Project Definition Document

Module: IN3007 – BSc Computer Science Individual Project (2025–26)

Student Name: Thi Minh Anh Vo

Student ID: 230044396

Supervisor / Consultant: Michal Krol

Project Title: Investigating YouTube Self-censorship practices.

Word Count: 1631 words

---

1. Overview and Purpose

YouTube's monetisation system enforces compliance with advertising policies by removing or restricting revenue. Rather than deleting videos, the platform limits ads or visibility, directly affecting creators' income.

The demonetisation process is opaque: explanations are rare and appeals are automated, prompting creators to form "folk theories" about restricted topics. To remain monetisable, many creators may adjust tone or avoid sensitive issues, creating a chilling effect beyond official policy.

This project investigates how monetisation influences creator language and engagement. Using public data from a carefully selected channel as a pilot case study, it aims to:

1. *Measure the extent of demonetisation using observable signals such as ad presence.*
2. *Quantify possible self-censorship by comparing linguistic and engagement patterns between monetised and demonetised videos.*

This overview remains conceptual, and implementation details are presented later in the Methodology section.

2. Problem Definition and Beneficiaries

2.1 Research Focus

The project explores the relationship between YouTube's monetisation rules and the observable linguistic or engagement changes made by creators. It examines whether measurable shifts in language or topic choice constitute a form of self-censorship driven by monetisation pressures.

2.2 Research Questions

- **RQ1:** How does the monetisation process influence creators' language, tone, and audience engagement?
- **RQ2:** How do these linguistic and monetisation patterns reflect public and creator perceptions of YouTube's demonetisation policies?
- **RQ3:** To what extent do demonetisation practices occur within the sampled channel(s), and how consistently are they applied?

2.3 Relevance and Beneficiaries

The study informs debates on algorithmic transparency and digital governance by evidencing how financial incentives shape online speech.

- **Academia:** Provides a reproducible framework for studying how algorithms affect language and monetisation.
- **Creators:** Informs language and topic choices while preserving authenticity and income.
- **YouTube & policymakers:** Supports fairer moderation and clearer communication of demonetisation rules.
- **Public:** Promotes transparency and trust in online information.

3. Objectives

Each objective is linked to the research questions: O1–O4 address RQ3, O4–O5 address RQ1, and O5 addresses RQ2.

ID	Objective
O1	Collect a bounded, verifiable sample of public YouTube data (metadata, transcripts, ad presence, and comments) from one pilot channel.
O2	Organise and document the dataset in a reproducible structure for independent validation.
O3	Develop a transparent rule-based classifier labelling videos as Likely Monetised, Likely Demonetised, or Uncertain using deterministic thresholds.

ID	Objective
O4	Test whether sensitive-term ratios and topics differ systematically between Likely Monetised and Likely Demonetised videos, using descriptive statistics and correlation analysis.
O5	Visualise the proportions of monetisation classes, distributions of sensitive-term ratios, and their relationships with engagement metrics over time.
O6	Interpret results ethically, acknowledging limitations and avoiding causal claims.

4. Methodology.

The analysis will follow a transparent, reproducible Python workflow using open-source tools. The pilot examines one English-language current-affairs channel where sensitive topics are common and may later extend to up to three channels subject to time and ethics approval.

Step 1 – Data gathering
Use the YouTube Data API v3 to collect public metadata (title, description, upload date, views), captions, and a bounded number of comments (e.g. first *N* top-level entries).
Record observable ad presence (pre-, mid-, or post-roll) using a consistent protocol.
All data will be stored in a reproducible folder structure organised by channel and video ID.

Step 2 – Pre-processing
Clean caption text (remove markup, normalise case and punctuation), then tokenise and lemmatise using NLP libraries such as NLTK or spaCy**.**
Create a manually reviewed lexicon of sensitive or advertiser-restricted terms (e.g. war-related or explicit language) and prepare text for feature extraction.

Step 3 – Feature engineering
For each video, compute:

- *Ad features*: number and type of ads (observed slots).
- *Linguistic features*: total word count, sensitive-term ratio, and broad syntactic or thematic indicators (e.g. tone or topic tags where feasible).
- *Meta features*: video duration, upload age, channel size.
- *Engagement features*: view count, like ratio, comment count, and approximate sentiment of comments.

Step 4 – Rule-Based Classification and Validation (RQ3)
Classify videos as *Likely Monetised*, *Likely Demonetised*, or *Uncertain* using ad presence and sensitive-term ratios:

- ad_count = 0 and sensitive_ratio ≥ T_1 → Likely Demonetised

- ad_count > 0 and sensitive_ratio < T_2 → Likely Monetised
- otherwise → Uncertain

Calibrate thresholds (T_1, T_2) on a small validation subset with known or observable monetisation status (e.g. visible ads or creator-confirmed cases).
Confirm consistency before applying rules to the full dataset to ensure results reflect observable reality and remain reproducible.

Step 5 – Quantitative Analysis of Demonetisation Patterns (RQ3)

Apply the validated classifier to estimate the proportion of monetised and demonetised videos.
Perform descriptive and correlation analyses linking monetisation class with engagement metrics (views, likes, comments).
Visualise distributions to illustrate the extent and consistency of demonetisation across videos and over time.

Step 6 – Linguistic and Perception Analysis (RQ1 & RQ2)

6.1 Self-Censorship Indicators (Transcripts):
Analyse transcripts for linguistic adaptation such as substitution of sensitive terms, tonal moderation, or topic avoidance.
Compare sensitive-term ratios and provide illustrative examples showing how creators adjust content between monetised and demonetised videos.

6.2 Perception Indicators (Comments):
Identify and thematically code viewer or creator comments mentioning "ads", "monetisation", "yellow icons", or "can't talk about X".
Group these into themes reflecting perceived limits on speech and relate them to observed transcript patterns.

Step 7 – Visualisation and Reporting

Generate summary tables and visualisations (bar, scatter, and time-series charts)
using **Matplotlib** and other open-source Python tools.
Summarise findings in written reports supported by visuals to explain quantitative results and qualitative themes.
Version-control all scripts and include an audit narrative detailing methodology, reporting structure, limitations, and ethical considerations.

5. Expected Outputs & Measure of Success

5.1 Expected Outputs

- Pilot dataset (≈20+ videos) containing metadata, transcripts, comments, and ad-presence records.

- Feature tables summarising linguistic, engagement, and meta metrics.
- Classifier report with defined rules, thresholds, and validation outcomes.
- Analytical summary comparing language and engagement across monetisation classes and time.
- Visual outputs showing monetisation proportions, sensitive-term patterns, and engagement correlations.
- Ethical reflection outlining key decisions, biases, and limitations.

Python is chosen for its strong open-source support and transparent data-analysis ecosystem. All code will be version-controlled in a private repository and documented for replication.

5.2 Measure of Success

Success is judged by clarity and reproducibility rather than numerical performance alone. Indicative targets:

- A cleanly structured dataset sufficient for descriptive analysis.
- A classifier showing plausible agreement with observed ads (≈ 70 % accuracy, ≈ 80 % agreement).
- Analyses revealing meaningful patterns in language and monetisation, even if inconclusive.
- A balanced discussion that reports limitations and avoids causal claims.

6. Legal, Social, Ethical and Professional Issues (LSEPI)

Uses only public YouTube data in line with API terms and UK GDPR; no personal or sensitive data are collected. The analysis is aggregate and anonymised, focused on transparency rather than individual creators.

Ethical risks are minimal: no interaction with participants, and results are reported descriptively. The project upholds the BCS Code of Conduct by ensuring integrity, honesty, and transparency.

7. Risks and Mitigation

The main risks to the project and their mitigation strategies are:

Risk	Likelihood	Impact	Mitigation
API quota limits delay collection	Low	Medium	Plan and batch API requests; cache and reuse stored data. The dataset is small relative to YouTube's quota, so interruptions are unlikely.

Risk	Likelihood	Impact	Mitigation
Weak or noisy ad signals	Low	High	Manually verify ad presence and cross-check with public tools (e.g. ytlarge.com). Repeated checks minimise misclassification from ad rotation.
Bias from single-channel focus	Medium	Medium	Start with one channel and, if time and ethics allow, extend to 1–3 to reduce single-source bias. Frame findings as comparative case studies.
Time overruns in feature work	Low	Medium	Plan ahead with detailed weekly milestones, prioritise core objectives (O1–O4) before any extensions, and monitor progress against the Gantt chart.
Misinterpretation of results	Low	High	Report correlations only; state limitations explicitly. (High impact if readers mistake correlation for causation.)

These project risks are separate from research-ethics risks, which concern potential harm to participants and are handled through the ethics checklist. As this project only uses public data and does not recruit participants, it is expected to fall into a low-risk category, subject to consultant approval.

8. Project Timeline and Coding Plan

Phase	Period	Key Activities
Setup & Ethics	Oct–Nov 2025	Ethics approval, PDD submission, select pilot channel.
Data Collection	Nov – Jan 2026	Retrieve metadata, captions, comments, and log ads,...
Pre-processing & Features	Jan – Feb 2026	Clean text, lemmatise, compute metrics.
Classifier Development	Feb 2026	Implement rule-based classifier and validate.
Analysis & Visualisation	Mar 2026	Perform descriptive and correlation analysis; generate charts.
Reporting	Apr 2026	Compile findings, prepare final report and presentation.

9. AI Use Declaration

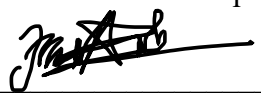
Large Language Models (LLMs) are used assistively to improve clarity and academic phrasing; debug Python scripts for data processing and visualisation; refine variable naming and chart labelling; and outline literature summaries that are manually verified.

No LLM output serves as ground truth, training data, or final analysis.
All code, results, and interpretations are independently produced and verified by the student.

This limited and transparent use complies with City, University of London's School of Science & Technology policy on the use of Generative AI in assessments and the BCS Code of Conduct (Sections 2a, 2c, 3a, 4a).

10. Declarations

Academic Integrity Declaration - I declare that this work is my own, produced in accordance with City University of London's policies and the BCS Code of Conduct. All data, code, and documentation have been independently verified.

Signature:  Date: 05/11/2025

Appendix A - Data & Labelling Protocol

Auto-labelling: detect ad presence/absence through page markers or manual observation.
Text pre-processing: remove stop-words, lemmatise tokens, compute sensitive-term ratios.
Reproducibility: record threshold values (T_1/T_2) and configuration choices; maintain version control. Optional validation may also use public monetisation checkers (e.g., ytlarge.com or captions.ai) to cross-verify ad presence.

Research Ethics Review Form for BSc and MSci Projects

Computer Science Research Ethics Committee (CSREC)
http://www.city.ac.uk/department-computer-science/research-ethics

Student Name: Thi Minh Anh Vo.

Student Number: 230044396.

Supervisor/Consultant: Michal Krol.

Project Title: Investigating YouTube Self-censorship practices.

Date: 29 October 2025.

Undergraduate students undertaking their final project in the Department of Computer Science must consider the ethics of their project work and ensure that it complies with research ethics guidelines and the law for data protection. In some cases, a project will need approval from an ethics committee before it can proceed. Usually, but not always, this will be because the student is involving other people ("participants") in the project.

To ensure that they give appropriate consideration to ethical issues, all students must complete this form and attach it to their project definition document (PDD). There are two parts:

PART A: Ethics Checklist. All students must complete this part.
The checklist identifies whether the project requires ethical approval and, if so, where to apply for approval.

PART B: Ethics Proportionate Review Form. Students who have answered "no" to all questions in A1, A2 and A3 and "yes" to question 4 in A4 in the ethics checklist must complete part B as well. The project supervisor or consultant has delegated authority to provide approval in such cases that are considered to involve MINIMAL risk. The approval may be ***provisional*** – *identifying the planned work with human end user participants as* likely to involve MINIMAL RISK. In such cases you must additionally seek ***full approval*** from the supervisor or consultant as the project progresses and details are established. You must obtain ***full approval*** in writing, before recruiting and engaging with human end users participants for your project.

A.1 If you answer YES to any of the questions in this block, your consultant/supervisor must have obtained approval for the project from an appropriate external ethics committee, and you need to have received written confirmation of this from him/her. Students cannot themselves apply for ethics approval in this case as the project is considered high risk". This type of research is not covered by City's process, and external approval from an appropriate institution is required.		*Delete as appropriate*
1.1	Does your research require approval from the National Research Ethics Service (NRES)?	**NO**
1.2	Will you recruit participants who are covered by the Mental Capacity Act 2005?	**NO**
1.3	Will you recruit any participants who are covered by the Criminal Justice System, for example, people on remand, prisoners and those on probation?	**NO**
A.2 If you answer YES to any of the questions in this block your consultant/supervisor must have obtained appropriate ethics committee approval		*Delete as appropriate*

2.1	Does your research involve participants who are unable to give informed consent? *For example, people who may have a degree of learning disability or mental health problem, that means they are unable to make an informed decision on their own behalf.*	**NO**
2.2	Is there a risk that your research might lead to disclosures from participants concerning their involvement in illegal activities?	**NO**
2.3	Is there a risk that obscene and or illegal material may need to be accessed for your research study (including online content and other material)?	**NO**
2.4	Does your project involve participants disclosing information about protected characteristics (as identified by the Equality Act 2010)? *For example: racial or ethnic origin; political opinions; religious beliefs; trade union membership; physical or mental health; sexual life; criminal offences and proceedings*	**NO**
2.5	Does your research involve you travelling to another country outside of the UK, where the Foreign & Commonwealth Office has issued a travel warning that affects the area in which you will study? *Please check the latest guidance from the FCO - http://www.fco.gov.uk/en/*	**NO**
2.6	Does your research involve invasive or intrusive procedures? *These may include, but are not limited to, electrical stimulation, heat, cold or bruising.*	**NO**
2.7	Does your research involve animals?	**NO**
2.8	Does your research involve the administration of drugs, placebos or other substances to study participants?	**NO**
A.3 If you answer YES to any of the questions in this block, then unless you are applying to an external ethics committee or the Senate Research Ethics Committee (SREC), you must apply for approval from the Computer Science Research Ethics Committee (CSREC) through Research Ethics Online - https://researchmanager.city.ac.uk/. Depending on the level of risk associated with your application, it may be referred to the Senate Research Ethics Committee (SREC).		*Delete as appropriate*
3.1	Does your research involve participants who are under the age of 18?	**NO**
3.2	Does your research involve adults who are vulnerable because of their social, psychological or medical circumstances (vulnerable adults)? *This includes adults with cognitive and / or learning disabilities, adults with physical disabilities and older people.*	**NO**
3.3	Are participants recruited because they are staff or students of City, University of London? *For example, students studying on a particular course or module.* *If yes, then approval is also required from the Head of Department or Programme Director.*	**NO**
3.4	Does your research involve intentional deception of participants?	**NO**
3.5	Does your research involve participants taking part without their informed consent?	**NO**
3.5	Is the risk posed to participants greater than that in normal working life?	**NO**
3.7	Is the risk posed to you, the researcher(s), greater than that in normal working life?	**NO**
A.4 If you answer YES to the following question and your answers to all other		*Delete as*

questions in sections A1, A2 and A3 are NO, then your project is deemed to be of MINIMAL RISK. **If this is the case, then you can apply for approval through your supervisor under PROPORTIONATE REVIEW. You do so by completing PART B of this form.** **If you have answered NO to all questions on this form, then your project does not require ethical approval. You should submit and retain this form as evidence of this.**		*appropriate*
4	Does your project involve human participants or their identifiable personal data? *For example, as interviewees, respondents to a survey or participants in testing.*	**NO**

PART B: Ethics Proportionate Review Form

If you answered YES to question 4 and NO to all other questions in sections A1, A2 and A3 in PART A of this form, then you may use PART B of this form to submit an application for a proportionate ethics review of your project. Your project supervisor has delegated authority to review and approve this application under proportionate review. You must receive final approval from your supervisor in writing before beginning the planned research.

However, if you cannot provide all the required attachments (see B.3) with your project proposal (e.g. because you have not yet written the consent forms, interview schedules etc), the approval from your supervisor will be ***provisional***. You **must** submit the missing items to your supervisor for approval prior to commencing these parts of your project. Once again, you must receive written confirmation from your supervisor that any provisional approval has been superseded by with ***full approval*** of the planned activity as detailed in the full documents. **Failure to follow this procedure and demonstrate that final approval has been achieved may result in you failing the project module and/or result in an academic misconduct investigation.**

Your supervisor may ask you to submit a full ethics application through Research Ethics Online, for instance if they are unable to approve your application, if the level of risks associated with your project change, or if you need an approval letter from the CSREC for an external organisation.

B.1 The following questions must be answered fully. All grey instructions must be removed.		*Delete as appropriate*
1.1.	Will you ensure that participants taking part in your project are fully informed about the purpose of the research?	**N/A**
1.2	Will you ensure that participants taking part in your project are fully informed about the procedures affecting them or affecting any information collected about them, including information about how the data will be used, to whom it will be disclosed, and how long it will be kept?	**N/A**
1.3	When people agree to participate in your project, will it be made clear to them that they may withdraw (i.e. not participate) at any time without any penalty?	**N/A**
1.4	Will consent be obtained from the participants in your project? Consent from participants **MUST** be obtained if you plan to involve them in your project or if you plan to use identifiable personal data from existing records. "Identifiable personal data" means data relating to a living person who might be identifiable if the record includes their name, username, student id, DNA, fingerprint, address, etc. *If YES, you must attach drafts of the participant information sheet(s) and consent form(s) that you will use in section B.3 or, in the case of an existing dataset, provide details of how consent has been obtained.* *You must also retain the completed forms for subsequent inspection. Failure to provide the completed consent request forms will result in withdrawal of any earlier ethical approval of your project.*	**N/A**
1.5	Have you made arrangements to ensure that material and/or private information obtained from or about the participating individuals will remain confidential?	**N/A**

B.2 If the answer to the following question (B2) is YES, you must provide details		*Delete as appropriate*
2	Will the research be conducted in the participant's home or other non-University location? *If **YES**, you must provide details of how your safety will be ensured.*	**N/A**

B.3 Attachments **ALL of the following documents MUST be provided to supervisors if applicable.** **All must be considered prior to final approval by supervisors.** **A written record of final approval must be provided and retained.**	***YES***	***NO***	***Not Applicable***
Details on how safety will be assured in any non-University location, including risk assessment if required (see B2)			**N/A**
Details of arrangements to ensure that material and/or private information obtained from or about the participating individuals will remain confidential (see B1.5) *Any personal data must be acquired, stored and made accessible in ways that are GDPR compliant.*			**N/A**
Full protocol for any workshops or interviews**			**N/A**
Participant information sheet(s)**			**N/A**
Consent form(s)**			**N/A**
Questionnaire(s)** *sharing a Qualtrics survey with your supervisor is recommended.*			**N/A**
Topic guide(s) for interviews and focus groups**			**N/A**
Permission from external organisations or Head of Department** *e.g. for recruitment of participants*			**N/A**

***If these items are not available at the time of submitting your project proposal, then **provisional approval** can still be given, under the condition that you must submit the final versions of all items to your supervisor for approval at a later date. **All** such items **must** be seen and approved by your supervisor before the activity for which they are needed begins. Written evidence of **final approval** of your planned activity must be acquired from your supervisor before you commence.*

Changes

If your plans change and any aspects of your research that are documented in the approval process change as a consequence, then any approval acquired is invalid. If issues addressed in Part A (the checklist) are affected, then you must complete the approval process again and establish the kind of approval that is required. If issues addressed in Part B are affected, then you must forward updated documentation to your supervisor and have received written confirmation of approval of the revised activity before proceeding.

Templates for Consent and Information

You must use the templates provided by the University as the basis for your participant information sheets and consent forms. You **must** adapt them according to the needs of your project before you submit them for consideration.

Participant Information Sheets, Consent Forms and Protocols must be consistent. Please ensure that this is the case prior to seeking approval. Failure to do so will slow down the approval process.

We strongly recommend using Qualtrics to produce digital information sheets and consent forms.

Further Information

https://www.city.ac.uk/about/governance/committees/cs-research-ethics

https://www.city.ac.uk/research/ethics/how-to-apply/participant-recruitment

https://www.city.ac.uk/research/ethics