

## Appendix

Table A1: This table functions as the primary codebook for the variables included in the dataset. It provides a comprehensive overview of the information collected for each deepfake entry. The following pages contain the coder form for new entries to the PDID that provides additional information on our taxonomy and descriptions of variables.

Variable	Field Name	Description
<b>Descriptors</b>		
Incident ID	incident_id	The number of the observation in the database
Report ID	report_id	The number of the incident in the database
Created By	creativity	Who created the row initially
Modified By	modified_by	Who last modified the row
Date Coded	date_coded	The date the row was coded
URL	url	The url of the deepfake
File	file	The image, video, audio, text
Date Posted	date_posted	When the deepfake was posted
Format	format	The format of the deepfake (video, image)
Transcript	transcript	Transcript of audio or text of deepfake
Summary Content	summary_content	A summary of what the deepfake is
Type	type_of_fake	Deepfake, cheap fake
Comments on Type	comment_type	A spot for coders to input comments about how they decided the type
<b>Verification</b>		
Sharer Fake Type	sharer_fake_type	Whether the sharer presents the information as fake or real
Verified Fake	external_verification	Is it verified to be fake
Verification of Fake	links_err_fraudulent	A place for coders to put evidence that it is fake
Identifiable Markings	evidence_fake	Are there any markings that indicate it is fake
Watermark	watermark	Does the deepfake have a watermark with metadata
<b>Social Media</b>		
Likes	likes	How many likes does the post have
Views	views	How many views does the post have
Comments	comments	How many comments does the post have
Shares	shares	How many shares does the post have
<b>Source</b>		
	original_source_type	Who is the original source of the deepfake, this is their 'type'. The options are the following: individual, group, unclear/unknown
	original_source_actor_type	Who is the original source of the deepfake, this is their 'type'. The options are the following: political leader or figure, government entity, private or NGO, unclear/unknown
	original_source_name	What is the original source's name
	original_source_country	Where is the original source from
	comments_coding_original_source	Coders can supply additional information about how they coded the original source information
<b>Sharer</b>		
	sharer_type	Who is the sharer of the deepfake, this is their 'type'. Options: non-identifiable, individual, group, bot, high impact user (i. 100,000 followers)
	sharer_job	What is the sharer's job, more detail of who they are. Options: politician, public figure, random user, journalist, researcher/academic, other, unclear/unknown
	high_impact_comment	If the coder selected high impact user for the job, they can add a comment to explain why
	sharer_name	What is the sharer's name
	sharer_location	Where is the sharer from, country
	sharer_city	Where is the sharer from, the city
	comments_sharer_coding	The coder can add additional comments about how they coded the sharer fields
<b>Targets</b>		
	target_type_macro	who is the target of the deepfake, what is their 'type'. Options: individual, group, unclear/unknown
	target_type_micro	who is the target of the deepfake, what is their 'type'. Options: politician, public figure, journalist, random user, government entity, private, NGO, marginalized group, unclear/unknown
	target_name	What is the target entity's name. Options: mussia, putin, senator, macro, police, donald trump
	target_location	Where is the target entity located
	target_responded	Did the target entity respond? Yes, no
	comments_target_coding	Coders can provide additional details about how they coded the targets
	target_response	If the target entity responded, what did they say?
<b>Context &amp; Content</b>		
	deepfake_content_depicts	What does the deepfake depict? Options: scandal, crime, sex-related, crisis event, election, government or legal system, war or violence, international relations, marginalized group, deepfake regulation
	harm_depicted	What kind of harm does the deepfake depict? Options: property, environmental, violence, human rights, discrimination, other, non-identifiable
	comments_content_type	Coders can add comments about how they coded the content of the deepfake
	communication_goal	What is the goal of the sharer? Options: satire, false information, entertainment, activism, education, diversion/flooding, harm property, harm political interference, harm environment, harm violence, harm sexual violence, harm human rights, harm emotional/psychological, harm reputation, harm other, non-identifiable
	comments_communication_goal	Coders can add comments about how they coded the communication goal
	core_frame	From Semelke and Volkenberg (2009), what is the framing of the deepfake. Options: human interest, conflict, economic consequences, morality, responsibility.
	comments_core_frame	Coders can add comments about how they coded the core frame
<b>Real-World Connections</b>		
	context_deepfake	What is the real-world context of what was happening when the deepfake was posted?
	text_around_deepfake	Copy and paste the text included with the deepfake
	harm_evidence	Is there evidence of real-world harms caused by the deepfake? If so, what type? Options: human rights, violence, property, financial loss, environment, political interference, discrimination, other, non-identifiable
	comments_harms_context	Coders can add comments about they chose the harms above
<b>Politics &amp; Policy</b>		
	social_policy_sector	medicine, life-sciences, agriculture, resource-extraction, transportation, manufacturing, process-automation, consumer-goods, energy-utilities, construction, field-services, security, biometrics, finance, sales, retail, customer-relations, business-services, analytics, broadcasting, media-production, networking, telecommunications, arts, sports, leisure, travel, lifestyle, education, military, public-safety, government, other, unclear
	comments_social_policy_sector	Coders can add comments about they chose the specific policy sector above
	hero	The hero represents the protagonist or desired outcome of a policy. It can be an individual, group, or nation. The hero is the central figure that policymakers want to highlight and promote as the positive force that will benefit society.
	villain	The villain represents the obstacle or challenge that the hero must overcome. It can be a specific problem, opposing group, policy failure, or any other negative force that stands in the way of achieving the desired outcome. The villain is often portrayed as the source of the problem the policy aims to address.
	plot	The storyline of narrative are that connects the hero and the villain. It outlines the sequence of events or actions that unfold as the hero confronts and overcomes the villain
	moral	The lesson or value that policymakers want to convey to the audience. It reflects the underlying principles or ideals that guide the policy's objectives. The moral aspect often emphasizes the importance of certain values like justice, equality, security, sustainability.
	comments_policy_narrative	Coders can add comments about how they coded the policy narrative.

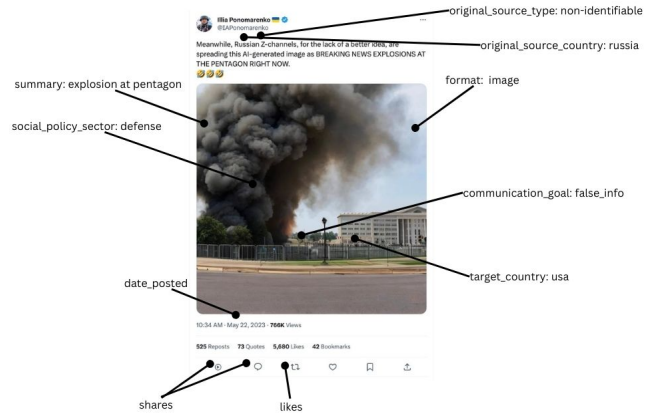


Figure A1: Annotated deepfake

# New Entry Political Deepfakes Incidents Database (GRAIL)

This is the form to add new entries to the PDID.

Today's date \*

11/30/2023

Deepfake ID number (from assigned spreadsheet) \*

Paste the URL of the deepfake (e.g., original social media post or news article) \*

Upload the deepfake file \*

📎 Drop files here or [browse](#)

What date was the deepfake posted? \*

mm/dd/yyyy

Is the deepfake an image, video, or audio? \*

Audio is not yet available

- ☐ image
- ☐ video
- ☐ audio

Provide a 1-2 sentence summary of the deepfake's depicted content. This is a summary of what is happening in the deepfake. \*

Is the content a deepfake, cheapfake, or is this unclear or unknown? \*

- ☐ deepfake
- ☐ cheapfake
- ☐ unclear/unknown

Optional: Why did you choose your answer above (cheapfake vs deepfake)?

Does the sharer seem to present the content as fake, real/authentic, or is this unclear? \*

- ☐ fake
- ☐ real\_authentic
- ☐ unclear

Is there external verification that the image/video/audio is fake or real? \*

(i.e., a news source, something that can be trusted as a fact checker)

**verified\_fake:** Verification that content is fake

**verified\_real:** Verification that content is real/authentic

**authenticity\_disputed:** Authenticity disputed or verification efforts have not established result

**unknown:** Evidence of verification attempts is lacking or unknown

- ☐ verified\_fake
- ☐ verified\_real
- ☐ authenticity\_disputed
- ☐ unknown

Share links and brief context about external verification efforts or indicate "NA" \*

Is there evidence that it is fake (the source, content, markings like blurred faces, hands, writing)?

Explain or indicate relevant source/s. If not, indicate "NA."

Is there a watermark, meta-data, or other labeling effort used to indicate the provenance of the content? \*

- ☐ yes
- ☐ no
- ☐ unclear\_unknown

How many likes does the post have? Enter "NA" if there is no way to "like" the post on the platform. \*

How many views does the post have? Enter "NA" if there is no way to see the "views" on the platform. \*

How many comments does the post have? Enter "NA" if there is no way to "comment" on the post on the platform. \*

How many shares does the post have? Enter "NA" if there is no way to "share" the post on the platform. \*

What is the original source's name? (indicate NA if unknown) \*

Is the original source (the creator of the deepfake) an individual, group, or unknown?

- ☐ individual
- ☐ group
- ☐ unclear\_unknown

What type of actor is the original source? \*

political\_leader\_figure: Political Leader or Figure

government\_entity: Government Entity (not individual)

private\_ngo: Private or Non-Governmental Organization

unclear\_unknown: Unclear/unknown

- ☐ political\_leader\_figure
- ☐ government\_entity
- ☐ private\_ngo
- ☐ unclear\_unknown
- ☐ Unclear/unknown
- ☐ Political Leader or Figure
- ☐ Private or Non-Governmental Organization
- ☐ Government Entity (not individual)

From what country is the original source? (indicate NA if unknown) \*

Optional: Do you have comments you would like to add about how you went about coding the above information?

(e.g., did you look at the user's profile)

What is the sharer's name? \*

Is the sharer any of the following? Check all that apply. \*

A high-impact user is someone with more than 100,000 followers

- ☐ non-identifiable
- ☐ individual
- ☐ group
- ☐ bot
- ☐ high\_impact\_user

What is the sharer's job? Check all that apply. \*

- ☐ politician
- ☐ public\_figure
- ☐ random\_user
- ☐ journalist
- ☐ researcher\_academic
- ☐ other
- ☐ unclear\_unknown

If you selected high impact user, why did you select that option? (e.g., indicate follower count)

From what country is the sharer? (indicate NA if unknown) \*

From what city is the sharer? (indicate NA if unknown) \*

Optional: Do you have comments you would like to add about how you went about coding this section?

What is the target entity's name? \*

What is the type of the target of the deepfake? Check all that apply. \*

- ☐ individual
- ☐ group
- ☐ unclear\_unknown

What is the type of the target of the deepfake? Check all that apply. \*

- ☐ politician
- ☐ public\_figure
- ☐ journalist
- ☐ random\_user
- ☐ government\_entity
- ☐ private
- ☐ ngo
- ☐ marginalized\_group
- ☐ unclear\_unknown

What is the target's country? \*

Did the target entity respond to the sharer's post? This should be a response directly in the thread/replies. If yes, check the box.

☐

Optional: Do you have comments you would like to add about how you went about coding the information about the target?

What content does the deepfake depict or closely pertain to? Check all that apply. \*

Scandal - The deepfake depicts or pertains to public controversies or scandals involving individuals or organizations.

Crime - The deepfake depicts or pertains to criminal activities including blackmail, extortion, fraud, or impersonation.

Sex Related - The deepfake depicts or pertains to explicit or sexually suggestive content.

Crisis Event - The deepfake depicts or pertains to a crisis event. Crises can include natural disasters, terrorist attacks, or public health emergencies.

Election - The deepfake depicts or pertains to electoral or other democratic processes.

Government or Legal System - The deepfake depicts or pertains to events related to legal proceedings, government hearings, or other non-election concepts.

War or Violence - The deepfake depicts or pertains to information related to war, terrorism, or military conflict.

International Relations - The deepfake depicts or pertains to events relevant to relations between different countries or to their interactions.

Marginalized Group - The deepfake depicts or pertains to individuals or communities facing social, economic, or political marginalization.

Deepfake Regulation - The deepfake depicts or pertains to information related to legal, policy, or ethical considerations, including those pertaining to regulating the creation, distribution, or detection of deepfakes.

- ☐ scandal
- ☐ crime
- ☐ sex\_related
- ☐ crisis\_event
- ☐ election
- ☐ government\_legal\_system
- ☐ war\_violence
- ☐ international\_relations
- ☐ marginalized\_group
- ☐ deepfake\_regulation

What type of harm does the deepfake depict, if any? Check all that apply. \*

Property Harm - The deepfake depicts or pertains to the destruction of public or private property.

Environmental Harm - The deepfake depicts or pertains to environmental harms or degradation.

Violence - The deepfake depicts or pertains to violence or aggression toward individuals or groups.

Human Rights - The deepfake depicts or pertains to the violation of a fundamental human right (including the right to privacy, the right to autonomy, or the right to freedom of expression).

Discrimination - The deepfake either depicts or pertains to information about minority groups, such information detrimental to those groups, symbols of hate, or depiction of hate crimes or violence toward protected groups.

- ☐ property
- ☐ environmental
- ☐ violence
- ☐ human\_rights
- ☐ discrimination
- ☐ other
- ☐ non-identifiable

Optional: Do you have comments you would like to add about how you went about coding the above information?

The above content is the core content or topic. You can talk about any remaining uncertainties.



What is the social policy sector of the deepfake? Check all that apply. \*

- ☐ medicine\_life\_sciences
- ☐ agriculture\_resource\_extraction
- ☐ transporation
- ☐ manufacturing\_process\_automation
- ☐ consumer\_goods
- ☐ energy\_utilities
- ☐ construction\_field\_services
- ☐ security\_biometrics
- ☐ finance
- ☐ sales\_retail\_customer\_relations
- ☐ business\_services\_analytics
- ☐ broadcasting\_media\_production
- ☐ networking\_telecommunications
- ☐ arts\_sports\_leisure\_travel\_lifestyle
- ☐ education
- ☐ military\_public\_safety\_govt
- ☐ other
- ☐ unclear

Optional: Do you have comments you would like to add about how you went about coding the above information (the social policy sector)?

If available, share information about the context of the deepfake?

(i.e., consider the moment, content surrounding it, replies, what is going on in the world)

If there is relevant text surrounding the deepfake, please copy and paste it here for additional context.

(e.g., text in the tweet/post/thread you are coding)

Is there evidence or reporting of harms tied to the deepfake, even if prospective or disputed? Check all that apply. \*

**Human Rights** - There is evidence tying the deepfake to a violation of one or more fundamental human rights, such as privacy, dignity, and freedom of expression.

**Violence** - There is evidence tying the deepfake to an act of violence or aggression against an individual or group.

**Property** - There is evidence tying the deepfake to the destruction of private or public property.

**Financial Loss** - There is evidence tying the deepfake to financial harm or loss of income or livelihood.

**Environment** - There is evidence tying the deepfake to environmental degradation.

**Political Interference** - There is evidence tying the deepfake to election interference, or interference with other political processes.

**Discrimination** - There is evidence tying the deepfake to discrimination or harm based on race, ethnicity, gender, religion, or other characteristics, exacerbating social divisions and contributing to prejudice and bias.

**Other** - There is evidence tying the deepfake to a harm that does not fit neatly into the above categories. That might be because the harm is intangible or it might not be neatly categorizable.

- ☐ human\_rights
- ☐ violence
- ☐ property
- ☐ financial\_loss
- ☐ environment
- ☐ political\_interference
- ☐ discrimination
- ☐ other
- ☐ non\_identifiable

Optional: Do you have comments you would like to add about how you went about coding the above information?

The above information is the harm that could be caused.

What is the communication goal of the sharer? Check all that apply. \*

Satire - The deepfake is shared to mock or criticize individuals, events, or societal issues in a humorous or ironic manner, often offering social commentary.

False Information - The deepfake is shared to deceive an audience or individual, e.g., misinformation or disinformation.

Diversion/Flooding - The deepfake is shared meant to overwhelm media platforms with a large volume of deepfakes meant to distract or confuse the public and/or make it difficult to decipher between true and fake content.

Education - The deepfake was shared to inform and instruct audiences about various subjects, technologies, or historical events, they may be used for instructional purposes.

Entertainment - The deepfake is shared for amusement, enjoyment, or artistic purposes.

Activism - The person sharing the deepfake does so to advocate for a cause, raise awareness about social or political issues, or mobilize support for specific movements or initiatives.

Harm, property - The deepfake is shared to glorify the destruction of private or public property.

Harm, political interference - The deepfake is shared for the sake of interfering with an election or another political process.

Harm, environment - The deepfake is shared to glorify or cause environmental degradation.

Harm, violence - The deepfake is shared to glorify violence or is meant to encourage violence against an individual or group.

Harm, sexual violence - The deepfake is shared to glorify rape, sexual assault, or sexual harassment, or is meant to normalize or encourage such or intends to harm a person by portraying that person in a sexual manner without her consent.

Harm, human rights - The deepfake is shared to encourage or normalize violation of one or more fundamental human rights, such as privacy, dignity, and freedom of expression.

Harm, emotional/psychological - The deepfake is shared to induce fear, panic, anger, or concern among the public.

Harm, reputation - The deepfake is intended to damage the reputation, standing, following, credibility, or character of the target individual or group.

Harm, other - The deepfake is shared to cause harm to an individual or group, but that harm is either intangible or does not neatly fit into the above categories.

- ☐ satire
- ☐ false\_info
- ☐ entertainment
- ☐ activism
- ☐ education
- ☐ diversion\_flooding
- ☐ harm\_property
- ☐ harm\_political\_interference
- ☐ harm\_environment
- ☐ harm\_violence
- ☐ harm\_sexual\_violence
- ☐ harm\_human\_rights
- ☐ harm\_emotional\_psychological
- ☐ harm\_reputation
- ☐ harm\_other
- ☐ non\_identifiable

Optional: Do you have comments you would like to add about how you went about coding the above information (the communication goal)?

What is the core frame of the deepfake? Check all that apply. \*

Human Interest - The deepfake beings a human face or an emotional angle to the presentation of an event, issue, or problem.

Conflict - The deepfake emphasizes conflict between individuals, groups, or institutions.

Economic Consequences - The deepfake reports an event, problem, or issue in terms of the consequences it will have economically on an individual, group, institution, region, or country.

Morality - The deepfake puts the event, problem, or issue in the context of religious tenets or moral prescriptions.

Responsibility - The deepfake presents an issue or problem in such a way as to attribute responsibility for its cause or solution to either the government or to an individual/group.

☐ human\_interest

☐ conflict

☐ economic\_consequences

☐ morality

☐ responsibility

Optional: Do you have comments you would like to add about how you went about coding the above information (the framing)?

Who was/is the hero (the protagonist or desired outcome of a policy)? \*

Who was/is the villain (the obstacle or challenge that the hero must overcome)? \*

What is the plot (the narrative that connects the hero and villain)? \*

What is the moral (the lesson or value to convey to the audience)? \*

Optional: Do you have comments you would like to add about how you went about coding the above information (the policy narrative)?