Appendix A: Codebook

<u>Taxonomy and Definitions</u>
The indentation may be a bit off, but now the hierarchy is: I, 1, A, a, i. This helps get us shorthand unique IDs for each of the variables. Variables are in plain text; definitions are italicized. All variables are marked with one of these "data types":
- binary
- decimal (ie, 2.4)
- integer (ie, 2 or -4)
- free text
- country (all countries are options in a drop-down)
- language (all languages are options in a drop-down)
- bloc (all regional blocs are options in a drop-down)
- year (relevant years are options in a drop-down)
- month (all months are option in a drop-down)
- range (of years -- pre-determined from data source)

Meta

- Serial # (integer)
- Campaign ID (<STARTING YEAR> <ATTACKER> against <TARGET>). *This is intended to aid categorization and reconciliation of overlapping campaigns.*
- Long Description (free text). *Summary of the campaign in one to two sentences.*
- Dates
  - Start
    - Starting Year (year). *Starting year of when the campaign first began. The starting year can be inferred from context or be explicitly stated in text.*
    - Starting Month (month).
    - Certainty Start (binary). *Input 1 when the source document clearly indicates that the campaign began in this time frame.*
  - End
    - Ending Year (year)
    - Ending Month (month). *In the event of a takedown, the ending month is the month of the takedown (unless it is demarcated that the assets stopped being active prior to the takedown date). For*

cases that are still occuring at the time of coding, the ending year and month should be left blank.
- ■ Certainty End (binary). *Input 1 when the source document clearly indicates that the campaign ended in this time frame.*
- ● Sources
  - ○ Year of Report (year). *Year of Citation 1.*
  - ○ Month of Report (month)
  - ○ Citation 1 (free text). *Chicago style.*
  - ○ Link 1 (free text)
  - ○ Citation 2 (free text). *Chicago style.*
  - ○ Link 2 (free text)
- ● Notes

I. Target

1. Primary Target
- A. Nation of Origin (country). *This should be filled out even when the target is not a nation. When a campaign targets a non-state political actor, the nation of origin of that non-state political actor is filled in this field, if that information is available. Distinguishable territories are nations.*
- B. Regional Bloc (bloc) *When a single nation of origin cannot be determined.*
- C. Other (free text)
- D. Notes (free text)

2. Target Category. *Categories are not mutually exclusive. All relevant categories can be added.*
- A. Government. *If a determination cannot be made between Civilian and Military categories, the Civilian category takes precedence.*
  - a. Civilian (binary). *The governing body and functions of a state, including national leaders, institutions, and non-military departments and agencies. Includes incumbent politicians running for re-election.*
  - b. Military (binary). *Military departments and agencies which enjoy the sanctioned use of force.*
- B. Political Party (binary). *Organized competitors for political power who can obtain or wield power directly. Includes politicians currently in office, as well as non-incumbent politicians running for office who are associated with a political party. Can also be an individual working for a party.*

C. Non-State Political Actor (binary). *Organized competitors for political power who can obtain or wield power, even if indirectly; not necessarily enfranchised. Non-state political actors are formally organized, coordinated, and cohesive. e.g. Greenpeace, the NRA, or the KKK.*

D. Business (binary). *Includes groups that contract out to the government, individuals looking for financial gain, and mercenaries.*

E. Influential Individuals (binary). *Individuals who are influential but who do not belong to a ruling government coalition. Includes groups of individuals who are not formally organized but work together. e.g. journalists, former politicians, or organized 4channers. For individuals who operate their own charitable foundations (and thus could be placed in Non-State Political Actor), coding depends on whether or not the disinformation is foremost targeting the individual, their foundation, or both.*

F. Electorate (binary). *The enfranchised population in a specific country or within a demarcated boundary.*

G. Racial, Ethnic, Religious, or Sexual Identity Group (binary). *A specific minority/majority group.*

H. Other (free text)

I. Notes (free text)

3. Quantitative Measures. *If nation of origin is filled out, this category should be filled out.*

    A. Political Stability. *Use the Estimate figure of the* WGI *data, "Political Stability and Absence of Violence/Terrorism." For 2019 and onwards, use most recently available data.*

        a. Political Stability Data. *Year prior to campaign start.*

            i. Political Stability (decimal).

            ii. Data Year (year)

        b. Political Stability Data. *Campaign start.*

            i. Political Stability (decimal).

            ii. Data Year (year)

        c. Political Stability Data. *Year after campaign start.*

            i. Political Stability (decimal).

            ii. Data Year (year)

    B. Refugee % Change

        a. Refugee % Change (decimal). *Use* UNHCR *data, table 6, "Annual rate of change of refugee stock."*

        b. Data Years (range). *Use "2010-2015" for 2015, to show historical progression. For 2018 and onwards, use most recently available data.*

C. Voice and Accountability. *Use the Estimate figure of the <u>WGI</u> data, "Voice and Accountability." For 2019 and onwards, use most recently available data.*
- a. Voice and Accountability Data. *Year prior to campaign start.*
    - i. Voice and Accountability (decimal).
    - ii. Data Year (year)
- b. Voice and Accountability Data. *Campaign start.*
    - i. Voice and Accountability(decimal).
    - ii. Data Year (year)
- c. Voice and Accountability Data. *Year after campaign start.*
    - i. Voice and Accountability (decimal).
    - ii. Data Year (year)

D. Internet Freedom. *Use <u>FH</u> data. Input number only, not ranking out of 100.*
- a. Internet Freedom. *Year prior to campaign start.*
    - i. Internet Freedom (integer).
    - ii. Data Year (year)
- b. Internet Freedom. *Campaign start year.*
    - i. Internet Freedom (integer).
    - ii. Data Year (year)
- c. Internet Freedom. *Year after campaign start.*
    - i. Internet Freedom (integer).
    - ii. Data Year (year)

4. Concurrent Events

A. Inter-state war (binary). *Threshold is 1,000 conflict deaths. Use <u>COW</u> data for 2007 and before. 2008 and after, supplement with research.*

B. Extra-state war (binary). *Threshold is 1,000 conflict deaths. Use <u>COW</u> data. For data after 2007, see below, Appendix I: Extra-state war.*

C. Intra-state wars (binary). *Threshold is 1,000 conflict deaths. Use <u>COW</u> data for 2007 and before. 2008 and after, supplement with research.*

D. Non-state wars (binary). *Wars in non-state territory or across state borders. Threshold is 1,000 conflict deaths. Use <u>COW</u> data for 2007 and before. 2008 and after, supplement with research.*

E. Federal Election (binary)

F. State Election (binary). *Includes elections at province, municipality, administrative region, department, prefecture, and local levels.*

G. Other (free text)

H. Notes (free text)

5. Secondary Target. *This should rarely be used.*

A. Nation of Origin (country)

B. Regional Bloc (bloc)

       C. Other (free text)
       D. Notes (free text)

6. Tertiary Target. *This should rarely be used. When there are more than three targets, define a regional bloc or code as different cases.*
       A. Nation of Origin (country)
       B. Regional Bloc (bloc)
       C. Other (free text)
       D. Notes (free text)

II. Platforms

1. Open Web
       A. State Media (binary). *Includes "state-adjacent" media, operated by government proxies or otherwise beholden to the state.*
       B. Independent Media (binary). *Media institutions that are not beholden to the government and can be reasonably assessed to score >60 by the NewsGuard rating process.*
       C. "Junk News" Website (binary). *A website that trafficks in deceptive headlines, fails to correct errors, avoids disclosure of funding sources, and avoids labeling advertisements. One that can be reasonably assessed to score <60 by the NewsGuard rating process.*

2. Social Media. *Social media accounts created by the disinformants for deceptive purposes.*
       A. Facebook (binary)
       B. Instagram (binary)
       C. Twitter (binary)
       D. YouTube (binary)
       E. LinkedIn (binary)
       F. Reddit (binary)
       G. VK (binary)
       H. Forum Board (binary)
       I. Other (free text)

3. Messaging Platforms
       A. WhatsApp (binary)
       B. Telegram (binary)

        C. Signal (binary)
        D. Line (binary)
        E. WeChat (binary)
        F. SMS (binary)
        G. Other (free text)

4. Advertisement (binary). *Advertisements purchased by disinformants to disseminate a message of disinformation. Includes ads on social media and the open web.*

5. Email (binary)

6. Other (free text)

7. Notes (free text). *Add more detailed information on social media platforms used (retweets, groups, etc), if applicable. Add specific numbers of social media metrics, if applicable.*

III. Content

1. Language (language). *The language of the disinformation. In coding, separate languages with a comma; do not use "and."*

2. Topics. *Subjects evident in the campaign.*
        A. Government (binary). *Includes international governing bodies.*
        B. Military (binary)
        C. Political Party (binary)
        D. Elections (binary)
        E. Nonstate Political Actor (binary)
        F. Business (binary)
        G. Influential Individuals (binary)
        H. Racial, Ethnic, Religious, or Sexual Identity Group (binary)
        I. Terrorism (binary)
        J. Immigration (binary)
        K. Economic Issue (binary)

        L. Other (free text)

    3. Notes (free text)

IV. Methods

1. Tactics

        A. Brigading (binary). *Patriotic trolls or organic coordination in which disinformants seemingly operate under their real identities. A concentrated effort by one online group to manipulate another, e.g. through mass-commenting a certain message.*

        B. Sockpuppets (binary). *Inauthentic social media accounts used for the purpose of deception which evidence a high likelihood of human operation. This includes catfishing and other highly tailored operations conducted under inauthentic personas.*

        C. Botnets (binary). *Inauthentic social media accounts used for the purpose of deception which evidence a high likelihood of automation. These accounts evidence no sustained human intervention beyond the effort necessary to program them initially. They often form large networks for the purpose of inauthentic amplification. This includes both fresh and repurposed accounts.*

        D. Search Engine Manipulation (binary) *Undermining search engine optimization techniques with the intention of creating an inorganic correlation of search queries and results. Often realized by way of cooperative efforts by online communities. e.g. "Google Bombing." May also include typosquatting with the intention to mislead or redirect to another URL.*

        E. Hacking

                a. DDoS (binary). *Distributed denial-of-service. Malicious attempt to disrupt server traffic. In the context of political disinformation campaigns, this is intended to make it more difficult for the target to launch an effective counter-messaging effort.*

                b. Data Exfiltration (binary). *The unauthorized movement of data. In the context of political disinformation campaigns, this is the acquisition of sensitive information through spearphishing or similar techniques that can be subsequently released by the disinformant to boost their messaging effort.*

        F. Deceptive Content Manipulation (binary). *Any content that has been deceptively edited by use of Photoshop or similar software. This includes*

the deceptive co-option and re-use of extant media branding and style guides. This does not include the use of deep learning processes.
   a. Deep learning processes (binary). *Augmented or fabricated content produced using deep learning processes. Includes "deep fakes," "deep voice," and textual generation.*
G. Other (free text)
H. Notes (free text)

2. Narrative Techniques

A. Constructive
   a. Activate (binary). *Bandwagon, pander, ignite. e.g., "If you love Mr. Trump, RT this."*
   b. Astroturf (binary). *Artificial consensus-building, inflation, or amplification. Also called a "Potemkin Village." e.g., "The #1 trending hashtag can't be wrong."*
B. Destructive
   a. Suppress (binary). *Harass, intimidate, exhaust. Often targets influential individuals.*
   b. Discredit (binary). *Libel, leak, tarnish. Often targets government, political parties, elections, or other institutions.*
C. Oblique
   a. Troll (binary). *Confusion by way of discourse infiltration and targeted distraction. Conscious efforts by disinformants to derail political movements through tailored engagement.*
   b. Flood (binary). *Confusion by way of hashtag invasion and mass noise generation. The hijacking of an online political movement through appropriation of an existing hashtag and addition of large quantity of unrelated material.*
D. Notes (free text)

3. Notes

V. Attribution

1. Primary Disinformant

A. Nation of Origin (country). *This should be filled out even when the attacker is not a national government. When a campaign is run by a non-state political actor, the nation of origin of that non-state political actor is filled in this field, if that information is available. Likewise, this should be filled out*

if the preponderance of attacker activity originates from within a single nation. Distinguishable territories are nations.

B. Regional Bloc (bloc)

C. Other (free text)

D. Notes (free text)

2. Disinformant Category. *Categories are not mutually exclusive. All relevant categories can be added.*

A. Government.

a. Direct Attribution (binary). *Public, definitive attribution to a national government by a social media platform or trusted government entity. These entities have access to signals intelligence and other publicly unavailable information.*

b. Proxy/Inferred Attribution (binary). *Informed attribution to a government or government-adjacent proxy in which definitive proof is absent. Such attribution is based on open-source data and inference. This includes attribution to political parties, non-state political actors, businesses, and influential individuals who are suspected to be working at the government's direction.*

B. Political Party (binary). *Organized competitors for political power who can obtain or wield power directly. Includes politicians currently in office, as well as non-incumbent politicians running for office who are associated with a political party. Can also be an individual working for a party.*

C. Non-State Political Actor (binary). *Organized competitors for political power who can obtain or wield power, even if indirectly; not necessarily enfranchised. Non-state political actors are formally organized, coordinated, and cohesive. e.g. Greenpeace, the NRA, or the KKK.*

D. Business (binary). *Includes groups that contract out to the government, individuals looking for financial gain, and mercenaries.*

E. Influential Individuals (binary). *Individuals who are influential but who do not belong to a ruling government coalition. Includes groups of individuals who are not formally organized but work together. e.g. journalists, former politicians, or organized 4channers. For individuals who operate their own charitable foundations (and thus could be placed in Non-State Political Actor), coding depends on whether or not the disinformation is foremost targeting the individual, their foundation, or both.*

F. Electorate (binary). *The enfranchised population in a specific country or within a demarcated boundary.*

G. Racial, Ethnic, Religious, or Sexual Identity Group (binary). *A specific minority/majority group.*

H. Other (free text)

        I.   Notes (free text)

3. Quantitative Measures. *If nation of origin is filled out, this category should be filled out.*

    A. Political Stability *Use the Estimate figure of the <u>WGI</u> data, "Political Stability and Absence of Violence/Terrorism." For 2019 and onwards, use most recently available data.*

        a. Political Stability Data. *Year prior to campaign start.*
            i.   Political Stability (decimal).
            ii.   Data Year (year)

        b. Political Stability Data. *Campaign start year.*
            i.   Political Stability (decimal).
            ii.   Data Year (year)

        c. Political Stability Data. *Year after campaign start.*
            i.   Political Stability (decimal).
            ii.   Data Year (year)

    B. Refugee % Change

        a. Refugee % Change (decimal). *Use <u>UNHCR</u> data, table 6, "Annual rate of change of refugee stock."*

        b. Data Years (range). *Use "2010-2015" for 2015, to show historical progression. For 2018 and onwards, use most recently available data.*

    C. Voice and Accountability. *Use the Estimate figure of the <u>WGI</u> data, "Voice and Accountability." For 2019 and onwards, use most recently available data.*

        a. Voice and Accountability Data. *Year prior to campaign start.*
            i.   Voice and Accountability (decimal).
            ii.   Data Year (year)

        b. Voice and Accountability Data. *Campaign start year.*
            i.   Voice and Accountability(decimal).
            ii.   Data Year (year)

        c. Voice and Accountability Data. *Year after campaign start.*
            i.   Voice and Accountability (decimal).
            ii.   Data Year (year)

    D. Internet Freedom. *Use <u>FH</u> data. Input number only, not ranking out of 100.*

        a. Internet Freedom. *Year prior to campaign start.*
            i.   Internet Freedom (integer).
            ii.   Data Year (year)

        b. Internet Freedom. *Campaign start year.*
            i.   Internet Freedom (integer).
            ii.   Data Year (year)

      c. Internet Freedom. *Year after campaign start.*
          i.   Internet Freedom (integer).
          ii.   Data Year (year)

4. Concurrent Events

A. Inter-state war (binary). *Threshold is 1,000 conflict deaths. Use COW data for 2007 and before. 2008 and after, supplement with research.*
B. Extra-state war (binary). *Threshold is 1,000 conflict deaths.Use COW data. For data after 2007, see below, Appendix I: Extra-state war.*
C. Intra-state wars (binary). *Threshold is 1,000 conflict deaths. Use COW data for 2007 and before. 2008 and after, supplement with research.*
D. Non-state wars (binary). *Wars in non-state territory or across state borders. Threshold is 1,000 conflict deaths. Use COW data for 2007 and before. 2008 and after, supplement with research.*
E. Federal Election (binary). National level election. Supra-national bodies are included.
F. State Election (binary). *Includes elections at province, municipality, administrative region, department, prefecture, and local levels.*
G. Other (free text)
H. Notes (free text)

5. Secondary Disinformant. *This should rarely be used.*

A. Nation of Origin (country)
B. Regional Bloc (bloc)
C. Other (free text)
D. Notes (free text)

6. Tertiary Disinformant. *This should rarely be used. When there are more than three disinformants, code as different cases.*

A. Nation of Origin (country)
B. Regional Bloc (bloc)
C. Other (free text)
D. Notes (free text)

VI. Intent. *This is the intent of the primary disinformant and any other disinformants coded.*

1. Objective (free text). *One or two succinct sentences.*
2. Category. *A clearly definable intent may not exist.*
    a. Civil (binary). *To include electoral interference, policy change.*
    b. Social (binary). *To include marginalization of majority/minority groups and general social fissure.*
    c. Economic (binary). *To include suppression of economic activity and destruction of capital.*
    d. Military (binary). *To include complement to offensive military campaign, or information paralysis of an adversary's military institutions.*

Notes (free text)