet censorship changed in Russia during the 1st year of military conflict in Ukraine

Translations: [ru](#)

24 Feb 2023

Ethiopia: Ongoing blocking of social media

15 Feb 2023



GUIDE

## KeptOn: Fighting internet shutdowns around the world

In 2021, Access Now's Shutdown Tracker Optimization Project (STOP), in collaboration with the #KeptOn coalition, recorded at least 182 internet shutdowns in 34 countries worldwide, including the longest internet shutdowns on record.

[LEARN MORE](#)

## RESOURCES



FREEDOM OF EXPRESSION

### #KeptOn: 2022 elections and internet shutdowns watch

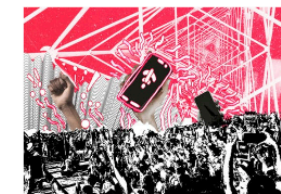
Elections and internet shutdowns are a disaster for democracy. See which...



DIGITAL SECURITY

### Internet shutdowns in 2021: the return of digital authoritarianism

In 2021, Access Now and the #KeptOn coalition documented 18...



DIGITAL SECURITY

### Taxonomy of a shutdown: 8 ways governments restrict access to the internet, and how to #KeptOn

Our new guide explains the technology behind network...



INTERNET SHUTDOWNS

### Internet shutdowns and elections handbook

Internet shutdowns and elections handbook A guide for election...



INTERNET SHUTDOWNS

### KeptOn: frequently asked questions

#KeptOn: A global movement to end



FREEDOM OF EXPRESSION

### Have questions about internet shutdowns? Kill Switch has answers

With this guide, you can learn...



OTHER PODCASTS

### BBC podcast on shutdowns Human rights talks podcast on shutdowns

Human rights talks podcast on shutdowns

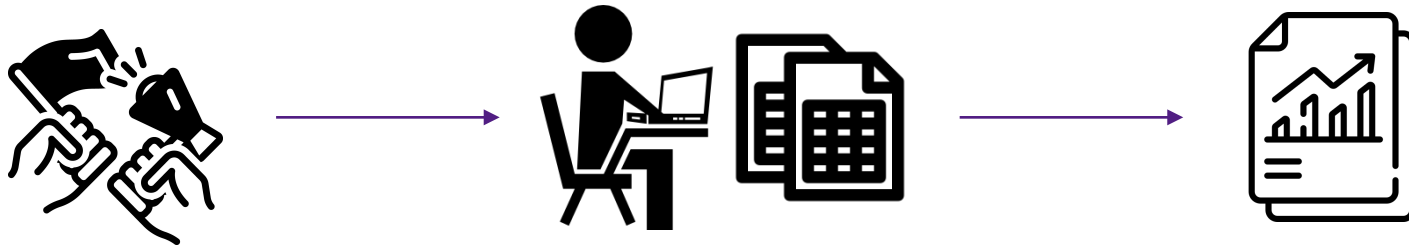


### Digital safety tips if you are disconnected — a framework for individuals to get prepared

Digital safety tips if you are disconnected — a framework for individuals to get prepared

# Internet censorship event detection

## Classical Internet censorship event detection pipeline



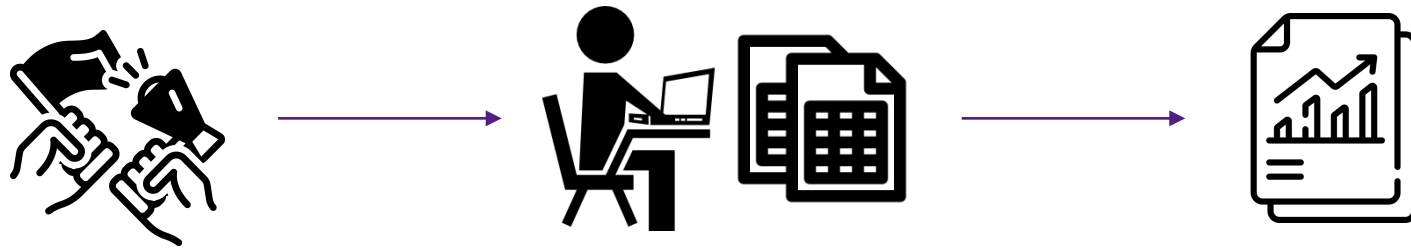
Journalists or on-the-ground activists reach out

Researchers investigate data from recent measurements in the country of interest

Generate reports

# Internet censorship event detection

## Classical Internet censorship event detection pipeline



Journalists or on-the-ground activists reach out

Researchers investigate data from recent measurements in the country of interest

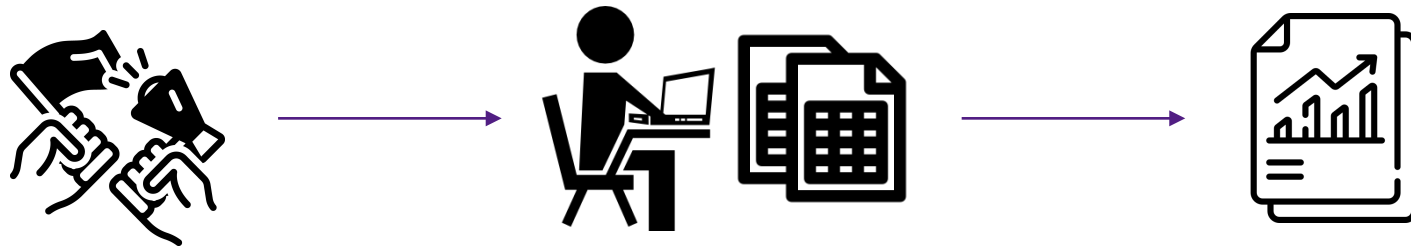
Generate reports



Not scalable on a global scale

# Internet censorship event detection

## Classical Internet censorship event detection pipeline



Journalists or on-the-ground activists reach out

Researchers investigate data from recent measurements in the country of interest

Generate reports



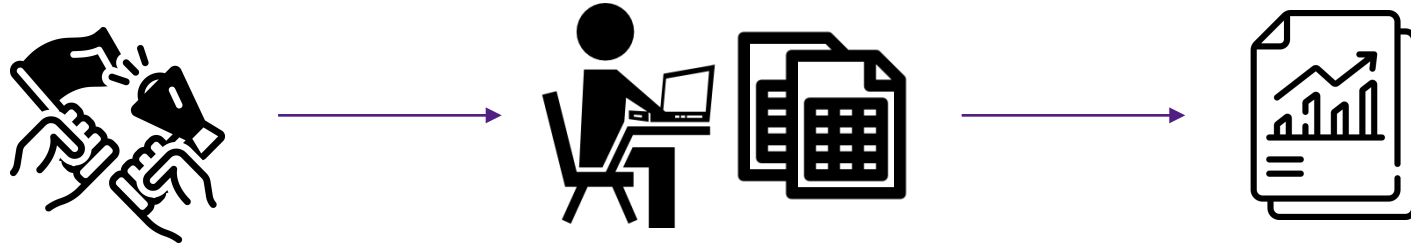
Not scalable on a global scale



Data is wasted if no one reaches out

# Internet censorship event detection

## Classical Internet censorship event detection pipeline



Journalists or on-the-ground activists reach out

Researchers investigate data from recent measurements in the country of interest

Generate reports



Not scalable on a global scale



Data is wasted if no one reaches out

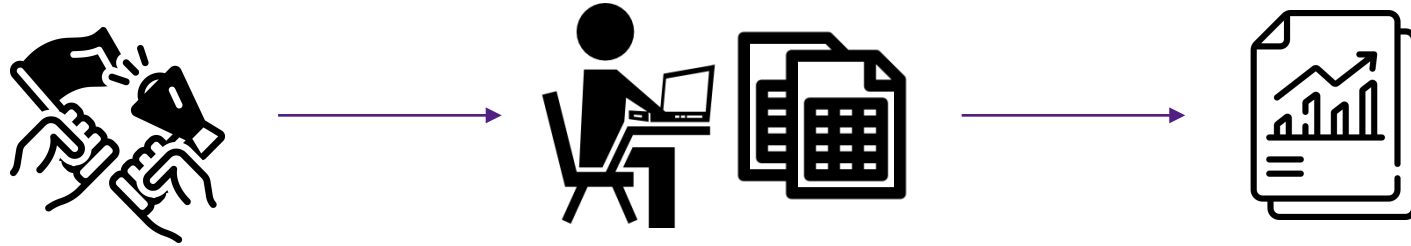


Automated method is needed



# Internet censorship event detection

## Classical Internet censorship event detection pipeline



Journalists or on-the-ground activists reach out

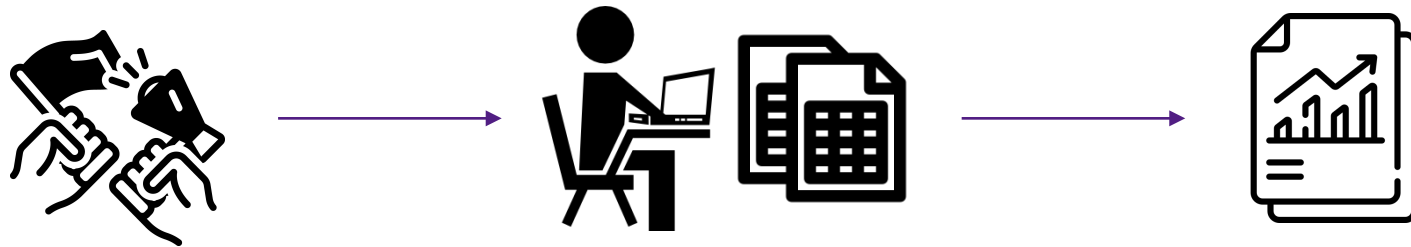
Researchers investigate data from recent measurements in the country of interest

Generate reports

## Other proposed methods

# Internet censorship event detection

## Classical Internet censorship event detection pipeline



Journalists or on-the-ground activists reach out

Researchers investigate data from recent measurements in the country of interest

Generate reports

## Other proposed methods

- **Time-series anomaly detection** techniques: bitmap-based detection. Sundara Raman et al.
- **Supervised Learning** on DNS measurement. Brown et al.

# Internet censorship event detection challenges

## 1. Lack of ground truth.

- **Censorship observatories** stopped labelling “normal/ blocked” and now report various network anomalies.

# Internet censorship event detection challenges

## 1. Lack of ground truth.

- **Censorship observatories** stopped labelling “normal/ blocked” and now report various network anomalies.

⚠ Binary classification based on a fixed set of heuristics is error-prone.

# Internet censorship event detection challenges

## 1. Lack of ground truth.

- **Censorship observatories** stopped labelling “normal/ blocked” and now report various network anomalies.

# Internet censorship event detection challenges

## 1. Lack of ground truth.

- **Censorship observatories** stopped labelling “normal/ blocked” and now report various network anomalies.
- **Censors** rarely disclose their blocking policies.

# Internet censorship event detection challenges

## 1. Lack of ground truth.

- **Censorship observatories** stopped labelling “normal/ blocked” and now report various network anomalies.
- **Censors** rarely disclose their blocking policies.



Unsupervised  
Learning

# Internet censorship event detection challenges

## 1. Lack of ground truth.

- **Censorship observatories** stopped labelling “normal/ blocked” and now report various network anomalies.
- **Censors** rarely disclose their blocking policies.

## 2. Measurement volatility.

- Domain test list (Tranco and the Citizen Lab Test List)
- Vantage point list

Unsupervised  
Learning



# Internet censorship event detection challenges

## 1. Lack of ground truth.

- **Censorship observatories** stopped labelling “normal/ blocked” and now report various network anomalies.
- **Censors** rarely disclose their blocking policies.

## 2. Measurement volatility.

- Domain test list (Tranco and the Citizen Lab Test List)
- Vantage point list

## 3. Large volume of data (78 billion data points from Censored Planet from 2019-2022).

Unsupervised Learning

# Internet censorship event detection goal

## 1. Efficiency

# Internet censorship event detection goal

## 1. Efficiency

## 2. Explainability

- Where (geolocation of censorship)?
- When (timespan of censorship)?
- How (blocking method)?
- What (blocklist)?

# Internet censorship event detection goal

## 1. Efficiency

## 2. Explainability

- **Where (geolocation of censorship)?**
- When (timespan of censorship)?
- How (blocking method)?
- What (blocklist)?

# Internet censorship event detection goal

## 1. Efficiency

## 2. Explainability

- Where (geolocation of censorship)?
- **When (timespan of censorship)?**
- How (blocking method)?
- What (blocklist)?

# Internet censorship event detection goal

## 1. Efficiency

## 2. Explainability

- Where (geolocation of censorship)?
- When (timespan of censorship)?
- **How (blocking method)?**
- What (blocklist)?

# Internet censorship event detection goal

## 1. Efficiency

## 2. Explainability

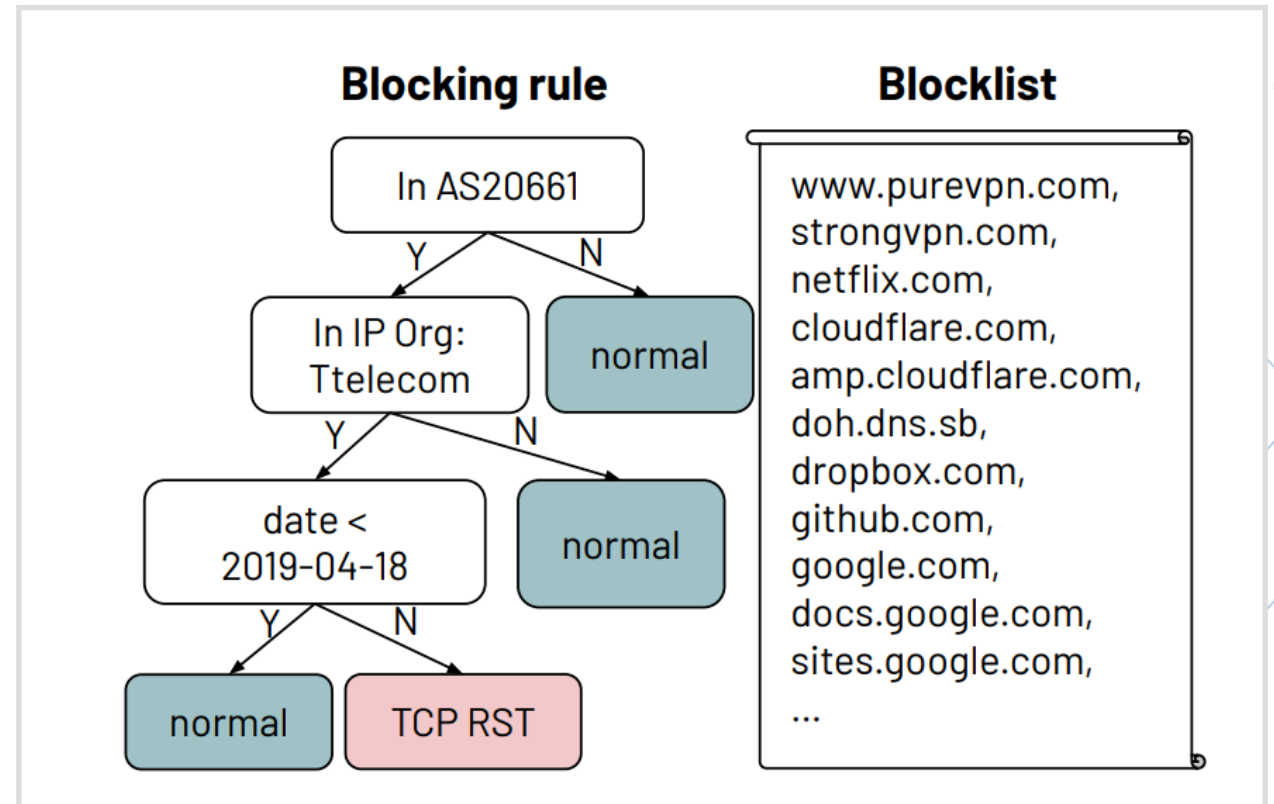
- Where (geolocation of censorship)?
- When (timespan of censorship)?
- How (blocking method)?
- **What (blocklist)?**

# Internet censorship event detection goal

## 1. Efficiency

## 2. Explainability

- Where (geolocation of censorship)?
- When (timespan of censorship)?
- How (blocking method)?
- What (blocklist)?



**Where:** Turkmenistan, AS20661, IP organization Ttelecom

**When:** After 2019 April

**How:** TCP reset

**Blocklist:** domains in the blocklist

**#NDSSSymposium2024**

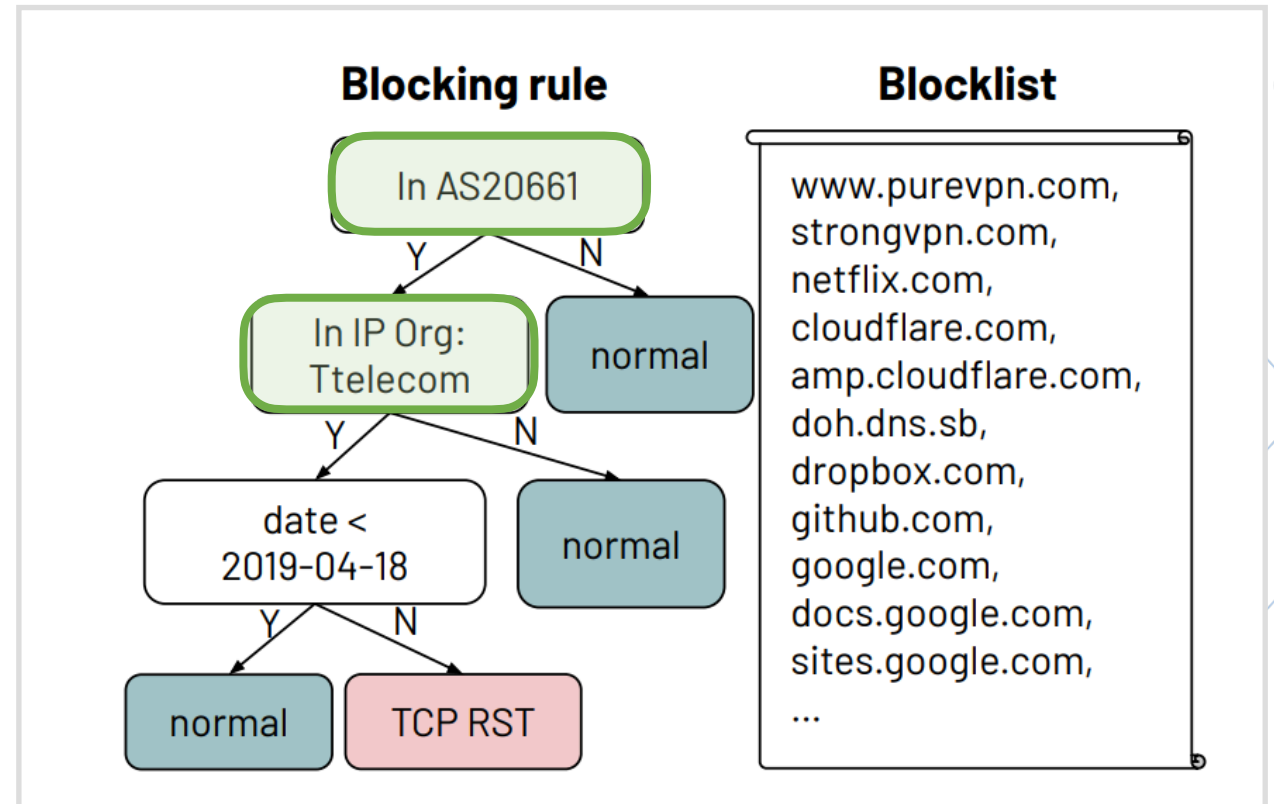


# Internet censorship event detection goal

## 1. Efficiency

## 2. Explainability

- Where (geolocation of censorship)?
- When (timespan of censorship)?
- How (blocking method)?
- What (blocklist)?



**Where:** AS20661 (Turkmenistan), IP organization Ttelecom

**When:** After 2019 April

**How:** TCP reset

**Blocklist:** domains in the blocklist

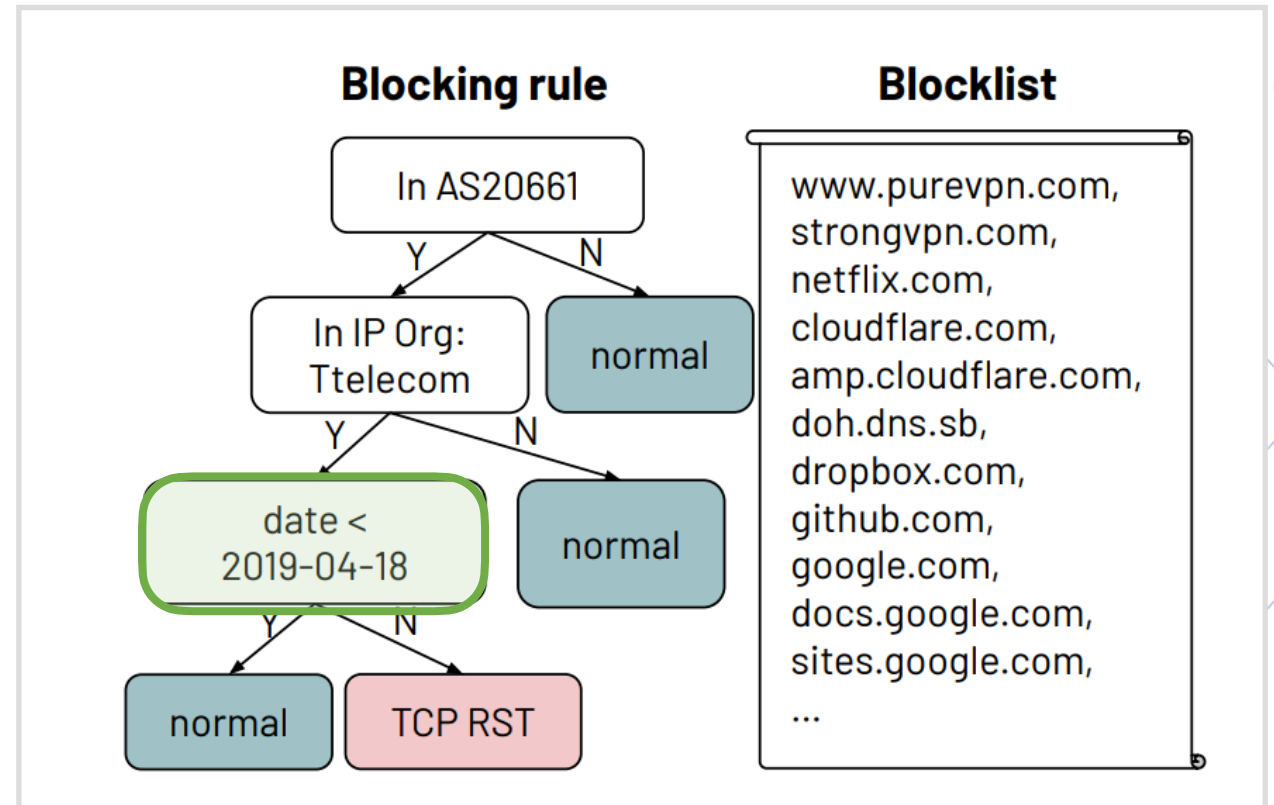
#NDSSSymposium2024

# Internet censorship event detection goal

## 1. Efficiency

## 2. Explainability

- Where (geolocation of censorship)?
- When (timespan of censorship)?
- How (blocking method)?
- What (blocklist)?



**Where:** AS20661 (Turkmenistan), IP organization Ttelecom

**When:** After 2019 April

**How:** TCP reset

**Blocklist:** domains in the blocklist

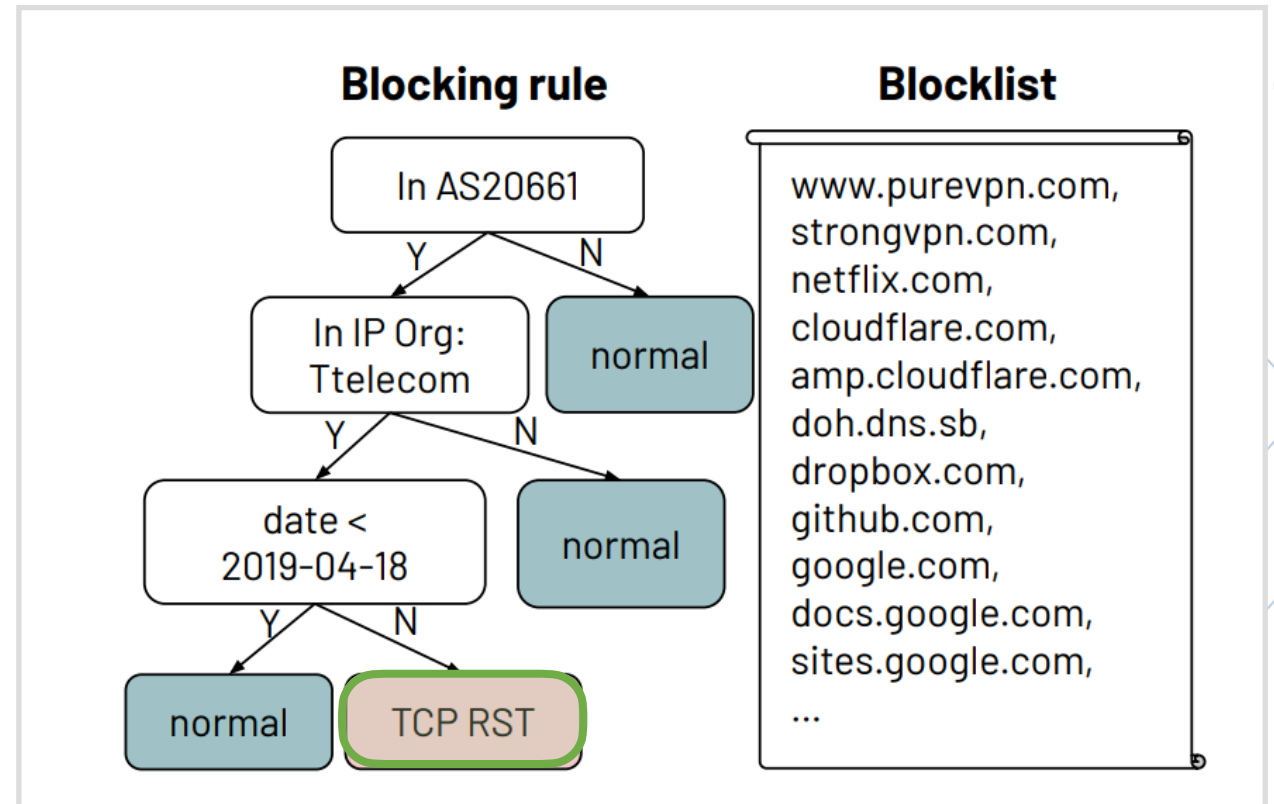
#NDSSSymposium2024

# Internet censorship event detection goal

## 1. Efficiency

## 2. Explainability

- Where (geolocation of censorship)?
- When (timespan of censorship)?
- How (blocking method)?
- What (blocklist)?



**Where:** AS20661 (Turkmenistan), IP organization Ttelecom

**When:** After 2019 April

**How:** TCP reset

**Blocklist:** domains in the blocklist

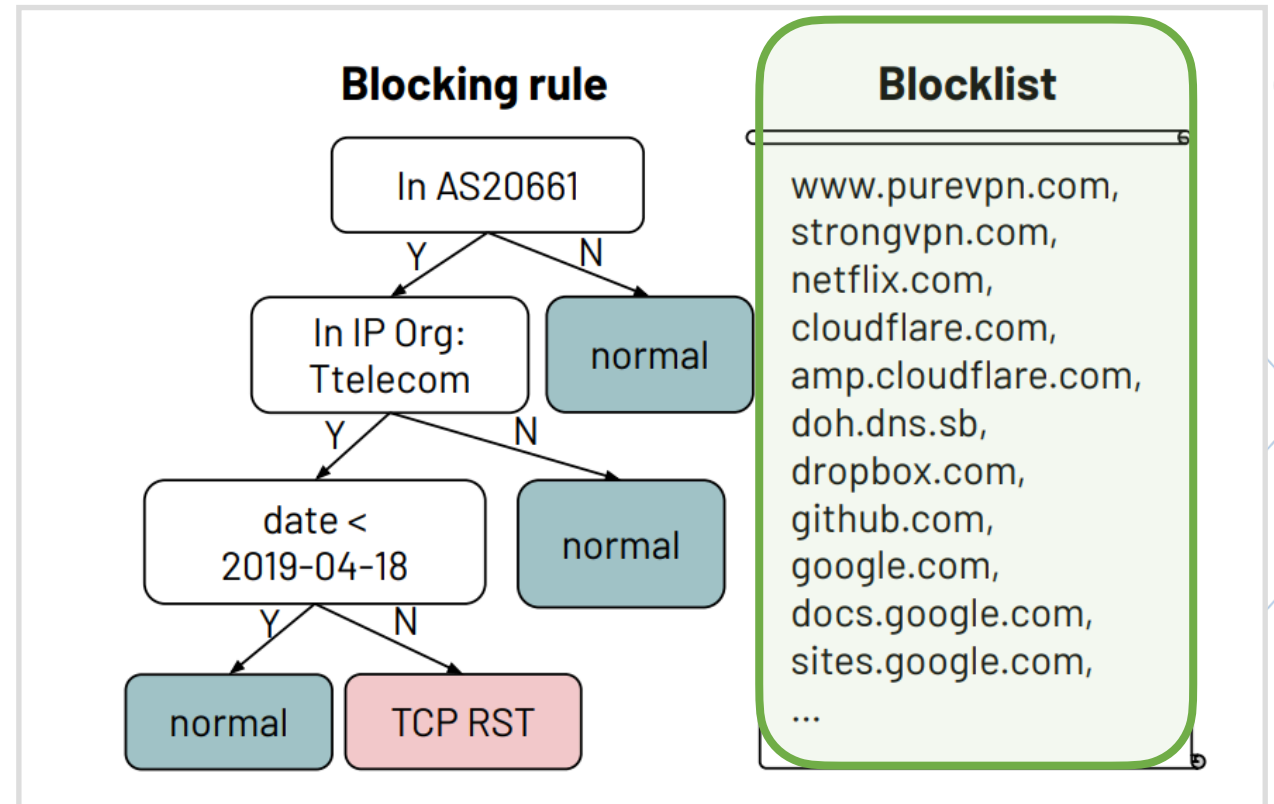
#NDSSSymposium2024

# Internet censorship event detection goal

## 1. Efficiency

## 2. Explainability

- Where (geolocation of censorship)?
- When (timespan of censorship)?
- How (blocking method)?
- What (blocklist)?



**Where:** AS20661 (Turkmenistan), IP organization Ttelecom

**When:** After 2019 April

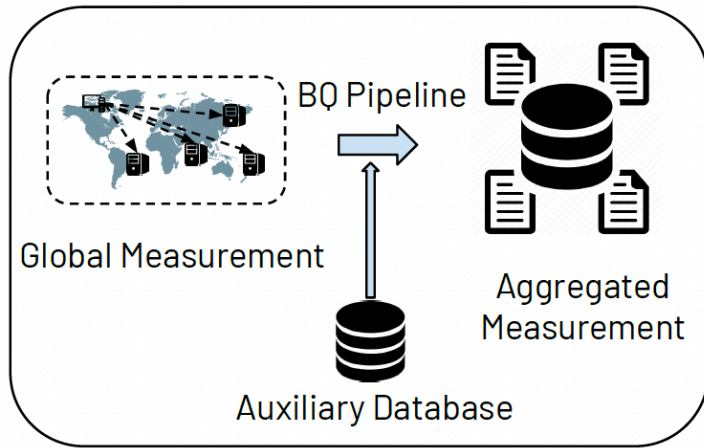
**How:** TCP reset

**Blocklist:** domains in the blocklist

#NDSSSymposium2024

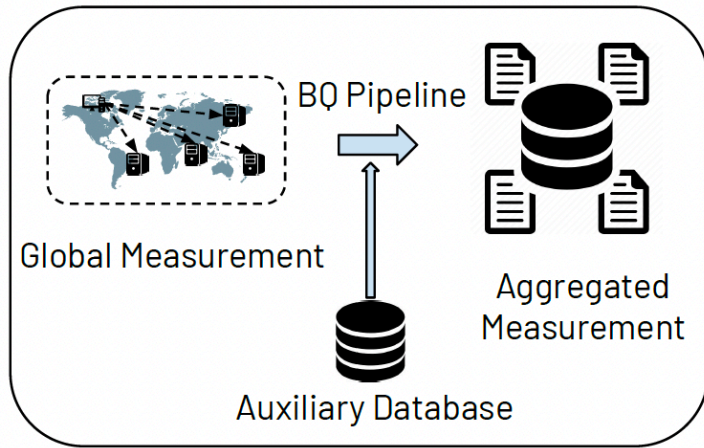
# CenDTect: an unsupervised learning censorship event detection system

(1) CP Raw Data Preprocessing



# CenDTect: an unsupervised learning censorship event detection system

(1) CP Raw Data Preprocessing



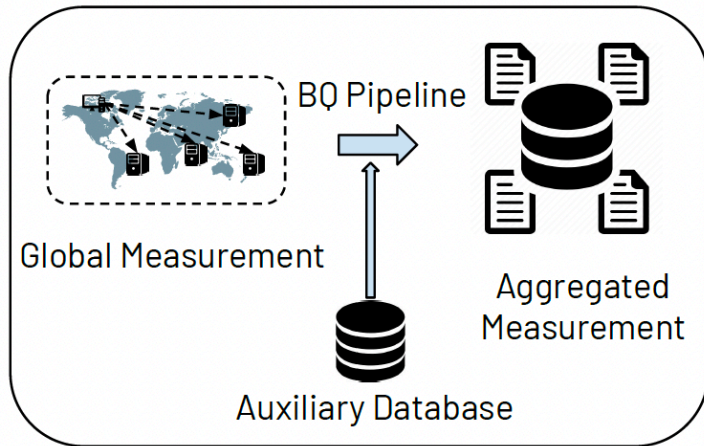
## Assumptions

IP Organization is the unit of censorship

- IP Organizations: the organization listed in the WHOIS record for an IP

# CenDTect: an unsupervised learning censorship event detection system

## (1) CP Raw Data Preprocessing



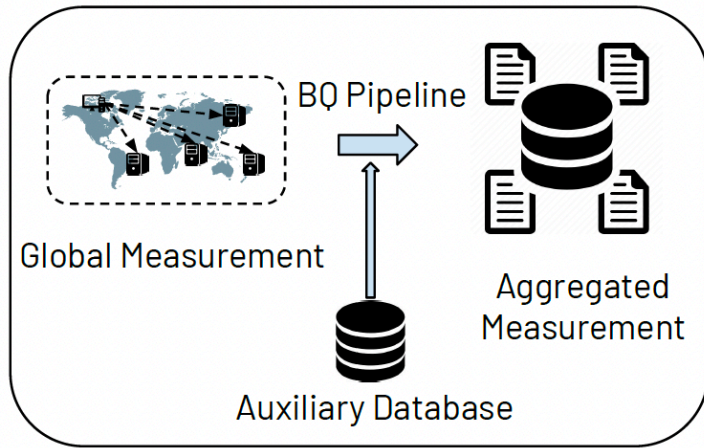
## Assumptions

### IP Organization is the unit of censorship

- IP Organizations: the organization listed in the WHOIS record for an IP
- Network admins of ISPs, university networks, company networks are able to implement and update blocking policies

# CenDTect: an unsupervised learning censorship event detection system

(1) CP Raw Data Preprocessing



## Assumptions

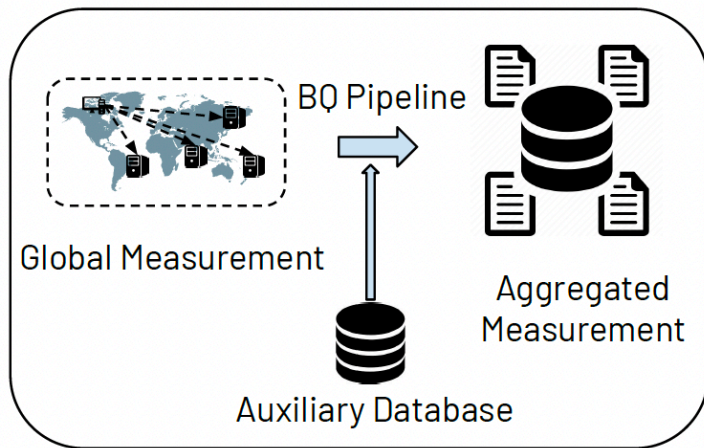
IP Organization is the unit of censorship

→ annotate measurement with IP org information in preprocessing



# CenDTect: an unsupervised learning censorship event detection system

(1) CP Raw Data Preprocessing



## Assumptions

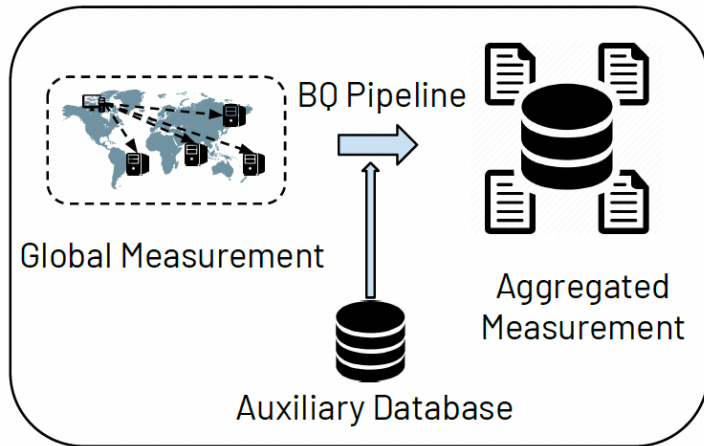
IP Organization is the unit of censorship

→ annotate measurement with IP org information in preprocessing

Censorship events involves simultaneous blocking of multiple domains

# CenDTect: an unsupervised learning censorship event detection system

## (1) CP Raw Data Preprocessing



## Assumptions

IP Organization is the unit of censorship

→ annotate measurement with IP org information in preprocessing

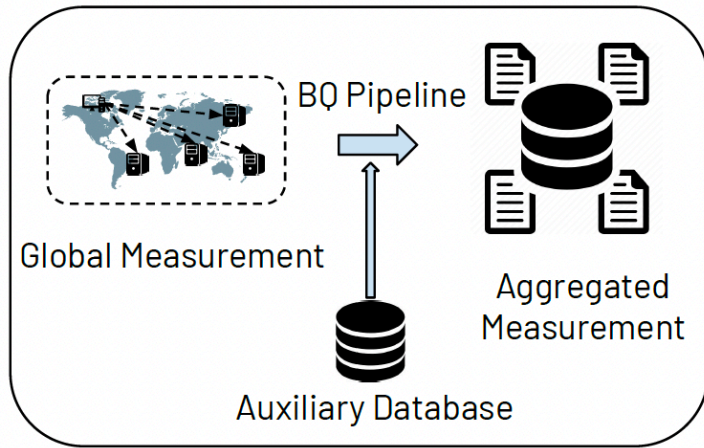
Censorship events involves simultaneous blocking of multiple domains

→ use decision tree to describe blocking behaviors on a set of vantage points

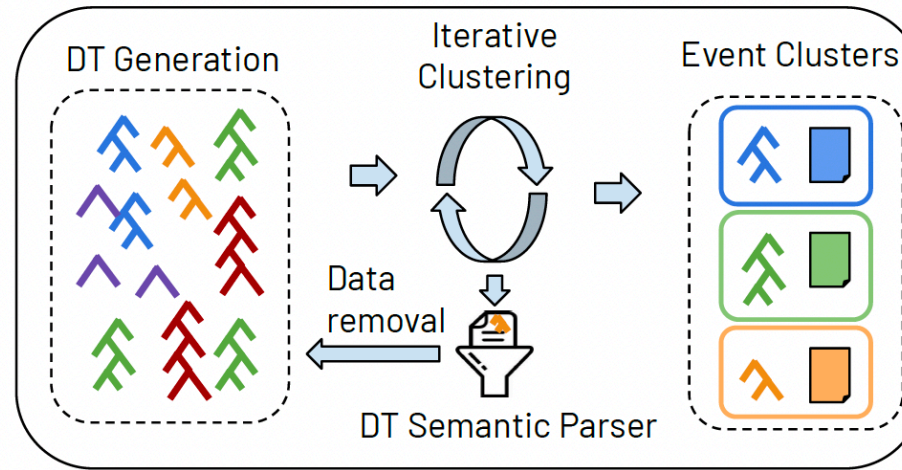
→ cluster domains with the similar blocking behaviors

# CenDTect: an unsupervised learning censorship event detection system

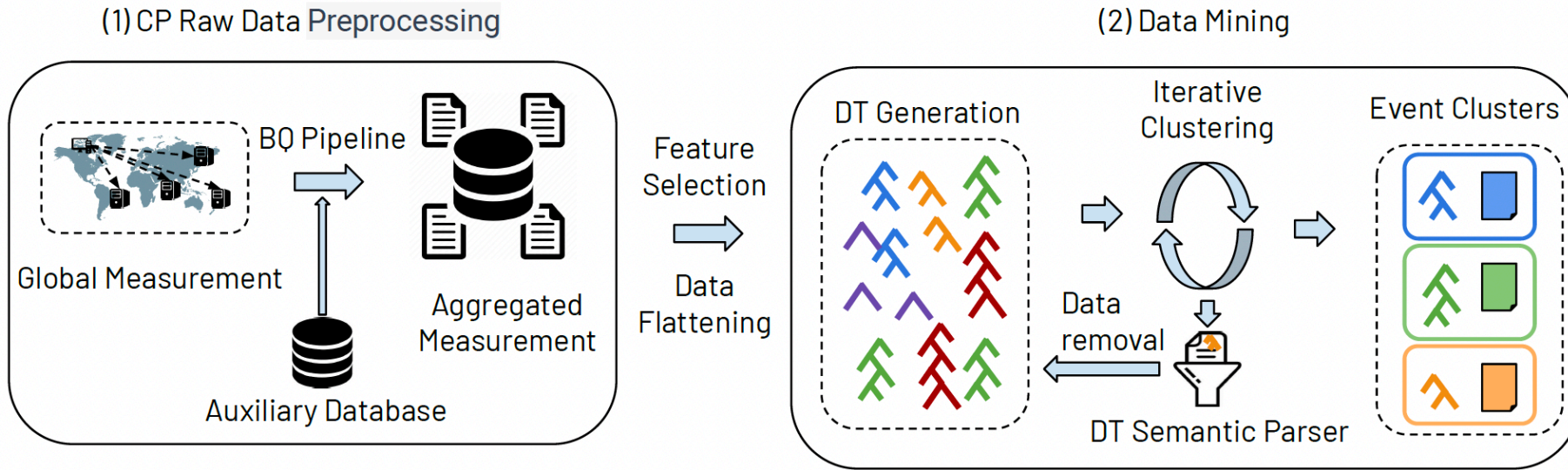
(1) CP Raw Data Preprocessing



(2) Data Mining



# CenDTect: an unsupervised learning censorship event detection system

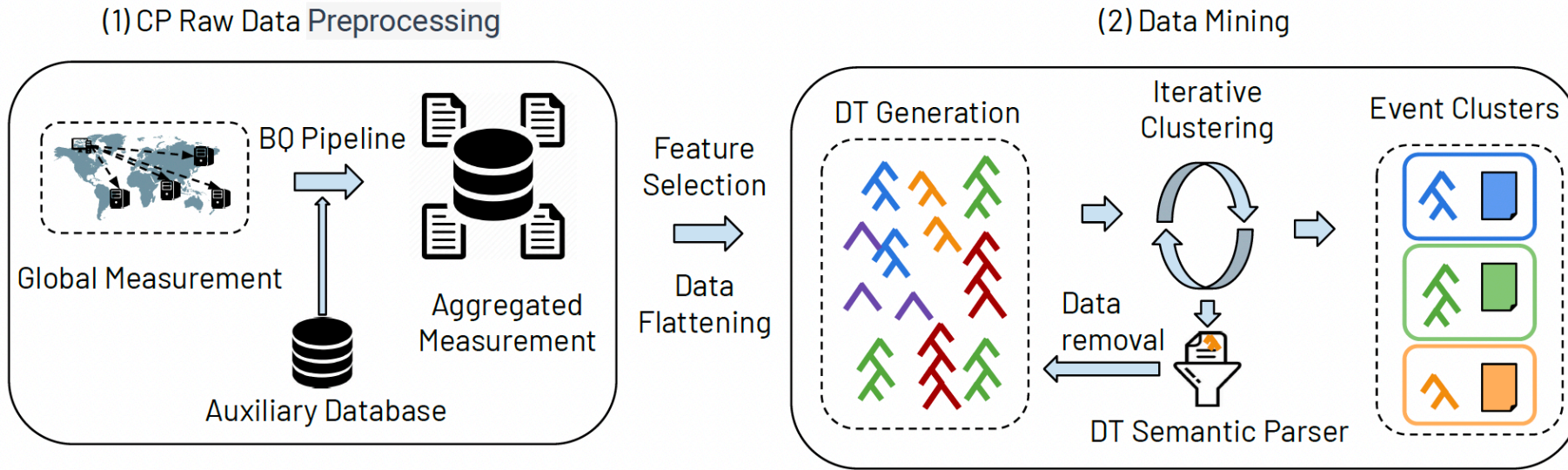


## Implementation

Iterative DBSCAN on a customized distance metrics (“cross-classification accuracy”):

$$f : d_i \rightarrow DT_i, d_j \rightarrow DT_j$$
$$DTDIST(d_i, d_j) = 1 - \max(DT_i.Pred(RES_{d_j}), DT_j.Pred(RES_{d_i}))$$

# CenDTect: an unsupervised learning censorship event detection system



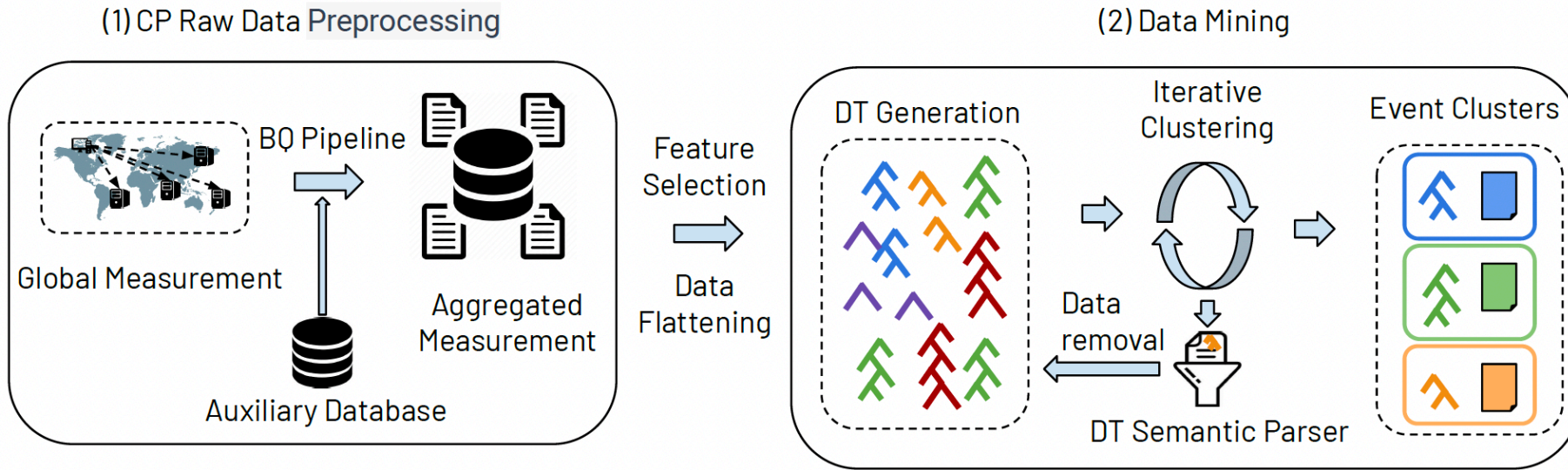
## Implementation

Iterative DBSCAN on a customized distance metrics (“cross-classification accuracy”):

$$f : d_i \rightarrow DT_i, d_j \rightarrow DT_j$$
$$DTDIST(d_i, d_j) = 1 - \max(DT_i.Pred(RES_{d_j}), DT_j.Pred(RES_{d_i}))$$



# CenDTect: an unsupervised learning censorship event detection system



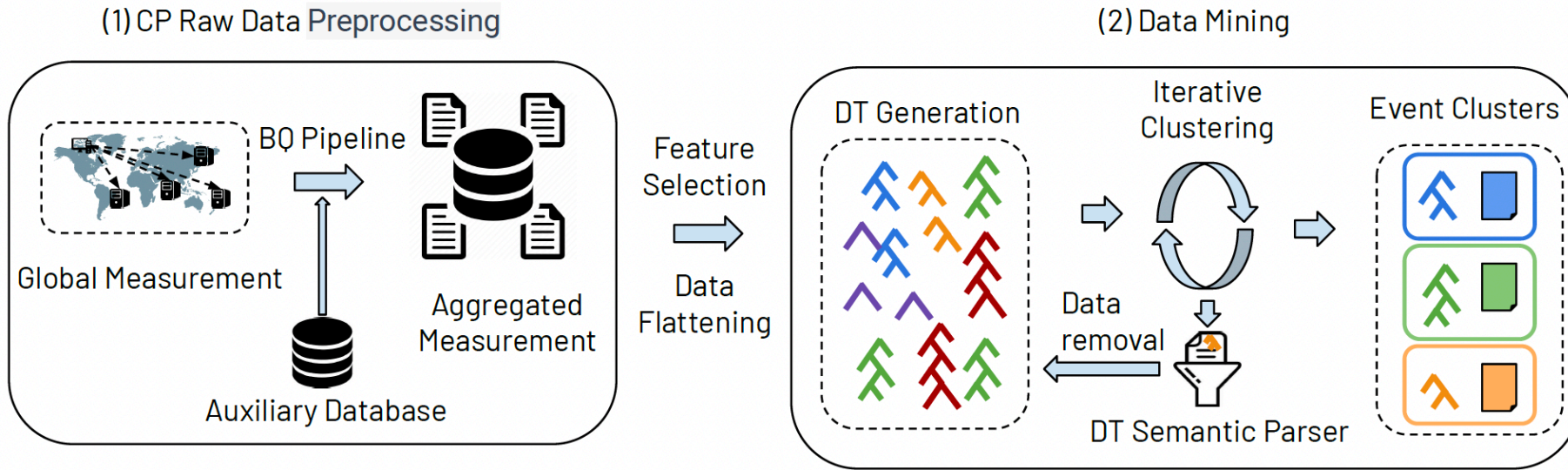
## Implementation

Iterative clustering with adapted prediction:

True Label	$p_1$	$p_1$	$p_1$	0	0
Pred Label	$p_1$	$p_2$	0	0	$p_2$
Pred Success	True	False	<b>True</b>	True	False

0 indicates the domain is accessible, and positive integers indicate different kinds of blocking.  $p_1$  and  $p_2$  are positive integers ( $p_1 \neq p_2$ )

# CenDTect: an unsupervised learning censorship event detection system



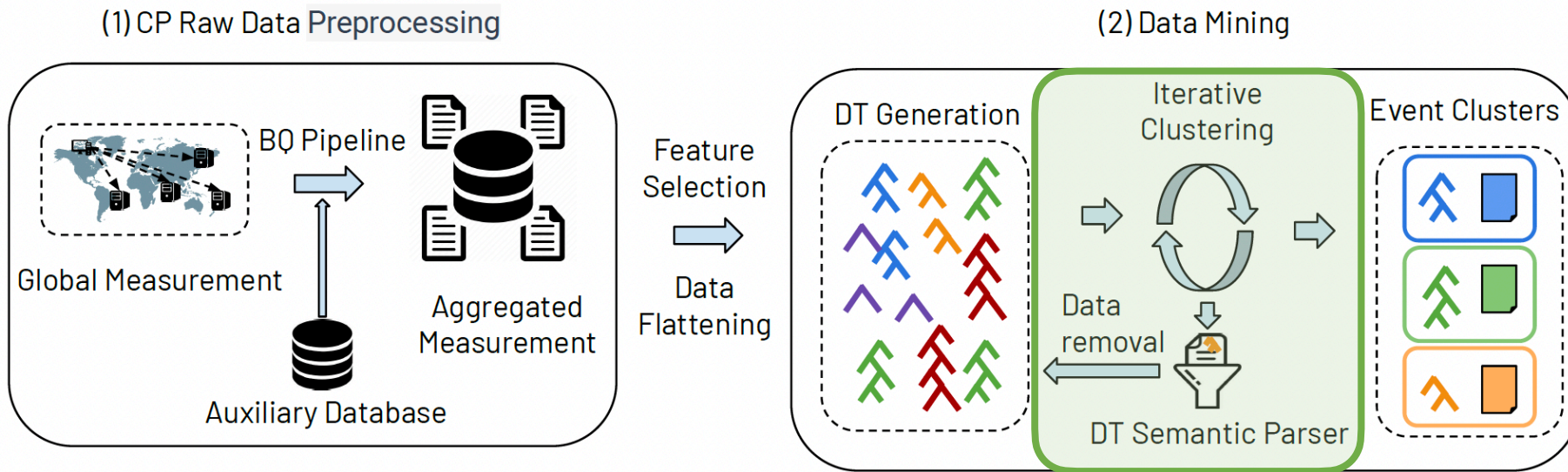
## Implementation

Iterative clustering with adapted prediction:

True Label	$p_1$	$p_1$	$p_1$	0	0
Pred Label	$p_1$	$p_2$	0	0	$p_2$
Pred Success	True	False	<b>True</b>	True	False

0 indicates the domain is accessible, and positive integers indicate different kinds of blocking.  $p_1$  and  $p_2$  are positive integers ( $p_1 \neq p_2$ )

# CenDTect: an unsupervised learning censorship event detection system



## Implementation

Iterative clustering with adapted prediction:

True Label	$p_1$	$p_1$	$p_1$	0	0
Pred Label	$p_1$	$p_2$	0	0	$p_2$
Pred Success	True	False	<b>True</b>	True	False

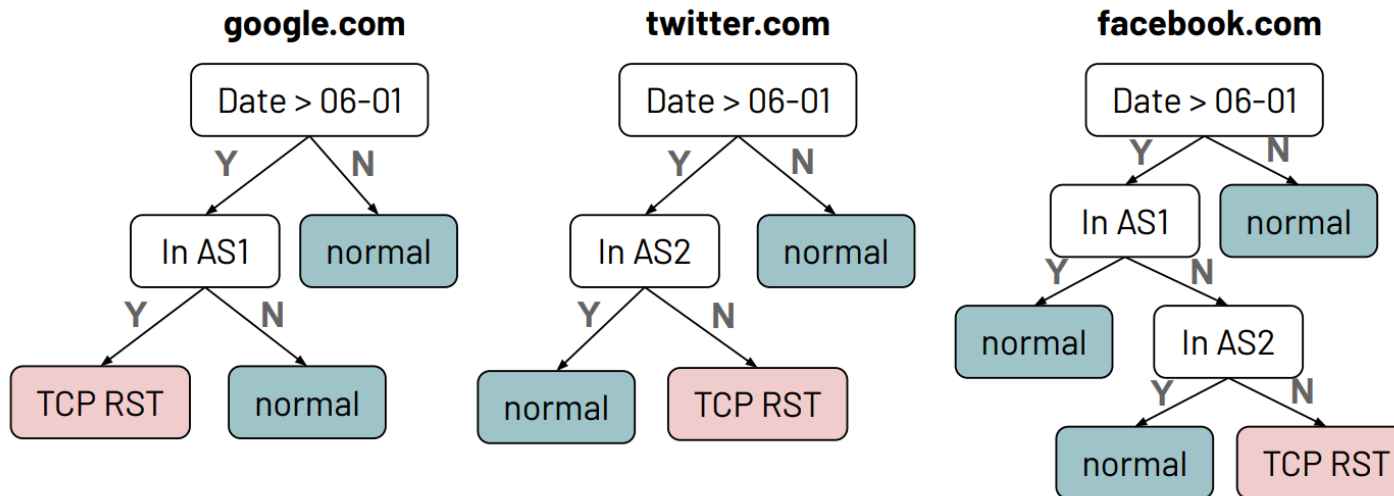
0 indicates the domain is accessible, and positive integers indicate different kinds of blocking.  $p_1$  and  $p_2$  are positive integers ( $p_1 \neq p_2$ )



# CenDTect: an unsupervised learning censorship event detection system

Clustering toy example:

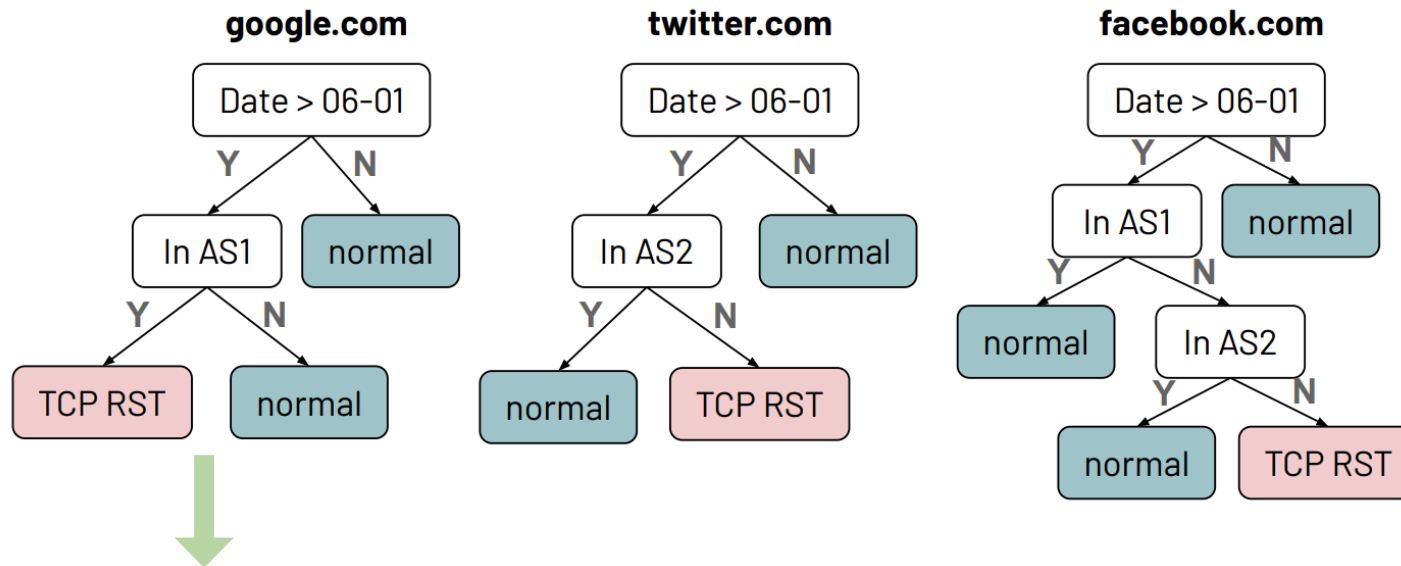
Decision trees for 3 domains in a region that only has 3 ASes: AS1, AS2, AS3.



# CenDTect: an unsupervised learning censorship event detection system

Clustering toy example:

Decision trees for 3 domains in a region that only has 3 ASes: AS1, AS2, AS3.

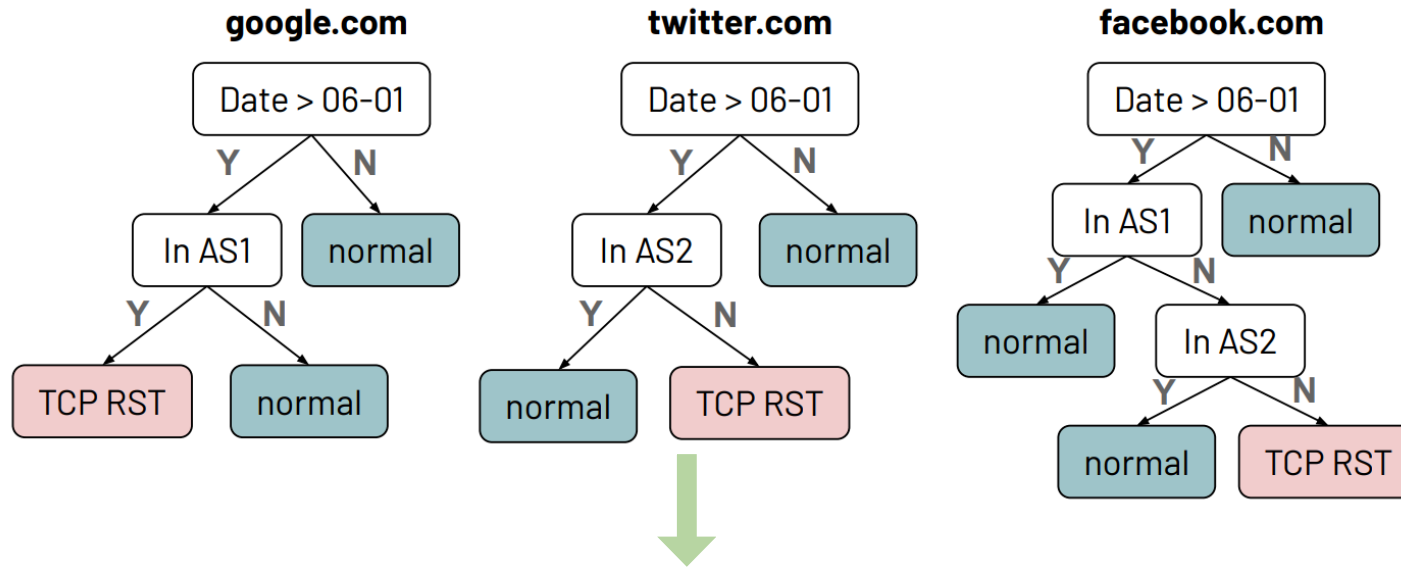


google.com is blocked  
in AS1 via TCP reset  
after June 1.

# CenDTect: an unsupervised learning censorship event detection system

Clustering toy example:

Decision trees for 3 domains in a region that only has 3 ASes: AS1, AS2, AS3.



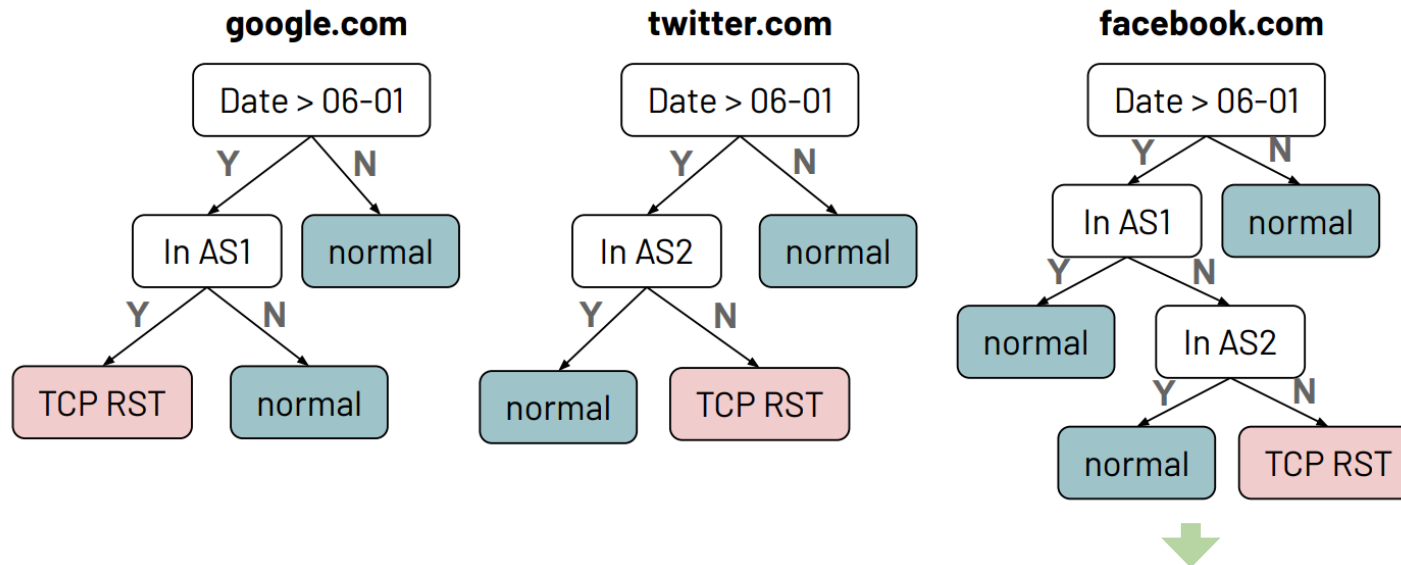
google.com is blocked in AS1 via TCP reset after June 1.

twitter.com is blocked via TCP reset after June 1 in AS1 and AS3.

# CenDTect: an unsupervised learning censorship event detection system

Clustering toy example:

Decision trees for 3 domains in a region that only has 3 ASes: AS1, AS2, AS3.



**google.com** is blocked in AS1 via TCP reset after June 1.

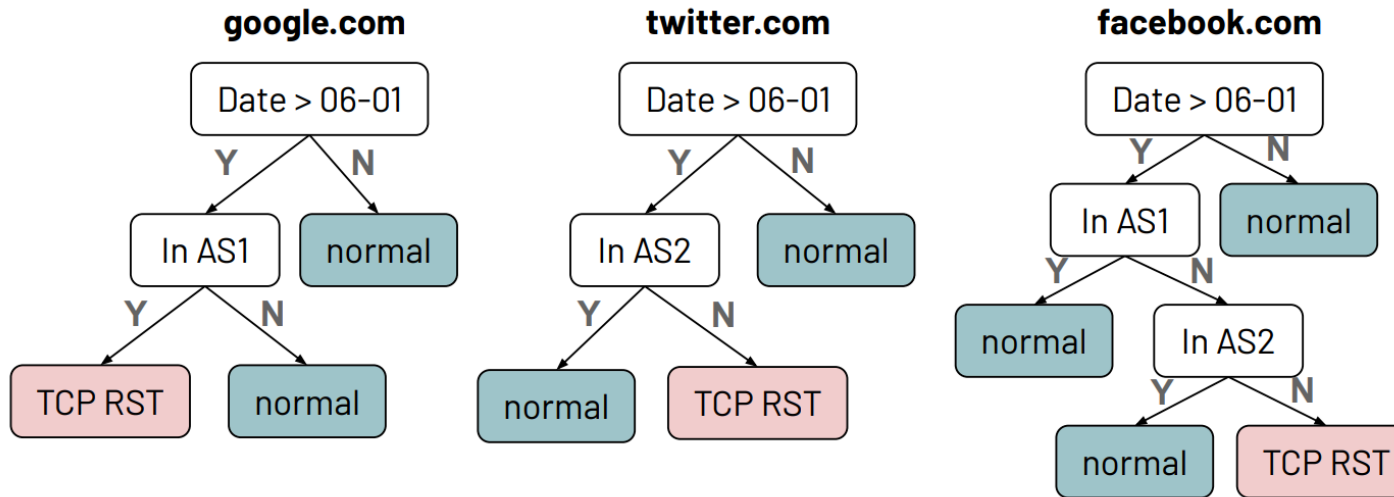
**twitter.com** is blocked via TCP reset after June 1 in AS1 and AS3.

**facebook.com** is blocked via TCP reset after June 1 in AS3.

# CenDTect: an unsupervised learning censorship event detection system

Clustering toy example:

Decision trees for 3 domains in a region that only has 3 ASes: AS1, AS2, AS3.



**google.com** is blocked in AS1 via TCP reset after June 1.

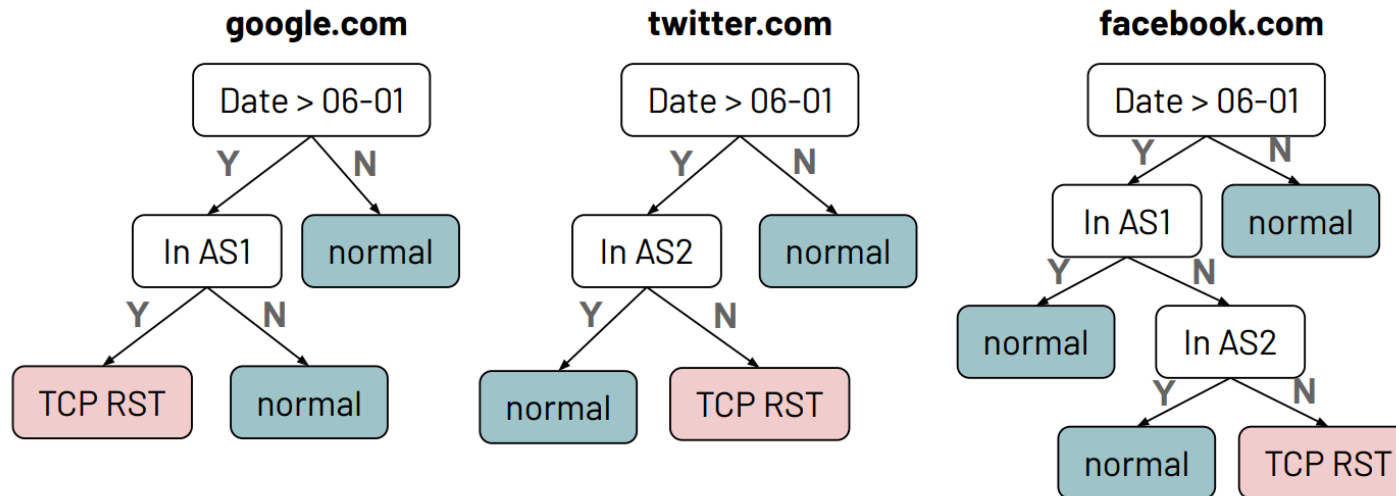
**twitter.com** is blocked via TCP reset after June 1 in AS1 and AS3.

**facebook.com** is blocked via TCP reset after June 1 in AS3.

# CenDTect: an unsupervised learning censorship event detection system

Clustering toy example:

Decision trees for 3 domains in a region that only has 3 ASes: AS1, AS2, AS3.

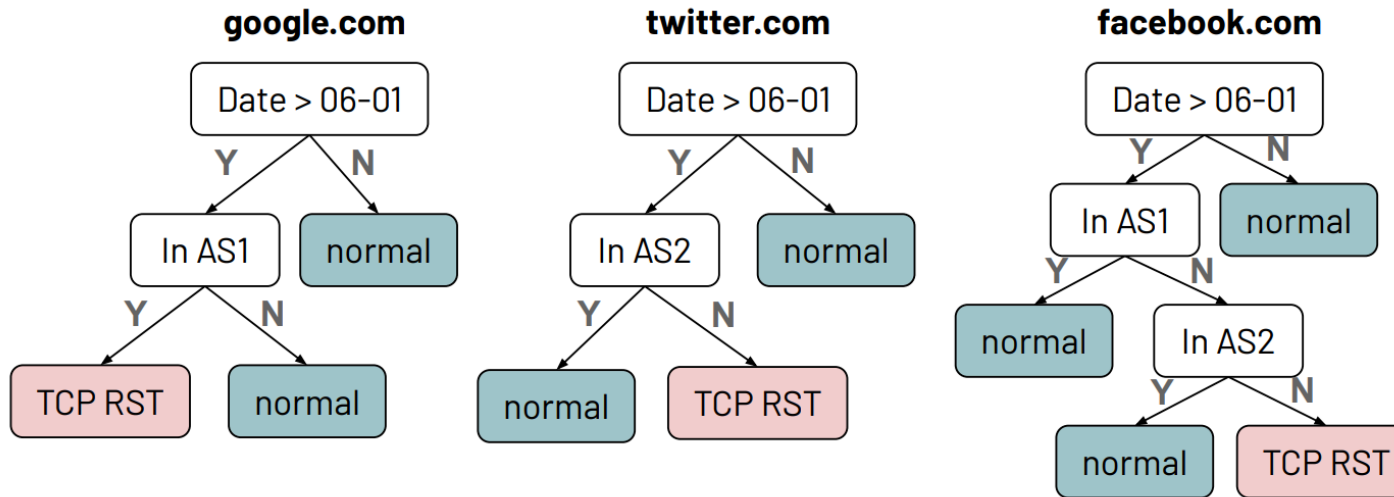


Round 1: Output cluster  $C_1 = \{\text{rule}=\text{DT}_{\text{google}}, \text{domains}=\{\text{google}, \text{twitter}\}\}$

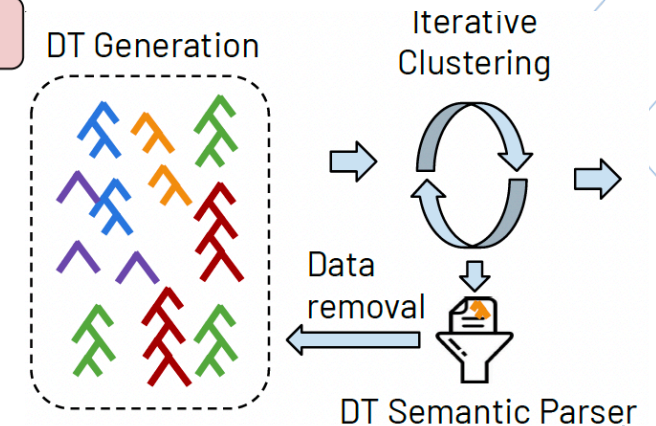
# CenDTect: an unsupervised learning censorship event detection system

Clustering toy example:

Decision trees for 3 domains in a region that only has 3 ASes: AS1, AS2, AS3.



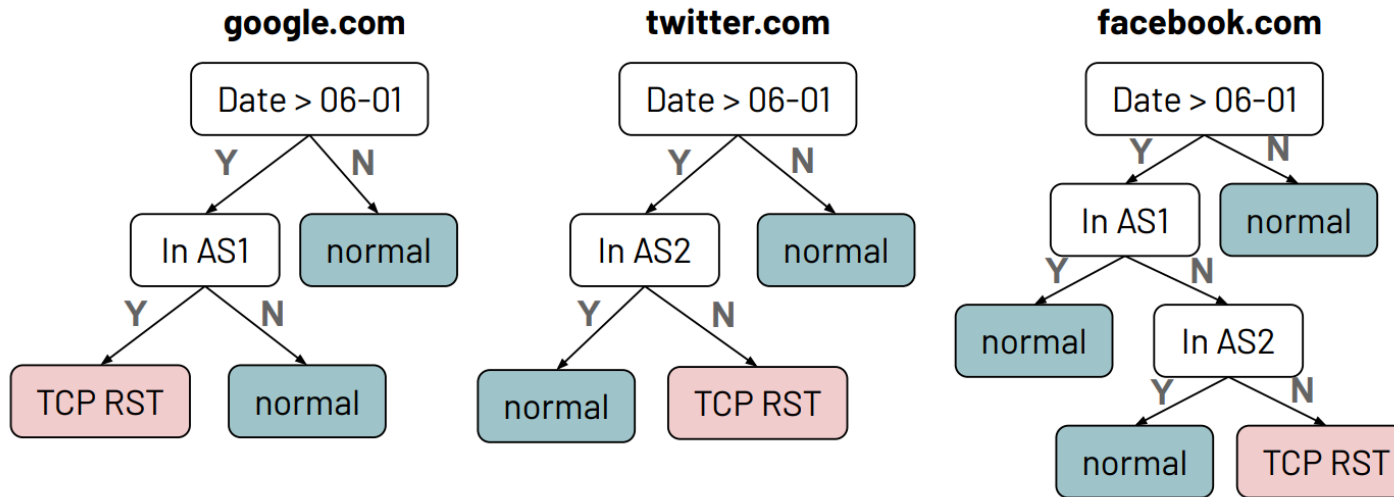
Round 1: Output cluster  $C_1 = \{\text{rule}=\text{DT}_{\text{google}}, \text{domains}=\{\text{google}, \text{twitter}\}\}$



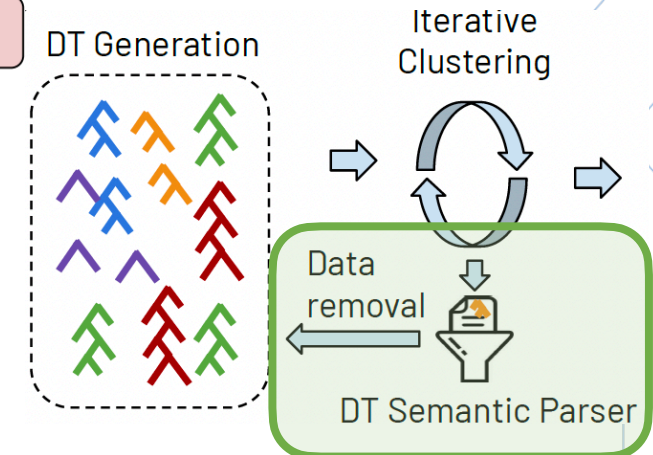
# CenDTect: an unsupervised learning censorship event detection system

Clustering toy example:

Decision trees for 3 domains in a region that only has 3 ASes: AS1, AS2, AS3.



Round 1: Output cluster  $C_1 = \{\text{rule}=\text{DT}_{\text{google}}, \text{domains}=\{\text{google}, \text{twitter}\}\}$

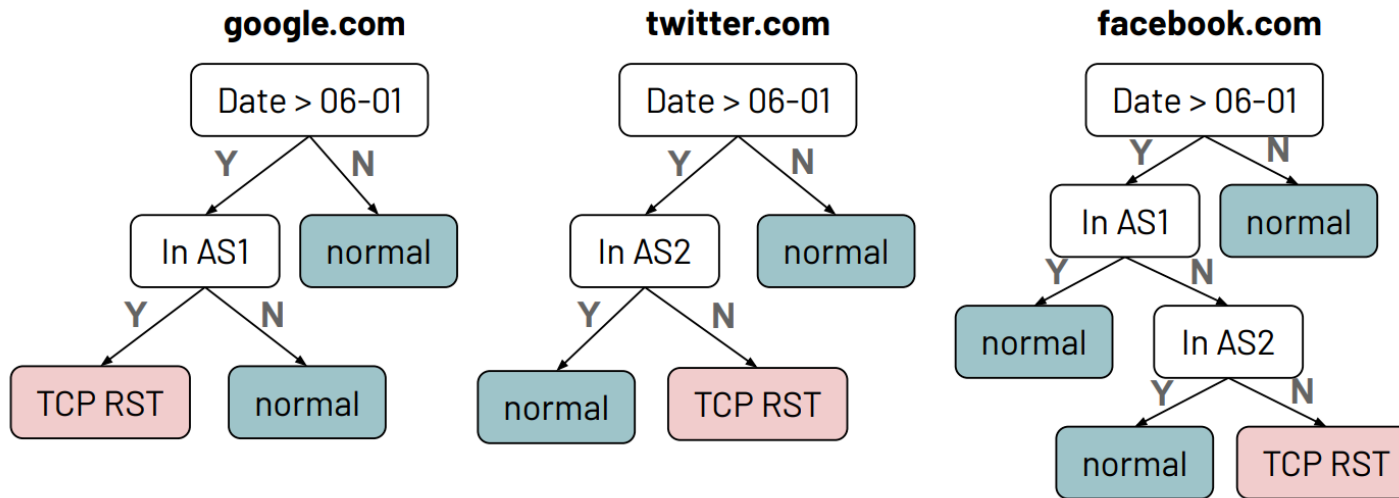




# CenDTect: an unsupervised learning censorship event detection system

Clustering toy example:

Decision trees for 3 domains in a region that only has 3 ASes: AS1, AS2, AS3.

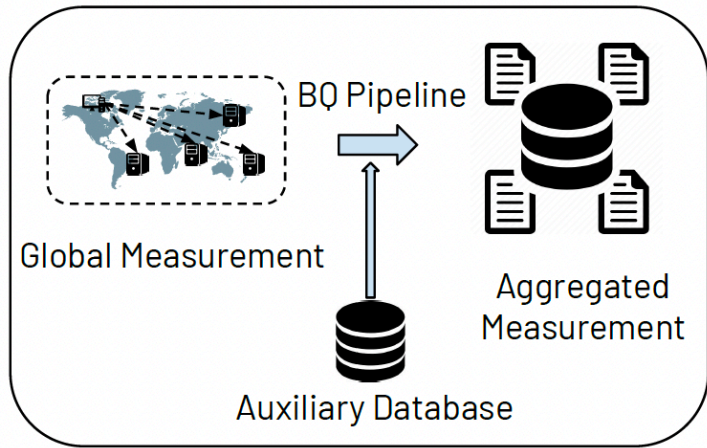


Round 1: Output cluster  $C_1 = \{\text{rule}=\text{DT}_{\text{google}}, \text{domains}=\{\text{google}, \text{twitter}\}\}$

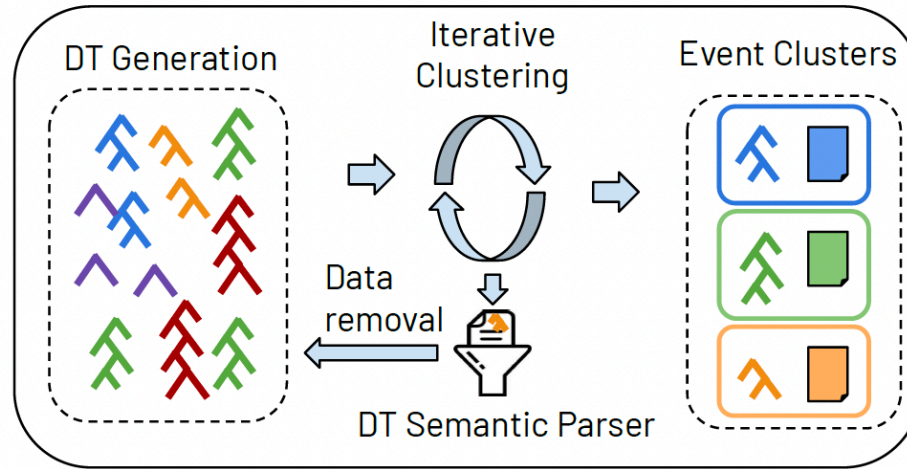
Round 2: Output cluster  $C_2 = \{\text{rule}=\text{DT}_{\text{facebook}}, \text{domains}=\{\text{twitter}, \text{facebook}\}\}$

# CenDTect: an unsupervised learning censorship event detection system

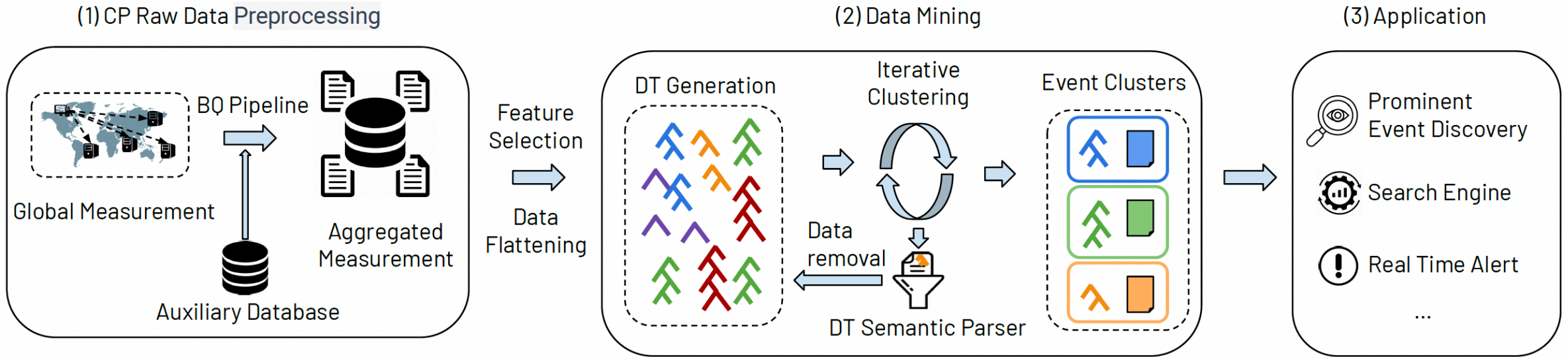
(1) CP Raw Data Preprocessing



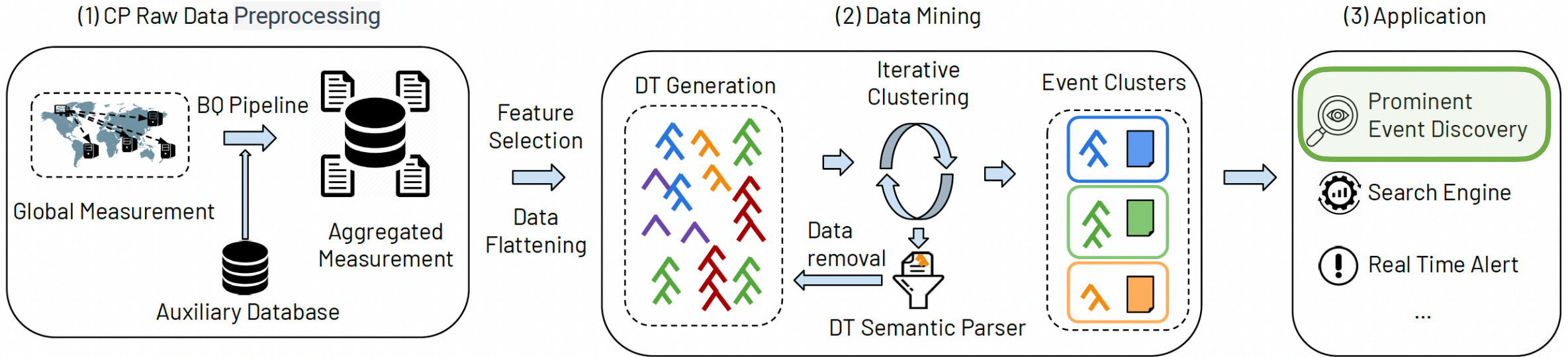
(2) Data Mining



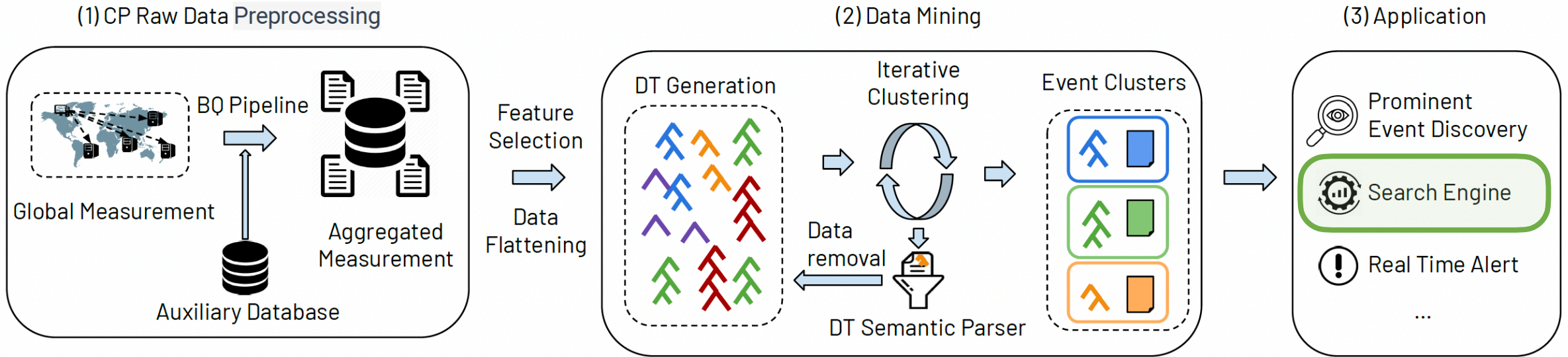
# CenDTect: an unsupervised learning censorship event detection system



# CenDTect: an unsupervised learning censorship event detection system

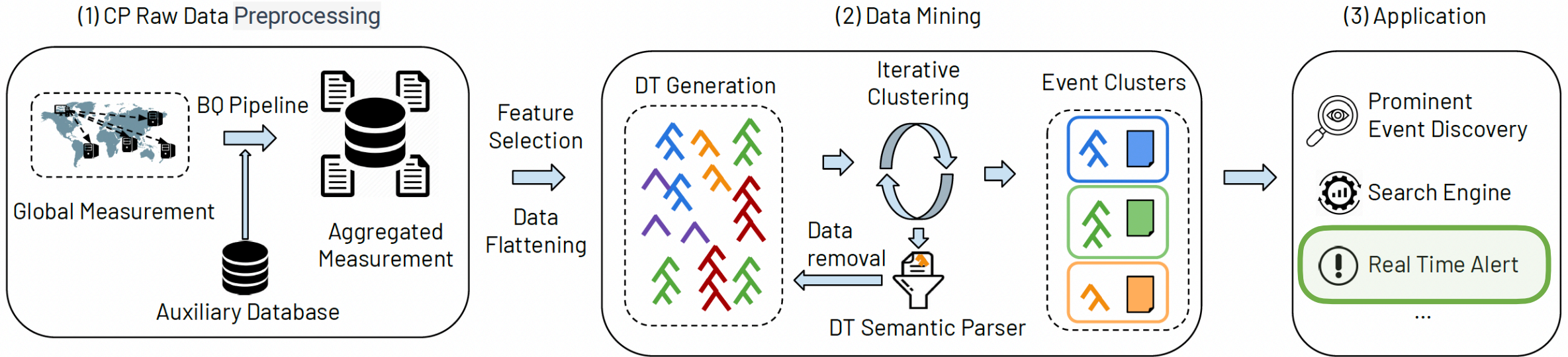


# CenDTect: an unsupervised learning censorship event detection system

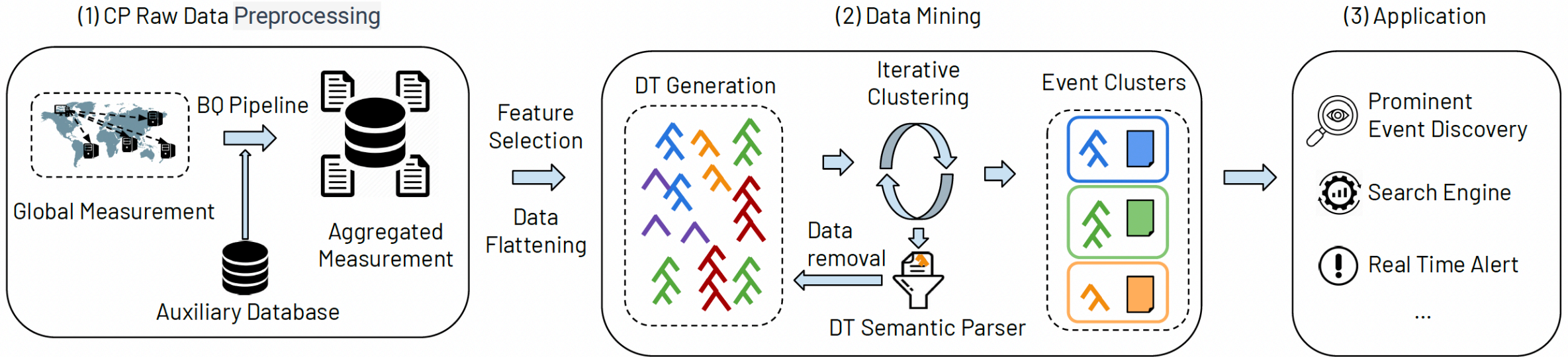




# CenDTect: an unsupervised learning censorship event detection system

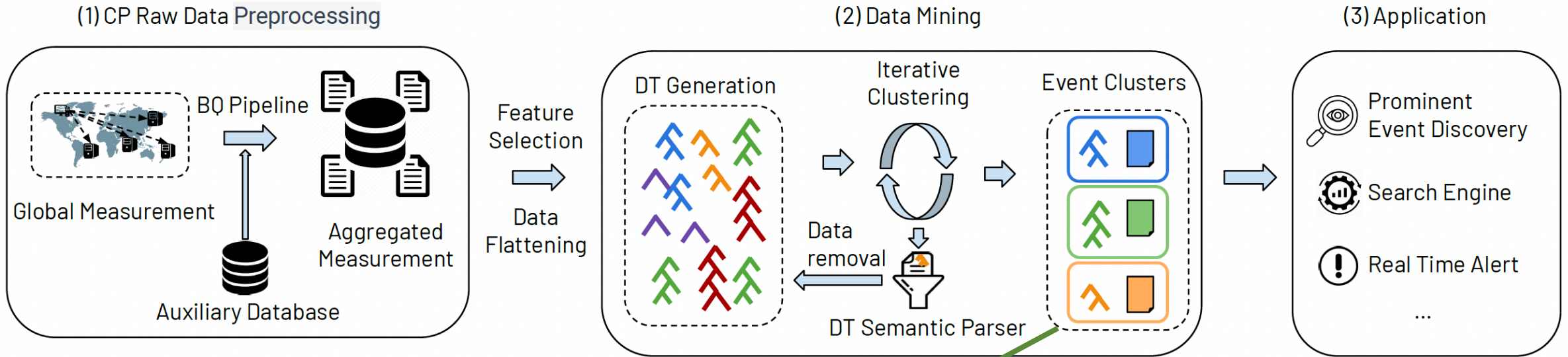


# CenDTect: an unsupervised learning censorship event detection system



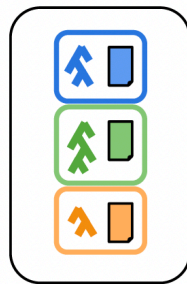
Prominent event discovery pipeline

# CenDTect: an unsupervised learning censorship event detection system



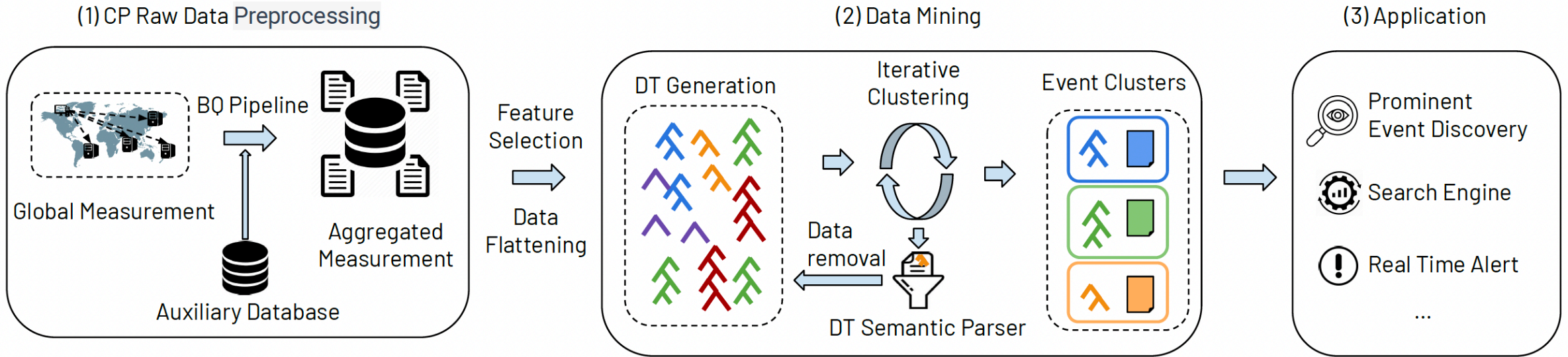
## Prominent event discovery pipeline

Event Clusters

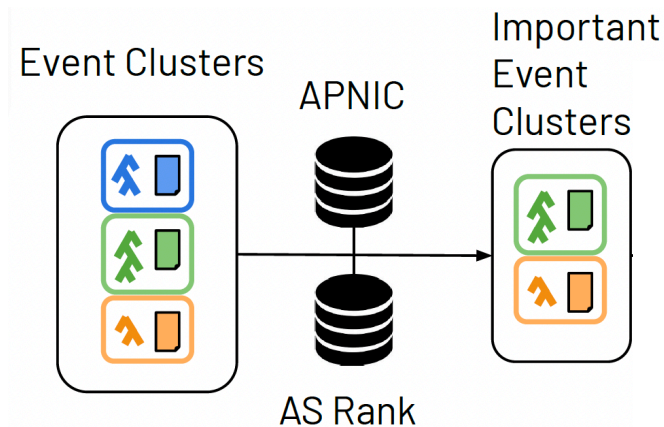




# CenDTect: an unsupervised learning censorship event detection system

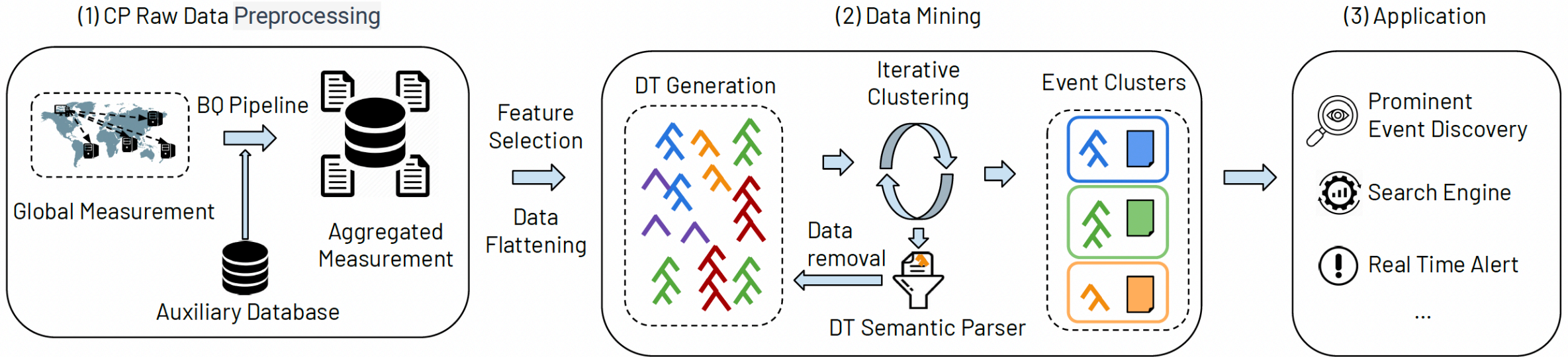


## Prominent event discovery pipeline

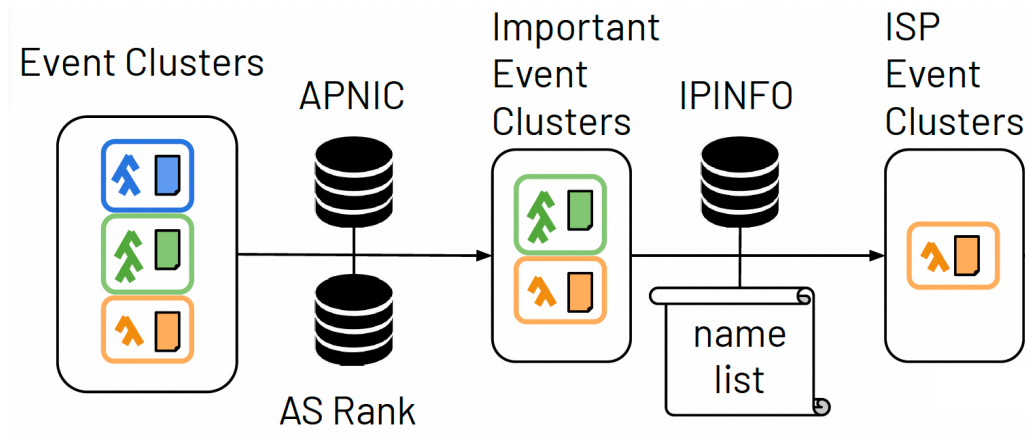


Top 20 ASes

# CenDTect: an unsupervised learning censorship event detection system

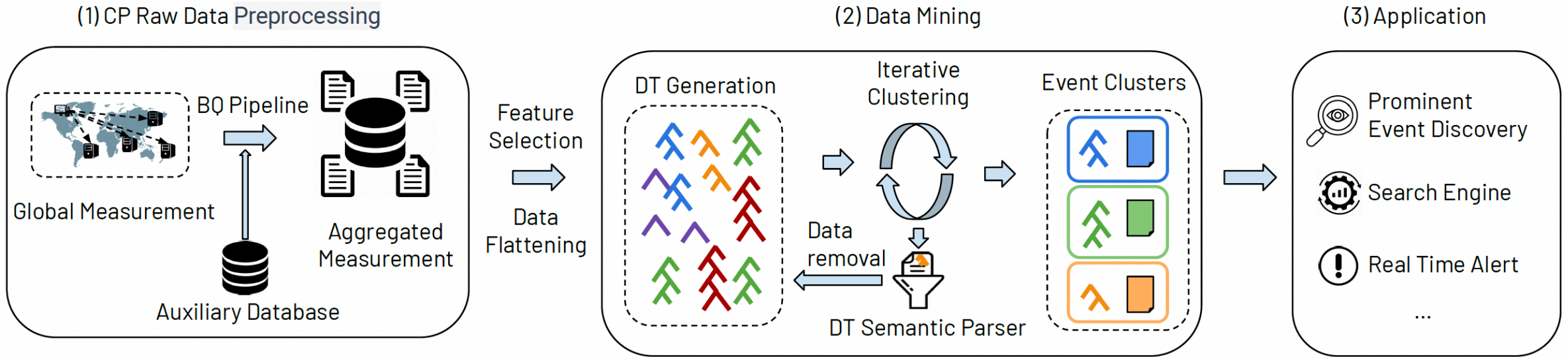


## Prominent event discovery pipeline

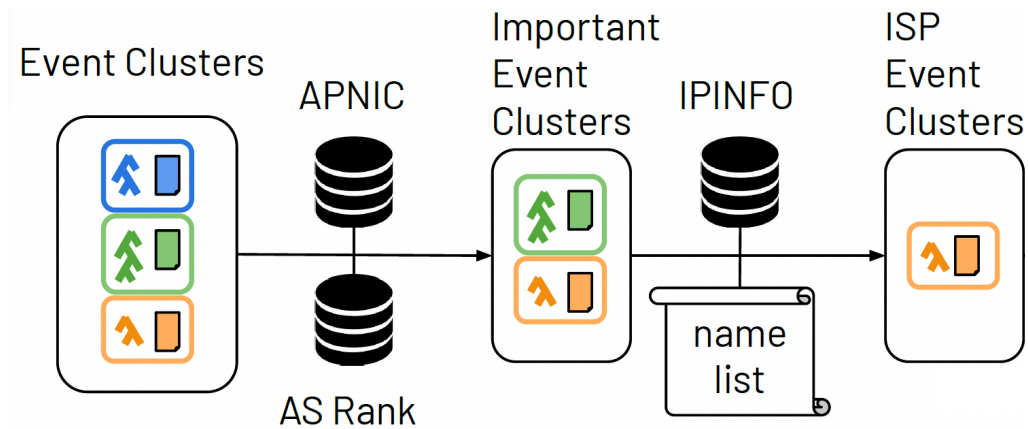


**ISP blocking:** detected within IPs owned by Internet Service Providers.

# CenDTect: an unsupervised learning censorship event detection system



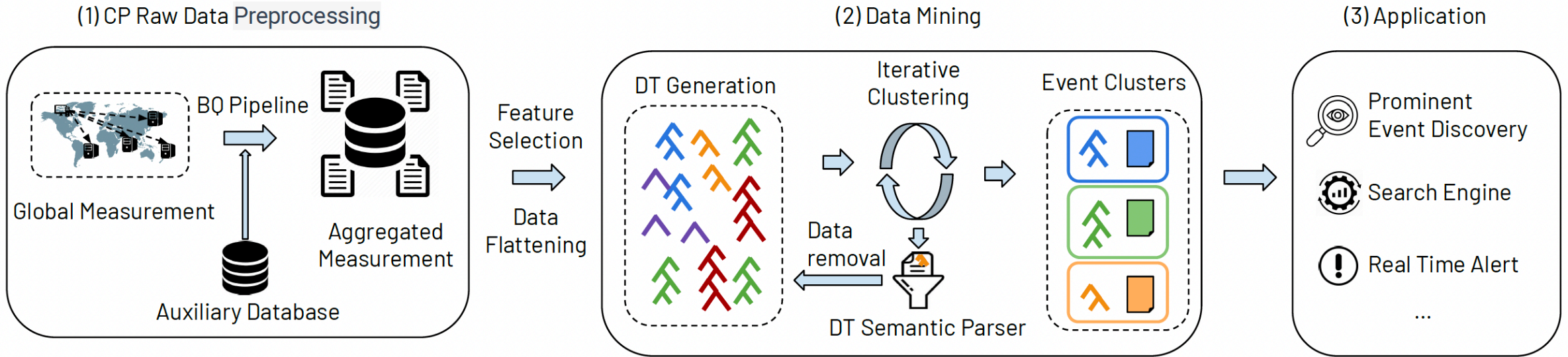
## Prominent event discovery pipeline



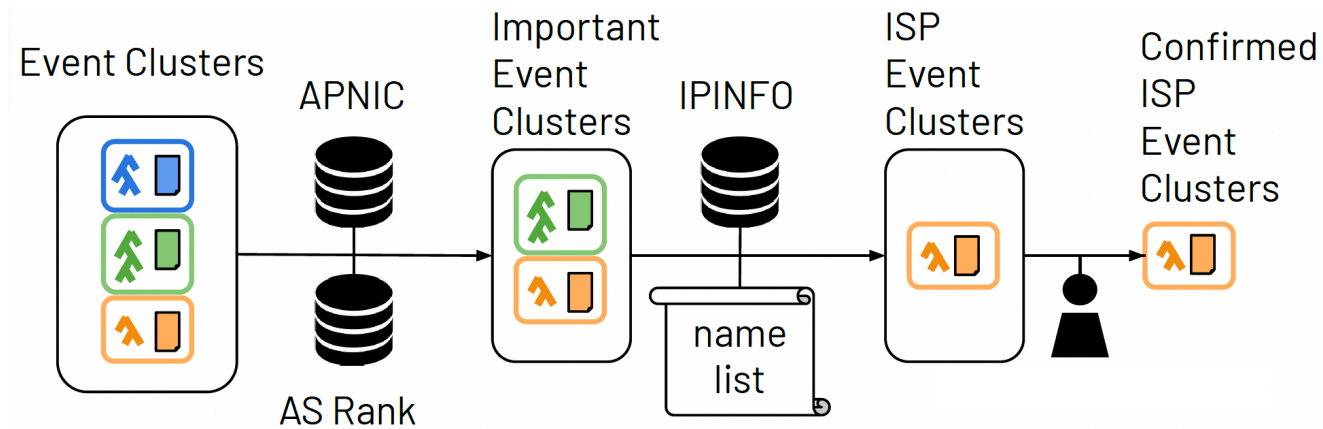
**Organizational blocking:** detected within corporations or educational institutions.



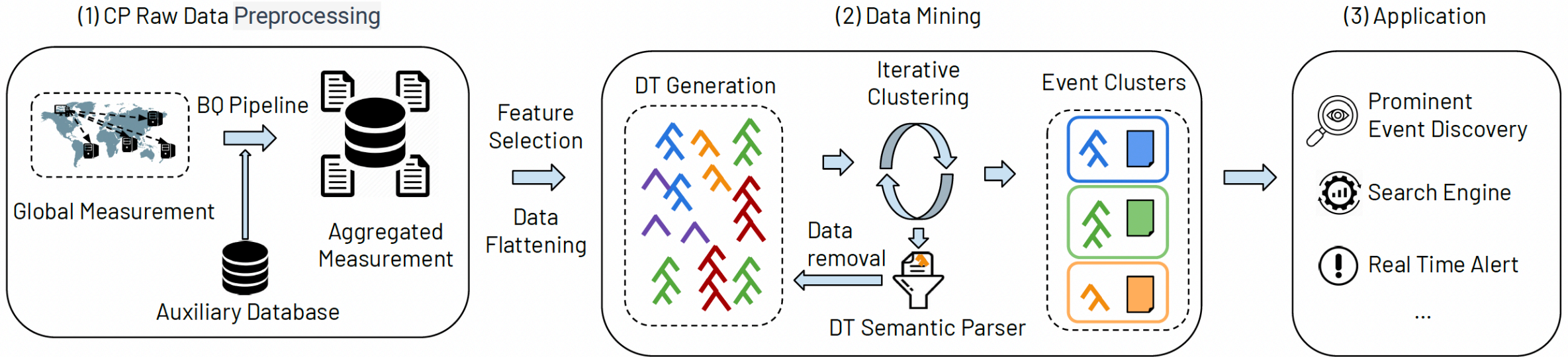
# CenDTect: an unsupervised learning censorship event detection system



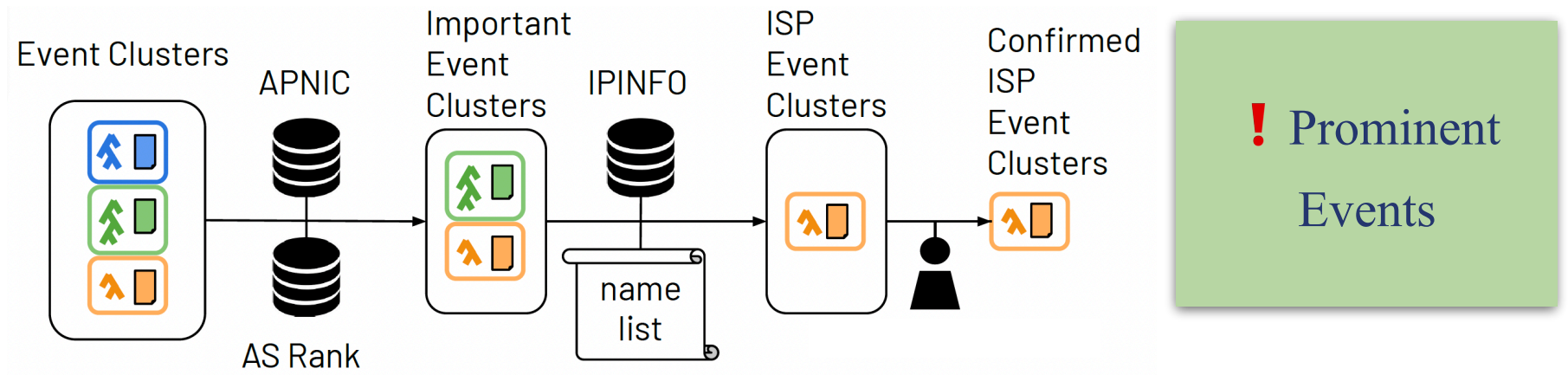
## Prominent event discovery pipeline



# CenDTect: an unsupervised learning censorship event detection system



## Prominent event discovery pipeline



# CenDTect findings

- 15,360 HTTP(S) clusters in 192 countries in 4 years (Jan 2019 to Dec 2022).
- 1,166 DNS clusters in 77 countries in 4-month (Aug 2022 to Dec 2022).

# CenDTect findings

- 15,360 HTTP(S) clusters in 192 countries in 4 years (Jan 2019 to Dec 2022).
- 1,166 DNS clusters in 77 countries in 4-month (Aug 2022 to Dec 2022).

## Prominent Event Discovery





# CenDTect findings



- 15,360 HTTP(S) clusters in 192 countries in 4 years (Jan 2019 to Dec 2022).
- 1,166 DNS clusters in 77 countries in 4-month (Aug 2022 to Dec 2022).

## Prominent Event Discovery

- **Persistent ISP-level blocking:** 32 countries and regions.
  - Verified with research papers, independent studies, OONI data and news reports.
  - Confirmed well-documented countries.
  - Discovered 16 unreported countries.
- 📖 Please refer to our paper for more details

# CenDTect findings

- 15,360 HTTP(S) clusters in 192 countries in 4 years (Jan 2019 to Dec 2022).
- 1,166 DNS clusters in 77 countries in 4-month (Aug 2022 to Dec 2022).

## Prominent Event Discovery

- **Persistent ISP-level blocking:** 32 countries and regions.
- **Temporal blocking:** 11 events from measurements in 2022.

# Temporal blocking events

Country	ASes	Count	Main Categories	Date	Evidence
Nepal	AS4007	15	SOCIAL, REL, GOV	Jan	
Venezuela	AS8048	8	ANON, NEWS, MEDIA	Feb	
Russia	AS28891 AS3216★ AS60459*	77	NEWS, SOCIAL	Mar	War [6]
Sri Lanka	AS18001	7	SOCIAL, NEWS	Apr-May	Protect [30]
Estonia	AS3327 AS3249	5	NEWS, HUMANR, FILE	May	Russia War [29]
Zimbabwe	AS37204★	17	NEWS, GOV, HUMANR	May-June	
Uganda	AS20294	5	NEWS, HUMANR, FILE	July	
Burkina Faso	AS37721	7	NEWS, SOCIAL, ANON	Aug	
Zambia	AS36959★ AS37146	20	COMM,SOCIAL,ANON	Aug	Election [87]
Armenia	AS43733★ AS12297★	38	NEWS, COMM, ANON	Sep	Conflict [6]
Iran	AS42337★ AS25184*	22	SOCIAL, ANON, BLOG	Sep-Oct	Protest [7]

Temporal blocking detected in HTTP(S) 2022 data.

# Temporal blocking events

- Happened during periods of **election, political unrest, protest, and war.**

Country	ASes	Count	Main Categories	Date	Evidence
Nepal	AS4007	15	SOCIAL, REL, GOV	Jan	
Venezuela	AS8048	8	ANON, NEWS, MEDIA	Feb	
Russia	AS28891 AS3216★ AS60459*	77	NEWS, SOCIAL	Mar	War [6]
Sri Lanka	AS18001	7	SOCIAL, NEWS	Apr-May	Protect [30]
Estonia	AS3327 AS3249	5	NEWS, HUMANR, FILE	May	Russia War [29]
Zimbabwe	AS37204★	17	NEWS, GOV, HUMANR	May-June	
Uganda	AS20294	5	NEWS, HUMANR, FILE	July	
Burkina Faso	AS37721	7	NEWS, SOCIAL, ANON	Aug	
Zambia	AS36959★ AS37146	20	COMM,SOCIAL,ANON	Aug	Election [87]
Armenia	AS43733★ AS12297★	38	NEWS, COMM, ANON	Sep	Conflict [6]
Iran	AS42337★ AS25184*	22	SOCIAL, ANON, BLOG	Sep-Oct	Protest [7]

Temporal blocking detected in HTTP(S) 2022 data.

# Temporal blocking events

- Happened during periods of election, political unrest, protest, and war.
- Often targets **news media, social networks, and anonymization tools.**

Country	ASes	Count	Main Categories	Date	Evidence
Nepal	AS4007	15	SOCIAL, REL, GOV	Jan	
Venezuela	AS8048	8	ANON, NEWS, MEDIA	Feb	
Russia	AS28891 AS3216★ AS60459*	77	NEWS, SOCIAL	Mar	War [6]
Sri Lanka	AS18001	7	SOCIAL, NEWS	Apr-May	Protect [30]
Estonia	AS3327 AS3249	5	NEWS, HUMANR, FILE	May	Russia War [29]
Zimbabwe	AS37204★	17	NEWS, GOV, HUMANR	May-June	
Uganda	AS20294	5	NEWS, HUMANR, FILE	July	
Burkina Faso	AS37721	7	NEWS, SOCIAL, ANON	Aug	
Zambia	AS36959★ AS37146	20	COMM, SOCIAL, ANON	Aug	Election [87]
Armenia	AS43733★ AS12297★	38	NEWS, COMM, ANON	Sep	Conflict [6]
Iran	AS42337★ AS25184*	22	SOCIAL, ANON, BLOG	Sep-Oct	Protest [7]

Temporal blocking detected in HTTP(S) 2022 data.

# Temporal blocking events

- Happened during periods of election, political unrest, protest, and war.
- Often targets news media, social networks, and anonymization tools.

Iran, September 2022:

- Multiple **anonymization tools** were blocked.

Country	ASes	Count	Main Categories	Date	Evidence
Nepal	AS4007	15	SOCIAL, REL, GOV	Jan	
Venezuela	AS8048	8	ANON, NEWS, MEDIA	Feb	
Russia	AS28891 AS3216★ AS60459*	77	NEWS, SOCIAL	Mar	War [6]
Sri Lanka	AS18001	7	SOCIAL, NEWS	Apr-May	Protect [30]
Estonia	AS3327 AS3249	5	NEWS, HUMANR, FILE	May	Russia War [29]
Zimbabwe	AS37204★	17	NEWS, GOV, HUMANR	May-June	
Uganda	AS20294	5	NEWS, HUMANR, FILE	July	
Burkina Faso	AS37721	7	NEWS, SOCIAL, ANON	Aug	
Zambia	AS36959★ AS37146	20	COMM,SOCIAL,ANON	Aug	Election [87]
Armenia	AS43733★ AS12297★	38	NEWS, COMM, ANON	Sep	Conflict [6]
Iran	AS42337★ AS25184*	22	SOCIAL, ANON, BLOG	Sep-Oct	Protest [7]

Temporal blocking detected in HTTP(S) 2022 data.

# Temporal blocking events

- Happened during periods of election, political unrest, protest, and war.
- Often targets news media, social networks, and anonymization tools.

Iran, September 2022:


- Multiple **anonymization tools** were blocked.
- Several DNS over HTTPS canary domains were blocked
  - doh.opendns.com
  - doh-fi.blahdns.com
  - mozilla.cloudflare-dns.com

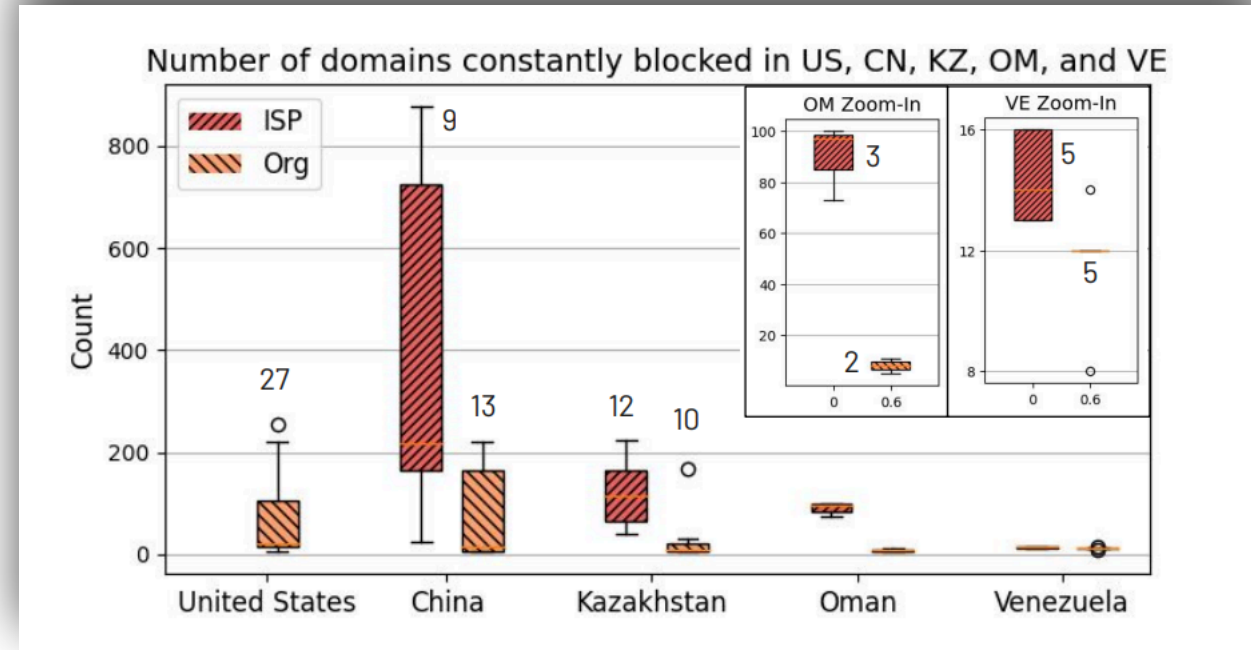
Country	ASes	Count	Main Categories	Date	Evidence
Nepal	AS4007	15	SOCIAL, REL, GOV	Jan	
Venezuela	AS8048	8	ANON, NEWS, MEDIA	Feb	
Russia	AS28891 AS3216★ AS60459*	77	NEWS, SOCIAL	Mar	War [6]
Sri Lanka	AS18001	7	SOCIAL, NEWS	Apr-May	Protect [30]
Estonia	AS3327 AS3249	5	NEWS, HUMANR, FILE	May	Russia War [29]
Zimbabwe	AS37204★	17	NEWS, GOV, HUMANR	May-June	
Uganda	AS20294	5	NEWS, HUMANR, FILE	July	
Burkina Faso	AS37721	7	NEWS, SOCIAL, ANON	Aug	
Zambia	AS36959★ AS37146	20	COMM,SOCIAL,ANON	Aug	Election [87]
Armenia	AS43733★ AS12297★	38	NEWS, COMM, ANON	Sep	Conflict [6]
Iran	AS42337★ AS25184*	22	SOCIAL, ANON, BLOG	Sep-Oct	Protest [7]

Temporal blocking detected in HTTP(S) 2022 data.



# Heterogeneous blocking within countries

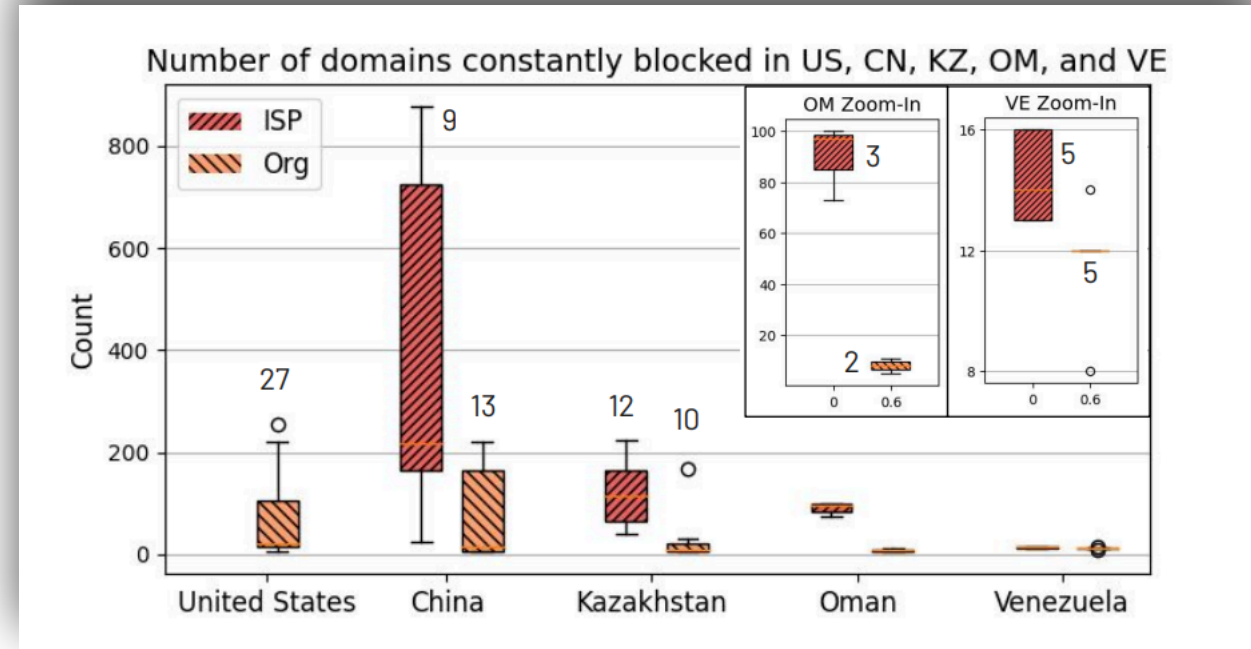
-  Many prior works report censorship based on the blocking ratio of the entire country/ASes.





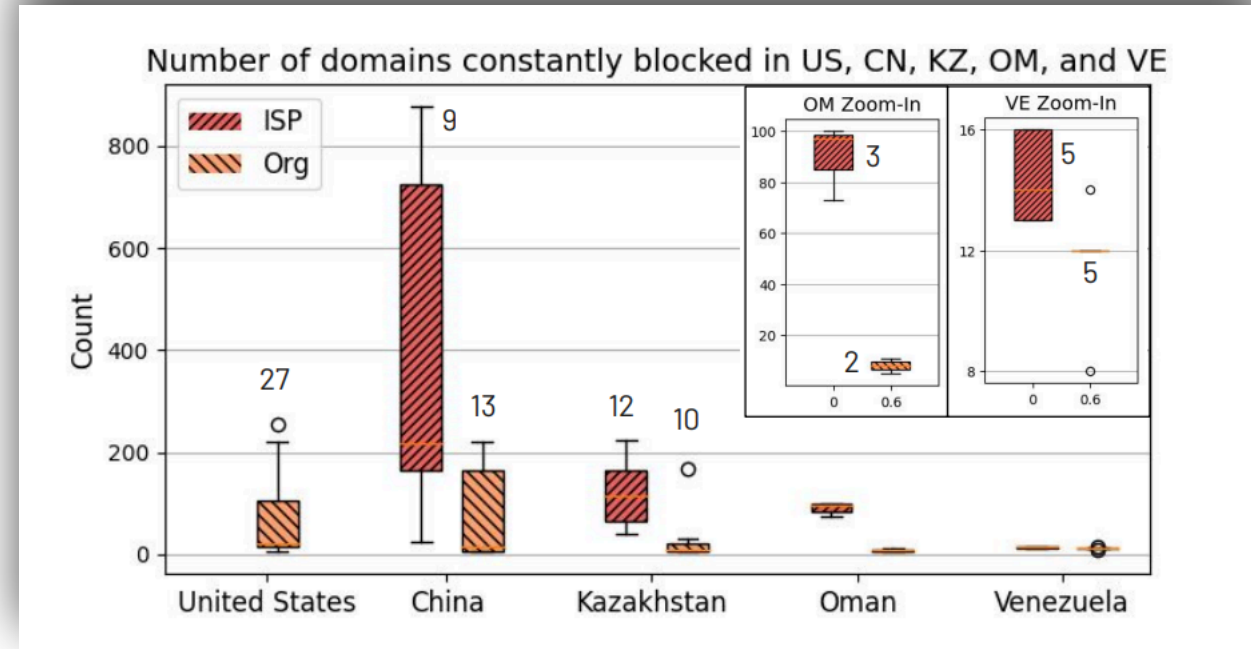
# Heterogeneous blocking within countries

- ⚠ Many prior works report censorship based on the blocking ratio of the entire country/ASes.
- Many countries have persistent, ISP-level blocking, significantly distinguishable from organizational blocking.



# Heterogeneous blocking within countries

- ⚠ Many prior works report censorship based on the blocking ratio of the entire country/ASes.
- Many countries have persistent, ISP-level blocking, significantly distinguishable from organizational blocking.



## 💡 Tips

- Recognize the diversity in blocking behavior within countries and Autonomous Systems (ASes) and avoid generalizations.

# Organizational blocking in free countries

- **ISP blocking and organizational blocking** affect different population segments and have different impacts on global Internet censorship.

Andrews  University

**143.207.2.29**

**is blocked.**

If you think this is in error please contact [filteradmin@andrews.edu](mailto:filteradmin@andrews.edu)

# Organizational blocking in free countries

- **ISP blocking and organizational blocking** affect different population segments and have different impacts on global Internet censorship.
- Discovered 14 free countries with organizational blocking, such as the United States, Australia, Canada, and Japan.
  - In the United States alone, we identify 59 clusters for organizational blocking

Andrews  University

**143.207.2.29**

**is blocked.**

If you think this is in error please contact [filteradmin@andrews.edu](mailto:filteradmin@andrews.edu)

# Organizational blocking in free countries

- **ISP blocking** and **organizational blocking** affect different population segments and have different impacts on global Internet censorship.
- Discovered 14 free countries with organizational blocking, such as the United States, Australia, Canada, and Japan.
  - In the United States alone, we identify 59 clusters for organizational blocking



## Tips

- Differentiate between ISP and organizational blocking practices.

# Contributions

**CenDTect:** an automated system for detecting Internet censorship events.

 Decision trees

 Iterative clustering

We report more than 100 ASes in 32 countries with persistent ISP blocking.

We identify 11 temporary blocking events in clusters discovered in 2022

Observed during periods of election, political unrest, protest, and war.

Our findings provide insights on

 The heterogeneity of blocking practices within countries.

 The importance of distinguish ISP blocking and organizational blocking



**Thank you!**

