Automatically Generating a Large, Culture-Specific Blocklist for China

Austin Hounsel, Prateek Mittal, Nick Feamster
Princeton University

Internet censorship is regularly measured





HERDICTWEB

NETBLOCKS
MAPPING NET FREEDOM

Test lists are used as input

- Manually curated
  - Citizen Lab Block List
- Automatically generated
  - Darer 2017, “FilteredWeb”
  - Sfakianakis 2011, “CensMon”
- Crowdsourced
  - Herdict
  - GreatFire.org

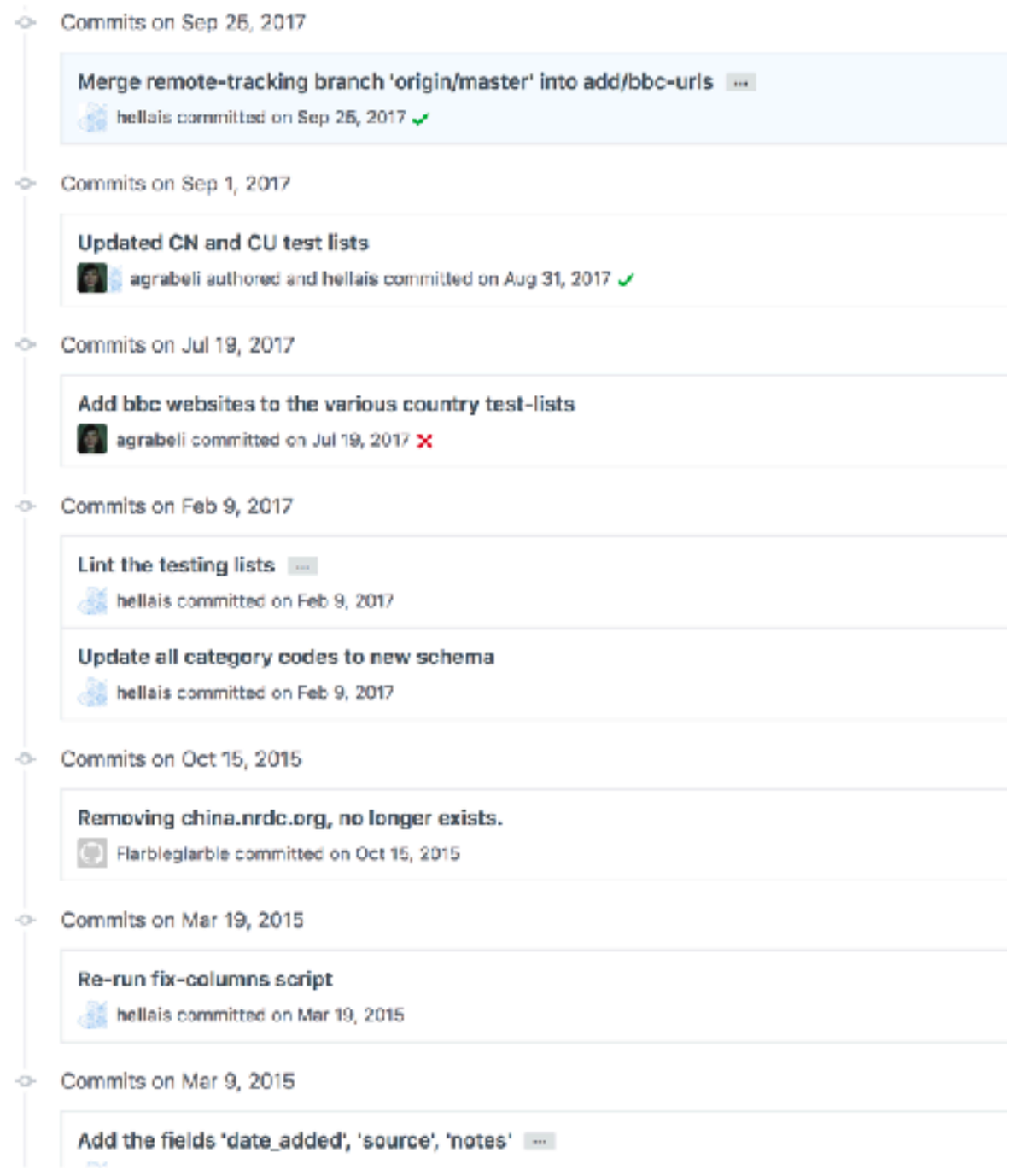




However, these lists are…

Not frequently updated

- Manually curated:
  - CLBL has 8 commits in past 3 years
- Automatically generated:
  - CensMon published in 2011
- Crowdsourced lists limited to what users submit



Not all publicly available

- FilteredWeb
  - We were shown the results for validation
- CensMon
- Complete block list from Chinese government



Mainly focused on English websites

- CLBL has limited # of China-specific websites
  - Mostly human rights orgs
- FilteredWeb doesn't parse Chinese websites







Our contributions

- Large, up-to-date, publicly available list of censored domains in China
  - None of these websites are in the largest blocklist
- Qualitative analysis of China-specific websites
- Pull-request into CLBL (in progress)

“Censorship findings are only as interesting as the sites and services that you test.”

– *Open Observatory of Network Interference*

So.. what are we missing?

THE CHURCH OF ALMIGHTY GOD

We welcome all truth seekers to contact us.

Search

Language Switch

Home | About Us | News | Books | Hymns | Videos | Testimonies | Readings | Sermons | Pictures

God's Management Plan for Saving Mankind—The Three Stages of Work



Jehovah Decrees Laws to Lead Man's Life



Jesus Redeems Man Through Crucifixion



Almighty God Expresses the Truth to Judge and Purify Man

Christ of the Last Days Has Appeared in China

"For as the lightning comes out of the east, and shines even to the west; so shall also the coming of the Son of man be" (Matthew 24:27). **"And at midnight there was a cry made, Behold, the bridegroom comes; go you out to meet him"** (Matthew 25:6). **"Behold, I stand at the door, and knock: if any man hear my voice, and open the door, I will come in to him, and will sup with him, and he with me"** (Revelation 3:20).

Utterances of Christ of the Last Days (The Words of the Spirit to the Churches)

More ›

Utterances of Christ | Books on Life Entry | Gospel | Testimonies



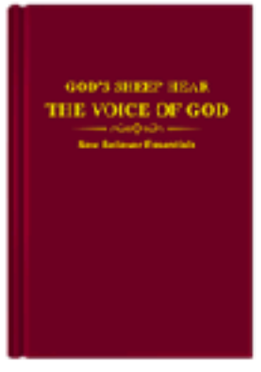
The Word Appears in the Flesh



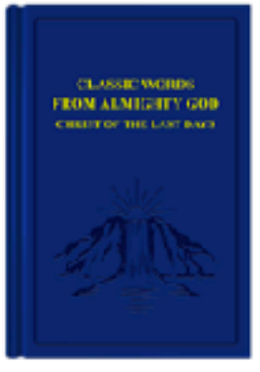
Continuation of The Word Appears



Utterances of Christ of the Last



God's Sheep Hear the Voice of God



Classic Words From Almighty

Classic Words From Almighty God



How to Know Reality



When It Comes to God, What Is Your Understanding



How Peter Came to Know Jesus



In Your Faith in God Y

Online

ENGLISH 简体 繁體

中国劳工论坛

CHINAWORKER.INFO · COMMITTEE FOR A WORKERS' INTERNATIONAL · CHINESE SECTION

search...

HOME | ABOUT US | NEWS | THEORY | DONATE! | STORE | LINKS | JOIN US | 中文



Hong Kong: Banning of political party marks serious escalation of political repression

3 weeks ago

Tiny Hong Kong National Party is being used as a test case for a wider political crackdown

China: Problems are piling up for Xi Jinping



The system of world capitalism faces an unprecedented crisis and turmoil in international relations

July 13, 2018, 12:20 pm

Hong Kong: 50,000 march to "end one-party dictatorship"



Lower numbers reflect need for clear strategy of mass action against authoritarian rule

July 2, 2018, 8:52 am

Hong Kong's June 4 vigil: What did the 'student boycott' achieve?



Yet again Hong Kong organised a powerful act of defiance on 1989 massacre anniversary

July 1, 2018, 2:12 am

SUBSCRIBE SOCIALIST MAGAZINE ENGLISH.



Socialist magazine can be ordered from socialist.hk@gmail.com

CHINA

te From: Detect language To: English View: Translation Original

Set as homepage to join favorite

Member Centre | Sponsorship | Soliciting volunteers | Online radio | Offer | Partner | about us | Contact us

Taiwanus.net 台灣海外網

首頁 | 新聞網 | 影音網 | 活動看板 | 線上論壇 | 自由圖片 | 台灣歌樂

Taiwan news | Comment column | US News | Community news | international News | China News | economic news | sports news | Technology news | Health news

News | Search for | Political society | Cultural life | Overseas Network | Special | Art movement | Technology Finance

萬家香 WAN JA SHAN
Handcrafted and Naturally Brewed
健康・天然・美味



www.wanjashan.com
inquiry@wanjashan.com
Visit Us
Email us

Recommended audio and video

- 7/29 AV-IRVINE Taiwan Overseas Network and Wang Mingzhe
- 7/29 AV-IRVINE Taiwan Overseas Network and Wang Mingzhe
- 8/4 Saturday Salt Lake City - Wang Mingzhe music master whirlwind; Taiwan overseas network
- 8/8 Portland--Wang Mingzhe's music performance banquet; Taiwan's overseas network entrepreneurship
- 8/12 Vancouver--Wang Mingzhe music singing feast
- Cao Changqing: The three strategies of President Tsai Ing-wen are great! Ye Junrong is incompetent
- 8/11 Morning-Seattle-Wang Mingzhe Concert & Taiwan Overseas Network
- Cao Changqing: I have ten reasons for Guo Wengui! (video + text draft)
- [Political and economical to see people] Trump told Tsai Ing-wen how to be president; Xi Jinping is a pseudo-person
- Big Age 15-24 episode
- 2018 8.4 [Political and Economical Views]

Taiwan news

- Taiwan United Nations Association for Advancement
- How to end the Chinese defeated exile group (Republic of China)
- Ye Wanghui: If Chiang Kai-shek voluntarily withdraws from the United Nations, the United States will not
- 30 years ago, the Taiwanese took care of / let this Ado Tsai "American A Hui
- In the management of the government, greedy Chen Shimeng's tip-end Zhou Ji: reversing the right and wrong
- China's bully Yadong Qingyun, why can't the Kuomintang candidates scream?
- Don't like "Chinese Taipei"? In fact, the Kuomintang rejected the international nuclear
- Wang Mingzhe, July 29, Irvine Taiwan Christ Elder
- Chuan Ma Ying-jeou intends to fight for another bid



李銘正 Lees Accounting, CPA, P.C.
註冊會計師事務所 Ming J. Lee, CPA



李致一醫學博士
美容整形 重整修復手術

Advertisements
O'r

桃園市政府 Taoyuan City Government

臺中市政府

臺南市政府 Tainan City Government

TAITRA 外貿發展協會

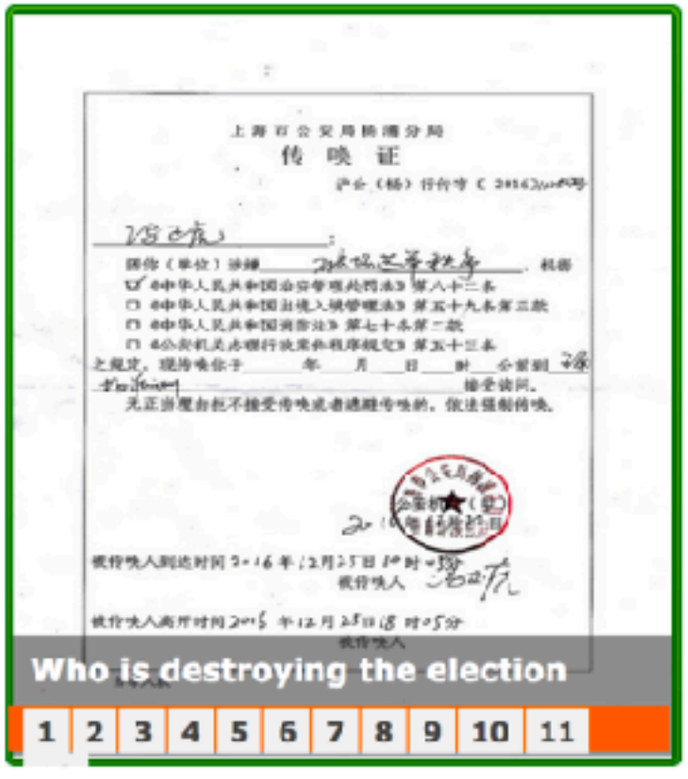
n: Detect language To: English View: Translation

Constitutional rights, happy life

冯正虎 de 博客



Setup as front page
Add to

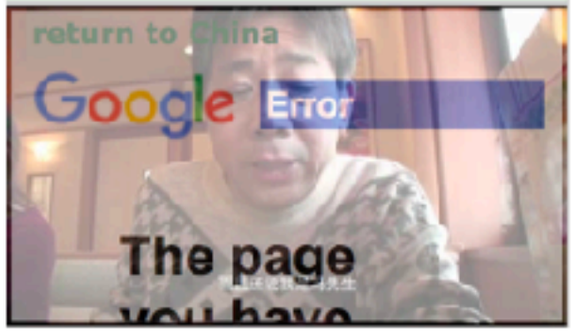


Feng Zhenghu went to

最近更新

- The 10th blocked notice of the essay | citizenshi...
- Litigation case | guard Feng Zhenghu's financial ...
- Announcement of the 8th blocked notice of the es...
- Miscellaneous essays to | protect the civil rights l...
- Miscellaneous essays | Feng Zhenghu went to Su...
- The painstaking work of | Feng Zhenghu's haircu...
- Li Zhengzheng and others sued Cui Yadong, Presi...
- The painstaking work of the | illegal police escort...
- The painstaking work of | Feng Zhenghu was ille...
- The painstaking work of | Feng Zhenghu was det...

Good life - Feng Zhenghu's



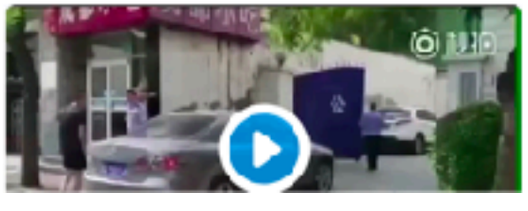
Feng Zhenghu de Twitter

Tweets by @fzhenghu

冯正虎 Retweeted

46664
@xingfuhul

终于北京人上街了，不是为了要命的雾霾，也不是为了致死的假疫苗，而是因为被骂"穷逼"。士可杀，不可辱，真是应景啊。2018年8月3日，曹远航被迫到丰台公安分局右安门派出所自首。但北京人仍然不依不饶。那会北京李棒棒骂外地人，可比这狠啊，可好像也没被刑事拘留也没见怎么着。



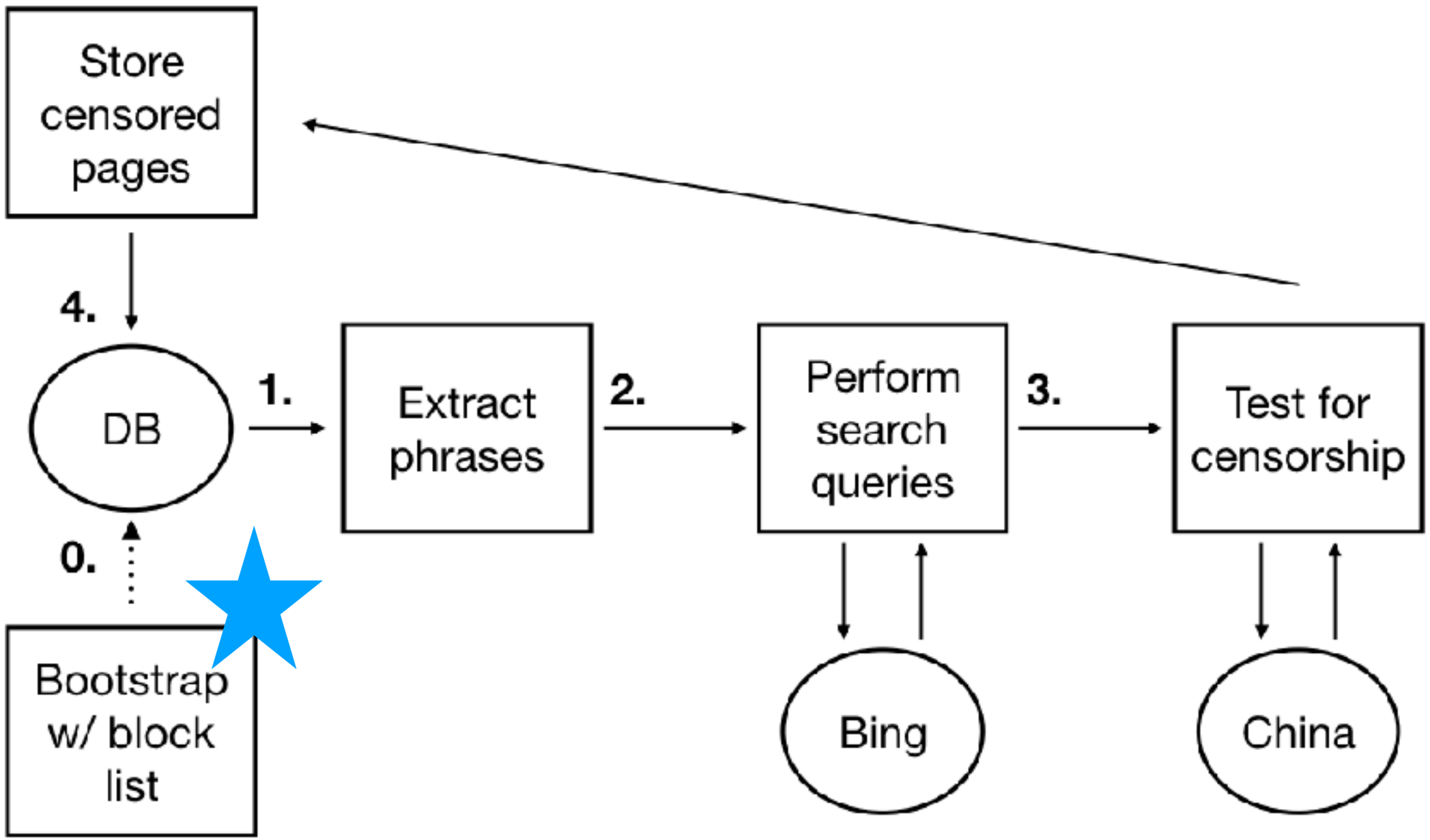
>> More Talk essays

- Feng Zhenghu went to Shangjiyuan Zhouz...
- Citizen Supervisory Group's 10th blocked ...
- The 8th notice of the citizen supervision gr...
- Notice of the 5th time that the civil rights l...
- Feng Zhenghu went to Suzhou to sweep th...
- Feng Zhenghu sent a letter to Li Qiang for ...
- Feng Zhenghu sent a letter to Li Qiang for ...

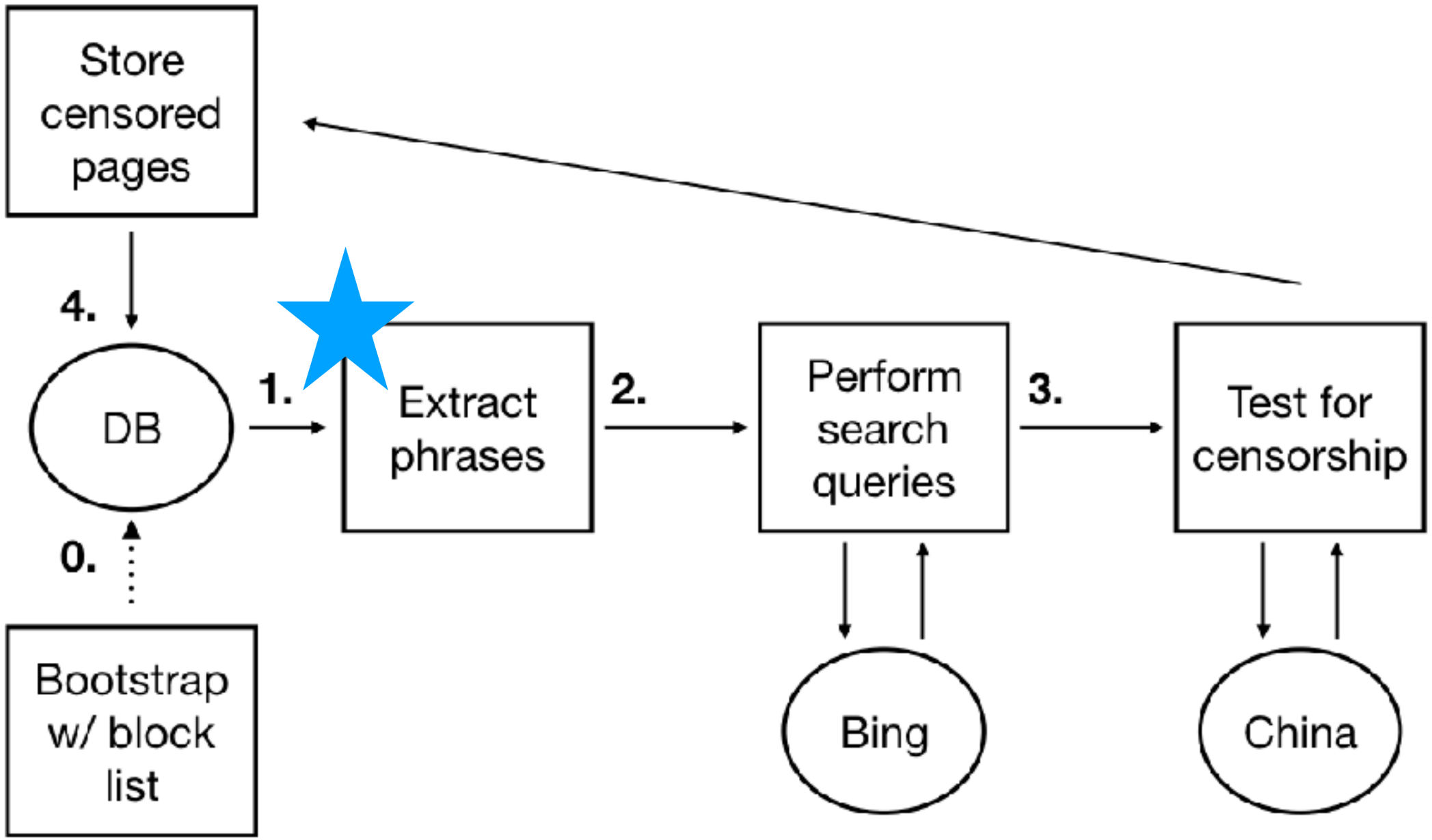
>> More Constitutional Rights

- Feng Zhenghu's election for the Shanghai People's ...
- The opening of personal information of the National ...
- Who can find the NPC deputies in China? (March 4, ...
- Establish legal authority and abolish "black jails"
- What should China's NPC deputies do?
- Who can find the NPC deputies in China? (March 6, ...
- Constitutional rights

Our approach



Our approach



What do we mean by “phrases”?

- We use “n-grams” of Chinese words/phrases as search terms
- Example: “Chinese human rights violation”
  - Bigrams: Chinese human, human rights, rights violation
  - Trigrams: Chinese human rights, human rights violation

What do we mean by “phrases”?

- However, because Chinese is written without spaces, word boundaries are probabilistic

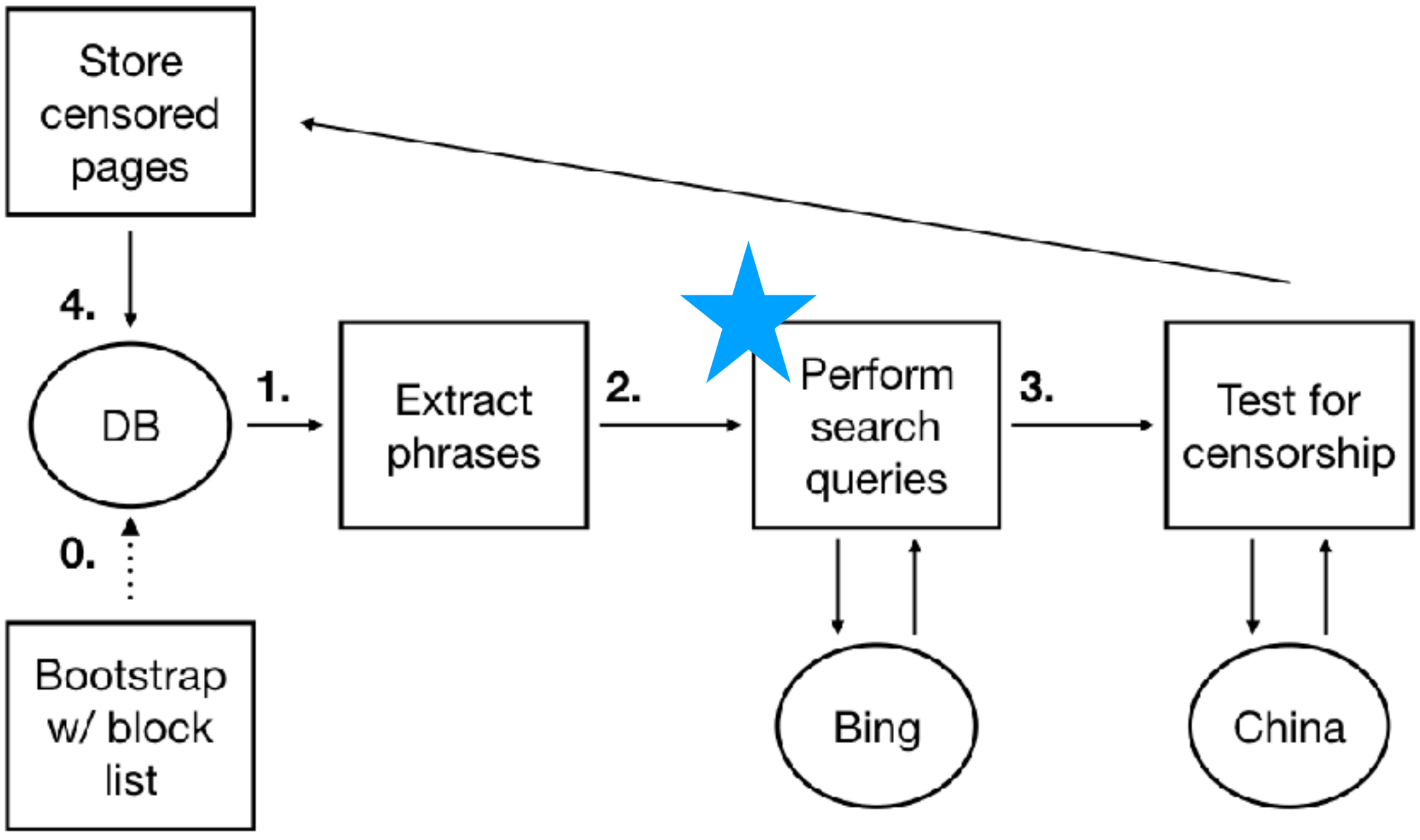
- Example: “特 首”

  - Bigram translation: Chief executive

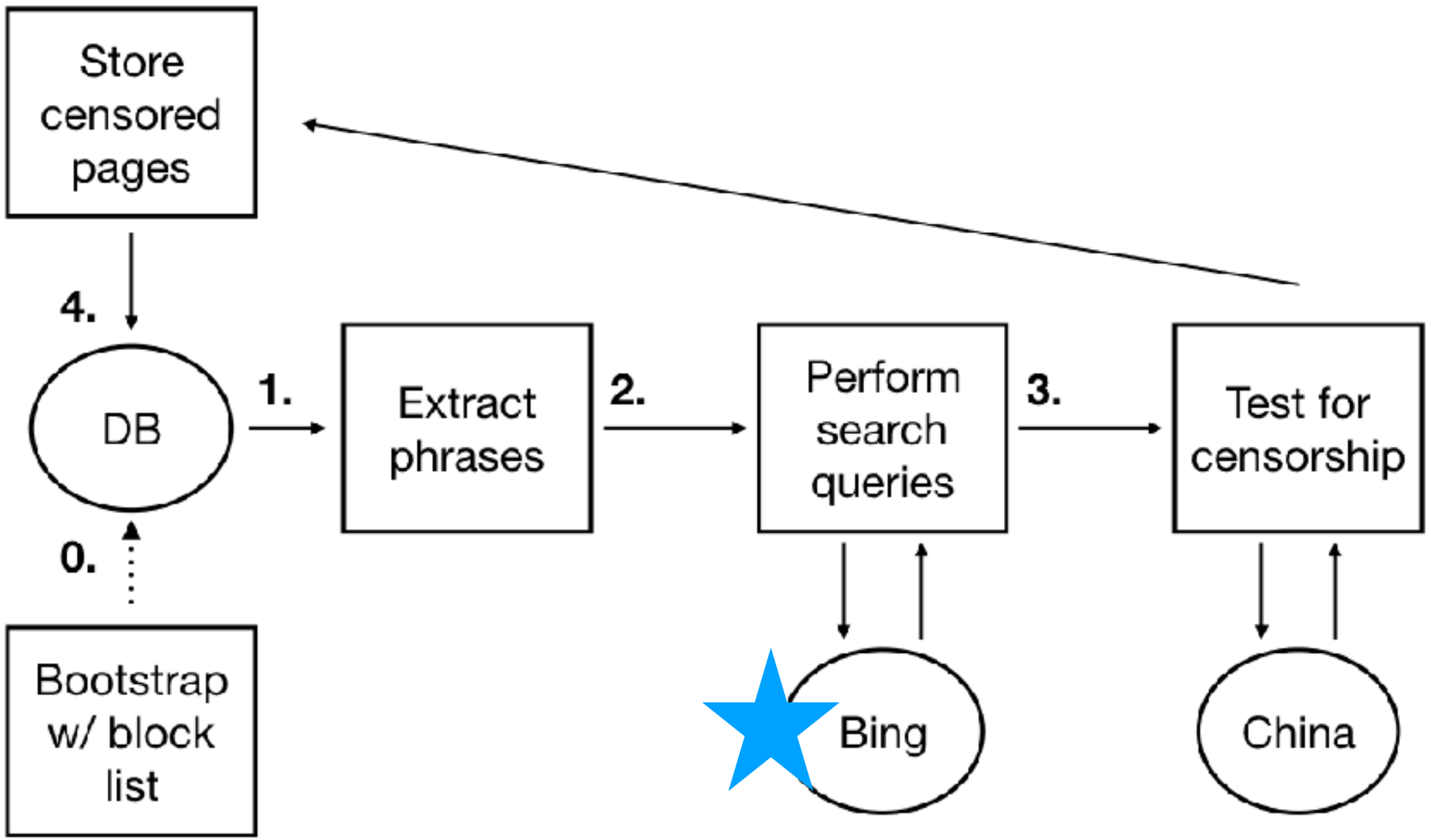
  - Unigram translations: Special, first

- We use whatever phrases Stanford CoreNLP outputs

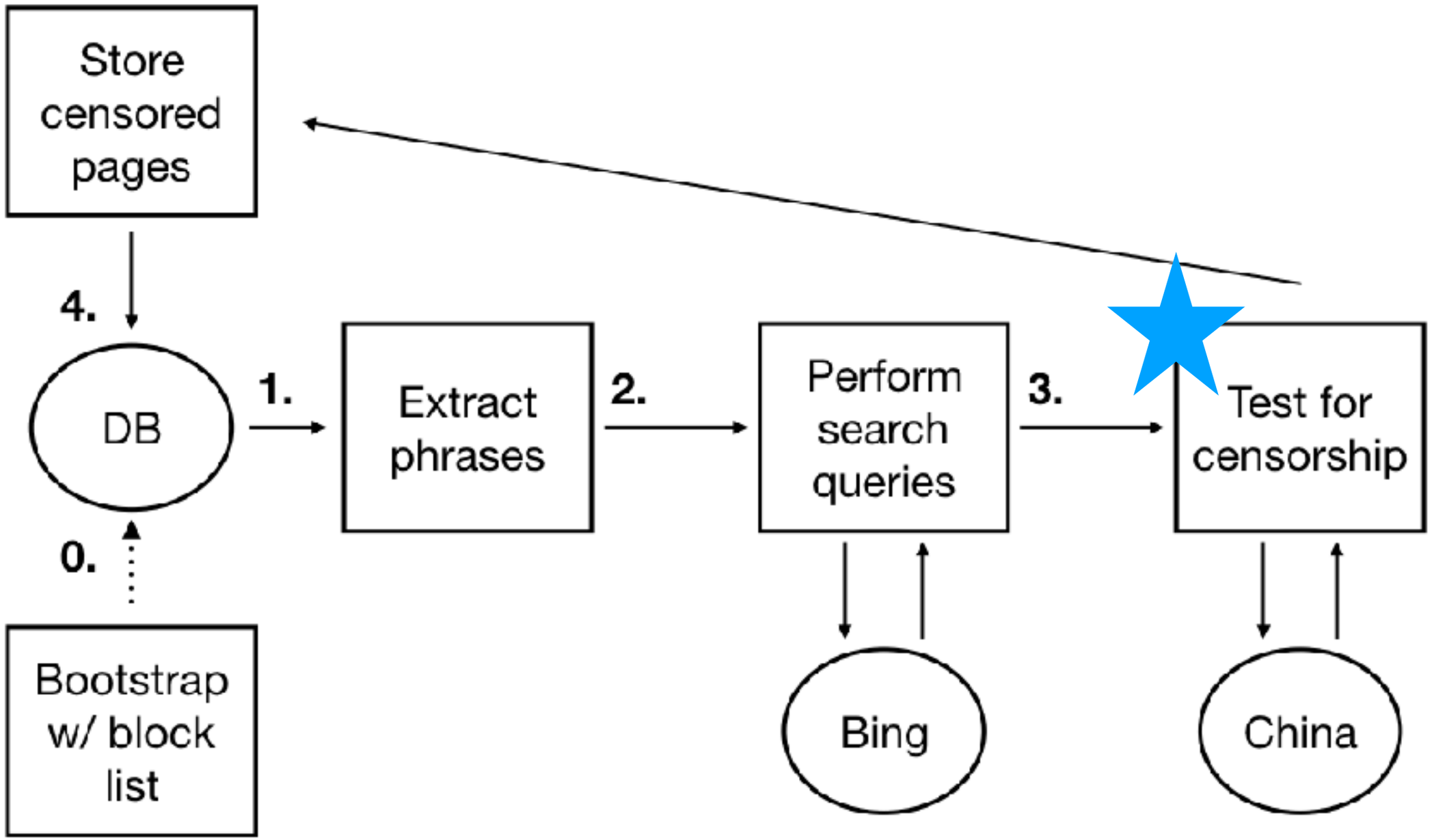
Our approach



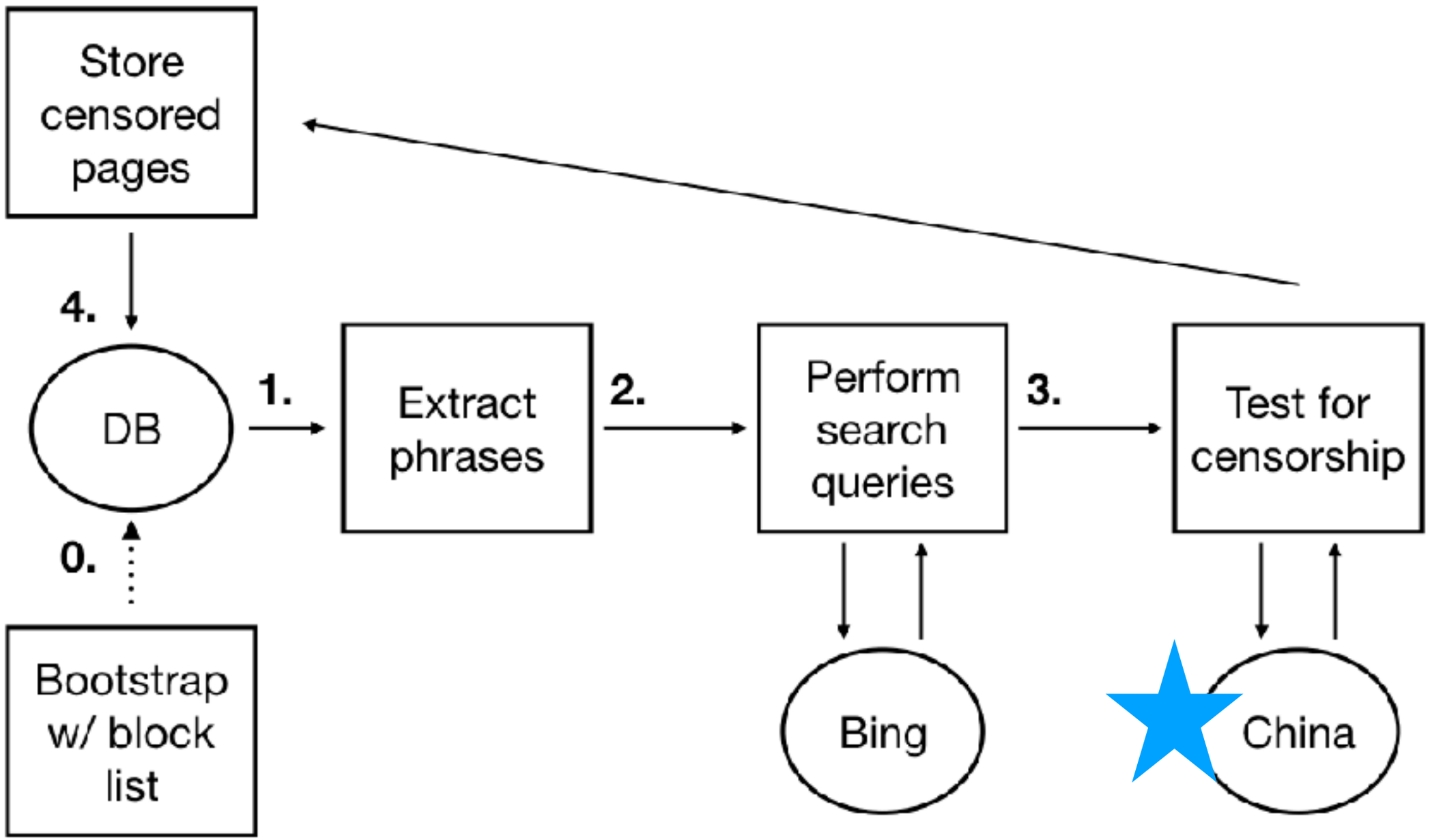
Our approach



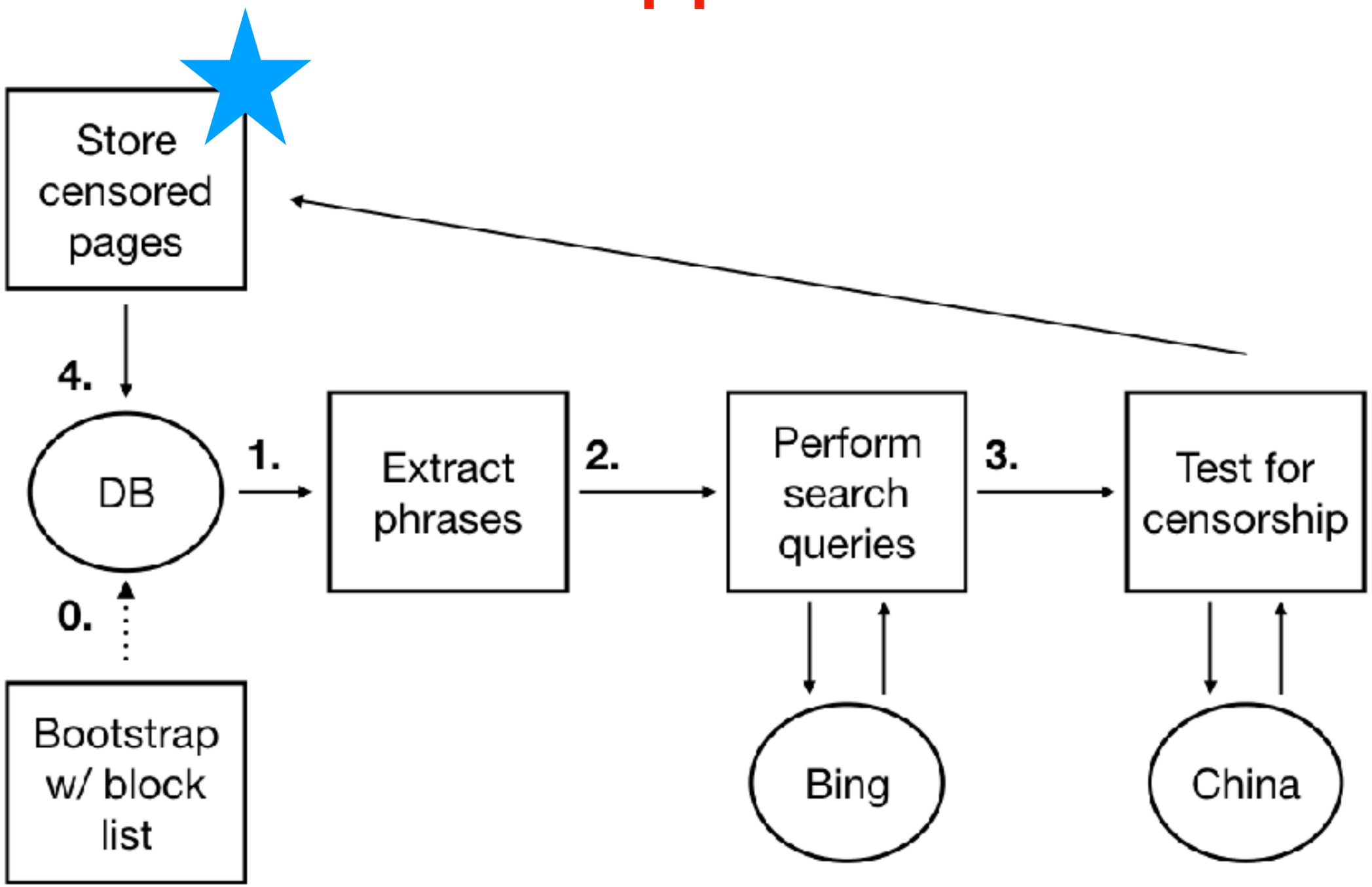
Our approach



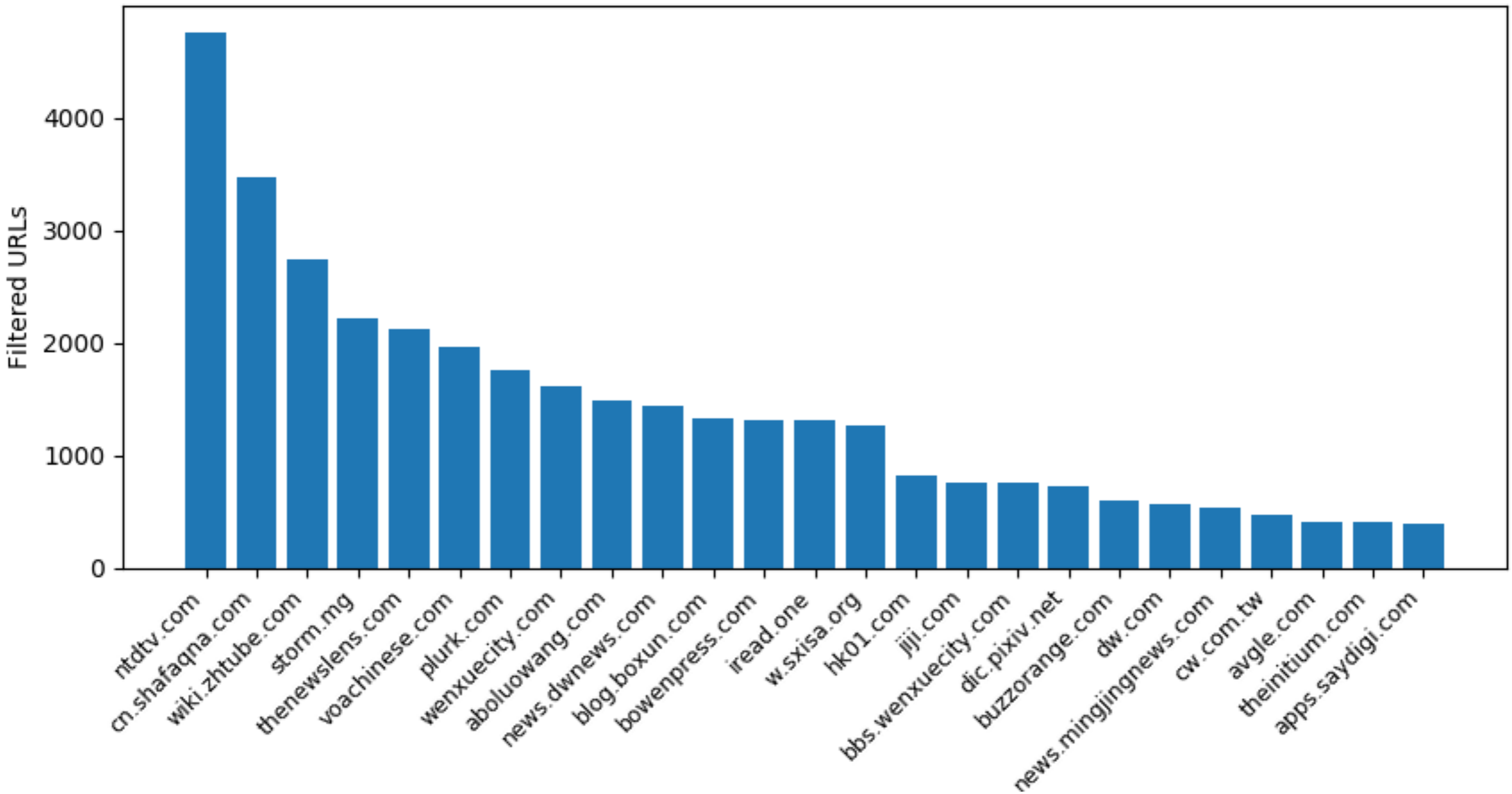
Our approach



Our approach



Finding #1: Blocklists are incomplete



風傳媒
THE STORM MEDIA

news comment life Special topic New news I am artichoke Hard to resist

Exclusive Wu Tong commits Taiwan.. Emperor's fear of eating evil fruit? US experts: China can't

latest news
Latest life

In the middle school, "Roric 6th Bureau has no responsibility to lose the Lions 10 to 6 defeat."
Maoxun 2018-08-10 23:49

Trump has approved the "chain immigration". The parents of the Yue are naturalized as US citizens.
2018-08-10 23:00

Even the US government has fallen? The Chinese map of the State Council website is included in the interpretation of Taiwan's Washington.
2018-08-10 22:45

The ruling of Qualcomm’ s



:r, TV announcer. **The mirror is pleasing!** You may find your stage here. Mirror fire recruiting: marketing, advertising, program host, TV reporte

Ningxia Tongxin Weizhou halal temple became a policeman worried about the stage public opinion

2018-08-10 10:31



Mirror Vo
Mirror Voic
Global broad
Google
The p
attem
User interac
苏冀苏冀
王峧山文
经济日报
Unknown
魏先生是
欢他的节
Unknown

(Faguang RFI Hill) The Weizheng Dasi Temple in Tongxin County, Ningxia Hui Autonomous Region, was triggered by the local government's order to dismantle the temple and the police confrontation with the temple and the suppression. Thousands of people have vowed to defend the mosque, and the police who have to execute the order today must have a bloody crisis.

According to a report by the Beijing-based reporter Haike, the Huizhou Great Mosque in Tongxin County, Ningxia Hui Autonomous Region was threatened with demolition and caused thousands of Hui people to defend. The mosque was recently renovated and expanded. A total of nine mosques have a new round of domes, and the official allegations of illegal construction of the mosque were ordered to be dismantled.

Chinese authorities are increasingly restricting religious supervision. The just-concluded National Religious Association meeting not only requires religious flags to fly the national flag, but also in the hall where the believers pray, and also hangs the axe and the Communist Party flag. But so far, the Hui people are slightly different from the Xinjiang Uygur residents, and they can still pray for relatively loose prayers. However, the Ningxia authorities ordered the rectification of the Weizhou Great Mosque, which was considered by the local Hui people to be a sign of the knife.

The local party and government departments all pointed out that the illegal buildings of Weizhou Dazhen Temple in Tongxin County of Ningxia Hui Autonomous Region were not allowed to persuade the local Hui people. Because the mosque was rebuilt for a long time, it was not accused by the government, and no government department allowed it to rebuild. It is impossible to start work on the project.

Residents of the Huizheng Temple in Weizhou, Tongxin County, Ningxia Hui Autonomous Region, said that they still support the government's policies. The slogans not only support the Chinese Communist Party, but also support the unity

China News

NEWS

Russian PM warns further sanctions by US would be 'act of economic war' - *Scmp*

NEWS

Taiwan s CPI climbs 1.75 pct in July - *en.gmw.cn*

NEWS

South Vietnam to experience electricity shortages - *en.gmw.cn*

NEWS

3,916 Chinese schools identified as specialist schools in youth football - *en.gmw.cn*

NEWS

Former Malaysian Prime Minister Najib Razak will go on trial in February, facing charges of financial corruption - *Scmp*

NEWS

US fears over 'unprecedented opponent' driving trade war, Chinese state media says - *Scmp*

NEWS

China helps Russia thwart US bid to

NEWS

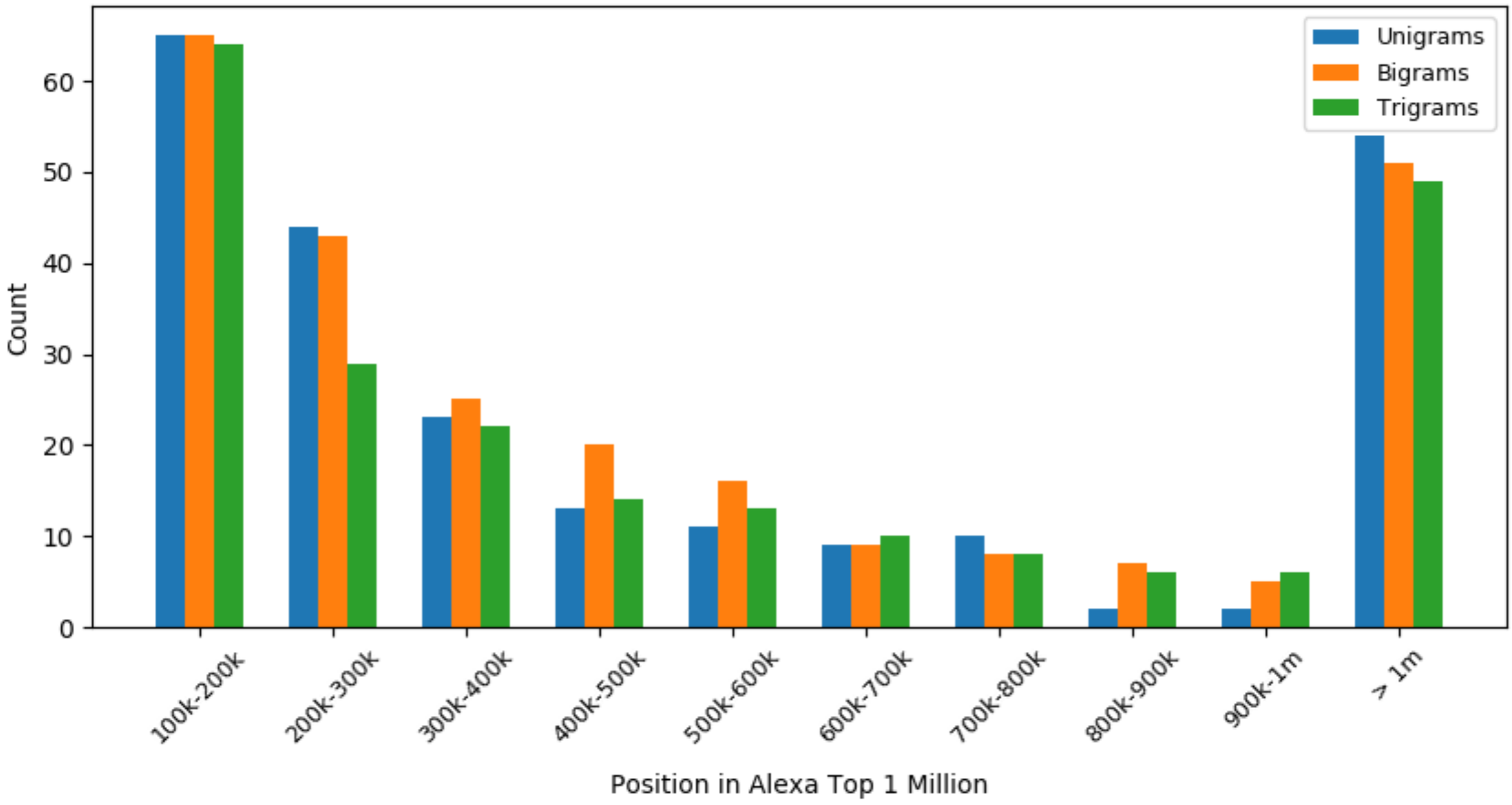
Judge to rule next week on injunction bid

Latest News

- Does Mount Everest's record wet summer point to a 'profound change' in the Earth's climate? - *scmp*
- 2018 Shanghai Isaac Stern International Violin Competition Launches - *Asia Today*
- Protest by Chinese Hui Muslims at Weizhou Grand Mosque goes on as sides seek compromise - *Scmp*
- Pachinko king Kazuo Okada denies charges in Hong Kong arrest - *Scmp*
- 'Asian August' comes to Hollywood but will it last? - *scmp*
- Bright industry outlook echoed by positive exhibitor recruitment at SPS – Industrial Automation Fair Guangzhou - *Asia Today*
- Ecoclean: The choice of the right cleaning technology decides Solvents are often a superior alternative in terms of cost efficiency and sustainability - *Asia Today*
- 'Just tell me why my father died.' Daughter of dead security guard lashes out at Hong Kong government for ignoring his family in their time of grief - *scmp*
- Chinese corruption inquiry finds US$39 million in cash at homes of ex-boss of Huarong Asset Management - *scmp*
- How Can a Small Business Sell Overseas - *Asia Today*
- In vaccination push, 184 Hong Kong primary schools to give free flu jabs - *Scmp*
- Cambodian PM Hun Sen's Facebook page was hacked to give away seats in


Plurk.
Find what's happening on Plurk.
DEERINGTON
Don't forget that applications close tonight at midnight EST!
3
FifteenStrangers
[Quick Attendance Check] We understand Otakon is happening, so we'd like to gauge who is going to be around at the investigation and/or trial this weekend. Please comment if you're going to be around, thank you!
9
– The Bay –
21
Top 90!
Top Plurk
View more
PrismEventsSL
Salem is now taking designer applications. A very limited quantity of standard stalls and
Social Light

Finding #2: Many unpopular websites are blocked



Orthodox Chinese | discuss | read | View source code | View

您好～歡迎來到**偽基百科**！我們的標語是：偽客觀、偽公正、偽持平、偽學術、專事挖苦。
若您有任何古靈精怪的惡搞點子，或可**提供您的內容**。若有興致也可**加入我們**，成為**偽基人大家庭**的一份子。

目前有站務公告，未登入的偽基人可按此查看。

由於2015年的系統更新，現在本站的 檔案： 、 協助： 、 論壇： 等命名空間內的頁面暫時無法正常運作，詳情請見

偽基百科
惡搞的百科全書

- Home
- news
- Anime game
- Book library
- Theme pavilion
- World Pseuco-base
- Random page
- Recent changes

Station service
- Instructions for use
- Editing manual
- community
- forum
- Voting area
- Destruction report
- Comments and suggestions

Toolbox
- Link to this page
- Related page revision record
- Special page
- Printable version
- Static link
- Page information
- Major contributor
- Quote this page

other languages

Home

Pseudo-objective , pseudo-fair , pseudo-flat , pseudo-academic , special purpose sarcasm

Pseudo-base

Everyone can edit the spoof encyclopedia

There are **6,567** bug entries

People | Humanities | History | Social Sciences | Social | Geography | Nature | Science | Technology | Exhibition | Mo

Featured entry

- **Tao Yuanming**

  Tao Yuanming, the world's most famous [redacted] writer, the world's poorest writer, once served as a small official, resigned because of dissatisfaction with official culture [redacted] .

- **Xi Jinping**



  Or **abide by the buns** , the master of the common horror empire Qingfeng emperor, referred to as Xi Di, surname Xi Wei Wei Ni, the full name of the curfew evaluation. It was inaugurated in the Western calendar year, and was the chairman of the Central Committee of the Communist Party of China and the Chairman of the Central Military

news

- **On July 31** , the **Public Security Bureau caught** which is famous for its rumors, recently promulgat burial and cremation. It is not only considered to b but even an incident in which the squadron forcibl protest against the government's push for cremati was burned down by the villagers [redacted]
- **On July 26** , **popcorn exploded?** At one o'cloc time, the US Embassy in Beijing , China , was atta white powder, and the ground was full of broken g glass was full of heart, showing the last piece of p

English | 中文 | Uygur | Français | Español | 日本語 | Indonesian | Русский new | Turkish | Hindi new | বাংলা new | German new | Portuguese new | فارسی new | اردو | عربي |

Islam Question and Answer

General Supervisor: Shaykh Muhammad Saalih al-Munajjid

Fri 28 DhQ 1439 - 10 August 2018

Check out the new version of Islam Q&A

New Fatwas

- He left Makkah before doing the tawaaf and sa'i of Hajj.
- If someone reaches Makkah at Zuhr on the day of 'Arafah and goes out to 'Arafaat and does not do tawaaf al-qudoom (tawaf of arrival).
- Is it prescribed to kiss the Black Stone without doing tawaaf?.
- He formed the intention to do sa'i during the first lap.
- Should a bald man who has no hair pass a razor over his head?.
- When he reaches Muzdalifah he should pray Maghrib and 'Ishaa', putting them together at the time of the earlier or later prayer.
- Do they have to stay in Mina after stoning the Jamaraat on the 13th of Dhu'l-Hijjah?.
- He is not sure whether he threw six or seven pebbles when stoning the Jamrah on behalf of someone else.
- Is it permissible to take stones or earth from Muzdalifah or from the Haram?.
- She was not sure how many circuits of tawaaf she had done, and she did another circuit after one



Categories

Articles & Books

Introduction to Islam

New Fatwas

Southern Mongolia Congress

Official website

A united organization
Working for
Southern Mongolian
democracy and
freedom

southmongolia.org

ABOUT SOUTHERN MONGOLIA CONGRESS REPORT ON HUMAN RIGHTS SITUATION IN SOUTHERN MONGOLIA THE COURSE OF ACTION

THE POLITICAL PLATFORM OF SOUTHERN MONGOLIA CONGRESS

OTHERS

SITUATION OF VIOLATION OF UNITED NATIONS CONVENTION ON THE ELIMINATION OF ALL FORMS OF DISCRIMINATION IN PEOPLE'S REPUBLIC OF CHINA

EBLOGADMIN AUGUST 1, 2018 NO COMMENTS

Original: https://tbinternet.ohchr.org/_layouts/treatybodyexternal /Download.aspx?symbolno=INT%2fCERD%2fNGO%2fCHN%2f31733&Lang=en Situation of Violation of United Nations Convention on the Elimination of All Forms of Discrimination in People's Republic of China July 7, 2018 Asian Solidarity Council for Freedom...

CONTINUE READING

from southmongolia.org...

LANGUAGE

日本語(apanese)
English
中文
Монгол хэл

PHOTOS OF THE FOUNDING CONGRESS

SITE NAVIGATION

- Home
- Important document
- Square circle
- Industrial transport literature
- Cultural Revolution Research
- Historical data
- Picture material
- Labour Party Forum
- Welcome to contact
- Friendly connection

WEBSITE LOGIN

account number

password

remember me

SIGN IN

- Forgot your password?
- Forgot your account?

China Labor Party - Chinese Labour Party

Chinese characteristics and Chinese model
Political Report of the Third National Congress of the Chinese Labor Party
square round

Guests, consultants, delegates: You are good!

The Third National Congress of the Chinese Labor Party is an important conference that will lead the way. Accordin to our two major assumptions, this conference should be held from Hong Kong, the second largest venue, to th mainland. Because as early as eleven years ago, we have already raised the issue of revolution and improvement. Th situation at the time was a reform and revolutionary race. We pointed out that the choice of improvement, or the choic of revolution, the initiative is not in the hands of the people, but in the hands of the rulers. We declare that if the rule choose to improve, the Chinese Labor Party is willing to give up the revolution and embrace reform. If the rulers give u the reform, the Chinese Labor Party will have to rise to the revolution.

Since Hu Wen's pro-government launched the scientific development concept of "people-oriented, harmoniou society", on the second day, we established the reconciliation route of casting swords into plows. At that time, w thought that after the improvement of Hu Wen, China should make great progress, and we can return to the mainland t hold the three major. It seems that our expectations for Hu Wen are too high.

Since the founding of the party on November 17 , 2000 , the Chinese Labor Party has been nearly fourteen years old Two years before the founding of the Party, that is, in 1998 , Dr. Wang Bingzhang had a sword in ten years, and h returned to China to establish the modern China Democracy Party. On the eve of the founding of the Democratic Party Bing Zhang brother personally flew to Australia to negotiate with me, hoping that the Labour Party in preparation for th

Finding #3: Political phrases are very effective as search terms

Chinese	English	Censored domains
王歧山	Wang Qishan	74%
李洪志	Li Hongzhi	64%
郭伯雄	Guo Boxiong	62%
胡锦涛	Hu Jintao	56%
胡平	Hu Ping	54%
Morty	Morty	52%
命案	Homicide	52%
特首	Chief executive	52%
Vimeo	Vimeo	50%
中情局	CIA	50%

Table 2: *Sample of unigrams with significant blockrates*

Finding #3: Political phrases are very effective as search terms

Chinese	English	Censored domains
中共 威胁	Chinese Communists threaten	50%
声明的 反共产主义	Declared anti-communist	44%
中国共 产党的公共安全	Public security of the CPC	42%
北京 清洁	Beijing clean-up	40%
江泽民 胡锦涛	Jiang Zemin Hu Jintao	40%
迫害 活动	Persecution activities	40%
官员 呼吁	Officials called on	36%
重的 公民	Heavy Citizen	36%
非法 拘留	Illegal detention	36%
不同 的民主	Different Democratic	34%

Table 3: *Sample of bigrams with significant block rates*

Finding #3: Political phrases are very effective as search terms

Chinese	English	Censored domains
北戴 河 会议	BEIDAIHE meeting	54%
中国 共产党 的宗教政策	The Chinese Communist Party's religious policy	42%
采取 暴力 镇压	To take a violent crackdown	38%
香港 政 治	Hong Kong Politics	34%
欧洲 议会 决议	European Parliament Resolution	32%
新 唐 王朝	New Tang Dynasty	32%
恐怖 事件 。	A terrorist event.	32%
天安 门广 场示威	Tienanmen Square Demonstrations	32%
敦促 美国 政府	Urging the US government	32%
1989 民主 运动	1989 democracy movement	30%

Table 4: *Sample of trigrams with significant blockrates*

Future work

- Add support for other languages to find censored websites pertaining to other countries

- Measure churn; do additional websites become (un)censored over time?

- Compare results across search engines

Conclusion



- We built a large, publicly-available, culture-specific list of censored domains in China
- It is important to keep testing lists up to date to understand how censorship in a given country changes over time
- https://github.com/ahounsel/auto-blocklist