

A wide-angle photograph of the main building of TU Wien, a large, white, classical-style building with a green roof and many windows. The building is surrounded by green trees in the foreground. The sky is clear and blue.

Propaganda detection in Online-News (Narrative Classification)

Team Backpropagandists

Members: Sergej Bensack, Jonas Kruse, Niklas Kessler, Leon Baumgärtner

Content

- Task introduction
- Data preprocessing
- Description of used methods
 - SVM
 - Keyword matching
 - BERT
 - LLAMA
- Results

Task Introduction



“Climate cycles are natural”

Basics

Which narratives are contained in a given text sequence?




→ Multilabel, Multiclass Classification Task

- Fixed set of narratives
 - Narrative - Subnarrative
 - Grouped by topics (Ukraine War, Climate Change)
- Can be used for propaganda detection

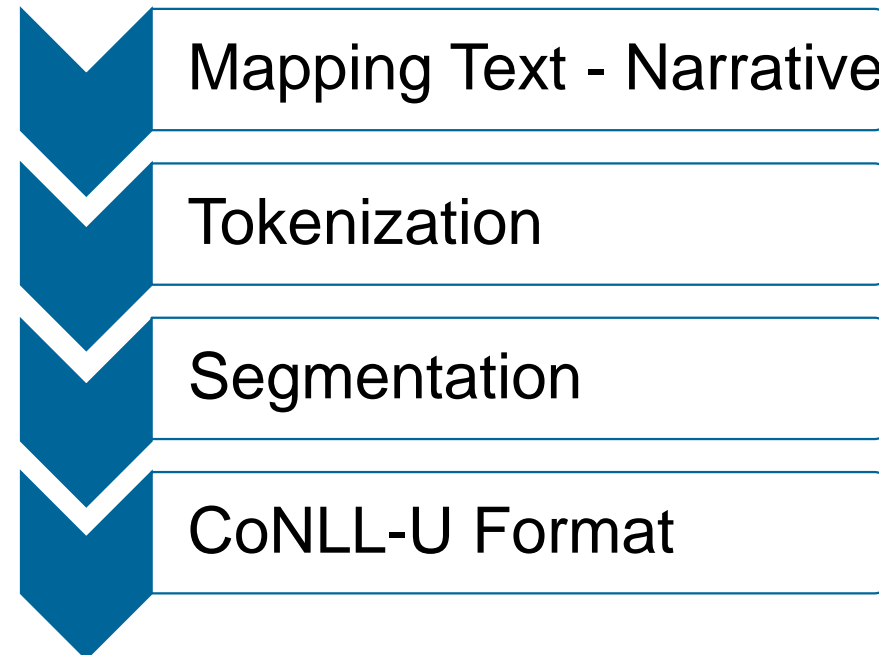
Dataset

- Languages:
 - English
 - Russian
 - Bulgarian
 - Portuguese
 - Hindi
- ~ 1700 instances
- 94 classes

Preprocessing

 EN_CC_100000	"The Washington Free Beacon reported that..."
 EN_CC_100002	...
 EN_CC_100003	...

annotations.txt	
EN_CC_100000.txt	CC: Criticism of climate movement
EN_CC_100002.txt	Other: Other
EN_UA_300209.txt	...



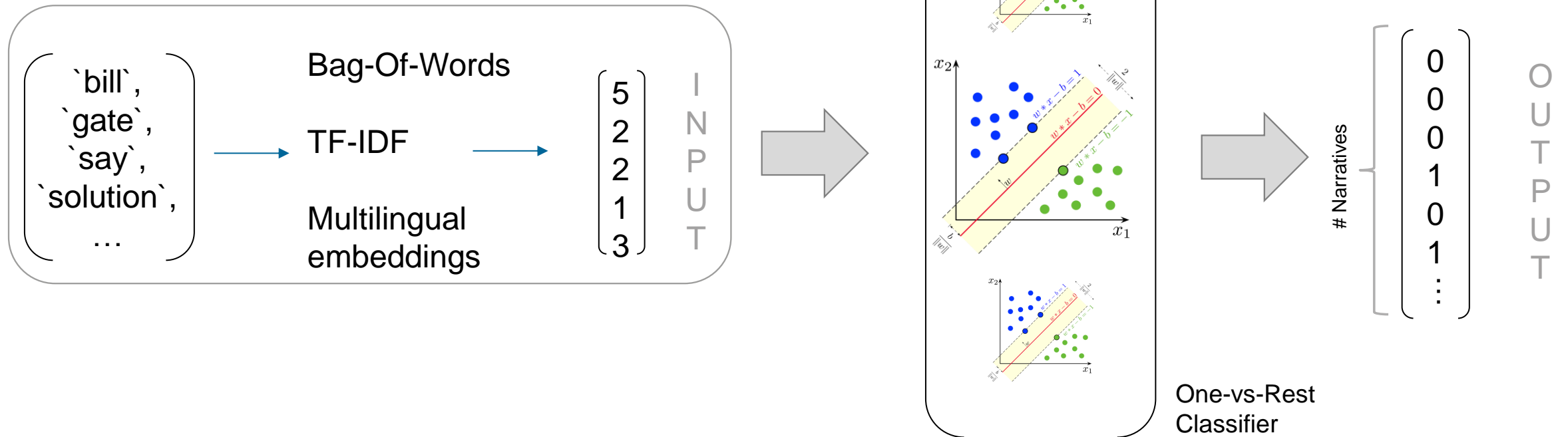
Processed Data

							INPUT	OUTPUT
	filename	language	content	topic	narrative_subnarrative_pairs	tokens	tokens_normalized	target_indices
0	EN_CC_100013.txt	EN	Bill Gates Says He Is 'The Solution' To Climat...	CC	[{'narrative': 'Criticism of climate movement'...	[['Bill', 'Gates', 'Says', 'He', 'Is', '', 'T...	['bill', 'gate', 'say', 'solution', 'climate'],...	[14]
1	EN_UA_300009.txt	EN	Russia: Clashes erupt in Bashkortostan as righ...	UA	[{'narrative': 'Other', 'subnarrative': 'Other'}]	[['Russia', ':', 'Clashes', 'erupt', 'in', 'Ba...	['russia', 'clash', 'erupt', 'bashkortostan', ...	[32]
2	EN_UA_300017.txt	EN	McDonald's to exit Russia, sell business in co...	UA	[{'narrative': 'Other', 'subnarrative': 'Other'}]	[['McDonald', '"s", 'to', 'exit', 'Russia', ', ...	['mcdonald', 'exit', 'russia', 'sell', 'busine...	[32]
3	EN_CC_100021.txt	EN	Collaborative plans, innovation keys to circul...	CC	[{'narrative': 'Other', 'subnarrative': 'Other'}]	[['Collaborative', 'plans', ':', 'innovation', ...	['collaborative', 'plan', 'innovation', 'key'],...	[42]
4	EN_UA_300041.txt	EN	Russia intends to supply light 'Mountain' tank...	UA	[{'narrative': 'Other', 'subnarrative': 'Other'}]	[['Russia', 'intends', 'to', 'supply', 'light'...	['russia', 'intend', 'supply', 'light', 'mount...	[32]
...

Methods used (1/2)

- SVM
- Keyword Matching
- BERT
- LLAMA

SVM



Vectorization Methods

■ Bag-of-Words

am:	3
diplomacy:	1
emissions:	0
is:	4
flooding:	1
that:	4

Custom Implementation

■ TF-IDF

$$TF = \frac{\text{word count}}{\text{total word count}}$$

$$IDF = \log \left(\frac{\text{total docs}}{\text{docs containing } w} \right)$$

$$TF-IDF = TF \times IDF$$

Sklearn

■ Multilanguage Embeddings

$\begin{bmatrix} 0, \\ 1, \\ 1, \\ 0, \\ 1 \end{bmatrix}$

Google: Bert (base, uncased)

Keyword Matching

Example Predictions:

- Test 1:
 - Actual: ukbased climate **activist** deflate tyre suv country group...
 - Predicted: west silence **activist** home promote right abroad report west silence activist...
- Test 2:
 - Actual: animal react total solar eclipse...
 - Predicted: molotov cocktail attack russian embassy moldova last day election...

Class	Precision	Recall	F1 Score	Accuracy
Keyword matching	0.0	0.0	0.0	0.0

Methods used (2/2)

- SVM
- Keyword Matching
- BERT
 - Input representation : texts + narratives + labels
 - Minimum examples per class
- LLAMA
 - Input representation : textxs + labels
 - Pre training of classification head

Results

LLAMA Quantative Results

Class	Precision	Recall	F1 Score	Accuracy
Other-Other	0.6393	0.9750	0.5670	0.5670
Any other class	0.0	0.0	0.0	0.0
Overall	0.5480	0.5480	0.5480	0.5480



Considerations for improvement

- Different approaches
 - All topics in one
 - Split by topic Climate Change/ Ukraine War
- Classes very imbalanced (only english texts)
 - Solution: adding further languages
- Analyzing English and Russian news articles

Class	Occurences
Other-Other	12,9 %
Narrative-subnarrative 1	2,08 %
Narrative-subnarrative 2	1,45 %

Results BERT: Quantitative overview

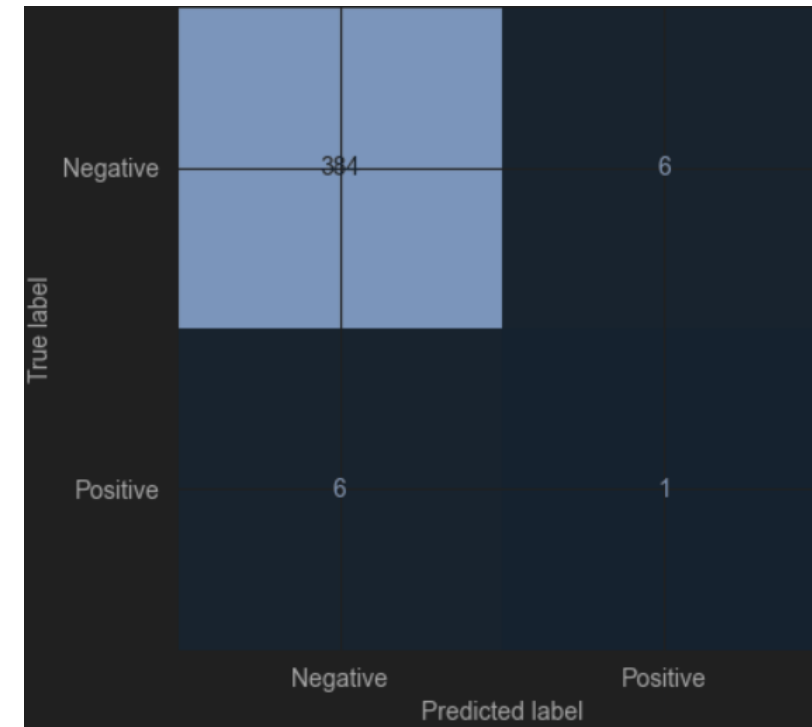
Used languages	Only English			
Metrics/ Class label	Accuracy	Precision	Recall	F1 Score
Other-Other	0,42	0,42	1	0,59
Other class	0,95	0	0	0

Narrative: Hidden plots by secret schemes of powerful groups

subnarrative: Climate agenda has hidden motives

Qualitative Analysis

- "Under the guise of 'fighting climate change,' the regime has now proposed new regulations [...] that target [...] dishwashers"
- "Even though most dishwashers available on the market already meet or exceed the [...] standards, the proposed rules have left some questioning why the Biden administration is focusing on suchs regulations rather than other priorities"

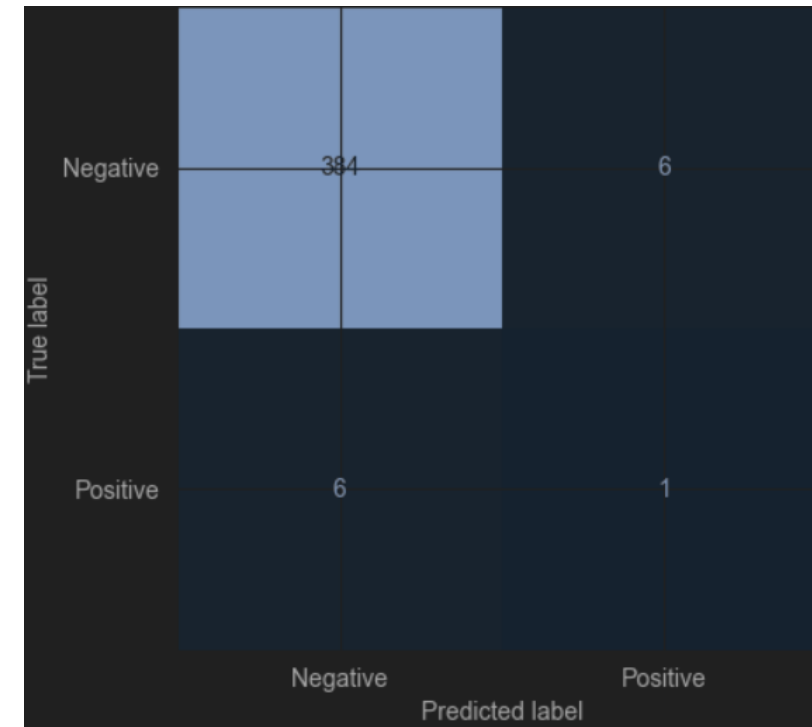


Narrative: Hidden plots by secret schemes of powerful groups

subnarrative: Climate agenda has hidden motives

Qualitative Analysis ✖

- "Caroline van der Plas founded this movement in 2019 and is going along with the nitrogen hoax, climate change lies and biodiversity made up bullshit [...]"
- "Accept Jesus Christ as our saviour"
- "Read the bible fast and pray"
- "Amen"
- " 🙏❤️ "



Results SVM: Quantitative Overview

Class occurrence weighted average on Test set

Method	Dataset	F1-Score	Precision	Recall
bow	cc	0.250	0.276	0.253
	compl	0.141	0.235	0.120
	ua	0.167	0.258	0.136
Mult. Embeddings	cc	0.285	0.335	0.267
	compl	0.206	0.351	0.166
	ua	0.207	0.365	0.157
tfidf	cc	0.264	0.381	0.231
	compl	0.225	0.438	0.171
	ua	0.209	0.436	0.146

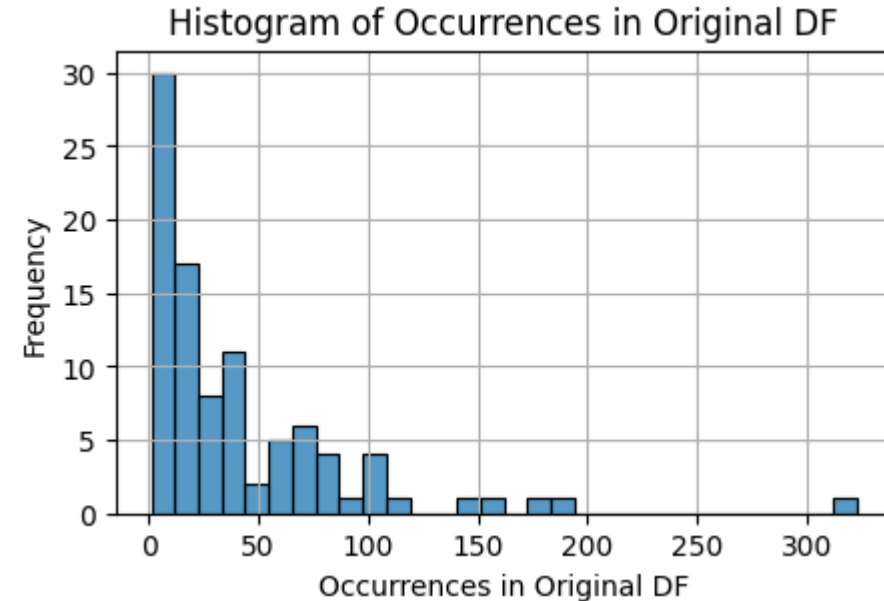


- Complete dataset approach is the most interesting to us (generalisability)
- Best method regarding f1-score for the complete dataset: tfidf
- Note that with macro average f1-scores, we are in the 10% range

SVM (tfidf, complete dataset): Class imbalance

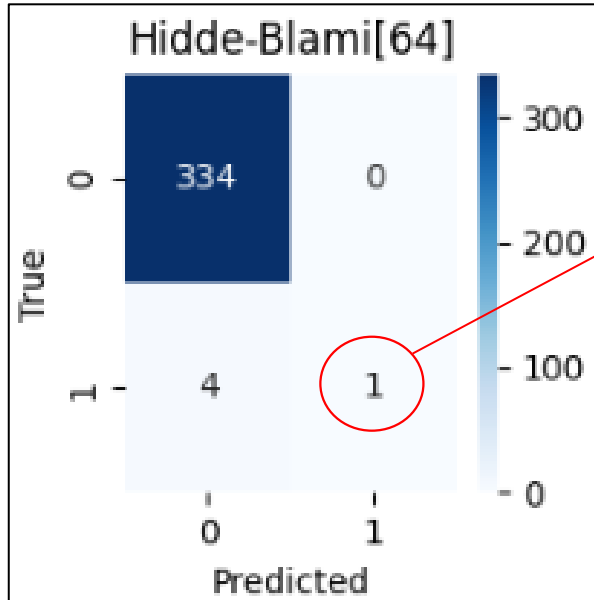
Narrative Index	Occurrences Train Set	Occurrences Test Set	Occurrences Preds
1	149	29	29
86	59	14	4
31	6	3	0

Median of class occurrence in total data: 22.5



- Model requires somewhere between 12 and 35 training examples of a class to start predicting it as positive
- For many classes, this is not given

Qualitative Analysis for a specific class



{*narrative*:

'Hidden plots by secret schemes of powerful groups',

{*subnarrative*:

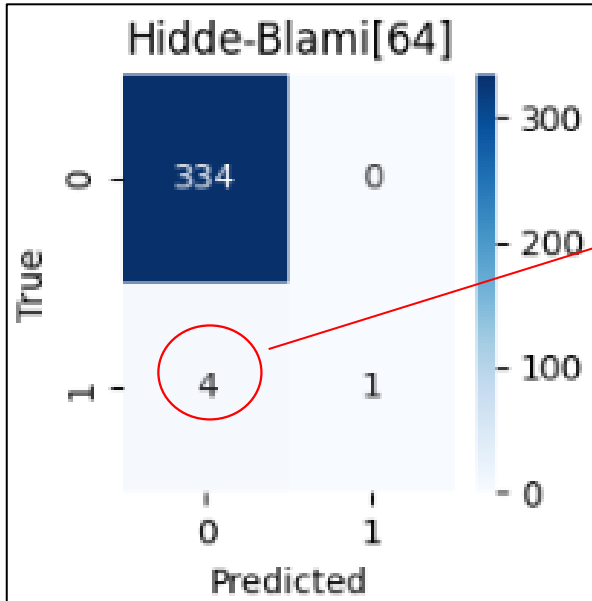
'Blaming global elites'}

NO MORE AIR TRAVEL

- Air travelers ... will need to take rail or a car instead, as these are considered by the **globalists** to be much better forms of transport for the climate.
- By 2029, these climate **extremists** want people ... to reduce their meat consumption by 50 percent.
- "It is possible that the events ... are part of a larger, more **sinister agenda** that is only just beginning,"....
- ... stopping climate change and global warming is code for mass **depopulation** and **genocide** ...

- Several potential high-impact words and explicit language: globalist, extremist, agenda, sinister, depopulation, genocide
- Globalists being the "powerful groups"/"global elites"

Qualitative Analysis for a specific class



{*narrative*:

'Hidden plots by secret schemes of powerful groups',

{*subnarrative*:

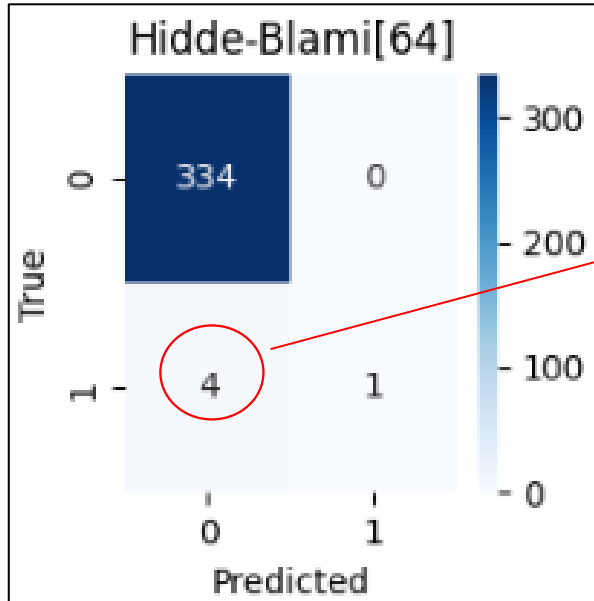
'Blaming global elites'}

CHICKEN FOOD POISONING

- If they are willing to **poison** the chickens to stop them from laying eggs, they are likely poisoning the HUMAN food supply to stop us from having children.
- Not only are hens unable to lay eggs ... after consuming **globalist**-manufactured tainted feed, but could it be possible those chickens and eggs are ... causing infertility issues for [humans] also? Local egg suppliers and farmers are blowing the whistle
- Every purchase at HealthRangerStore.com helps fund our efforts to build and share more tools for empowering humanity with knowledge and abundance.
- Ever since October, several American egg farmers have experienced ZERO egg production from their chickens, until they switched their feed to a different kind, some even going with goat feed

- Potential high-impact words: globalist, poison, ...
- Again, globalists being the "powerful groups"/"global elites"
- Article also contains unrelated content such as self-advertisement and much content about farming/food

Qualitative Analysis for a specific class



{*narrative*:

'Hidden plots by secret schemes of powerful groups',

{*subnarrative*:

'Blaming global elites'}

OXFORD RESIDENTS MOUNT RESISTANCE

- Oxford residents are taking matters into their own hands and destroying the street zone sectioning barriers. Oxford is going to be the centre of the resistance to the **climate change madness** in the UK.
- We predicted the UK was just getting started with their resistance Taking matters into their own hands, Oxford residents are destroying street zoning barriers.
- Further action: There are further planned dates for activists to take to the streets in various forms of protest. Piers Corbyn and friends are taking to the streets Then an event which has potential to draw a large crowd

- Potential high-impact words: climate change and madness
- Very short and no explicit mention of groups or elites, rather "the UK" is addressed
- Even for a human, it's debatable if global elites are being blamed (implicit)

A wide-angle photograph of the TU Wien main building, a large, white, classical-style structure with a green copper roof and many windows. The building is surrounded by lush green trees in the foreground. In the background, the city of Vienna and distant mountains are visible under a clear blue sky.

Thanks for your attention!

Discussion/Questions