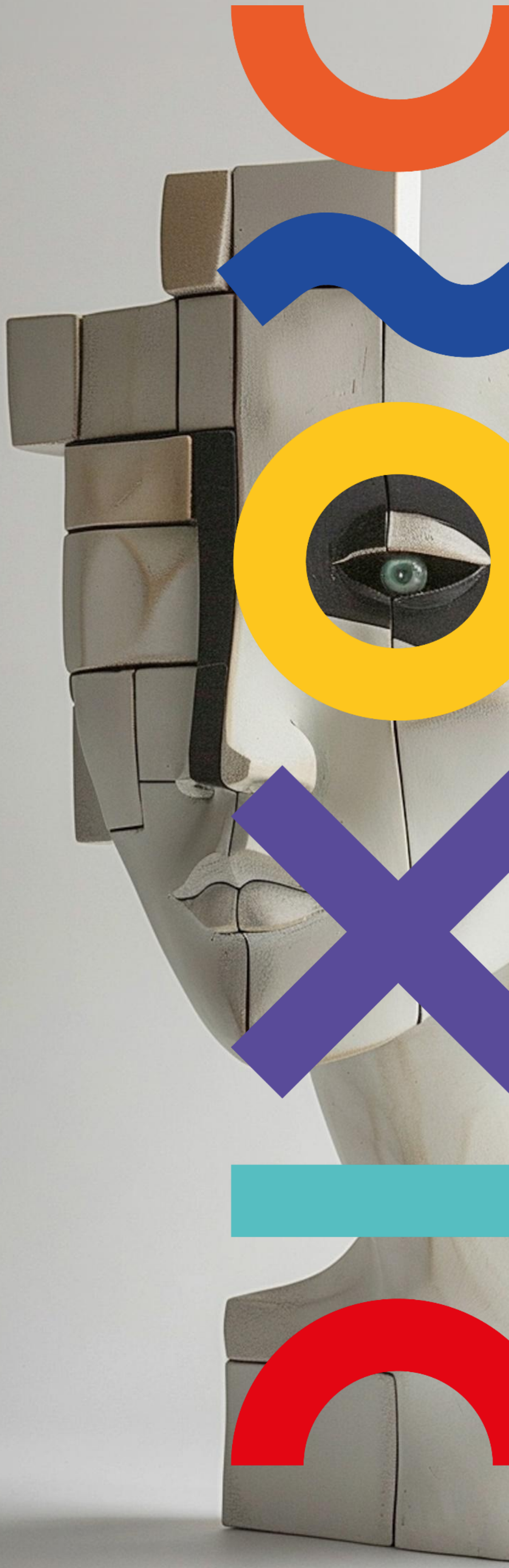
D1.2

Data Management Plan



D1.2 Data Management Plan

Dissemination Level: PU - Public
Lead Partner: ASM
Due date: 30.09.2024
Actual submission date: 30.09.2024
Actual submission date after review: 09.01.2026

PUBLISHED IN THE FRAMEWORK OF

ENCODE - Unveiling emotional dimensions of politics to foster European democracy consumers

AUTHORS

Małgorzata Walczak – Gomuła, ASM
Joanna Syrda, ASM
Łukasz Wilczyński, ASM

REVISION AND HISTORY CHART

VERSION	DATE	EDITORS		COMMENT
0.1	03.07.2024	Małgorzata Walczak – Gomuła, Joanna Syrda, Łukasz Wilczyński	ASM	Draft
0.2	10.09.2024	Agnieszka Kowalska	ASM	Internal review
0.3	19.09.2024	Łukasz Wilczyński	ASM	Second draft sent for review
0.4	26.09.2024	Seyma Celem	RIE	Review
0.5	27.09.2024	Łukasz Wilczyński	ASM	Update according to comments
1.0	30.09.2024	Aleksandra Oleksik	ASM	Submission to Participant Portal
1.1	09.01.2026	Agnieszka Kowalska, Łukasz Wilczyński	ASM	Update according to feedback provided in the 1st review meeting and ENCODE - Review Report
2.0	09.01.2026	Aleksandra Oleksik	ASM	Submission to Participant Portal

Funded by the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

DISCLAIMER

The information in this document is subject to change without notice. Company or product names mentioned in this document may be trademarks or registered trademarks of their respective companies.

All rights reserved

The document is proprietary of the ENCODE consortium members. No copying or distributing, in any form or by any means, is allowed without the prior written agreement of the owner of the property rights.
This document reflects only the authors' view. The European Community is not liable for any use that may be made of the information contained herein. Responsibility for the information and views expressed in the therein lies entirely with the author(s).

TABLE OF CONTENTS

Funded by the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

EXECUTIVE SUMMARY

The Data Management Plan (DMP) for the ENCODE project defines an operational framework for managing research data throughout its entire lifecycle, from collection and processing to storage, sharing, long-term preservation, and destruction. The DMP sets out concrete procedures for handling different categories of data in compliance with EU and national legislation, in particular the General Data Protection Regulation (GDPR), and in line with Horizon Europe requirements. It serves as a key instrument to ensure transparency, accountability, data security, and responsible data reuse across the consortium.

ENCODE generates different kinds of data, including social media-derived analytical outputs (like sentiment analysis), survey data, interview materials, biometric data, experimental data, and policy-related datasets. These data are collected using qualitative and quantitative methodologies to support the project's objective of analysing the emotional dimensions of political engagement and democratic processes.

The DMP operationalises the FAIR (Findable, Accessible, Interoperable, and Reusable) principles through project-specific procedures. Anonymised, aggregated, and derived datasets are made openly available via recognised research data repositories supporting secure HTTPS access, while raw or sensitive data (e.g. personal, biometric, or platform-restricted social media data) are stored securely on institutional servers and shared only under restricted conditions where justified. Clear decision rules govern data accessibility, following the principle "as open as possible, as closed as necessary".

Data storage arrangements are defined according to data type, sensitivity, and intended use, combining secure institutional servers, the ENCODE project SharePoint environment for controlled internal exchange, and trusted repositories for long-term preservation and dissemination. Backup procedures, access controls, and retention and destruction rules are specified.

The DMP is a living document and will be updated as necessary to reflect new data, methodological developments, or changes in the consortium, ensuring continued compliance and relevance throughout the project lifecycle. Through this structured approach, the DMP supports ENCODE's scientific objectives and maximises the long-term value, integrity, and impact of the project's research outputs.

Funded by the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

1 INTRODUCTION

1.1 THE ENCODE PROJECT

The ENCODE project, titled "Unveiling Emotional Dimensions of Politics to Foster European Democracy," aims to explore and decode the role of emotions in political discourse and their impact on democratic processes. Recognising that emotional appeals have significantly influenced political movements and voter behaviour, ENCODE seeks to understand the interplay between emotions, values, and identities. The project's primary goal is to create new positive narratives that can foster trust and engagement in European democratic processes, thereby counteracting the negative emotions that often dominate political discussions. Through innovative methodologies, including social media sentiment analysis, biometric research, and surveys, ENCODE aims to provide policymakers with tools and strategies to better incorporate the emotional needs of citizens into governance, ultimately enhancing democratic resilience and fostering a more inclusive political environment.

1.2 OBJECTIVES OF DELIVERABLE

The **Data Management Plan (DMP)** is a crucial aspect of effective data handling in any research project under Horizon Europe, ensuring the smooth integration and reuse of data and knowledge. In the ENCODE project, a dedicated task (Task 1.4) has been established to oversee data management. The DMP will detail the entire data lifecycle for the information collected, processed, and generated in ENCODE, particularly focusing on the handling of research data during and after the project, types of data collected, processed, or generated, methodologies and standards applied, open access data, and how the data will be maintained after the project concludes.

In line with the requirements of the Open Research Data Pilot (ORDP), one of the core objectives of this deliverable is to formalise the ENCODE project's data management strategy, while ensuring open, free-of-charge access to digital research data generated during the project[1].

While the DMP establishes core rules for data management, not all technical details can be fully specified at this stage, as data types, analytical approaches, and processing requirements are further refined during the initiation of individual work packages. For this reason, detailed data management arrangements - such as variable definitions, file structures, metadata granularity, and data processing workflows - are discussed and agreed during Work Package kick-off meetings and early implementation phases, in line with the overarching framework set out in D1.2.

1.3 STRUCTURE OF THE DOCUMENT

The deliverable is structured in the following sections:

- Chapter 1 – Introduction to the deliverable.
- Chapter 2 – Procedures for Data Management in the ENCODE project.

[1] H2020 Programme Guidelines on FAIR Data Management in Horizon 2020, Version 3.0, 26 July 2016, https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-datamgt_en.pdf

Funded by the European Union
This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

1.4 RELATION TO OTHER TASKS

Deliverable D1.2 (Data Management Plan) underpins the implementation of all scientific and technical tasks in the ENCODE project by providing a common framework for the responsible handling, storage, sharing, and reuse of research data. The procedures defined in this deliverable apply horizontally across work packages and guide data-related activities throughout the project lifecycle.

Scientific tasks involving data collection and analysis - such as social media analysis, surveys, interviews, biometric measurements, experiments, and co-creation activities - rely on the DMP to ensure compliance with ethical, legal, and FAIR principles from the outset. As work packages progress, refinements agreed at the WP level are incorporated into updates of the DMP, ensuring that it remains accurate, operational, and responsive to evolving project needs. Through this iterative approach, D1.2 supports methodological coherence and effective collaboration across tasks, while maintaining flexibility to accommodate the diverse and evolving data requirements of the ENCODE project.

Funded by the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

2 PROCEDURES FOR DATA MANAGEMENT IN ENCODE PROJECT

Data in ENCODE follow a defined lifecycle: (1) collection according to approved methodologies and ethical protocols; (2) secure storage on institutional or/and project-approved platforms; (3) processing and anonymisation; (4) internal use and analysis; (5) sharing via approved repositories with defined access conditions; and (6) long-term archiving or secure destruction.

2.1 DATA COLLECTION

To achieve its objectives, the project will collect a wide range of data using multiple complementary research methodologies, each contributing to a comprehensive analysis of emotional dynamics in political contexts. The types of data collected throughout the project include:

1. Social media data and sentiment analysis

ENCODE will analyse social media content to assess how public emotions are expressed and amplified in relation to key political issues, including the COVID-19 pandemic, the Russian invasion of Ukraine and other national and international topics. Social media data will be collected through systematically designed search queries based on language-specific keywords aligned with the project's theoretical framework. Data extraction will rely on application programming interfaces (APIs) from Twitter (X), enabling the collection of large-scale, issue-specific political discourse.

The collected data will be filtered, coded, and analysed using advanced sentiment and emotion analysis techniques. Machine learning and Natural Language Processing (NLP) methods will be applied to identify emotional expressions, dominant sentiment patterns, and key emotional drivers within political communication. By integrating sentiment analysis with comparative cross-country perspectives, the project will map how emotions circulate, intensify, and interact with democratic processes across different national contexts.

2. Survey data

ENCODE will deploy comprehensive surveys targeting European citizens, policymakers, and marginalised groups to capture a broad spectrum of emotional responses to democratic resilience and policy-making. The surveys will include validated questionnaires designed to assess the emotional and cognitive responses of individuals to political issues. In particular, the surveys will examine attitudes toward populism, conspiracy theories, and European identity.

3. Interview data

In-depth qualitative interviews will be conducted with key stakeholders, including policymakers, citizens, and activists. These interviews will explore emotions related to political events and policies, seeking to understand how emotional narratives influence decision-making processes. The interviews will be transcribed and coded to ensure the anonymity of participants while retaining the richness of emotional content.

Funded by the European Union
This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

4. Biometric data

To complement survey and interview data, biometric research will be conducted. ENCODE will use face-tracking technologies and other biometric methods to capture non-verbal emotional responses during experiments. This data will allow researchers to identify discrepancies between verbal declarations and emotional reactions, providing a deeper understanding of affective political responses.

5. Experimental data

ENCODE will run a series of experiments involving approximately 2,900 citizens from different European countries. These experiments are designed to observe emotional reactions to various political stimuli, including democratic resilience under stress, and how emotions influence political actions. The data from these experiments will be crucial in developing emotional maps and understanding how emotions shape political behaviour.

6. Heat maps and time series data

The project will produce heat maps to visualise emotional trends in relation to democratic resilience. These maps will be based on the survey and experimental data collected over time. Additionally, time-series data will be gathered to track the evolution of emotional responses to key political events, providing insights into how emotions fluctuate during crises such as the COVID-19 pandemic and the Russian invasion of Ukraine.

7. Policy-Related Data

Finally, ENCODE will collect data related to the policymaking process, focusing on how emotions shape the interactions between citizens and policymakers. This data will help formulate new policy guidelines aimed at addressing the emotional needs of European citizens, ultimately fostering a more resilient democratic environment.

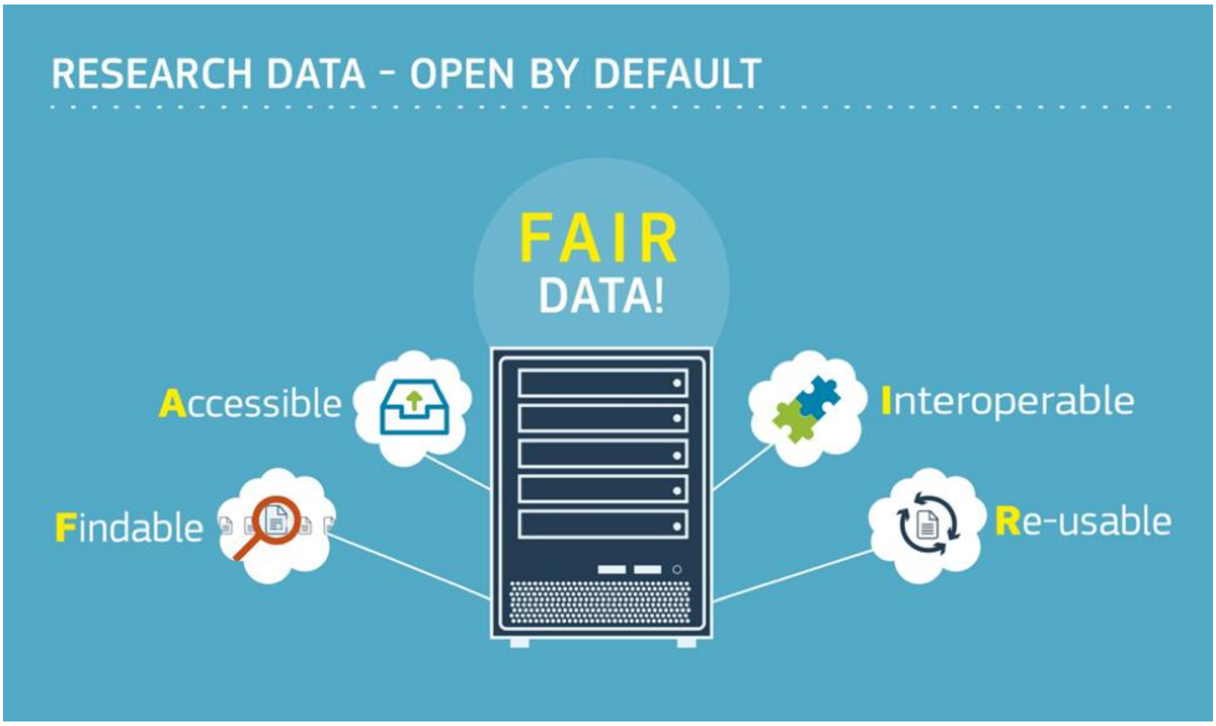
2.2 FAIR DATA PRINCIPLES

The FAIR principles provide a set of rules and criteria to enhance the discoverability, accessibility, interoperability, and reusability of digital research data. Research data management is not an objective in its own right, but rather a critical foundation for effective knowledge sharing, enabling further scientific discovery, innovation, and the meaningful integration and reuse of research data.

Funded by the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

Figure 1 FAIR Data (Source: https://www.openaire.eu/how-to-make-your-data-fair)



Compliance with FAIR principles will be verified by the Project Coordinator during deliverable review.

Below, a summary of the FAIR principles is provided[2], together with practical general guidelines for their implementation within the ENCODE project.

1. Findable

F1. (Meta)data are assigned a globally unique and persistent identifier.
F2. Data are described with rich metadata (defined by R1 below)
F3. Metadata clearly and explicitly include the identifier of the data they describe
F4. (Meta)data are registered or indexed in a searchable resource

To ensure that the data generated during the ENCODE project is findable, we will implement the following provisions:

- Data will be recorded in a predetermined structure and with agreed data formats. This aspect will be discussed at GA/WP meetings while analysing the data in different tasks. Data structure and format will ensure interoperability and ease of use.
- In line with FAIR guidelines, data will be assigned unique identifiers to enable easy identification and tracking, to be agreed upon by the project Partners.
- Each dataset generated within the project is assigned a persistent identifier.
- Selected datasets generated within ENCODE will be deposited in a repository approved by the Partners upon finalisation (after quality assurance) or, at the latest, by the end of the project. A general overview of the datasets generated is provided in Table 1 below. A more detailed list of datasets to be deposited will be identified for

[2] *Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18.*

Funded by the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

each Work Package (WP) at the start of the respective WP and updated upon its completion.

2. Accessible

A1. (Meta)data are retrievable by their identifier using a standardised communications protocol

A1.1 The protocol is open, free, and universally implementable

A1.2 The protocol allows for an authentication and authorisation procedure, where necessary

A2. Metadata are accessible, even when the data are no longer available.

The data generated by the ENCODE project will be made accessible through several channels, ensuring broad dissemination and visibility. Primarily, the ENCODE project website will serve as a central hub for accessing the data. In addition, a variety of promotion activities are planned throughout the project's duration, including newsletters, the sharing of results at events and conferences, presentations at project meetings, and dedicated webinars. The project will also produce videos and brochures to communicate findings more effectively, along with professional and scientific publications aimed at both academic and industry audiences. All publications of ENCODE will be published as open access.

Throughout the project, ENCODE partners will actively explore different open repositories, supporting standard access protocols (HTTPS), where the research data can be deposited, further enhancing accessibility and long-term availability of the data. These repositories could include platforms like Zenodo, Figshare, or institutional repositories, ensuring compliance with FAIR (Findable, Accessible, Interoperable, and Reusable) principles.

Regarding the specific research focus on politics and emotions, the ENCODE project presents several opportunities for publishing and sharing results. As an example:

- Academic journals specializing in political science, such as *Political Psychology*, *Journal of Politics*, and *European Journal of Political Research*.
- Journals focused on emotions in social and behavioural sciences, like *Emotion*, *Journal of Experimental Social Psychology*, and *Emotion Review*.
- Conferences such as the *International Society of Political Psychology (ISPP)*, *American Political Science Association (APSA)*, and *European Consortium for Political Research (ECPR)*.
- Open access repositories like *SSRN* for preprints, or discipline-specific platforms like *OpenEmotions* for interdisciplinary research on emotions and politics.

Table 1 below presents the datasets that will be made available in ENCODE.

Table 1 Overview of datasets to be generated in ENCODE project

Data type	Openly available	Additional information and justification
Social Media data	No	Raw database obtained with X is restricted; however, analysis performed based on these data will be openly available (see below Sentiment and Emotion Analysis Data).
Survey data	Yes	Fully anonymised survey datasets (CSV), Questionnaires and variable codebooks, Sampling and weighting documentation. Survey data will be anonymised to remove

Funded by the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

		any personal identifiers, allowing open sharing in line with GDPR and FAIR principles.
Interview data	Yes	Thematic coding frameworks, analytical outputs, and metadata describing interview context and methodology will be made available without risking participant identification.
Biometric data	Yes	Aggregated and processed biometric indicators, statistical summaries and visualisations (heatmaps and/or timestamps), methodological descriptions of biometric measurements.
Experimental data	Yes	Anonymised experimental datasets, experimental stimuli (where permitted), protocols and analytical outputs. Experimental data will be anonymised and documented to allow reuse and replication while protecting participants' identities.
Heat Maps and Time Series Data	Yes	Aggregated heat maps and time-series datasets, visual outputs. These datasets are produced through aggregation and do not contain personal data, making them suitable for open dissemination.
Sentiment and Emotion Analysis Data	Yes	Aggregated datasets (counts, frequencies), codebooks and methodological documentation.
Policy-related data	Yes	Policy briefs, roadmaps and guidelines and synthesised policy-relevant datasets. Policy-related outputs are designed for dissemination and reuse by policymakers, stakeholders, and the wider public.

Moreover, access conditions will follow the principle *"as open as possible, as closed as necessary"*. Fully anonymised datasets and aggregated results will be made openly available. Data containing personal, biometric, or potentially identifiable information will be either anonymised before sharing or retained under restricted access where anonymisation is not feasible. Decisions on access level will be taken jointly by the data-producing partner and the Project Coordinator, in consultation with the Data Protection Officer, and are documented in the dataset metadata.

3. Interoperable

I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (Meta)data use vocabularies that follow FAIR principles
I3. (Meta)data include qualified references to other (meta)data

In ENCODE, data will be stored in a non-proprietary and widely used format wherever possible (e.g. CSV, TXT, JSON, MP4). Controlled vocabularies and harmonised variable naming are applied across work packages, as agreed in WP/task meetings. In addition, documentation describing data structure, variables, and processing steps is provided alongside each dataset to facilitate integration with other ENCODE data and external research. The goal is to implement an approach to data exchange and reuse across diverse applications.

Funded by the European Union
This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

4. Reusable

R1. (Meta)data are richly described with a plurality of accurate and relevant attributes
R1.1. (Meta)data are released with a clear and accessible data usage license
R1.2. (Meta)data are associated with detailed provenance
R1.3. (Meta)data meet domain-relevant community standards

In ENCODE, datasets will be accompanied by clear documentation on data provenance, collection methods, processing steps, and limitations. To permit the widest re-use of data, licences will be assigned at the dataset level, typically using Creative Commons licences (e.g. CC BY or CC BY-NC), unless restrictions apply.

CC BY[3]

This license allows others to distribute, remix, and build upon the data, as long as they credit the source. The license allows for commercial use. CC BY includes the following elements:

BY: credit must be given to the creator.

CC BY-NC[4]

Under this licence, the material may be reused, adapted, and shared for non-commercial purposes, in any medium or format, as long as the original creator is properly credited. The CC BY-NC licence includes the following elements:

BY: credit must be given to the creator.

NC: Only non-commercial uses of the work are permitted.

Regarding the availability of data, the following content-specific reports have been designated with a "public" status. As a result, these reports will be actively promoted and made accessible to a broad audience. This ensures that key findings and insights generated throughout the project are disseminated widely, fostering transparency, knowledge sharing, and collaboration within the broader research community and among stakeholders. Additionally, these publicly available reports will contribute to the project's goal of enhancing the visibility and impact of its outcomes, ensuring that the data and insights produced are utilised for further research, development, and practical applications.

Table 2 Public content-related deliverables

Del No	Title	Nature	Dissemination level
D2.1	Key ENCODE's concepts and their intersections	R	PU
D2.2	Theories of emotional politics	R	PU
D2.3	Emotion-related drivers of politics	R	PU
D3.1	Overview of the state-of the art	R	PU
D3.2	Detailed methodology of the social networking analyses	R	PU
D3.3	Sentiment analysis	R	PU

[3] About CC Licenses - Creative Commons
[4] Ibidem

D3.4	Catalogue of best practices	R	PU
D4.1	Methodology for the elicitation of emotions	R	PU
D4.2	Generating emotional responses	R	PU
D4.3	Emotional maps	R	PU
D5.1	Emotions and cognitive and learning effects	R	PU
D5.2	Experimentally-validated survey questions	R	PU
D5.3	Proof of concept: democratic resilience	R	PU
D5.4	Emotions and threats to democracy	R	PU
D5.5	Emotions and mobilisation	R	PU
D5.6	Emotions, gender, and intersectionality	R	PU
D6.1	Co-Creation Methodology	R	PU
D6.2	Co-Creation Report	R	PU
D6.5	Summary report of the co-creation evaluation	R	PU
D6.3	Anonymized dataset	DATA	PU
D6.4	Handbook of emotional politics of the future narratives	R	PU
D7.1	Workshops including guidelines and materials	R	PU
D7.2	Policy brief based on the 6 workshops	R	PU
D7.3	Future scenarios and the roadmap for policy recommendations	R	PU
D7.4	Report on Creating the Alumni group	R	PU

2.3 INFORMATION FOR THE PROCEDURES FOR DATA PROTECTION

In the ENCODE project, data protection is ensured through a combination of robust security measures. Databases containing personal data, such as email addresses, will be secured with strong passwords. Access to these databases is restricted exclusively to authorised ENCODE researchers and project managers. Additionally, any results from interviews will be stored under password protection, with responses coded to anonymise identities. The mapping between names and codes will be maintained by the project partner responsible for analysis in a secure offline environment or on a digital device (MS SharePoint).

All computers storing project data will be safeguarded by login credentials and password protection. These devices will be further protected with up-to-date antivirus software and an activated firewall. For additional security, after periods of inactivity, a screen saver with password reactivation will be enabled to prevent unauthorised access. This requirement applies to all relevant partners.

Beyond these measures, the ENCODE project follows a layered security approach that incorporates encryption, access controls, and continuous monitoring to protect personal data. Data is encrypted both at rest and during transmission, ensuring that unauthorised access is effectively blocked. Access to data is role-based, meaning that only those with explicit authorisation can view or process specific categories of information. To enhance security further, regular audits and assessments are conducted to identify and address potential vulnerabilities. Moreover, as mentioned several times, compliance with GDPR regulations is integral to the project, particularly regarding the rights of data subjects, such as access, rectification, and erasure.

2.4 INFORMATION FOR THE PROCEDURES FOR DATA STORAGE

Data storage in the ENCODE project is organised to ensure data security, integrity, accessibility, and compliance with GDPR and Horizon Europe requirements throughout the data lifecycle. ENCODE applies a differentiated storage approach depending on data type:

- Secure institutional servers are used for raw and sensitive data during the research phase.
- The ENCODE project collaborative SharePoint platform is used for controlled internal data exchange between partners.
- Open research data repositories will be used for long-term storage and dissemination of anonymised, aggregated, or derived datasets intended for public access.

In terms of Partners' responsibilities, data-producing partners are responsible for the secure storage of data during collection and analysis and oversee compliance with the Data Management Plan. They also register the data in open repositories for long-term storage and post-project availability. Whereas, the Project Coordinator is responsible for the secure storage of data in ENCODE SharePoint and verifies that appropriate storage solutions are used across the consortium.

2.4.1 STORAGE DURING DATA COLLECTION AND ANALYSIS

During data collection and analysis, research data are stored on secure institutional servers operated by the respective consortium partners and, where appropriate, on the ENCODE project's secure SharePoint environment managed by ASM. These storage environments implement technical and organisational safeguards, including access control and regular system maintenance.

- Personal and sensitive data (e.g. survey responses, interview transcripts, biometric data) are stored on secured servers with access restricted to authorised project staff only.
- Raw social media data obtained via platform APIs are stored in compliance with platform terms of service on the PBY server and are accessible exclusively to authorised researchers involved in the analysis (PBY).
- Working datasets used for analysis are stored in secure project environments (e.g. institutional servers for internal analysis or protected ENCODE SharePoint space to be used by all project Partners) with access rights.
- Project deliverables (draft and final versions) are stored internally on the ENCODE project SharePoint during preparation and review. Final, submitted deliverables are additionally uploaded to the European Commission Funding & Tenders Portal, which serves as the official contractual repository. Public deliverables and dissemination materials approved for open access are made available via the ENCODE project website and, where applicable, through open research repositories (e.g. Zenodo) to ensure long-term accessibility and visibility.

All partners ensure that access to stored data is limited to personnel with a legitimate research role and that storage systems are protected by strong authentication mechanisms. To prevent data loss and ensure continuity of research activities, all storage systems used in ENCODE implement regular automated backup procedures according to institutional policies, which are stored in secure environments. Responsibility for backup implementation lies with the data-hosting partner.

Funded by the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

2.4.2 LONG-TERM STORAGE AND POST-PROJECT AVAILABILITY

After completion of the project, datasets designated for preservation are transferred to trusted long-term repositories that ensure persistent access, stable identifiers (DOIs), and compliance with FAIR principles. These datasets remain accessible via secure HTTPS connections.

Sensitive data that cannot be made publicly available is either:

- retained in secure institutional archives with restricted access, or
- destroyed in accordance with the procedures described in Section 2.5.

Retention periods are defined based on ethical approvals, legal requirements, and scientific relevance.

2.5 INFORMATION FOR THE PROCEDURES FOR DATA DESTRUCTION

At the conclusion of the ENCODE project, or when data is no longer needed, a comprehensive process for data removal will be implemented. Databases containing organisational information will be deleted either when partners decide or, at the latest, at the project's end. Since these databases do not include personal data, rigorous procedures are not required. Contact databases will be securely deleted after the project's conclusion. For files containing interview responses, responses will be coded to prevent the identification of companies or organisations. To maintain methodological credibility and allow scientific verification, these coded responses will be stored for an extended period. However, the list connecting specific organisations/respondents to their codes will be deleted once the research phase is complete, or at the latest, by the project's end.

When data is transmitted via email, it will be saved, digitised, and then removed from mailboxes to prevent unauthorised access. In alignment with ENCODE's data management practices, any remaining records will be securely destroyed using recognised methods such as cryptographic erasure for digital data and physical destruction for paper records. In cases where data must be archived, strict access controls will ensure that only authorised personnel have access. Moreover, participants retain the right to request the deletion of their data, which will result in immediate and irreversible removal from all storage systems.

Funded by the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

2.6 DATA PROTECTION OFFICER

The Data Protection Officer (DPO), Katarzyna Walburg, has been appointed from the Project Coordinator's organisation. Her contact details (iod@asmresearch.pl) will be made available to all data subjects involved in the research via the project website. The DPO is responsible for several key functions, including informing project staff of their rights and obligations related to data protection, ensuring that the ENCODE project complies with GDPR and relevant national regulations when processing personal data, and investigating any data protection concerns that arise. The DPO's role also involves monitoring data processing activities, providing guidance on data protection impact assessments, and serving as the primary point of contact for both data subjects and regulatory authorities.

In the event of a data breach, a predefined procedure will be initiated to evaluate and resolve the incident, focusing on assessing risks to the rights and freedoms of affected individuals. Any security incident posing a high risk will be promptly reported to the DPO, who will take all necessary actions to minimize potential harm. Affected individuals will receive timely email notifications detailing the nature of the breach, the type of information compromised, and the steps being taken to address the issue. Additionally, the DPO may offer training materials for project members and stakeholders to ensure ongoing compliance with data protection standards, while also maintaining a clear and efficient channel for reporting any data breaches and implementing corrective measures.

Funded by the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

ACRONYM	FULL NAME
CA	Consortium Agreement
CSV	Comma-Separated Values
D	Deliverable
DMP	Data Management Plan
DoA	Description of Action
DOI	Digital Object Identifier
DPO	Data Protection Officer
EC	European Commission
EU	European Union
FAIR	Findable, Accessible, Interoperable, Reusable
GA	Grant Agreement
GDPR	General Data Protection Regulation
HTTPS	Hypertext Transfer Protocol Secure
JSON	JavaScript Object Notation
M	Month
MP4	MPEG-4 Part 14 (digital multimedia format)
NLP	Natural Language Processing
ORDP	Open Research Data Pilot
PC	Project Coordinator
PM	Person month
WP	Work Package

Funded by the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).

ENCODE

Funded by
the European Union

This project has received funding from the European Union under the Horizon Europe Research & Innovation Programme (Grant Agreement no. 101132698 ENCODE).