Codebook *Heidelberg Cyber Conflict Dataset (HD-CY.CON) 1.0*
Institute of Political Science, Heidelberg University

December 2021

General remarks:
Codes and sub-codes are separated by ",".
In case of missing values, "na" is coded.

1. Incident no.

Consecutive numbers for incidents in total, sorted in ascending order according to the start year of the operation.

2. Inclusion criteria / *multiple codings possible (I - II)*

Code	**Subcode**	**Description**
1		Attack conducted by nation state *(generic "state-attribution" or direct attribution towards specific state-entities, e.g., intelligence agencies)*
2		Attack conducted by non-state group / non-state actor with political goals (religious, ethnic, etc. groups) / undefined actor with political goals
	1	Attack conducted by a state-sponsored group *("cyber-proxies")*
3		Targeted attack on political target; politicized
4		Targeted attack on political target; not politicized
5		Untargeted attack on political target; politicized
6		Targeted attack on non-political target; politicized
7		Untargeted attack on non-political target; politicized
8		Targeted attack on various targets; politicized
9		Untargeted attack on various targets; politicized
10		Targeted attack on (amongst others political) targets; not politicized.

Remarks:
"Political targets" are receiver categories 1 and 2 (see below).
Ransomware-operations are only included if they have been attributed to a state (affiliated) actor and / or have been politicized.

3. Source incident detection / disclosure / *multiple codings possible (I - II)*

Actor that was the first to disclose the incident

Code	**Description**
0	Incident disclosed by media *(without further information on source)*
1	Incident disclosed by victim
2	Incident disclosed by IT-security company
3	Incident disclosed by attacker
4	Incident disclosed by third-party-actor *(e.g., Citizen Lab, Amnesty International) or authorities of another state*
5	Incident disclosed by authorities of victim state

4. Start date (year)

5. State date (month)

6. State date (date)

7. End date (year)

8. End date (month)

9. End date (date)

10. Receiver category / *multiple codings possible (I - X)*

Code	**Subcode**	**Description**
0		Unknown
1		State institutions / political system
	1	Government / ministries
	2	Legislative
	3	Civil service / administration
	4	Judiciary
	5	Military
	6	Police
	7	Intelligence agencies
	8	Political parties
	9	Election infrastructure / related systems
	11	Other (e.g., embassies)
2		International / supranational organization
3		Critical infrastructure
	1	Energy
	2	Water
	3	Transportation
	4	Health
	5	Chemicals
	6	Telecommunications
	7	Food
	8	Finance
	9	Defence industry
4		Social groups
	1	Ethnic
	2	Religious
	3	Hacktivist
	4	Criminal
	5	Terrorist
	6	Human rights organizations / activists
	7	Political opposition / dissidents / expats
	8	Other (e.g., NGOs with political goals)
5		Commercial targets
6		End user(s)
7		Media
8		Science
9		Other

11. Receiver country / *multiple codings possible (I - X)*

ISO 3166 ALPHA-3 country codes, see: https://en.wikipedia.org/wiki/ISO_3166-1_alpha-3

Additional institutional or regional codes:

CAU --> Caucasus
NATO --> NATO
EU --> EU
EUROPE --> Europe
EASTEU --> Eastern Europe
WESTEU --> Western Europe
NORTHEU --> Northern Europe
SCS --> South Chinese Sea
MENA --> Mena Region
IAAF --> International Association of Athletics Federations
ISIS --> ISIS
MEA --> Middle East
ASIA --> Asia
SEA --> Southeast Asia
SA --> South Asia
EASIA --> Eastern Asia
SA --> South Asia
CENTAS --> Central Asia
GLOBAL --> Worldwide
WADA --> World Anti-Doping Agency
WTO --> World Trade Organization
UN --> United Nations
UNICEF --> UNICEF
ESA --> European Space Agency
IAEA --> International Atomic Energy Agency
OSCE --> Organization for Security and Cooperation in Europe
IKPO --> Interpol
IMF --> International Monetary Fund
OPCW --> Organization for the Prohibition of Chemical Weapons
CENTAM --> Central America
GULFC --> Gulf Countries
BALKANS --> Balkan States
NAF --> North Africa

12. Receiver regime type / *multiple codings possible (I - III)*

According to Freedom House “Freedom in the World” report for the respective countries in the respective year.

Code	**Description**
0	Country unknown
1	Freedom House (respective year) “free”
2	Freedom House (respective year) “partly free”

3	Freedom House (respective year) “not free”

13. Initiator category

Code	**Subcode**	**Description**
0		Unknown - not attributed
1		State
2		Non-state actor, state sponsorship suggested
	1	Non-state-group, state-involvement suggested *(widely held view, but not invoked in this case)*
3		Non-state-group
	1	Ethnic actors
	2	Religious actors
	3	Hacktivist(s)
	4	Criminal(s)
	5	Terrorist(s)
	6	Other actors *(e.g., private technology companies, hacking for hire groups without state affiliation or research entities)*
4		Individual hacker(s)
5		Other

14. Initiator name(s) / *multiple codings possible (I - II)*

Name(s) of the initiating group (if known), according to the telemetry of private IT-companies and (if available) the designations of aligned offline-state-units, e.g., military unit names.

15. Initiator country / multiple codings possible *(I - V)*

ISO 3166 ALPHA-3 country codes, see: https://en.wikipedia.org/wiki/ISO_3166-1_alpha-3

16. Initiator regime type / *multiple codings possible (I - II)*

According to Freedom House “Freedom in the World” report for the respective countries in the respective year.

Code	**Description**
0	Country Unknown
1	Freedom House (respective year) “free”
2	Freedom House (respective year) “partly free”
3	Freedom House (respective year) “not free”

17. Attribution year

Only coded for operations with attributed state-involvement on the attacker-side (initiator categories 1, 2 & 2,1).

18. Political/legal response year (e.g., indictment, arrest, lawsuit, sanctions)

19. Attribution basis / *multiple codings possible (I - IV)*

Code	**Description**
0	Unknown - not attributed / Media-based attribution
1	Receiver attributes attacker
2	IT-security community attributes attacker
3	Attacker confirms
4	Contested attribution
5	Attribution by third-party
6	Attribution by receiver government / state entity

20. Attribution type / *multiple codings possible (I - IV)*

Remark: Attribution Type I refers to Attribution Basis I; Attribution Type II refers to Attribution Basis II etc.

Code	**Description**
0	Attribution type unclear
1	Self-attribution in the course of the attack *(e.g., via defacement statements on websites)*
2	Media report *(e.g., Reuters makes an attribution statement, without naming further sources)*
3	Direct statement in media report (*e.g., Reuters article cites the attribution statements by a person*) */* self-attribution via social media
4	Anonymous statement in media report *(e.g., Reuters article cites the attribution statements of unnamed officials, or persons with knowledge into the matter etc.)*
5	Technical report *(e.g., by IT-companies, Citizen Lab, EFF)*
6	Political statement / report *(e.g., on government / state agency websites)*
7	Indictment/sanctions/arrests (by state authorities) / lawsuit (by private actors)
8	Statement in media report and political statement/technical report
9	Statement in media report and indictment / sanctions
10	Political statement/report and indictment / sanctions

21. Attribution IT-security community (initiator category) / *multiple codings possible (I - II)*

Code	**Subcode**	**Description**
0		Unknown
1		State
2		Non-state actor, state sponsorship suggested
	1	Non-state-group, state-sponsorship suggested *(widely held view, but not invoked in this case)*
3		Non-state - group
	1	Ethnic
	2	Religious
	3	Hacktivist
	4	Criminal
	5	Other
4		Individual hacker
5		Other

Remark: *Only coded in case of available attribution by private IT-companies.*

22. Attribution IT-security community (initiator country) / *multiple codings possible (I - II)*

ISO 3166 ALPHA-3 country codes, see: https://en.wikipedia.org/wiki/ISO_3166-1_alpha-3

23. Country of origin: IT-Attribution / *multiple codings possible (I - II)*

ISO 3166 ALPHA-3 country codes for the geographic origin(s) (in general legal headquarters) of the IT-security company / companies, cited as attribution basis for encoding 21. and 22.

24. Temporal attribution sequence

Only relevant if Attribution Basis is 2 *(IT-security community attributes attacker)* AND 6 *(attribution by receiver government / state entity)*.

Code	**Description**
0	Temporal attribution sequence unclear
1	Political attribution before IT-security attribution
2	IT-security attribution before political attribution

25. Cyber-conflict issue / *if appropriate (multiple codings possible) (I - IV)*

*According to the Conflict Issues by the Conflict Barometer of the Heidelberg Institute for International Conflict Research (HIIK), see https://hiik.de/hiik/methodology/?lang=en.**

Code	**Description**
0	Unknown
1	System / ideology
2	National power
3	Autonomy
4	Territory
5	Subnational predominance
6	Resources
7	International power
8	Decolonization
9	Secession
10	Cyber-specific**
11	Other

Remarks:

**The conflict issue of "decolonization" was excluded from the list, because its saliency in the context of cyber-operations seemed unlikely / implausible.*

***The conflict issue "cyber-specific" was added, e.g., for cyber-operations concerning genuinely cyber-related events or issues (example: DDoS-operations by hacktivists against a national court authorizing online-censorship-measures).*

26. Offline-conflict (intensity) - HIIK

According to the intensity levels by the Conflict Barometer of the Heidelberg Institute for International Conflict Research (HIIK), see https://hiik.de/hiik/methodology/?lang=en.

Only relevant for cyber-operations that indicated direct links to the respective HIIK-offline-conflicts.

Code	**Subcode**	**Description**
0		No spillover
1		Yes / HIIK intensity
	1	HIIK 1
	2	HIIK 2

	3	HIIK 3
	4	HIIK 4
	5	HIIK 5

27. Offline-conflict (issue) - HIIK / *multiple codings possible (I - IV)*

According to the conflict issues by the Conflict Barometer of the Heidelberg Institute for International Conflict Research (HIIK), see https://hiik.de/hiik/methodology/?lang=en.

Only relevant for cyber-operations that indicated direct links to the respective HIIK-offline-conflicts.

Code	**Description**
0	No spillover
1	System/ideology
2	National power
3	Autonomy
4	Territory
5	Subnational predominance
6	Resources
7	International power
8	Decolonization
9	Secession
10	Other
11	Third-party intervention / third-party affection*

Remark:

**Additionally coded for cases in which the cyber-attacker conducts its operation because of a specific offline-conflict captured by the HIIK conflict barometer, without being officially part of it (e.g., patriotic hacktivists from one state act in solidarity with another state because of escalating dynamics in the offline-conflict-realm).*

28. Zero day(s)

Code	**Subcode**	**Description**
0		Unknown
1		No
2		Yes
	1	One
	2	multiple

29. Incident type(s) / *multiple codings possible (I - III)*

Code	**Subcode**	**Description**
1		Data theft / data loss
	1	Combined with doxing
2		Disruption
3		Hijacking / system misuse

Intensity / *calculation based on effects of incidents*

30. Data theft

Code	Description
0	none
1	For political / military targets: non-classified information *(incident scores 1 point in intensity)* For private / commercial targets: non-sensitive information *(incident scores 1 point in intensity)*
2	For political / military targets: classified information *(incident scores 2 points in intensity)* For private / commercial targets: sensitive information *(incident scores 2 points in intensity)*

31. Disruption*

Code	Description
0	none
1	Short-term disruption *(< 24h; incident scores 1 point in intensity)*
2	Long-term disruption *(> 24h; incident scores 2 points in intensity)*

**Remark:*

The disruption needs to be caused directly by the respective cyber-operation. Disruptions caused by precautionary measures are not included here.

32. Hijacking (accessing and controlling a system)

Code	Description
0	none
1	Hijacking, not used - empowerment *(incident scores 1 point in intensity)*
2	Hijacking, system misuse, e.g., through data theft and / or disruption *(incident scores 2 points in intensity)*

33. Physical effects (spatial)

Code	Description
0	none
1	Local effects, e.g., affecting only one restricted area of a country or region (*incident scores 1 point in intensity*)
2	Widespread effects, e.g., affecting different regions of country or a country as a whole *(incident scores 2 points in intensity)*

34. Physical effects (temporal)

Code	Description
0	none
1	Short duration *(< 24h; incident scores 1 point in intensity)*
2	Long lasting effects *(> 24h; incident scores 2 points in intensity)*

35. Unweighted cyber intensity

Sum of data theft + disruption + hijacking + physical effects spatial + physical effects temporal = maximum score 10

36. Target / effect multiplier

Code	Description
1	Moderate - high political importance
2	Very high political importance (e.g., critical infrastructure, military) - intensity multiplied by 1.5

The 1.5 multiplier is applied based on societal effects. It is only applied, when there is a widespread disruption of services that are of critical importance for the functioning of affected societies. In case of the military the respective social function (national defense) has to be affected, the same applies to critical infrastructures. Other attacks on critical infrastructure / military entities are multiplied by 1.0.

Respective mathematical caps: 5; 10; 15

37. Weighted cyber intensity

Unweighted cyber intensity multiplied with target multiplier / *brought up to round figure*

1	Low / moderate intensity
2	
3	
4	
5	
6	High intensity
7	
8	
9	
10	
11	Very high intensity
12	
13	
14	
15	

38. Casualties

Code	Description
0	No casualties
1	Injured (incident gets marked as category A)
2	< 25 (incident gets marked as category B)
3	>= 25 (incident gets marked as category C)

Additional Information

39. Sources

URLs

40. Sources attribution

URLs

41. Sources politicization

URLs

42. Incident name

Denominator that is mostly used to refer to the respective incident (e.g., Stuxnet), else descriptive denominator (e.g., Ukraine power outage 2015).

43. Description

Short description of the incident.