

# Aligning Computational Tools for the Political Science Research Lifecycle

A structured framework for analysing institutional texts

---

James P. Cross

23rd IOEA Session, Corsica, April 2026

Associate Professor, SPIRe, University College Dublin  
Connected.Politics Lab

## Workshop Overview

<b>Part 1</b> Framing the Research Lifecycle	15 min
<b>Part 2</b> Data Gathering and Corpus Curation	15 min
<b>Part 3</b> Ontology, Annotation, and Validation	20 min
<b>Part 4</b> LLMs as Research Tools	20 min
<b>Part 5</b> Reporting Standards	10 min
<b>Part 6</b> Horizon Scanning and Wrap-Up	10 min

## Workshop Overview

<b>Part 1</b> Framing the Research Lifecycle	15 min
<b>Part 2</b> Data Gathering and Corpus Curation	15 min
<b>Part 3</b> Ontology, Annotation, and Validation	20 min
<b>Part 4</b> LLMs as Research Tools	20 min
<b>Part 5</b> Reporting Standards	10 min
<b>Part 6</b> Horizon Scanning and Wrap-Up	10 min

### Hands-on exercise:

- Assessing voting intentions in Turnberry trade agreement debate
- From corpus to codebook to LLM benchmarking to results

## PART 1

# Framing the Research Lifecycle

## Starting point

**Computational tools are revolutionising how we study institutions**

## Starting point

**Computational tools are revolutionising how we study institutions**

- Institutions produce digital records
  - Multimodal accounts of how institutions deliberate, decide, and communicate

## Starting point

### Computational tools are revolutionising how we study institutions

- Institutions produce digital records
  - Multimodal accounts of how institutions deliberate, decide, and communicate
- Collecting them is easier than ever
  - APIs, Advanced web scrapers (Beautiful Soup; Selenium + coding assistants)

## Starting point

### Computational tools are revolutionising how we study institutions

- Institutions produce digital records
  - Multimodal accounts of how institutions deliberate, decide, and communicate
- Collecting them is easier than ever
  - APIs, Advanced web scrapers (Beautiful Soup; Selenium + coding assistants)
- Data extraction and processing is easier than ever
  - OCR
  - Image-to-text
  - Audio-to-text
  - Video-to-text

## Starting point

### Computational tools are revolutionising how we study institutions

- Institutions produce digital records
  - Multimodal accounts of how institutions deliberate, decide, and communicate
- Collecting them is easier than ever
  - APIs, Advanced web scrapers (Beautiful Soup; Selenium + coding assistants)
- Data extraction and processing is easier than ever
  - OCR
  - Image-to-text
  - Audio-to-text
  - Video-to-text
- Classification (measuring concepts of interest) at scale now matching/exceeding human performance on many tasks
  - Unsupervised, semi-supervised, supervised ML methods well established
  - LLMs with reasoning capabilities augment existing methods

3

## (Re-)Structuring the Research Process: A Five-Stage Workflow

These developments have introduced new workflows into the research process

4

## (Re-)Structuring the Research Process: A Five-Stage Workflow

These developments have introduced new workflows into the research process

- Given a **well-defined research question**, we can define five integrated workflow stages:
  1. Data gathering and corpus curation
  2. Ontology development and tasks
  3. Annotation and validation

## (Re-)Structuring the Research Process: A Five-Stage Workflow

These developments have introduced new workflows into the research process

- Given a **well-defined research question**, we can define five integrated workflow stages:
  1. Data gathering and corpus curation
  2. Ontology development and tasks
  3. Annotation and validation
  4. LLM applications
  5. Reporting standards

## (Re-)Structuring the Research Process: A Five-Stage Workflow

These developments have introduced new workflows into the research process

- Given a **well-defined research question**, we can define five integrated workflow stages:
  1. Data gathering and corpus curation
  2. Ontology development and tasks
  3. Annotation and validation
  4. LLM applications
  5. Reporting standards

### GARBAGE IN, GARBAGE OUT

**Each stage constrains the next**

Disconnected workflows create hidden problems downstream

## Introducing Our Example: The Turnberry Trade Agreement Debate

## Introducing Our Example: The Turnberry Trade Agreement Debate

- On 27/7/25, Trump and von der Leyen met at Trump's Turnberry golf resort in Scotland and reached a **political agreement** on tariff and trade issues

## Introducing Our Example: The Turnberry Trade Agreement Debate

- On 27/7/25, Trump and von der Leyen met at Trump's Turnberry golf resort in Scotland and reached a **political agreement** on tariff and trade issues
- The core terms of the deal were **structurally asymmetric**
  - The EU to eliminate tariffs on most US industrial goods
  - US to apply a 15% tariff on most EU goods — an increase from the pre-deal  $\pm 3\%$
  - Steel and aluminium subject to separate arrangements (50% tariffs above quotas)

## Introducing Our Example: The Turnberry Trade Agreement Debate

- On 27/7/25, Trump and von der Leyen met at Trump's Turnberry golf resort in Scotland and reached a **political agreement** on tariff and trade issues
- The core terms of the deal were **structurally asymmetric**
  - The EU to eliminate tariffs on most US industrial goods
  - US to apply a 15% tariff on most EU goods — an increase from the pre-deal  $\pm 3\%$
  - Steel and aluminium subject to separate arrangements (50% tariffs above quotas)
- The EU **legislative process** was repeatedly disrupted by **US actions**
  - Greenland, US Supreme Court ruling, Spain coercion, Middle East

5

## Introducing Our Example: The Turnberry Trade Agreement Debate

- On 27/7/25, Trump and von der Leyen met at Trump's Turnberry golf resort in Scotland and reached a **political agreement** on tariff and trade issues
- The core terms of the deal were **structurally asymmetric**
  - The EU to eliminate tariffs on most US industrial goods
  - US to apply a 15% tariff on most EU goods — an increase from the pre-deal  $\pm 3\%$
  - Steel and aluminium subject to separate arrangements (50% tariffs above quotas)
- The EU **legislative process** was repeatedly disrupted by **US actions**
  - Greenland, US Supreme Court ruling, Spain coercion, Middle East
- In the EP, the **INTA Committee** inserted **amendments** to protect EU interests:
  - Sunrise clause; Suspension clause; Sunset clause (ends 31/3/2028); EP monitoring powers

5

## Introducing Our Example: The Turnberry Trade Agreement Debate

- On 27/7/25, Trump and von der Leyen met at Trump's Turnberry golf resort in Scotland and reached a **political agreement** on tariff and trade issues
- The core terms of the deal were **structurally asymmetric**
  - The EU to eliminate tariffs on most US industrial goods
  - US to apply a 15% tariff on most EU goods — an increase from the pre-deal  $\pm 3\%$
  - Steel and aluminium subject to separate arrangements (50% tariffs above quotas)
- The EU **legislative process** was repeatedly disrupted by **US actions**
  - Greenland, US Supreme Court ruling, Spain coercion, Middle East
- In the EP, the **INTA Committee** inserted **amendments** to protect EU interests:
  - Sunrise clause; Suspension clause; Sunset clause (ends 31/3/2028); EP monitoring powers
- On 26/3/26, the amended Committee Report was debated and voted on in plenary

5

## Introducing Our Example: The Turnberry Trade Agreement Debate

- On 27/7/25, Trump and von der Leyen met at Trump's Turnberry golf resort in Scotland and reached a **political agreement** on tariff and trade issues
- The core terms of the deal were **structurally asymmetric**
  - The EU to eliminate tariffs on most US industrial goods
  - US to apply a 15% tariff on most EU goods — an increase from the pre-deal  $\pm 3\%$
  - Steel and aluminium subject to separate arrangements (50% tariffs above quotas)
- The EU **legislative process** was repeatedly disrupted by **US actions**
  - Greenland, US Supreme Court ruling, Spain coercion, Middle East
- In the EP, the **INTA Committee** inserted **amendments** to protect EU interests:
  - Sunrise clause; Suspension clause; Sunset clause (ends 31/3/2028); EP monitoring powers
- On 26/3/26, the amended Committee Report was debated and voted on in plenary

### Research Question:

Does debate rhetoric predict voting behaviour?

5

PART 2

## Data Gathering and Corpus Curation

*The question defines the corpus. Now: how to build one.*

### Data Sources and Access

## Data Sources and Access

### Parliamentary records

Plenary transcripts (VO)

### Audio/video recordings

Debates & deliberations

### Legislative records

Policy texts, minutes, votes

## Data Sources and Access

### Parliamentary records

Plenary transcripts (VO)

### Audio/video recordings

Debates & deliberations

### Legislative records

Policy texts, minutes, votes

**Today we will focus mostly on the plenary transcripts, with associated vote results**

## Data Sources and Access

### Parliamentary records

Plenary transcripts (VO)

### Audio/video recordings

Debates & deliberations

### Legislative records

Policy texts, minutes, votes

Today we will focus mostly on the plenary transcripts, with associated vote results

What are the main challenges associated with collecting these data?

## Data Sources and Access

### Parliamentary records

Plenary transcripts (VO)

### Audio/video recordings

Debates & deliberations

### Legislative records

Policy texts, minutes, votes

Today we will focus mostly on the plenary transcripts, with associated vote results

What are the main challenges associated with collecting these data?

- Access
- Format heterogeneity
- Temporal coverage
- Multilingual content

## Constructing the Corpus

What data do we need and how should we structure the plenary debate corpus?

## Constructing the Corpus

What data do we need and how should we structure the plenary debate corpus?

### Corpus design

- **Scope:** One debate session
- **Unit of analysis:** Individual speech?  
Vote?
- **Text:** Original + English translation

## Constructing the Corpus

What data do we need and how should we structure the plenary debate corpus?

### Corpus design

- **Scope:** One debate session
- **Unit of analysis:** Individual speech? Vote?
- **Text:** Original + English translation

### Metadata

- Speaker id
- Role in Parliament
- Party group
- Vote (outcome variable of interest)

8

## Constructing the Corpus

What data do we need and how should we structure the plenary debate corpus?

### Corpus design

- **Scope:** One debate session
- **Unit of analysis:** Individual speech? Vote?
- **Text:** Original + English translation

### Metadata

- Speaker id
- Role in Parliament
- Party group
- Vote (outcome variable of interest)

### Question:

What additional metadata might we be interested when predicting voting behaviour?

8

# The corpus

	A	B	C	D	E	F	G
1	speech_id	speech_id	speaker	role	party_group	speech_text_vo	speech_text_en
2	9	CRE-10-2026-03-26-ITM-003-08	Hermann Tertsch	on behalf of the PFE Group	PFE	Dar un voto afirmativo es aplaudir la ineptitud de Vo	Castling an affirmative vote is applauding Von der Ley
3	10	CRE-10-2026-03-26-ITM-003-09	Daniele Polato	on behalf of the ECR Group	ECR	Servono certezze, serve stabilità. Questa è la direzi	Certainty is needed, stability is needed. This is the di
4	13	CRE-10-2026-03-26-ITM-003-12	Martin Schirdewan	on behalf of The Left Group	The Left	Und auch wenn die Kommissionspräsidentin Ursula	And even if Commission President Ursula von der Le
5	16	CRE-10-2026-03-26-ITM-003-15	Christophe Bay	MEP	PfE	D'ailleurs, ce n'est pas tant l'administration Trump q	Moreover, it is not so much the Trump administration
6	17	CRE-10-2026-03-26-ITM-003-16	Billy Kelleher	MEP	Renew	On the issue of steel and aluminium, yes, there is a	On the issue of steel and aluminium, yes, there is a n
7	21	CRE-10-2026-03-26-ITM-003-20	Enikő Győri	MEP	PfE	Miért is ment bele ebbe a megalázó alkuba a Bizott	Why did the Commission agree to this humiliating de
8	25	CRE-10-2026-03-26-ITM-003-24	Maria Grapini	MEP (catch-the-eye)	S&D	Și de aceea, domnule comisar, vă rog să transmiteți	And that is why, Commissioner, I ask you to convey
9	26	CRE-10-2026-03-26-ITM-003-25	Sebastian Tynkkynen	MEP (catch-the-eye)	ECR	Sebastian Tynkkynen (European Conservatives and	Sebastian Tynkkynen (European Conservatives and
10	31	CRE-10-2026-03-26-ITM-003-30	Juan Fernando López	MEP (catch-the-eye)	S&D	Pero esta no es la última palabra, es una Resolución	But this is not the last word, it is a Resolution that set
11	33	CRE-10-2026-03-26-ITM-003-32	Valdis Dombrovskis	Member of the Commission	Commission	At the same time, we will continue to engage with o	At the same time, we will continue to engage with our

9

## PART 3

# Ontology, Annotation, and Validation

*The corpus defines what annotations are possible. Now: how to design them.*

10

## From Theory to Tasks

### Annotation is a theoretical step

- **Key question:** How does our concept of interest (expressed support for a proposal) map to observable features of our data (plenary speeches)
- Ontology development is where **domain expertise** matters most

## From Theory to Tasks

### Annotation is a theoretical step

- **Key question:** How does our concept of interest (expressed support for a proposal) map to observable features of our data (plenary speeches)
- Ontology development is where **domain expertise** matters most

### Key trade-offs when building an ontology

- Context specificity v generalisability
- Granular vs coarse categories
- Deductive vs inductive approaches
- Binary, categorical, or ordinal scales

## From Theory to Tasks

### Annotation is a theoretical step

- **Key question:** How does our concept of interest (expressed support for a proposal) map to observable features of our data (plenary speeches)
- Ontology development is where **domain expertise** matters most

### Key trade-offs when building an ontology

- Context specificity v generalisability
- Granular vs coarse categories
- Deductive vs inductive approaches
- Binary, categorical, or ordinal scales

### LLMs can assist in scaling the classification process

BUT 'Garbage in, garbage out' applies when LLMs make confident but incorrect judgements

11

## Practical Exercise: Building a Codebook

### Measurement Task: Does a speech express support for the draft position of the EP?

- Keep in mind context, speaker, and coder constraints

12

## Practical Exercise: Building a Codebook

**Measurement Task: Does a speech express support for the draft position of the EP?**

- Keep in mind context, speaker, and coder constraints

**What are the appropriate categories to capture levels of support expressed in speeches?**

## Practical Exercise: Building a Codebook

**Measurement Task: Does a speech express support for the draft position of the EP?**

- Keep in mind context, speaker, and coder constraints

**What are the appropriate categories to capture levels of support expressed in speeches?**

1 = For

2 = Against

## Practical Exercise: Building a Codebook

**Measurement Task: Does a speech express support for the draft position of the EP?**

- Keep in mind context, speaker, and coder constraints

**What are the appropriate categories to capture levels of support expressed in speeches?**

- 1 = For
- 2 = Against
- 3 = No clear position (abstain?)

## Practical Exercise: Building a Codebook

**Measurement Task: Does a speech express support for the draft position of the EP?**

- Keep in mind context, speaker, and coder constraints

**What are the appropriate categories to capture levels of support expressed in speeches?**

- 1 = For
- 2 = Against
- 3 = No clear position (abstain?)
- 4 = Procedural speeches

## Practical Exercise: Building a Codebook

**Measurement Task: Does a speech express support for the draft position of the EP?**

- Keep in mind context, speaker, and coder constraints

**What are the appropriate categories to capture levels of support expressed in speeches?**

- 1 = For
- 2 = Against
- 3 = No clear position (abstain?)
- 4 = Procedural speeches
- 9 = Don't know

## Practical Exercise: Building a Codebook

**Measurement Task: Does a speech express support for the draft position of the EP?**

- Keep in mind context, speaker, and coder constraints

**What are the appropriate categories to capture levels of support expressed in speeches?**

- 1 = For
- 2 = Against
- 3 = No clear position (abstain?)
- 4 = Procedural speeches
- 9 = Don't know

## Coding Exercise with CodeBook Studio

Divide into groups of 3-4



**CodeBook Studio**  
codebook.streamlit.app



**Workshop Materials**  
Google Drive folder

## Speech 2: Daniele Polato (ECR)

Mr President, dear colleagues, in this Chamber we have different sensibilities regarding the United States, we know that.

On one side we have ideology, polemics, criticism, 1970s anti-Americanism. On the other side there are us, who want instead to work on what unites us rather than what divides us.

We must send a message of responsibility that gives certainty to our businesses and our workers. Keep your ideologies, polemics and criticism. Those – as we have already seen with the Green Deal – do not take us anywhere, do not consolidate relationships, do not create development.

We want the vote on these agreements to be a signal of a new start [...] But to those who want to exploit or delay this vote, we want to remind them that businesses have no time to waste, especially in this delicate phase in which we live.

Certainty is needed, stability is needed. This is the direction we want to go. This is what businesses and workers are asking of us. This is the right choice for Europe.

### Speech 3: Martin Schirdewan (The Left)

Mr President! [...] This trade deal disadvantages European industry and it endangers jobs. And in addition to the one-sided tariffs you mentioned, European states are supposed to be obligated to buy massive amounts of liquefied gas and other fossil fuels from the US. So instead of European energy independence, we now have a new dependency on the US – I consider that total madness.

European companies are furthermore supposed to be obligated to invest massively in the US instead of in Europe. That would be like paying ransom to the extortionist [...] And just yesterday Trump made clear again that he has absolutely no intention of honouring the tariff part of this agreement.

Trump wants his big tech oligarchs not to have to pay taxes in the EU. He doesn't care about European data protection. He wants to annex Greenland. He threatens Spain. He carries out attacks in violation of international law against Venezuela and Iran.

And even if Commission President Ursula von der Leyen agreed to this act of submission, dear colleagues: you cannot seriously seal such a deal with the madman in the White House. A Parliament that takes itself seriously must never give in to this extortion.

15

### Speech 7: Maria Grapini (S&D)

Mr President, Commissioner, dear colleagues, the European Union's agreement with the United States is of course a necessity, but from my point of view, agreement also means respect between partners. To raise taxes fivefold, from 3 to 15, to want to impose 50 per cent for steel and aluminium and to believe you respect your partner, seems absolutely unacceptable to me.

That is why I think about what would have happened if the agreement had been signed and remained as the Commission conceived it.

Commissioner, it is not possible not to take into account the conditions and the market of the European Union. Fortunately, we, Parliament, have asserted ourselves and continue to assert ourselves with very clear requirements.

A trade agreement will work in the long term if there is respect and equity between both parties.

And yes, I cannot fail to say that the impact, had we gone with these tariffs, the impact on industry, especially the metals industry, would be disastrous.

And that is why, Commissioner, I ask you to convey and respect Parliament's position!

16

## Speech 8: Sebastian Tynkkynen (ECR)

Mr Chair, in how many ways does the left try to prevent Europe from succeeding? I don't know. I lost count a long time ago. But one perennial favourite of theirs has always been obstructing US trade. When we buy American fighter jets, it won't do, shouts the left. When we get a trade deal with Trump, it won't do, shouts the left. When Finland recently concluded a massive icebreaker deal where we sell icebreakers to the United States, that didn't seem to work for the left either. The left was already calling for pulling out of the deal.

Apparently we can't even sell our expertise when the counterpart is the US. How about it, left, should we soon also ban watching American films, like in the Soviet Union, whose teachings you still seem to follow in your US policy? But remember, the US won, the Soviet Union lost. It's time for you to switch sides too.

17

## Speech 10: Valdis Dombrovskis (European Commission)

Mr President, honourable Members, thank you for your interventions this morning.

We take note and appreciate a wide spectrum of views in this House. This is not an easy debate, but I welcome that there is an overall recognition of the importance of the EU-US trading relationship, and that moving ahead is the best step we can take for our citizens, workers and firms.

As we approach the vote shortly, I would once again underline that we recognise the validity of the points which have been taken up by the European Parliament and that we are ready to work constructively in the trilogues to see how to address them in the best manner possible.

At the same time, we will continue to engage with our US counterparts, raising our concerns where necessary, working together on issues of mutual concern where possible, and laying the foundations for further and future engagement.

18

## Discussion

### Which speeches were easy/difficult to code?

Disagreements reveal where the codebook needs sharper definitions

19

## Discussion

### Which speeches were easy/difficult to code?

Disagreements reveal where the codebook needs sharper definitions

	A	B	C	D	E	F	G	H
1	speech_number	speech_id	speaker	role	party_group	speech_text_vo	speech_text_en	human_coded_position
2	9	CRE-10-2026-03-26-ITM-003-08	Hermann Tertsch	on behalf of the P	PfE	Dar un voto afirmativo es aplaudir la inep	Casting an affirmative vote is applauding Von de	Against
3	10	CRE-10-2026-03-26-ITM-003-09	Daniele Polato	on behalf of the E	ECR	Servono certezze, serve stabilità. Questa	Certainty is needed, stability is needed. This is th	Support
4	13	CRE-10-2026-03-26-ITM-003-12	Martin Schirdew	on behalf of The U	The Left	Und auch wenn die Kommissionspräsident	And even if Commission President Ursula von de	Against
5	16	CRE-10-2026-03-26-ITM-003-15	Christophe Bay	MEP	PIE	D'ailleurs, ce n'est pas tant l'administratio	Moreover, it is not so much the Trump administr	Against
6	17	CRE-10-2026-03-26-ITM-003-16	Billy Kelleher	MEP	Renew	On the issue of steel and aluminium, yes,	On the issue of steel and aluminium, yes, there i	Support
7	21	CRE-10-2026-03-26-ITM-003-20	Enikő Győri	MEP	PIE	Miért is ment ebbe a negatív alku	Why did the Commission agree to this humiliat	Against
8	25	CRE-10-2026-03-26-ITM-003-24	Maria Grapini	MEP (catch-the-e	S&D	Si de aceea, domnule comisar, vă rog să	And that is why, Commissioner, I ask you to com	Against
9	26	CRE-10-2026-03-26-ITM-003-25	Sebastian Tympky	MEP (catch-the-e	ECR	Sebastian Tympkyinen (European Conserva	Sebastian Tympkyinen (European Conservatives	Don't know
10	31	CRE-10-2026-03-26-ITM-003-30	Juan Fernando L	MEP (catch-the-e	S&D	Pero esta no es la última palabra, es una	But this is not the last word, it is a Resolution th	Don't know
11	33	CRE-10-2026-03-26-ITM-003-32	Valdis Dombrovsk	Member of the Co	Commission	At the same time, we will continue to enga	At the same time, we will continue to engage wit	Procedural

We will return to these later. Please download your dataset

19

PART 4

## LLMs As Classification Tools

*Human annotations set the standard. Now: can LLMs meet it?*

## LLMs As Classification Tools

## LLMs As Classification Tools

### KEY DECISION POINT

Should I use LLMs for my classification task?

21

## LLMs As Classification Tools

### KEY DECISION POINT

Should I use LLMs for my classification task?

#### Where LLMs excel

- Zero/few-shot classification
- Summarisation of long documents
- Multilingual analysis
- Codebook-consistent annotation

21

## LLMs As Classification Tools

### KEY DECISION POINT

Should I use LLMs for my classification task?

#### Where LLMs excel

Zero/few-shot classification  
Summarisation of long documents  
Multilingual analysis  
Codebook-consistent annotation

#### Where caution is needed

- Hallucination in fact extraction
- Sensitivity to prompt framing
- Systematic biases in summarisation
- Reproducibility challenges
- Energy use and cost at scale

21

## LLMs As Classification Tools

### KEY DECISION POINT

Should I use LLMs for my classification task?

#### Where LLMs excel

Zero/few-shot classification  
Summarisation of long documents  
Multilingual analysis  
Codebook-consistent annotation

#### Where caution is needed

- Hallucination in fact extraction
- Sensitivity to prompt framing
- Systematic biases in summarisation
- Reproducibility challenges
- Energy use and cost at scale

**IF you choose to use LLMs for classification you face the following choices:**

- Model choice (Open-weight vs closed-source models)
- Prompt style (Basic, persona, CoT)
- Learning approach (Zero v few-shot learning)

21

## McLaren et al. (2026): LLM Performance When Annotating at Scale

### Experimental set-up

22

## McLaren et al. (2026): LLM Performance When Annotating at Scale

### Experimental set-up

- 4x annotation tasks
  - Approval; Psychological distance; Economic sentiment; Manifesto topic

22

## McLaren et al. (2026): LLM Performance When Annotating at Scale

### Experimental set-up

- 4x annotation tasks
  - Approval; Psychological distance; Economic sentiment; Manifesto topic
- 6x open-weight models
  - Deepseek R1; Gemma 3; GPT OSS; LLaMA 3.1; Qwen 2.5; Qwen 3

## McLaren et al. (2026): LLM Performance When Annotating at Scale

### Experimental set-up

- 4x annotation tasks
  - Approval; Psychological distance; Economic sentiment; Manifesto topic
- 6x open-weight models
  - Deepseek R1; Gemma 3; GPT OSS; LLaMA 3.1; Qwen 2.5; Qwen 3
- 4x hyperparameters varied
  - Model choice
  - Model size
  - Zero v few-shot
  - Prompt (standard, persona, CoT)

## McLaren et al. (2026): LLM Performance When Annotating at Scale

### Experimental set-up

- 4x annotation tasks
  - Approval; Psychological distance; Economic sentiment; Manifesto topic
- 6x open-weight models
  - Deepseek R1; Gemma 3; GPT OSS; LLaMA 3.1; Qwen 2.5; Qwen 3
- 4x hyperparameters varied
  - Model choice
  - Model size
  - Zero v few-shot
  - Prompt (standard, persona, CoT)
- Identical quantisation, hardware, prompts

22

## McLaren et al. (2026): LLM Performance When Annotating at Scale

### Experimental set-up

- 4x annotation tasks
  - Approval; Psychological distance; Economic sentiment; Manifesto topic
- 6x open-weight models
  - Deepseek R1; Gemma 3; GPT OSS; LLaMA 3.1; Qwen 2.5; Qwen 3
- 4x hyperparameters varied
  - Model choice
  - Model size
  - Zero v few-shot
  - Prompt (standard, persona, CoT)
- Identical quantisation, hardware, prompts

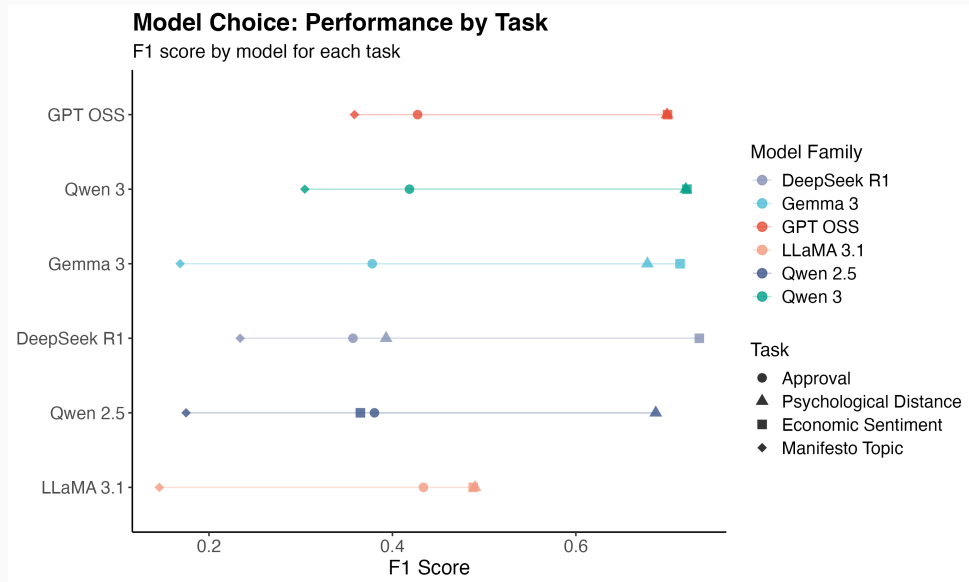
### Central finding

Interaction effects dominate. No single model or technique is uniformly superior.

22

## Model Choice: No Single Model Dominates

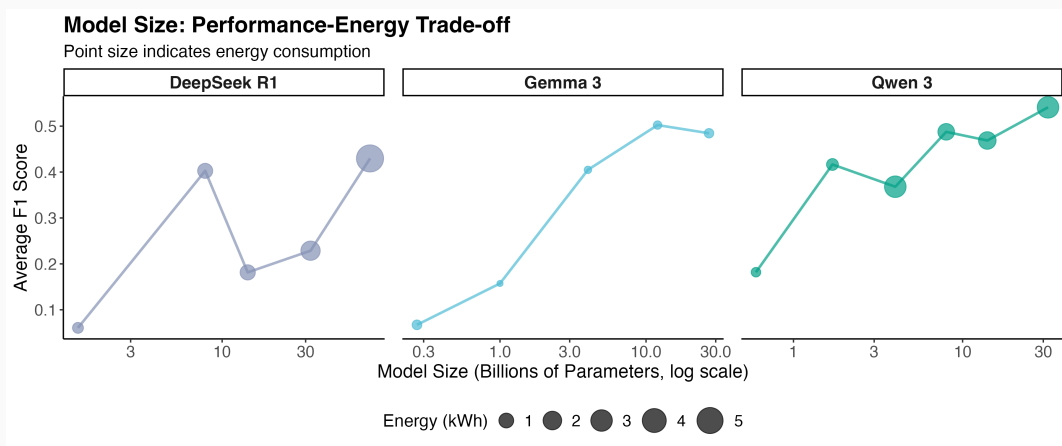
### Cross-task performance varies



23

## Model Size: Bigger Is Not Always Better

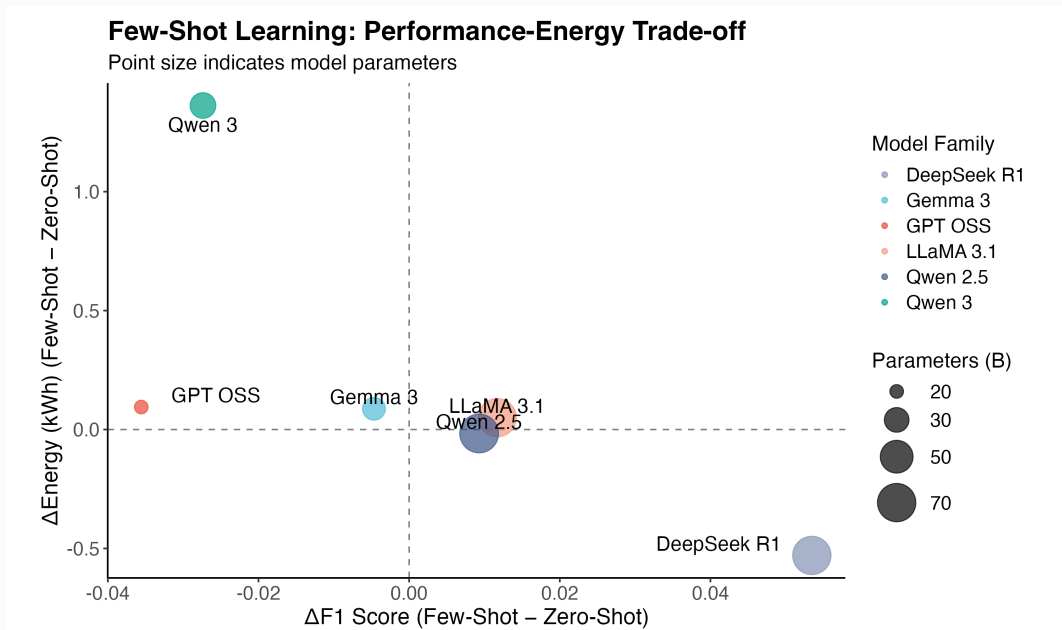
### Larger models from the same family do not always perform better



24

## Few-Shot Learning: Not Always Helpful

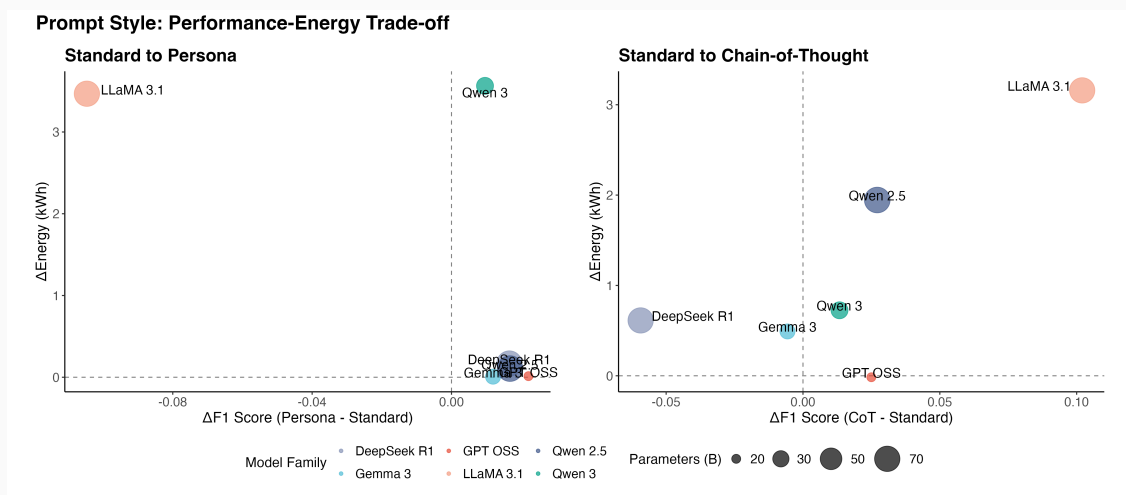
Strong zero-shot performers may degrade under few-shot conditions



25

## Prompt Style: Erratic, Sometimes Harmful

Neither persona nor chain-of-thought reliably improves performance over standard prompt



26

## Returning To Our Example: From Human Labels to LLM Benchmarking

### Comparing model performance to ground truth

- Take human-coded ground truth of 36 speeches

## Returning To Our Example: From Human Labels to LLM Benchmarking

### Comparing model performance to ground truth

- Take human-coded ground truth of 36 speeches
- Feed same codebook, instructions, and speeches to LLMs
  - Closed source model: Claude Opus 4.6 (API)
  - Open source models: Gemma 3 (4B; 27B); Qwen 3.5 (27B; 35B MoE); Nemotron (120B MoE)

## Returning To Our Example: From Human Labels to LLM Benchmarking

### Comparing model performance to ground truth

- Take human-coded ground truth of 36 speeches
- Feed same codebook, instructions, and speeches to LLMs
  - Closed source model: Claude Opus 4.6 (API)
  - Open source models: Gemma 3 (4B; 27B); Qwen 3.5 (27B; 35B MoE); Nemotron (120B MoE)
- Compare results using standard metrics
  - Accuracy (% correct), F1, Cohen's Kappa, Krippendorff's Alpha

27

## Results: How well do LLMs replicate human judgement?

Model	Params	Accuracy	F1
Claude Opus 4.6	API	88.6%	<b>0.66</b>
Gemma3	4B	88.6%	0.61
Gemma3	27B	88.6%	<b>0.66</b>
Qwen3.5	27B	88.6%	<b>0.66</b>
Qwen3.5	35B <sup>MoE</sup>	88.6%	<b>0.66</b>
Nemotron	120B <sup>MoE</sup>	57.1%	0.35

**Table 1:** Classification: Position on the Deal (zero-shot, standard prompt)

28

## Results: How well do LLMs replicate human judgement?

Model	Params	Accuracy	F1
Claude Opus 4.6	API	88.6%	<b>0.66</b>
Gemma3	4B	88.6%	0.61
Gemma3	27B	88.6%	<b>0.66</b>
Qwen3.5	27B	88.6%	<b>0.66</b>
Qwen3.5	35B <sup>MoE</sup>	88.6%	<b>0.66</b>
Nemotron	120B <sup>MoE</sup>	57.1%	0.35

**Table 1:** Classification: Position on the Deal (zero-shot, standard prompt)

Six very different models converge on the same accuracy. Why?

28

## The Edge Cases

Speaker	Human	Gemma 4B	Gemma 27B	Qwen 27B	Qwen 35B	Nemotron	Claude Opus
Grapini (Speech 7)	Against	✗	✗	✗	✗	✓	✗
Tynkkynen (Speech 8)	Don't know	✗	DK	✗	✗	✗	✗
López Aguilar (Speech 9)	Don't know	✗	✗	✗	✗	✗	✗
Dombrovskis (Speech 10)	Procedural	✗	✗	✗	✗	✗	✗

✓ = matches human GT ✗ = disagrees DK = Don't know - = no response

29

## Speech 7: Maria Grapini (S&D)

Mr President, Commissioner, dear colleagues, the European Union's agreement with the United States is of course a necessity, but from my point of view, agreement also means respect between partners. To raise taxes fivefold, from 3 to 15, to want to impose 50 per cent for steel and aluminium and to believe you respect your partner, seems absolutely unacceptable to me.

That is why I think about what would have happened if the agreement had been signed and remained as the Commission conceived it.

Commissioner, it is not possible not to take into account the conditions and the market of the European Union. Fortunately, we, Parliament, have asserted ourselves and continue to assert ourselves with very clear requirements.

A trade agreement will work in the long term if there is respect and equity between both parties [...] And that is why, Commissioner, I ask you to convey and respect Parliament's position!

30

## Speech 7: Maria Grapini (S&D)

Mr President, Commissioner, dear colleagues, the European Union's agreement with the United States is of course a necessity, but from my point of view, agreement also means respect between partners. To raise taxes fivefold, from 3 to 15, to want to impose 50 per cent for steel and aluminium and to believe you respect your partner, seems absolutely unacceptable to me.

That is why I think about what would have happened if the agreement had been signed and remained as the Commission conceived it.

Commissioner, it is not possible not to take into account the conditions and the market of the European Union. Fortunately, we, Parliament, have asserted ourselves and continue to assert ourselves with very clear requirements.

A trade agreement will work in the long term if there is respect and equity between both parties [...] And that is why, Commissioner, I ask you to convey and respect Parliament's position!

### Human GT

Against

### LLMs

Support (5/6)

Against (1/6)

30

## Speech 7: Maria Grapini (S&D)

Mr President, Commissioner, dear colleagues, the European Union's agreement with the United States is of course a necessity, but from my point of view, agreement also means respect between partners. To raise taxes fivefold, from 3 to 15, to want to impose 50 per cent for steel and aluminium and to believe you respect your partner, seems absolutely unacceptable to me.

That is why I think about what would have happened if the agreement had been signed and remained as the Commission conceived it.

Commissioner, it is not possible not to take into account the conditions and the market of the European Union. Fortunately, we, Parliament, have asserted ourselves and continue to assert ourselves with very clear requirements.

A trade agreement will work in the long term if there is respect and equity between both parties [...] And that is why, Commissioner, I ask you to convey and respect Parliament's position!

*Critical of the terms, but frames it as improving the deal rather than rejecting it outright. Five of six models read Support; the human coder focused on "absolutely unacceptable".*

**Human GT**

Against

**LLMs**

Support (5/6)

Against (1/6)

30

## Speech 8: Sebastian Tynkynen (European Conservatives and Reformists)

Mr Chair, in how many ways does the left try to prevent Europe from succeeding? I don't know. I lost count a long time ago. But one perennial favourite of theirs has always been obstructing US trade. When we buy American fighter jets, it won't do, shouts the left. When we get a trade deal with Trump, it won't do, shouts the left. When Finland recently concluded a massive icebreaker deal where we sell icebreakers to the United States, that didn't seem to work for the left either. The left was already calling for pulling out of the deal.

Apparently we can't even sell our expertise when the counterpart is the US. How about it, left, should we soon also ban watching American films, like in the Soviet Union, whose teachings you still seem to follow in your US policy? But remember, the US won, the Soviet Union lost. It's time for you to switch sides too.

31

## Speech 8: Sebastian Tynkkynen (European Conservatives and Reformists)

Mr Chair, in how many ways does the left try to prevent Europe from succeeding? I don't know. I lost count a long time ago. But one perennial favourite of theirs has always been obstructing US trade. When we buy American fighter jets, it won't do, shouts the left. When we get a trade deal with Trump, it won't do, shouts the left. When Finland recently concluded a massive icebreaker deal where we sell icebreakers to the United States, that didn't seem to work for the left either. The left was already calling for pulling out of the deal.

Apparently we can't even sell our expertise when the counterpart is the US. How about it, left, should we soon also ban watching American films, like in the Soviet Union, whose teachings you still seem to follow in your US policy? But remember, the US won, the Soviet Union lost. It's time for you to switch sides too.

**Human GT**

Don't know

**LLMs**

Support (5/6)

Don't know

(1/6)

31

## Speech 8: Sebastian Tynkkynen (European Conservatives and Reformists)

Mr Chair, in how many ways does the left try to prevent Europe from succeeding? I don't know. I lost count a long time ago. But one perennial favourite of theirs has always been obstructing US trade. When we buy American fighter jets, it won't do, shouts the left. When we get a trade deal with Trump, it won't do, shouts the left. When Finland recently concluded a massive icebreaker deal where we sell icebreakers to the United States, that didn't seem to work for the left either. The left was already calling for pulling out of the deal.

Apparently we can't even sell our expertise when the counterpart is the US. How about it, left, should we soon also ban watching American films, like in the Soviet Union, whose teachings you still seem to follow in your US policy? But remember, the US won, the Soviet Union lost. It's time for you to switch sides too.

**Human GT**

Don't know

**LLMs**

Support (5/6)

Don't know

(1/6)

*No explicit position on the deal. The speech attacks opponents of US trade, which most LLMs read as implicit support. The human coder required a direct statement and coded Don't know.*

31

## Speech 9: Juan Fernando López Aguilar (S&D)

Mr President, Commissioner, it is true that the relationship with the United States that we know as transatlantic is important for workers and businesses, because it is after all the European Union's first trading partner. And it is also true that the United States is not Trump: it pre-existed him and will survive him.

But it was Trump at Turnberry, at his golf course in Scotland, who tried to impose on the European Union through President Von der Leyen a humiliating agreement. And if President von der Leyen accepts it, she certainly does not represent us, because this Parliament has the last word, according to Article 218 of the Treaty on the Functioning.

But this is not the last word, it is a Resolution that sets the position of the European Parliament binding on the Council and the Commission to guarantee legal certainty and, above all, reciprocity [. . .] Therefore, either legal certainty and mutual respect and reciprocity are guaranteed, or the European Parliament will reject it with the vote against of the Socialist Group.

32

## Speech 9: Juan Fernando López Aguilar (S&D)

Mr President, Commissioner, it is true that the relationship with the United States that we know as transatlantic is important for workers and businesses, because it is after all the European Union's first trading partner. And it is also true that the United States is not Trump: it pre-existed him and will survive him.

But it was Trump at Turnberry, at his golf course in Scotland, who tried to impose on the European Union through President Von der Leyen a humiliating agreement. And if President von der Leyen accepts it, she certainly does not represent us, because this Parliament has the last word, according to Article 218 of the Treaty on the Functioning.

But this is not the last word, it is a Resolution that sets the position of the European Parliament binding on the Council and the Commission to guarantee legal certainty and, above all, reciprocity [. . .] Therefore, either legal certainty and mutual respect and reciprocity are guaranteed, or the European Parliament will reject it with the vote against of the Socialist Group.

### Human GT

Don't know

### LLMs

Against (5/6)

Support (1/6)

32

## Speech 9: Juan Fernando López Aguilar (S&D)

Mr President, Commissioner, it is true that the relationship with the United States that we know as transatlantic is important for workers and businesses, because it is after all the European Union's first trading partner. And it is also true that the United States is not Trump: it pre-existed him and will survive him.

But it was Trump at Turnberry, at his golf course in Scotland, who tried to impose on the European Union through President Von der Leyen a humiliating agreement. And if President von der Leyen accepts it, she certainly does not represent us, because this Parliament has the last word, according to Article 218 of the Treaty on the Functioning.

But this is not the last word, it is a Resolution that sets the position of the European Parliament binding on the Council and the Commission to guarantee legal certainty and, above all, reciprocity [. . .] Therefore, either legal certainty and mutual respect and reciprocity are guaranteed, or the European Parliament will reject it with the vote against of the Socialist Group.

*Conditional stance: supports the Parliament's resolution (which adds safeguards) while threatening rejection if conditions are unmet. Human coded Don't know; most LLMs read opposition.*

**Human GT**

Don't know

**LLMs**

Against (5/6)

Support (1/6)

32

## Speech 10: Valdis Dombrovskis (European Commission)

Mr President, honourable Members, thank you for your interventions this morning.

We take note and appreciate a wide spectrum of views in this House. This is not an easy debate, but I welcome that there is an overall recognition of the importance of the EU-US trading relationship, and that moving ahead is the best step we can take for our citizens, workers and firms.

As we approach the vote shortly, I would once again underline that we recognise the validity of the points which have been taken up by the European Parliament and that we are ready to work constructively in the trilogues to see how to address them in the best manner possible.

At the same time, we will continue to engage with our US counterparts, raising our concerns where necessary, working together on issues of mutual concern where possible, and laying the foundations for further and future engagement.

33

## Speech 10: Valdis Dombrovskis (European Commission)

Mr President, honourable Members, thank you for your interventions this morning.

We take note and appreciate a wide spectrum of views in this House. This is not an easy debate, but I welcome that there is an overall recognition of the importance of the EU-US trading relationship, and that moving ahead is the best step we can take for our citizens, workers and firms.

As we approach the vote shortly, I would once again underline that we recognise the validity of the points which have been taken up by the European Parliament and that we are ready to work constructively in the trilogues to see how to address them in the best manner possible.

At the same time, we will continue to engage with our US counterparts, raising our concerns where necessary, working together on issues of mutual concern where possible, and laying the foundations for further and future engagement.

**Human GT**  
Procedural  
**LLMs**  
Support (6/6)

33

## Speech 10: Valdis Dombrovskis (European Commission)

Mr President, honourable Members, thank you for your interventions this morning.

We take note and appreciate a wide spectrum of views in this House. This is not an easy debate, but I welcome that there is an overall recognition of the importance of the EU-US trading relationship, and that moving ahead is the best step we can take for our citizens, workers and firms.

As we approach the vote shortly, I would once again underline that we recognise the validity of the points which have been taken up by the European Parliament and that we are ready to work constructively in the trilogues to see how to address them in the best manner possible.

At the same time, we will continue to engage with our US counterparts, raising our concerns where necessary, working together on issues of mutual concern where possible, and laying the foundations for further and future engagement.

**Human GT**  
Procedural  
**LLMs**  
Support (6/6)

*Coded Procedural because the Commissioner is summarising the debate and signalling next steps rather than expressing a personal stance. All six LLMs read Support.*

33

## What the Edge Cases Teach Us

### 1. Rhetorical framing $\neq$ voting intention

- Models over-index on positive language; miss diplomatic opposition (Grapini) and procedural framing (Dombrovskis)

## What the Edge Cases Teach Us

### 1. Rhetorical framing $\neq$ voting intention

- Models over-index on positive language; miss diplomatic opposition (Grapini) and procedural framing (Dombrovskis)

### 2. Meta-commentary is not a position

- Tynkkynen attacks the left for opposing US trade without stating his own stance. Models hallucinate a position from context

## What the Edge Cases Teach Us

### 1. Rhetorical framing $\neq$ voting intention

- Models over-index on positive language; miss diplomatic opposition (Grapini) and procedural framing (Dombrovskis)

### 2. Meta-commentary is not a position

- Tynkkynen attacks the left for opposing US trade without stating his own stance. Models hallucinate a position from context

### 3. Conditional stances are genuinely ambiguous

- López Aguilar's "if conditions are met . . . otherwise we reject" is hard even for humans. Codebook disambiguation rules earn their keep

34

## What the Edge Cases Teach Us

### 1. Rhetorical framing $\neq$ voting intention

- Models over-index on positive language; miss diplomatic opposition (Grapini) and procedural framing (Dombrovskis)

### 2. Meta-commentary is not a position

- Tynkkynen attacks the left for opposing US trade without stating his own stance. Models hallucinate a position from context

### 3. Conditional stances are genuinely ambiguous

- López Aguilar's "if conditions are met . . . otherwise we reject" is hard even for humans. Codebook disambiguation rules earn their keep

### 4. The ceiling effect (and what breaks it)

- Five of six models hit  $\sim 89\%$  regardless of size (4B to API-scale).
- Nemotron (120B MoE) scores only 57%, showing that scale alone does not guarantee competence.
- The question is not *what accuracy* but *where and why* a model fails

34

## What updates might the codebook need before automating further?

### Procedural role $\neq$ substantive position

- Identify and pre-filter procedural speeches if feasible
- Upsample for clearer signal if not

## What updates might the codebook need before automating further?

### Procedural role $\neq$ substantive position

- Identify and pre-filter procedural speeches if feasible
- Upsample for clearer signal if not

### Disambiguation of stated intent:

- Upsample speeches with voting intention; downsample tone

## What updates might the codebook need before automating further?

### Procedural role $\neq$ substantive position

- Identify and pre-filter procedural speeches if feasible
- Upsample for clearer signal if not

### Disambiguation of stated intent:

- Upsample speeches with voting intention; downsample tone

### Use “Don’t know” only when no signal exists

- Clarify interpretation of weak signal cases in codebook
- This is tricky and somewhat context dependent (vibes)

35

## PART 4.1

# Answering the Research Question

*LLM classifications validate. Now: can we predict voting behaviour?*

36

## RQ: Do expressions of support/opposition predict how MEPs actually voted?

Source	Vote prediction accuracy
Human (ground truth)	80.0%
Claude Opus 4.6	86.7%
Gemma 3 4B	86.7%
Gemma 3 27B	83.3%
Qwen 3.5 27B	86.7%
Qwen 3.5 35B <sup>MoE</sup>	86.7%
Nemotron 120B <sup>MoE</sup>	60.0%

### Key findings on

- 5x LLMs outperform the human coder predicting observed votes
- Human's caution for "Don't know" for 2x speakers who voted *For* costs accuracy
- The LLMs correctly inferred support

37

## Speech 2: Daniele Polato (ECR)

Mr President, dear colleagues, in this Chamber we have different sensibilities regarding the United States, we know that.

On one side we have ideology, polemics, criticism, 1970s anti-Americanism. On the other side there are us, who want instead to work on what unites us rather than what divides us.

We must send a message of responsibility that gives certainty to our businesses and our workers. Keep your ideologies, polemics and criticism. Those – as we have already seen with the Green Deal – do not take us anywhere, do not consolidate relationships, do not create development.

We want the vote on these agreements to be a signal of a new start [...] But to those who want to exploit or delay this vote, we want to remind them that businesses have no time to waste, especially in this delicate phase in which we live.

Certainty is needed, stability is needed. This is the direction we want to go. This is what businesses and workers are asking of us. This is the right choice for Europe.

**Human GT**  
Support

**LLMs**  
Support (6/6)

38

## Speech 2: Daniele Polato (ECR)

Mr President, dear colleagues, in this Chamber we have different sensibilities regarding the United States, we know that.

On one side we have ideology, polemics, criticism, 1970s anti-Americanism. On the other side there are us, who want instead to work on what unites us rather than what divides us.

We must send a message of responsibility that gives certainty to our businesses and our workers. Keep your ideologies, polemics and criticism. Those – as we have already seen with the Green Deal – do not take us anywhere, do not consolidate relationships, do not create development.

We want the vote on these agreements to be a signal of a new start [...] But to those who want to exploit or delay this vote, we want to remind them that businesses have no time to waste, especially in this delicate phase in which we live.

Certainty is needed, stability is needed. This is the direction we want to go. This is what businesses and workers are asking of us. This is the right choice for Europe.

**Human GT**

Support

**LLMs**

Support (6/6)

**Vote**

For

38

## Speech 3: Martin Schirdewan (The Left)

Mr President! [...] This trade deal disadvantages European industry and it endangers jobs. And in addition to the one-sided tariffs you mentioned, European states are supposed to be obligated to buy massive amounts of liquefied gas and other fossil fuels from the US. So instead of European energy independence, we now have a new dependency on the US – I consider that total madness.

European companies are furthermore supposed to be obligated to invest massively in the US instead of in Europe. That would be like paying ransom to the extortionist. And yet the Supreme Court, the highest court in the US, itself ruled that this trade agreement has no valid legal basis. And just yesterday Trump made clear again that he has absolutely no intention of honouring the tariff part of this agreement.

Trump wants his big tech oligarchs not to have to pay taxes in the EU. He doesn't care about European data protection. He wants to annex Greenland. He threatens Spain. He carries out attacks in violation of international law against Venezuela and Iran.

And even if Commission President Ursula von der Leyen agreed to this act of submission, dear colleagues: you cannot seriously seal such a deal with the madman in the White House. A Parliament that takes itself seriously must never give in to this extortion.

**Human GT**

Against

**LLMs**

Against (6/6)

39

### Speech 3: Martin Schirdewan (The Left)

Mr President! [...] This trade deal disadvantages European industry and it endangers jobs. And in addition to the one-sided tariffs you mentioned, European states are supposed to be obligated to buy massive amounts of liquefied gas and other fossil fuels from the US. So instead of European energy independence, we now have a new dependency on the US – I consider that total madness.

European companies are furthermore supposed to be obligated to invest massively in the US instead of in Europe. That would be like paying ransom to the extortionist. And yet the Supreme Court, the highest court in the US, itself ruled that this trade agreement has no valid legal basis. And just yesterday Trump made clear again that he has absolutely no intention of honouring the tariff part of this agreement.

Trump wants his big tech oligarchs not to have to pay taxes in the EU. He doesn't care about European data protection. He wants to annex Greenland. He threatens Spain. He carries out attacks in violation of international law against Venezuela and Iran.

And even if Commission President Ursula von der Leyen agreed to this act of submission, dear colleagues: you cannot seriously seal such a deal with the madman in the White House. A Parliament that takes itself seriously must never give in to this extortion.

**Human GT**

Against

**LLMs**

Against (6/6)

**Vote**

Against

39

### Speech 6: Enikő Győri (P/E)

Mr President! EUR 4 billion worth of goods change hands between the United States and the Union every day. This also shows how important a balanced partnership based on mutual respect is. However, Ursula von der Leyen struck a humiliating deal with the US last summer. She proved that Europe is weak and can be blackmailed. A severely unequal agreement was reached, which is far from bringing mutual benefits. We used to pay 2–3 per cent tariffs on our exports, now we will pay 15, while the US brings in industrial goods duty-free.

Why did the Commission agree to this humiliating deal? So that America would continue to support Ukraine. [...] All that remains is damage limitation, because European companies need predictability and de-escalation. The Commission should learn from the American President how to effectively stand up for our own interests.

**Human GT**

Against

**LLMs**

Against (6/6)

40

## Speech 6: Enikő Győri (P/E)

Mr President! EUR 4 billion worth of goods change hands between the United States and the Union every day. This also shows how important a balanced partnership based on mutual respect is. However, Ursula von der Leyen struck a humiliating deal with the US last summer. She proved that Europe is weak and can be blackmailed. A severely unequal agreement was reached, which is far from bringing mutual benefits. We used to pay 2–3 per cent tariffs on our exports, now we will pay 15, while the US brings in industrial goods duty-free.

Why did the Commission agree to this humiliating deal? So that America would continue to support Ukraine. [...] All that remains is damage limitation, because European companies need predictability and de-escalation. The Commission should learn from the American President how to effectively stand up for our own interests.

**Human GT**

Against

**LLMs**

Against (6/6)

**Vote**

Abstain

40

## Speech 9: Juan Fernando López Aguilar (S&D)

Mr President, Commissioner, it is true that the relationship with the United States that we know as transatlantic is important for workers and businesses, because it is after all the European Union's first trading partner. And it is also true that the United States is not Trump: it pre-existed him and will survive him.

But it was Trump at Turnberry, at his golf course in Scotland, who tried to impose on the European Union through President Von der Leyen a humiliating agreement. And if President von der Leyen accepts it, she certainly does not represent us, because this Parliament has the last word, according to Article 218 of the Treaty on the Functioning.

But this is not the last word, it is a Resolution that sets the position of the European Parliament binding on the Council and the Commission to guarantee legal certainty and, above all, reciprocity [...]. Therefore, either legal certainty and mutual respect and reciprocity are guaranteed, or the European Parliament will reject it with the vote against of the Socialist Group.

**Human GT**

Don't know

**LLMs**

Against (5/6)

Support (1/6)

41

## Speech 9: Juan Fernando López Aguilar (S&D)

Mr President, Commissioner, it is true that the relationship with the United States that we know as transatlantic is important for workers and businesses, because it is after all the European Union's first trading partner. And it is also true that the United States is not Trump: it pre-existed him and will survive him.

But it was Trump at Turnberry, at his golf course in Scotland, who tried to impose on the European Union through President Von der Leyen a humiliating agreement. And if President von der Leyen accepts it, she certainly does not represent us, because this Parliament has the last word, according to Article 218 of the Treaty on the Functioning.

But this is not the last word, it is a Resolution that sets the position of the European Parliament binding on the Council and the Commission to guarantee legal certainty and, above all, reciprocity [. . .] Therefore, either legal certainty and mutual respect and reciprocity are guaranteed, or the European Parliament will reject it with the vote against of the Socialist Group.

**Human GT**  
Don't know

**LLMs**  
Against (5/6)  
Support (1/6)

**Vote**  
For

41

### PART 5

## Reporting Standards

*Pipeline choices interact unpredictably. Documenting them is not optional.*

42

## Reporting Standards for LLM Annotation

### Report at minimum:

- Model identity and parameters
  - Quantisation method if relevant
  - Sampling hyper-parameters

## Reporting Standards for LLM Annotation

### Report at minimum:

- Model identity and parameters
  - Quantisation method if relevant
  - Sampling hyper-parameters
- Full prompt text or template
  - Learning approach

## Reporting Standards for LLM Annotation

### Report at minimum:

- Model identity and parameters
  - Quantisation method if relevant
  - Sampling hyper-parameters
- Full prompt text or template
  - Learning approach
- Hardware specification
- Efficiency metrics if available

## Reporting Standards for LLM Annotation

### Report at minimum:

- Model identity and parameters
  - Quantisation method if relevant
  - Sampling hyper-parameters
- Full prompt text or template
  - Learning approach
- Hardware specification
- Efficiency metrics if available
- Multiple evaluation metrics
  - Precision, accuracy, F1, intercoder reliability metrics

## Reporting Standards for LLM Annotation

### Report at minimum:

- Model identity and parameters
  - Quantisation method if relevant
  - Sampling hyper-parameters
- Full prompt text or template
  - Learning approach
- Hardware specification
- Efficiency metrics if available
- Multiple evaluation metrics
  - Precision, accuracy, F1, intercoder reliability metrics
- Reproducibility provisions
  - Data pipeline, codebook, code repository etc.

43

## Reporting Standards for LLM Annotation

### Report at minimum:

- Model identity and parameters
  - Quantisation method if relevant
  - Sampling hyper-parameters
- Full prompt text or template
  - Learning approach
- Hardware specification
- Efficiency metrics if available
- Multiple evaluation metrics
  - Precision, accuracy, F1, intercoder reliability metrics
- Reproducibility provisions
  - Data pipeline, codebook, code repository etc.

### Why this matters

- Pipeline choices interact unpredictably
- Omitting details makes cross-study comparison impossible
- **Reproducibility**

43

PART 6

## Horizon Scanning and Wrap-Up

### The Bleeding Edge: Agentic Workflows

**Autonomous multi-step AI pipelines that plan, execute, evaluate, and iterate autonomously**

## The Bleeding Edge: Agentic Workflows

Autonomous multi-step AI pipelines that plan, execute, evaluate, and iterate autonomously

### An agent might:

- Build a corpus
- Develop a draft codebook
- Annotate at scale
- Evaluate its own consistency
- Flag uncertain cases

45

## The Bleeding Edge: Agentic Workflows

Autonomous multi-step AI pipelines that plan, execute, evaluate, and iterate autonomously

### An agent might:

- Build a corpus
- Develop a draft codebook
- Annotate at scale
- Evaluate its own consistency
- Flag uncertain cases

### The framework introduced today becomes even more critical

Without checkpoints, agentic pipelines compound errors

### Open questions:

- Rigour under automation?
- Reporting for agent decisions?
- Reproducibility norms when pipeline generated on the fly?

45

## To conclude

Computational tools are changing how we study institutions

## To conclude

Computational tools are changing how we study institutions

- Significant new opportunities across all stages of the research process
  - Data collection; extraction; cleaning; classification; analysis
  - All of this scales **VERY** quickly

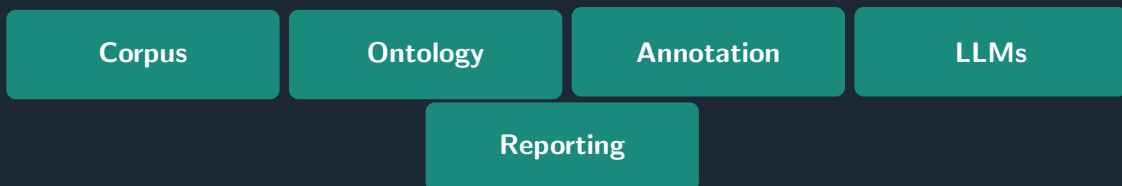
## To conclude

### Computational tools are changing how we study institutions

- Significant new opportunities across all stages of the research process
  - Data collection; extraction; cleaning; classification; analysis
  - All of this scales **VERY** quickly
- BUT also significant new challenges
  - Errors compound over the pipeline
  - Replication and reporting standards must evolve to keep up
  - Human oversight still required (for now)

46

# Thank You



James P. Cross — Connected Politics Lab — UCD

[ucd.ie/connected\\_politics](http://ucd.ie/connected_politics)

46

## APPENDIX

# Selected Speech Classifications

Examples of mismatched and correctly classified cases from the corpus

### Christophe Bay (PfE)

Mr President, the agreement signed by Mrs von der Leyen asks us to endorse a new series of trade concessions in favour of the United States. The elimination of customs duties on American goods, versus 15 per cent customs duties for European products, inevitably leads to disadvantaging the European market, without any serious guarantee of reciprocity and without a safety net.

All this occurs in a perfectly well-known context. The United States has continued to use the tariff weapon as an instrument of political and economic pressure. And faced with this, what does the Commission do? It rushes to implement commitments that weaken our position, our customs revenues and potentially several of our productive sectors.

Moreover, it is not so much the Trump administration that should be blamed, but rather the lack of fighting spirit of the Von der Leyen Commission. We are told that we must accept these concessions to preserve the transatlantic relationship. But preserving a relationship does not mean capitulating. Cooperating does not mean falling into line.

**Human GT**

Against

**LLMs**

Against (6/6)

## Christophe Bay (P/E)

Mr President, the agreement signed by Mrs von der Leyen asks us to endorse a new series of trade concessions in favour of the United States. The elimination of customs duties on American goods, versus 15 per cent customs duties for European products, inevitably leads to disadvantaging the European market, without any serious guarantee of reciprocity and without a safety net.

All this occurs in a perfectly well-known context. The United States has continued to use the tariff weapon as an instrument of political and economic pressure. And faced with this, what does the Commission do? It rushes to implement commitments that weaken our position, our customs revenues and potentially several of our productive sectors.

Moreover, it is not so much the Trump administration that should be blamed, but rather the lack of fighting spirit of the Von der Leyen Commission. We are told that we must accept these concessions to preserve the transatlantic relationship. But preserving a relationship does not mean capitulating. Cooperating does not mean falling into line.

*Entirely critical rhetoric; no signal of voting intention. Voted For despite opposing the deal's terms.*

**Human GT**

Against

**LLMs**

Against (6/6)

**Vote**

For

## Hermann Tertsch (P/E)

Mr President, Commissioner, let me say upfront that VOX does not participate in the anti-American phobia of this Chamber. Europe needs a stable and loyal trade alliance with the United States. There is no real alternative to the great transatlantic pact. Throwing ourselves into China's arms, as many here propose, would be suicide. Unfortunately, what Von der Leyen presents to us as an agreement with the United States cannot be approved: 15 per cent tariffs versus 0 per cent would mean, for Spain, another nail in the coffin of the primary sector [...]

President Trump knows how to negotiate, always with his country's interests in mind. This Commission and its president Von der Leyen do not know how to negotiate and are not concerned with the interests of European nations. [...]

Casting an affirmative vote is applauding Von der Leyen's ineptitude. And that is what we are gambling with.

**Human GT**

Against

**LLMs**

Against (6/6)

## Hermann Tertsch (P/E)

Mr President, Commissioner, let me say upfront that VOX does not participate in the anti-American phobia of this Chamber. Europe needs a stable and loyal trade alliance with the United States. There is no real alternative to the great transatlantic pact. Throwing ourselves into China's arms, as many here propose, would be suicide. Unfortunately, what Von der Leyen presents to us as an agreement with the United States cannot be approved: 15 per cent tariffs versus 0 per cent would mean, for Spain, another nail in the coffin of the primary sector [...]

President Trump knows how to negotiate, always with his country's interests in mind. This Commission and its president Von der Leyen do not know how to negotiate and are not concerned with the interests of European nations. [...]

Casting an affirmative vote is applauding Von der Leyen's ineptitude. And that is what we are gambling with.

*“Cannot be approved” and “casting an affirmative vote is applauding ineptitude” read as Against, but the MEP abstained.*

**Human GT**

Against

**LLMs**

Against (6/6)

**Vote**

Abstain

## Enikő Győri (P/E)

Mr President! EUR 4 billion worth of goods change hands between the United States and the Union every day. This also shows how important a balanced partnership based on mutual respect is. However, Ursula von der Leyen struck a humiliating deal with the US last summer. She proved that Europe is weak and can be blackmailed. A severely unequal agreement was reached, which is far from bringing mutual benefits. We used to pay 2–3 per cent tariffs on our exports, now we will pay 15, while the US brings in industrial goods duty-free.

Why did the Commission agree to this humiliating deal? So that America would continue to support Ukraine. [...] All that remains is damage limitation, because European companies need predictability and de-escalation. The Commission should learn from the American President how to effectively stand up for our own interests.

**Human GT**

Against

**LLMs**

Against (6/6)

## Enikő Győri (P/E)

Mr President! EUR 4 billion worth of goods change hands between the United States and the Union every day. This also shows how important a balanced partnership based on mutual respect is. However, Ursula von der Leyen struck a humiliating deal with the US last summer. She proved that Europe is weak and can be blackmailed. A severely unequal agreement was reached, which is far from bringing mutual benefits. We used to pay 2–3 per cent tariffs on our exports, now we will pay 15, while the US brings in industrial goods duty-free.

Why did the Commission agree to this humiliating deal? So that America would continue to support Ukraine. [...] All that remains is damage limitation, because European companies need predictability and de-escalation. The Commission should learn from the American President how to effectively stand up for our own interests.

*“Humiliating deal” signals opposition, but “damage limitation” and “predictability” hint at reluctant acceptance. The MEP abstained.*

**Human GT**

Against

**LLMs**

Against (6/6)

**Vote**

Abstain

## Juan Fernando López Aguilar (S&D)

Mr President, Commissioner, it is true that the relationship with the United States that we know as transatlantic is important for workers and businesses, because it is after all the European Union's first trading partner. And it is also true that the United States is not Trump: it pre-existed him and will survive him.

But it was Trump at Turnberry, at his golf course in Scotland, who tried to impose on the European Union through President Von der Leyen a humiliating agreement. And if President von der Leyen accepts it, she certainly does not represent us, because this Parliament has the last word, according to Article 218 of the Treaty on the Functioning.

But this is not the last word, it is a Resolution that sets the position of the European Parliament binding on the Council and the Commission to guarantee legal certainty and, above all, reciprocity [...]. Therefore, either legal certainty and mutual respect and reciprocity are guaranteed, or the European Parliament will reject it with the vote against of the Socialist Group.

**Human GT**

Don't know

**LLMs**

Against (5/6)

Support (1/6)

## Juan Fernando López Aguilar (S&D)

Mr President, Commissioner, it is true that the relationship with the United States that we know as transatlantic is important for workers and businesses, because it is after all the European Union's first trading partner. And it is also true that the United States is not Trump: it pre-existed him and will survive him.

But it was Trump at Turnberry, at his golf course in Scotland, who tried to impose on the European Union through President Von der Leyen a humiliating agreement. And if President von der Leyen accepts it, she certainly does not represent us, because this Parliament has the last word, according to Article 218 of the Treaty on the Functioning.

But this is not the last word, it is a Resolution that sets the position of the European Parliament binding on the Council and the Commission to guarantee legal certainty and, above all, reciprocity [. . .] Therefore, either legal certainty and mutual respect and reciprocity are guaranteed, or the European Parliament will reject it with the vote against of the Socialist Group.

*Conditional stance: supports the Parliament's resolution (which adds safeguards) while threatening rejection if conditions are unmet. Human coded Don't know; most LLMs read opposition; the MEP voted For.*

### Human GT

Don't know

### LLMs

Against (5/6)

Support (1/6)

### Vote

For

## Sebastian Tynkkynen (ECR)

Mr Chair, in how many ways does the left try to prevent Europe from succeeding? I don't know. I lost count a long time ago. But one perennial favourite of theirs has always been obstructing US trade. When we buy American fighter jets, it won't do, shouts the left. When we get a trade deal with Trump, it won't do, shouts the left. When Finland recently concluded a massive icebreaker deal where we sell icebreakers to the United States, that didn't seem to work for the left either. The left was already calling for pulling out of the deal.

Apparently we can't even sell our expertise when the counterpart is the US. How about it, left, should we soon also ban watching American films, like in the Soviet Union, whose teachings you still seem to follow in your US policy? But remember, the US won, the Soviet Union lost. It's time for you to switch sides too.

### Human GT

Don't know

### LLMs

Support (5/6)

Don't know

(1/6)

## Sebastian Tynkkynen (ECR)

Mr Chair, in how many ways does the left try to prevent Europe from succeeding? I don't know. I lost count a long time ago. But one perennial favourite of theirs has always been obstructing US trade. When we buy American fighter jets, it won't do, shouts the left. When we get a trade deal with Trump, it won't do, shouts the left. When Finland recently concluded a massive icebreaker deal where we sell icebreakers to the United States, that didn't seem to work for the left either. The left was already calling for pulling out of the deal.

Apparently we can't even sell our expertise when the counterpart is the US. How about it, left, should we soon also ban watching American films, like in the Soviet Union, whose teachings you still seem to follow in your US policy? But remember, the US won, the Soviet Union lost. It's time for you to switch sides too.

*No explicit position on the mandate. The speech attacks opponents of US trade, which most LLMs read as implicit support. The human coder required a direct statement and coded Don't know. The MEP voted For.*

**Human GT**  
Don't know

**LLMs**  
Support (5/6)  
Don't know  
(1/6)

**Vote**  
For

## APPENDIX

### Correctly Classified Speeches

Speech texts where all six LLMs matched the human-coded position and the observed vote

## Daniele Polato (ECR)

Mr President, dear colleagues, in this Chamber we have different sensibilities regarding the United States, we know that.

On one side we have ideology, polemics, criticism, 1970s anti-Americanism. On the other side there are us, who want instead to work on what unites us rather than what divides us.

We must send a message of responsibility that gives certainty to our businesses and our workers. Keep your ideologies, polemics and criticism. Those – as we have already seen with the Green Deal – do not take us anywhere, do not consolidate relationships, do not create development.

We want the vote on these agreements to be a signal of a new start [...] But to those who want to exploit or delay this vote, we want to remind them that businesses have no time to waste, especially in this delicate phase in which we live.

Certainty is needed, stability is needed. This is the direction we want to go. This is what businesses and workers are asking of us. This is the right choice for Europe.

**Human GT**  
Support

**LLMs**  
Support (6/6)

## Daniele Polato (ECR)

Mr President, dear colleagues, in this Chamber we have different sensibilities regarding the United States, we know that.

On one side we have ideology, polemics, criticism, 1970s anti-Americanism. On the other side there are us, who want instead to work on what unites us rather than what divides us.

We must send a message of responsibility that gives certainty to our businesses and our workers. Keep your ideologies, polemics and criticism. Those – as we have already seen with the Green Deal – do not take us anywhere, do not consolidate relationships, do not create development.

We want the vote on these agreements to be a signal of a new start [...] But to those who want to exploit or delay this vote, we want to remind them that businesses have no time to waste, especially in this delicate phase in which we live.

Certainty is needed, stability is needed. This is the direction we want to go. This is what businesses and workers are asking of us. This is the right choice for Europe.

**Human GT**  
Support

**LLMs**  
Support (6/6)

**Vote**  
For

*Clear pro-deal framing throughout: "certainty is needed", "the right choice for Europe". All six models and the human coder agreed on Support.*

## Billy Kelleher (Renew)

Mr President, in a very volatile world, and particularly in a volatile relationship between the United States and the European Union, it is critically important that we play our part in trying to stabilise that relationship.

While we do hear ideological debates inside this Chamber from the left and from the right, in the context of President Trump, I think we should be looking at it through the eyes of workers and businesses across the European Union. They need certainty in a very uncertain time.

EUR 4.6 billion is traded daily across the Atlantic in goods and services, and anything that gives confidence to this particular trade agreement I very much welcome.

I also believe that we have to be firm in saying to President Trump and the US administration that respect is critically important [ . . . ] Time will tell, but in the meantime, we must play our part to ensure that we stand firm with regard to what was agreed in Turnberry and that we insist, from the United States's perspective, that they also honour the commitments [ . . . ]

On the issue of steel and aluminium, yes, there is a need for us to continue to negotiate in those particular areas, but at the moment certainty is what is required for businesses and workers.

**Human GT**

Support

**LLMs**

Support (6/6)

## Billy Kelleher (Renew)

Mr President, in a very volatile world, and particularly in a volatile relationship between the United States and the European Union, it is critically important that we play our part in trying to stabilise that relationship.

While we do hear ideological debates inside this Chamber from the left and from the right, in the context of President Trump, I think we should be looking at it through the eyes of workers and businesses across the European Union. They need certainty in a very uncertain time.

EUR 4.6 billion is traded daily across the Atlantic in goods and services, and anything that gives confidence to this particular trade agreement I very much welcome.

I also believe that we have to be firm in saying to President Trump and the US administration that respect is critically important [ . . . ] Time will tell, but in the meantime, we must play our part to ensure that we stand firm with regard to what was agreed in Turnberry and that we insist, from the United States's perspective, that they also honour the commitments [ . . . ]

On the issue of steel and aluminium, yes, there is a need for us to continue to negotiate in those particular areas, but at the moment certainty is what is required for businesses and workers.

**Human GT**

Support

**LLMs**

Support (6/6)

**Vote**

For

*Supportive framing anchored in certainty for workers and businesses, with explicit welcome for the agreement. Unambiguous for both human coders and LLMs.*

## Martin Schirdewan (The Left)

Mr President! [...] This trade deal disadvantages European industry and it endangers jobs. And in addition to the one-sided tariffs you mentioned, European states are supposed to be obligated to buy massive amounts of liquefied gas and other fossil fuels from the US. So instead of European energy independence, we now have a new dependency on the US – I consider that total madness.

European companies are furthermore supposed to be obligated to invest massively in the US instead of in Europe. That would be like paying ransom to the extortionist. And yet the Supreme Court, the highest court in the US, itself ruled that this trade agreement has no valid legal basis. And just yesterday Trump made clear again that he has absolutely no intention of honouring the tariff part of this agreement.

Trump wants his big tech oligarchs not to have to pay taxes in the EU. He doesn't care about European data protection. He wants to annex Greenland. He threatens Spain. He carries out attacks in violation of international law against Venezuela and Iran.

And even if Commission President Ursula von der Leyen agreed to this act of submission, dear colleagues: you cannot seriously seal such a deal with the madman in the White House. A Parliament that takes itself seriously must never give in to this extortion.

**Human GT**

Against

**LLMs**

Against (6/6)

## Martin Schirdewan (The Left)

Mr President! [...] This trade deal disadvantages European industry and it endangers jobs. And in addition to the one-sided tariffs you mentioned, European states are supposed to be obligated to buy massive amounts of liquefied gas and other fossil fuels from the US. So instead of European energy independence, we now have a new dependency on the US – I consider that total madness.

European companies are furthermore supposed to be obligated to invest massively in the US instead of in Europe. That would be like paying ransom to the extortionist. And yet the Supreme Court, the highest court in the US, itself ruled that this trade agreement has no valid legal basis. And just yesterday Trump made clear again that he has absolutely no intention of honouring the tariff part of this agreement.

Trump wants his big tech oligarchs not to have to pay taxes in the EU. He doesn't care about European data protection. He wants to annex Greenland. He threatens Spain. He carries out attacks in violation of international law against Venezuela and Iran.

And even if Commission President Ursula von der Leyen agreed to this act of submission, dear colleagues: you cannot seriously seal such a deal with the madman in the White House. A Parliament that takes itself seriously must never give in to this extortion.

**Human GT**

Against

**LLMs**

Against (6/6)

**Vote**

Against

*Unambiguous opposition: "total madness", "act of submission", "must never give in to this extortion". No room for interpretive disagreement.*