Lorcan McLaren
PhD Researcher
Connected_Politics Lab
School of Politics and International Relations
University College Dublin
Belfield, Dublin 4, Ireland
lorcan.mclaren@ucdconnect.ie
https://lorcanmclaren.com/

POL42560 AI and Large Language Models

Spring Trimester 2026

Version: January 22, 2026

Time:	Fri, 11:00 - 12:50	Module instructor:	Lorcan McLaren
Format:	In-person lectures and labs		lorcan.mclaren@ucdconnect.ie
Credits:	10, Level 4	Office hours:	Wed, 15:00 - 16:00 (sign up)

1 Overview

Large language models (LLMs), such as those behind tools like ChatGPT, have garnered significant attention in the media for their ability to generate human-like text, sparking both enthusiasm and debate about their implications for various fields. For social scientists, LLMs offer a potentially useful approach to processing and interpreting vast amounts of data, enhancing our ability to study complex societal issues. This course offers an interdisciplinary approach to understanding and applying LLMs in social and political science, with a focus on text analysis. It combines theoretical foundations with practical, hands-on experience in applying LLMs to address substantive social science questions. Students will explore the capabilities and limitations of LLMs and engage critically with issues such as bias, environmental impact, misinformation, and intellectual property rights. They will also become familiar with essential research practices including documentation, reproducibility, and validation.

2 Learning Outcomes

By the end of the course, students will be able to:

1. Understand the fundamentals of large language models and their applications in social sciences.
2. Implement LLMs for various tasks relevant to political and social science research using Python.
3. Critically evaluate the ethical and societal implications of AI technologies, particularly LLMs, in the context of social science.
4. Apply best practices in research documentation, reproducibility, and validation when using AI tools.
5. Engage in informed discussions about the impact of AI on society, including issues of bias, sustainability, and concentration of power.

3 Prerequisites

This course is designed for students with or without experience in Python, but a willingness to learn quickly and engage with technical content is essential. It aims to equip students with the skills and knowledge needed to effectively apply AI in social science research.

If students do not have prior programming experience, they would benefit from taking POL42340 Programming for Social Scientists alongside this course to familiarise themselves with Python fundamentals.

POL42050 Quantitative Text Analysis provides a broader overview of text-as-data approaches, complementary to the content of this course.

4 Structure

Each class will involve a lecture introducing the topic, followed by a hands-on lab session where we will work on exercises related to the lecture content. Students are expected to complete the readings before class each week, and to submit their lab work via Brightspace. See Section 11 for an overview of expected content.

5 Assessment

Table 1: Assessment components

Component	Deadline	Grade %
Lab participation	Continuous	20
Group presentations	April 24	20
Group project	May 1	60

5.1 Lab participation

Each lab session includes a Jupyter notebook that must be submitted through Brightspace for participation credit. While your solutions don't need to be perfect, they should demonstrate a genuine attempt at completing all exercises. Include clear comments throughout your code to show your understanding of the concepts.

You should **execute all code cells prior to saving**, so that the output for each cell is visible in the submitted version.

You will have **one week to submit the notebook from each lab session**, up to the start of the subsequent class (e.g., the notebook for Week 2 must be submitted by 08:59 on Friday in Week 3). You may miss up to one submission without prior approval. Late submissions will not be accepted unless an extension has been granted, as exercise solutions may be discussed in class.

5.2 Group presentation

Each group must deliver a presentation lasting **no more than 12 minutes**, using either LaTeX Beamer, PowerPoint or Google Slides. Submit your slides on Brightspace **by 23:59 on Wednesday, April 22**. Groups should divide the presentation between 2-3 speakers, with non-presenting members taking responsibility for answering questions from peers and the instructor during the Q&A session.

You should prepare together by practising the presentation and anticipating potential questions. The same grade will be awarded to all group members, unless one or more members fail to participate in both the presentation and Q&A.

Your presentation should cover:

1. Research question and importance
2. Theoretical expectations and main hypothesis
3. Data description
4. Methodology for hypothesis testing
5. Initial results (if available)
6. Challenges encountered and potential risks for next steps (if applicable)

5.3 Group project

Each group must submit a **4,000-5,000 word** research paper (**excluding references and appendices**) as well as the code used to conduct the analysis. All students in a group will be awarded the same grade, but you may be asked to indicate who contributed to which parts of the project.

You have free choice in selecting your research question, but you must justify why it is relevant and important for understanding or addressing a meaningful social science question. You may either collect original data appropriate for your research question or identify and use existing datasets (see Appendix A for some suggestions), clearly explaining your data source selection. Your analysis must incorporate **at least one** of the methods covered in this course - embeddings, transformers, or generative models - in a way that meaningfully advances your research objectives.

Your final paper and code should be submitted via Brightspace in the following format:

1. A research paper in PDF format. The paper must be written in LaTeX using the collaborative Overleaf document that will be created for each group
2. A GitHub repository containing all code needed to reproduce your analysis

5.3.1 Research paper structure

The paper should contain the following sections. Further subsections can be added as necessary to organise the content. Expected content for each section is also outlined:

1. **Introduction**
    - Research question/puzzle
    - Significance
    - Main hypothesis
2. **Theory**
    - Theoretical framework
    - Prior findings on DV-IV relationship
    - Expected effects
3. **Data & Methods**
    - Dataset overview
    - Unit of analysis
    - Sample size & missing data
    - Analysis approach
4. **Results**
    - Empirical findings
    - Statistical tests
    - Key patterns
5. **Conclusion**
    - Link to research question
    - Main insights
    - Limitations
    - Broader implications

6 Grading

The group presentations and project will be graded according to the criteria below:

Table 2: Grade descriptors

Grade	**Criteria**
A+	Outstanding: An exceptionally deep and systematic engagement with the assessment task, with consistently impressive demonstration of a comprehensive mastery of the subject matter and discerning judgement, reflecting: • a deep and broad knowledge and highly-developed critical insight, as well as effective synthesis of extensive reading • a critical comprehensive and perceptive appreciation of the relevant literature or theoretical, technical or professional framework • an exceptional ability to organise, analyse and succinctly present arguments fluently and lucidly with a high level of critical analysis, supported by very convincingly deployed evidence, citation or quotation • a highly-developed capacity for original, creative and logical thinking
A	Excellent: A deep and systematic engagement with the assessment task, with consistently impressive demonstration of a comprehensive mastery of the subject matter, reflecting: • a deep and broad knowledge and critical insight as well as extensive reading • a critical and comprehensive appreciation of the relevant literature or theoretical, technical or professional framework • an exceptional ability to organise, analyse and present arguments fluently and lucidly with a high level of critical analysis, amply supported by evidence, citation or quotation • a substantial capacity for original, creative and logical thinking
B	Very Good: A substantial engagement with the assessment task, demonstrating: • a thorough familiarity with the relevant literature or theoretical, technical or professional framework • well-developed capacity to analyse issues, organise material, present arguments clearly and cogently well supported by evidence, citation or quotation • some original insights and capacity for creative and logical thinking
C	Good: An intellectually competent and factually sound answer, marked by: • evidence of a reasonable familiarity with the relevant literature or theoretical, technical or professional framework • good developed arguments, but more statements of ideas • arguments or statements adequately but not well supported by evidence, citation or quotation • some critical awareness and analytical qualities • some evidence of capacity for original and logical thinking
D	Satisfactory: An acceptable level of intellectual engagement with the assessment task showing: • some familiarity with the relevant literature or theoretical, technical or professional framework • mostly statements of ideas, with limited development of argument • limited use of evidence, citation or quotation • limited critical awareness displayed • limited evidence of capacity for original and logical thinking

Table 2: Grade descriptors

Grade	Criteria
D-	Acceptable: The minimum acceptable level of intellectual engagement with the assessment task with: • the minimum acceptable appreciation of the relevant literature or theoretical, technical or professional framework • ideas largely expressed as statements, with little or no developed or structured argument • minimum acceptable use of evidence, citation or quotation • little or no analysis or critical awareness displayed or is only partially successful • little or no demonstrated capacity for original and logical thinking
FM	Unacceptable: An unacceptable level of intellectual engagement with the assessment task, with: • no appreciation of the relevant literature or theoretical, technical or professional framework • no developed or structured argument • no use of evidence, citation or quotation • no analysis or critical awareness displayed or is only partially successful • no demonstrated capacity for original and logical thinking

7 Office Hours

I will run weekly office hours during the teaching term on Wednesdays from 15:00-16:00. If you wish to attend, either in person or online, you should book a slot via Calendly.

8 Textbooks and Readings

All mandatory readings are indicated in the course schedule and are open access or available through the UCD Library. Readings should be completed ahead of class each week.

While this course will focus on the practical use of LLMs for social science research, the following textbooks have been used to develop the content and should be consulted if you want a more technical understanding of how these models work under the hood. Some of these are available through the Library, while others are not:

8.1 Approachable

- Alammar, J. and Grootendorst, M. (2024) *Hands-On Large Language Models: Language Understanding and Generation*. 1st edition. Beijing Boston Farnham: O'Reilly Media.
- Grimmer, J., Roberts, M.E. and Stewart, B.M. (2022) *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton: Princeton University Press.

8.2 Deeper dive

- Raschka, S. (2024) *Build a Large Language Model (From Scratch)*. Manning.
- Kamath, U. et al. (2024) *Large Language Models: A Deep Dive: Bridging Theory and Practice*. 2024 edition. Cham: Springer.

9 Software

All code for this course should be written in Python. For those who haven't worked with this language previously, POL42340 Programming for Social Scientists runs in parallel to this course and provides a comprehensive introduction to Python.

We will use VS Code to write and execute code locally. For some lab sessions and the group project, where your personal device may not be powerful enough or where long execution times are expected, Google Colab should be used. Colab provides limited free access to T4 GPUs, which will enable you to run code that cannot be executed locally without a GPU[1].

We will use `venv` and `pip` to manage our Python environment and install packages (though feel free to use Anaconda if you are more familiar with this platform). I strongly encourage the use of Git for version control, particularly when working on your group project; however, we will not cover this extensively in class.

Useful Python packages for dealing with textual data:

- spaCy
- NLTK
- Gensim
- SentenceTranformers
- transformers (HuggingFace)
- LangChain

10 Policies

10.1 Plagiarism

Academic integrity is one of the core values of the UCD Education Strategy and includes adherence to the highest ethical and academic standards. Students, researchers and staff achieve academic integrity through sound academic writing, avoiding plagiarism, and use of appropriate referencing and citation.

Plagiarism is the inclusion, in any form of assessment, of material without due acknowledgement of its original source. Plagiarism is a form of academic dishonesty and may include, but is not limited to, the following:

- Presenting in your own name, work authored by a third party, such as other students, friends or family (with or without permission), or work purchased, via organisations such as essay mills. The original source may be in written form or in any other media (for example, audio or video)
- Presenting ideas, theories, concepts, methodologies or data from the work of another without due acknowledgement
- Incorrect paraphrasing, presenting text, digital work, music, video recordings or images of others with only minor changes (e.g., using synonyms or changing the sentence structure) from an original source; the inclusion of a citation does not eliminate this. Correct paraphrasing in your own words must also include appropriate citation of the original source material
- Representing collaborative work as solely your own, including colluding with or copying from others during examinations
- Presenting work for an assignment which has also been submitted (in part or whole) for another assignment at UCD or another institution (i.e., self-plagiarism)

Plagiarism can be either intentional or unintentional. In both instances it is a serious academic offence and may be subject to Student Discipline Procedure. All students are responsible for being familiar with the University's policy statement on plagiarism and are encouraged, if in doubt, to seek guidance from an academic member of staff.

[1] Running HuggingFace pipelines on Colab:
https://dev.to/ajmal_hasan/using-hugging-face-models-in-google-colab-a-beginners-35ll

10.2 AI Use

I encourage the use of generative AI tools when completing the assignments for this module, but all work relying on AI-generated content must adhere to the highest academic standards. Users of this technology must be aware of what it can and, more importantly, what it cannot do well. It is crucial for you to exercise judgement when evaluating the quality and reliability of content generated through AI platforms. AI is not a panacea for all writing challenges; it will not automatically generate a flawless, logically coherent, and factually correct assignment. Instead, use AI as a tool to tackle specific issues such as brainstorming and idea formation, literature discovery, and text drafting issues. View your preferred AI platform(s) as useful but imperfect tools that can offer inspiration, new perspectives, and supplementary areas for research for your own work. In-depth research on your part remains essential to ensure coherent, factual, and scientifically informed perspectives in your assignment. Always cross-reference the information AI offers against other independent and reliable sources.

AI use must be in line with UCD's policies on academic integrity and adhere to the highest academic standards. See here for details: https://libguides.ucd.ie/academicintegrity.

Failure to adhere to the guidelines set out below is a violation of Section 4.3 of the UCD Academic Integrity Policy.

10.2.1 Documenting AI Use

Generative AI is a novel and rapidly evolving tool in an academic context. What is impossible today might well become trivial tomorrow (keeping in mind the academic standards mentioned above). In order to address this, you are expected to provide an account of the tools used and how they were used in a mandatory appendix to your group project. This appendix will be assessed as part of each assignment, with grade points awarded for effective communication of the methods used to generate content. For each instance where a generative AI tool is used, you need to provide:

1. An in-text citation or footnote. For example:
   - "Some AI-generated text (*OpenAI*, 2023)"
2. A bibliographic reference to the tool used and the date of access
3. An entry in the mandatory AI appendix detailing how the tool was used. For example:

Table 3: Example table demonstrating how generative AI was used for an assignment

AI tool	**Explanation**	**Prompt**
Perplexity	Literature search	"Find peer-reviewed papers comparing traditional machine learning vs LLM approaches for political text classification, particularly for legislative text categorization."
ChatGPT	Structure planning	"Outline a methods section for a paper using GPT embeddings to classify political manifestos into policy areas. Include key considerations about training data, validation, and limitations."
Claude	Draft refinement	"Review this draft section on LLM-based classification of parliamentary speeches. Suggest improvements to the methodology section, focusing on model validation and bias mitigation techniques."

For code, while I do not expect you to keep a log of all prompts used, I do expect you to add comments indicating where code has fully or partially been generated by a tool such as ChatGPT or GitHub Copilot.

1. **Code generation**

```python
# GitHub Copilot generated this function to clean and preprocess text data.
# I modified the function to include stemming and fixed an issue with
# handling NaN values.
import re
from nltk.stem import PorterStemmer

def preprocess_text(text):
    """
    Cleans and preprocesses text by lowercasing, removing special characters,
    and stemming words.
    """
    ps = PorterStemmer()
    if pd.isnull(text):  # Added to handle NaN values
        return ""
    text = text.lower()
    text = re.sub(r'[^a-z\s]', '', text)  # Remove special characters
    words = text.split()
    words = [ps.stem(word) for word in words]  # Apply stemming
    return ' '.join(words)
```

2. **Debugging assistance**

```python
# I encountered an issue where the tokenizer failed on texts containing emojis.
# ChatGPT suggested using `emoji` library to remove emojis before tokenization.
import emoji
from nltk.tokenize import word_tokenize

def tokenize_text(text):
    """
    Tokenizes text after removing emojis.
    """
    # ChatGPT suggested using emoji.demojize to handle emojis
    text = emoji.demojize(text)
    return word_tokenize(text)
```

3. **Optimisation and refactoring**

```python
# GitHub Copilot suggested a single function for sentiment analysis.
# I refactored it for modularity, allowing reuse of the tokenizer and model.
from transformers import pipeline

# Suggested by Copilot: Combined everything into one function.
# Refactored by me to separate model initialization and inference.
def initialize_sentiment_model():
    """
    Initializes the sentiment analysis model and tokenizer.
    """
    return pipeline("sentiment-analysis")

def analyze_sentiment(model, texts):
    """
    Analyzes the sentiment of a list of texts using a preloaded model.
    """
    return model(texts)

# Usage
sentiment_model = initialize_sentiment_model()
results = analyze_sentiment(sentiment_model,
    ["I love climate action!", "This is a disaster."])
```

10.2.2 Generative AI tools

Below I have listed some AI tools that might help you complete your assignment(s):

- ChatGPT – good at writing code
- Claude – better than ChatGPT at writing text
- Gemini – fast and useful integrations
- Perplexity – great for finding relevant web pages and documents
- HuggingChat – open source, flexible, and free!

10.3 Late Submission

Lab notebooks submitted late will not be accepted unless an extension has been granted, as exercise solutions may be discussed in class. For group projects submitted past the due date, the following penalties will be applied:

- Coursework received at any time within 10 working days of the due date will be graded, but a penalty will apply.
  - Coursework submitted at any time up to and including five working days after the due date will have the grade awarded reduced by one grade point (for example, from B- to C+).
  - Coursework submitted more than five working days but up to and including 10 working days after the due date will have the grade reduced by two grade points (for example, from B- to C).
  - Where a student finds they have missed a deadline for submission, they should be advised that they may use the remainder of the week to improve their submission without additional penalty.
- Coursework received more than 10 working days after the due date will not be accepted. Regulations regarding extenuating circumstances apply.

10.4 Dignity and Respect

UCD is committed to the promotion of an environment for work and study which upholds the dignity and respect of all members of the UCD community and which supports your right to study and/or work in an environment which is free of any form of bullying, harassment or sexual misconduct (including sexual harassment and sexual violence).

There are a number of supports in place if you are experiencing bullying, harassment or sexual misconduct and you are strongly encouraged to come forward to seek confidential support and guidance on the range of informal options and formal options for resolving issues as appropriate. Reports of bullying, harassment or sexual misconduct can also be made anonymously through UCD's Report and Support tool.

UCD is actively promoting a culture where bullying, harassment and sexual misconduct is not tolerated, where everyone is respected and feels valued, included and that they belong in UCD.

Further information can be found on the dedicated Dignity and Respect website, which provides a central hub to bring together information on the supports available if you are experiencing issues of a bullying, harassment or sexual misconduct.

10.5 Syllabus Modification Rights

I reserve the right to reasonably alter the elements of the syllabus at any time to keep pace with the course schedule. Moreover, I may change the content of specific sessions, depending on the participants' prior knowledge and research interests. If I make adjustments, I will upload the revised syllabus to Brightspace.

11 Course Schedule

Week 1: Introduction

Lecture: Module overview and organisation
Lab: Setting up a Python development environment
Key reading(s):

- Grossmann, I. et al. (2023) 'AI and the transformation of social science research', *Science*, 380(6650), pp. 1108–1109. https://doi.org/10.1126/science.adi1778
- Ziems, C. et al. (2023) 'Can Large Language Models Transform Computational Social Science?' *arXiv*. http://arxiv.org/abs/2305.03514

Week 2: Text-as-Data

Lecture: Tasks in text-as-data (scaling, topic modelling, classification, clustering, keyword extraction, NER, sentiment analysis)
Lab: Dealing with textual data in Python
Key reading(s):

- Grimmer, J. and Stewart, B.M. (2013) 'Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts', *Political Analysis*, 21(3), pp. 267–297. https://doi.org/10.1093/pan/mps028

Week 3: Embeddings

Lecture: Word embeddings, sequence embeddings, and document embeddings
Lab: Measuring text similarity using embeddings
Key reading(s):

- Rheault, L. and Cochrane, C. (2020) 'Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora', *Political Analysis*, 28(1), pp. 112–133. https://doi.org/10.1017/pan.2019.26
- Rodriguez, P.L. and Spirling, A. (2022) 'Word Embeddings: What Works, What Doesn't, and How to Tell the Difference for Applied Research', *Journal of Politics*, 84(1), pp. 101–115. https://doi.org/10.1086/715162

Week 4: Transformers

Lecture: Basics of transformer models
Lab: Fine tuning transformers for text classification
Key reading(s):

- Widmann, T. and Wich, M. (2023) 'Creating and Comparing Dictionary, Word Embedding, and Transformer-Based Models to Measure Discrete Emotions in German Political Text', *Political Analysis*, 31(4), pp. 626–641. https://doi.org/10.1017/pan.2022.15
- Laurer, M. et al. (2024) 'Less Annotating, More Classifying: Addressing the Data Scarcity Issue of Supervised Machine Learning with Deep Transfer Learning and BERT-NLI', *Political Analysis*, 32(1), pp. 84–100.https://doi.org/10.1017/pan.2023.20

Week 5: Validation and Performance Measurement

Lecture: Performance metrics, human coding, and cross-validation
Lab: Measuring model performance with `scikit-learn`
Key reading(s):

- Birkenmaier, L., Lechner, C.M. and Wagner, C. (2024) 'The Search for Solid Ground in Text as Data: A Systematic Review of Validation Practices and Practical Recommendations for Validation', *Communication Methods and Measures*, 18(3), pp. 249–277. https://doi.org/10.1080/19312458.2023.2285765

Week 6: Generative Language Models I

Lecture: Basics of decoder-only models
Lab: Introduction to HuggingFace `transformers`
Key reading(s):

- Heseltine, M. and Hohenberg, B.C. von (2024) 'Large language models as a substitute for human experts in annotating political text', *Research & Politics*. https://doi.org/10.1177/20531680241236239
- Gilardi, F., Alizadeh, M. and Kubli, M. (2023) 'ChatGPT outperforms crowd workers for text-annotation tasks', *Proceedings of the National Academy of Sciences*, 120(30), p. e2305016120. https://doi.org/10.1073/pnas.2305016120

Week 7: Generative Language Models II

Lecture: Advanced prompting
Lab: Introduction to `langchain`
Key reading(s):

- Abdurahman, S. et al. (2024) 'Perils and opportunities in using large language models in psychological research', *PNAS Nexus*, 3(7), p. 245. https://doi.org/10.1093/pnasnexus/pgae245
- Barrie, C., Palmer, A. and Spirling, A. (2024) 'Replication for Language Models'. https://arthurspirling.org/documents/BarriePalmerSpirling_TrustMeBro.pdf

---

Study Break

Week 8: Interpretability, Explainability and Bias in LLMs

Lecture: Techniques for interpretability in text classification; sources of bias and implications
Lab: Applying `transformers-interpret` and `ferret` for interpreting model predictions; visualising bias
Key reading(s):

- Wan, Y. et al. (2023) "'Kelly is a Warm Person, Joseph is a Role Model": Gender Biases in LLM-Generated Reference Letters'. *arXiv*. https://doi.org/10.48550/arXiv.2310.09219
- Rossi, L., Harrison, K. and Shklovski, I. (2024) 'The Problems of LLM-generated Data in Social Science Research', *Sociologica*, 18(2), pp. 145–168. https://doi.org/10.6092/issn.1971-8853/19576

Good Friday

Week 10: Ethical Use of LLMs?

Lecture: Ethical concerns in developing and deploying LLMs
Lab: Measuring environmental costs of LLM use
Key reading(s):

- Williams, A., Miceli, M. and Gebru, T. (2022) 'The Exploited Labor Behind Artificial Intelligence', *Noema*, 13 October. https://www.noemamag.com/the-exploited-labor-behind-artificial-intelligence (Accessed: 26 September 2024).
- Bender, E.M. et al. (2021) 'On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?', in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. New York, NY, USA: Association for Computing Machinery (FAccT '21), pp. 610–623. https://doi.org/10.1145/3442188.3445922

- Bucher, M.J.J. and Martini, M. (2024) 'Fine-Tuned "Small" LLMs (Still) Significantly Outperform Zero-Shot Generative AI Models in Text Classification'. *arXiv*. https://doi.org/10.48550/arXiv.2406.08660

Week 11: Beyond Text

Lecture: Social science applications for multimodal data
Lab: Working with images, audio and video in Python
Key reading(s):

- Lüken, M. et al. (2024) 'MEXCA - A Simple and Robust Pipeline for Capturing Emotion Expressions in Faces, Vocalization, and Speech'. *OSF*. https://doi.org/10.31234/osf.io/56svb

---

Week 12: Project Presentations

A Datasets

Explore the following datasets for use in your group project. These datasets vary in format and accessibility - some are readily available in CSV format while others may require additional wrangling or scraping. You are also welcome to find other datasets to work with or to scrape your own data from news sites or other sources. While extra effort in data collection will be considered in grading, remember that the main focus is applying the methods covered in this module.

Party Manifestos and Political Programs

- **Manifesto Project**
- **The American Presidency Project**

Parliamentary Speeches

- **ParlEE**
  - V1
  - V2
  - V3
- **UN General Debate Corpus**
- **US Congressional Record**
- **Irish Oireachtas Record**

Political Advertising

- **Meta Ads Library API**

News and Legal Documents

- **LexisNexis**

Collections

The following links collect multiple datasets for you to explore. Some of these are textual data, while others may be useful to augment your analysis (e.g data on party ideology, vote share, government/opposition status and more).

- **Irish Politics Data**
- **PolData**
- **NLP Datasets**
- **Google Dataset Search - Political Science**