

PopAut: An Annotated Corpus for Populism Detection in Austrian News Comments

THE 2024 JOINT INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS, LANGUAGE
RESOURCES AND EVALUATION
20-25 MAY, 2024 / TORINO, ITALIA

Ahmadou Wagne
Julia Neidhardt
Thomas E. Kolb

CDL - RecSys @ TU Wien
recsys-lab@ec.tuwien.ac.at
<https://recsys-lab.at/>

Joint International Conference on Computational Linguistics, Language Resources and Evaluation 2024 | 20-25 May 2024, Torino, Italy

PopAut: An Annotated Corpus for Populism Detection in Austrian News Comments

Ahmadou Wagne^{1,2}, Julia Neidhardt^{1,2}, Thomas E. Kolb^{1,2}

Motivation

- Political parties sharing populist views have achieved electoral success in many countries worldwide
- Presence of populist talking points expressed by citizens during the COVID-19 pandemic
- Citizens as populist actors have received limited attention in research
- Providing automated methods for large scale content analysis of populist user comments

Problem

- Prior research mostly focuses on the detection of populist messages expressed by politicians
- Existing methods to detect populist user comments enclose dictionary-based methods [4], while machine-learning is underrepresented
- Lack of (German) annotated populist user comments

Populism				
Definition	Anti-Elitism	People-Sovereignty	People-Centrism	
<p>Ideational definition of populism Cas Mudde - The Populist Zeitgeist 2004 [8]:</p> <p>An ideology that considers society to be ultimately separated into two homogeneous and antagonistic groups, „the pure people“ versus „the corrupt elite, and which argues that politics should be an expression of the volonté général (general will) of the people.</p>	<p>Key messages [1]:</p> <ul style="list-style-type: none"> • Discrediting the elites • Blaming the elites (for negative developments) • Detaching the elite from the people 	<p>Key messages [1]:</p> <ul style="list-style-type: none"> • Demanding popular sovereignty • Denying elite sovereignty 	<p>Key messages [1]:</p> <ul style="list-style-type: none"> • Stressing the people's virtues • Praising the people's achievements • Stating a monolithic people 	<p>Operationalization of ideational populism [2]</p>

Populism

Definition

Ideational definition of populism
Cas Mudde - The Populist Zeitgeist 2004 [8]:

An ideology that considers society to be ultimately separated into two homogeneous and antagonistic groups, „the pure people“ versus „the corrupt elite, and which argues that politics should be an expression of the volonté général (general will) of the people.

Anti-Elitism

Key messages [1]:

- Discrediting the elites
- Blaming the elites (for negative developments)
- Detaching the elite from the people

People-Sovereignty

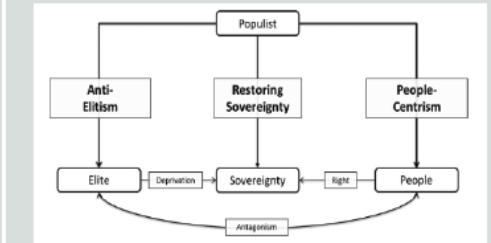
Key messages [1]:

- Demanding popular sovereignty
- Denying elite sovereignty

People-Centrism

Key messages [1]:

- Stressing the people's virtues
- Praising the people's achievements
- Stating a monolithic people



Operationalization of ideational populism [2]

Populism

Definition

Ideational definition of populism
Cas Mudde - The Populist Zeitgeist 2004 [8]:

An ideology that considers society to be ultimately separated into two homogeneous and antagonistic groups, „the pure people“ versus „the corrupt elite, and which argues that politics should be an expression of the volonté générale (general will) of the people.

Anti-Elitism

Key messages [1]:

- Discrediting the elites
- Blaming the elites (for negative developments)
- Detaching the elite from the people

People-Sovereignty

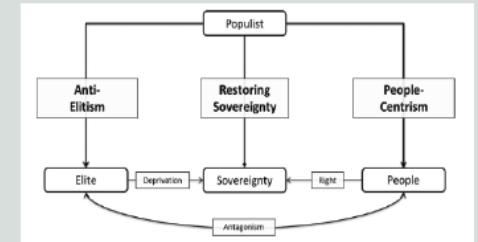
Key messages [1]:

- Demanding popular sovereignty
- Denying elite sovereignty

People-Centrism

Key messages [1]:

- Stressing the people's virtues
- Praising the people's achievements
- Stating a monolithic people



Operationalization of ideational populism [2]

Populism

Definition

Ideational definition of populism
Cas Mudde - The Populist Zeitgeist 2004 [8]:

An ideology that considers society to be ultimately separated into two homogeneous and antagonistic groups, „the pure people“ versus „the corrupt elite, and which argues that politics should be an expression of the volonté général (general will) of the people.

Anti-Elitism

Key messages [1]:

- Discrediting the elites
- Blaming the elites (for negative developments)
- Detaching the elite from the people

People-Sovereignty

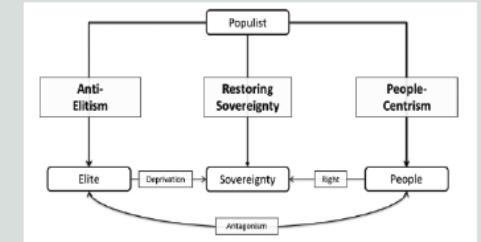
Key messages [1]:

- Demanding popular sovereignty
- Denying elite sovereignty

People-Centrism

Key messages [1]:

- Stressing the people's virtues
- Praising the people's achievements
- Stating a monolithic people



Operationalization of ideational populism [2]

Populism

Definition

Ideational definition of populism
Cas Mudde - The Populist Zeitgeist 2004 [8]:

An ideology that considers society to be ultimately separated into two homogeneous and antagonistic groups, „the pure people“ versus „the corrupt elite, and which argues that politics should be an expression of the volonté général (general will) of the people.

Anti-Elitism

Key messages [1]:

- Discrediting the elites
- Blaming the elites (for negative developments)
- Detaching the elite from the people

People-Sovereignty

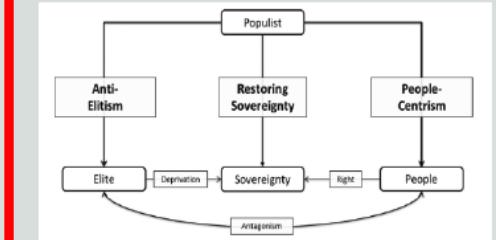
Key messages [1]:

- Demanding popular sovereignty
- Denying elite sovereignty

People-Centrism

Key messages [1]:

- Stressing the people's virtues
- Praising the people's achievements
- Stating a monolithic people

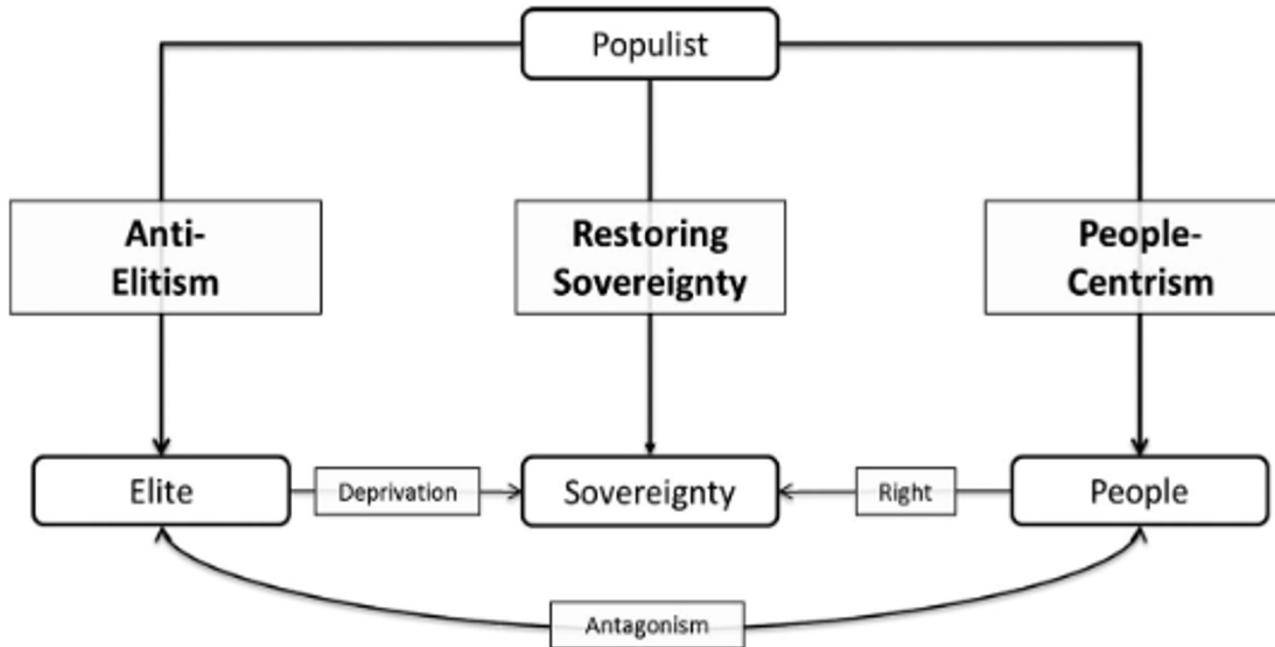


Populism

Definition

Ideational definition of populism
Cas Mudde - The Populist Zeitgeist 2004 [8]:

An ideology that considers society to be ultimately divided into two homogeneous and antagonistic groups: "the pure people" versus "the corrupt elite", and which claims that politics should be an expression of the volontariat (general will) of the people.



Operationalization of ideational populism [2]

Literature

Populism:

- Ideational populism as a „thin“ ideology that can coexist with and borrow from other ideologies [8]
- Other definitions: political strategy [12], communication style [7]
- Growing consensus on motives anti-elitism, people-centrism and people-sovereignty [1, 10, 13]
- Operationalization of populism by defining key messages conveyed by populist statements [1]

Citizens as Populist Actors:

- Most literature focuses on politicians or media as populist actors [1, 9]
- „Participatory populism“ by citizens observed in the online sphere and social media [3]

Populism Detection:

- Early studies rely on manual coding or expert-surveys [6]
- Automated methods based on dictionaries [3, 9]; application of machine-learning focus on party manifestos or political speeches [5, 11]

Data

- Data from January 2019 until November 2021
- Background: application in a case study that compares the prevalence of populist user comments in the context of the COVID-19 pandemic
- Sample is divided in a reference sample containing comments from 2019, a COVID-19 sample with comments under COVID-19 related articles and a non-COVID-19 article with comments posted under all other articles during the pandemic
- Populism dictionary [4] used to identify potentially populist comments (populism considered a rare phenomenon)

	Reference	COVID-19	Non-COVID
Comments	9,683,153	14,391,704	16,885,895
Articles	49,421	14,135	68,642
Users	55,780	64,401	82,354

Statistics of the full samples

Score	Number of comments
0	587
1	12
2	551
3	46
4	4

Number of dictionary terms in the annotation sample of size 1,200

DER STANDARD

<https://www.derstandard.at>

Annotation Study

- Five participants with different academic backgrounds
- Guidelines based on ideational definition [8] and prior work by Ernst et al. [1]
- Multiple choice annotation of populist motives: anti-elitism, people-sovereignty, people-centrism, none
- Binary populism label assigned, if one motive is present -> majority vote
- Metric: Krippendorff's alpha for inter-annotator agreement

Group 5

Trugschluß... das hat im Abendland Tradition und wir sind nie davon abgelenkt.
Die Menschheit kam nur zum Fehlschluß, dass, dadurch, dass sich gewisse Imperien durch die faschistischen Achsenmächte eingeschränkt und bedroht fühlten und sie letztlich aufreichten, wie eigentlich den Faschismus bekämpfen.
Dem war nur nie so.
Der Faschismus ist von oben herab immer gerne gesehen, passt er doch perfekt ins Konzept von Kapitalismus UND "divide et impera"

Check all that apply

None
 People-Sovereignty
 Anti-Elitism
 People-Centrism

Previous Next

Literature

Populism:

- Ideational populism as a „thin“ ideology that can coexist with and borrow from other ideologies [8]
- Other definitions: political strategy [12], communication style [7]
- Growing consensus on motives anti-elitism, people-centrism and people-sovereignty [1, 10, 13]
- Operationalization of populism by defining key messages conveyed by populist statements [1]

Citizens as Populist Actors:

- Most literature focuses on politicians or media as populist actors [1, 9]
- „Participatory populism“ by citizens observed in the online sphere and social media [3]

Populism Detection:

- Early studies rely on manual coding or expert-surveys [6]
- Automated methods based on dictionaries [3, 9]; application of machine-learning focus on party manifestos or political speeches [5, 11]

Data

- Data from January 2019 until November 2021
- Background: application in a case study that compares the prevalence of populist user comments in the context of the COVID-19 pandemic
- Sample is divided in a reference sample containing comments from 2019, a COVID-19 sample with comments under COVID-19 related articles and a non-COVID-19 article with comments posted under all other articles during the pandemic
- Populism dictionary [4] used to identify potentially populist comments (populism considered a rare phenomenon)

	Reference	COVID-19	Non-COVID
Comments	9,683,153	14,391,704	16,885,895
Articles	49,421	14,135	68,642
Users	55,780	64,401	82,354

Statistics of the full samples

Score	Number of comments
0	587
1	12
2	551
3	46
4	4

Number of dictionary terms in the annotation sample of size 1,200

DER STANDARD

<https://www.derstandard.at>

Annotation Study

- Five participants with different academic backgrounds
- Guidelines based on ideational definition [8] and prior work by Ernst et al. [1]
- Multiple choice annotation of populist motives: anti-elitism, people-sovereignty, people-centrism, none
- Binary populism label assigned, if one motive is present -> majority vote
- Metric: Krippendorff's alpha for inter-annotator agreement

Group 5

Trugschluß... das hat im Abendland Tradition und wir sind nie davon abgelenkt.
Die Menschheit kann nur zum Fehlschluß, dass, dadurch, dass sich gewisse Imperien durch die faschistischen Achsenmächte eingeschränkt und bedroht fühlen und sie letztlich ausradieren, wie eigentlich den Faschismus bekämpfen.
Dem war nur nie so.
Der Faschismus ist von oben herab immer gerne gesehen, passt er doch perfekt ins Konzept von Kapitalismus UND "divide et impera"

Check all that apply

None
 People-Sovereignty
 Anti-Elitism
 People-Centrism

Previous Next

Literature

Populism:

- Ideational populism as a „thin“ ideology that can coexist with and borrow from other ideologies [8]
- Other definitions: political strategy [12], communication style [7]
- Growing consensus on motives anti-elitism, people-centrism and people-sovereignty [1, 10, 13]
- Operationalization of populism by defining key messages conveyed by populist statements [1]

Citizens as Populist Actors:

- Most literature focuses on politicians or media as populist actors [1, 9]
- „Participatory populism“ by citizens observed in the online sphere and social media [3]

Populism Detection:

- Early studies rely on manual coding or expert-surveys [6]
- Automated methods based on dictionaries [3, 9]; application of machine-learning focus on party manifestos or political speeches [5, 11]

Data

- Data from January 2019 until November 2021
- Background: application in a case study that compares the prevalence of populist user comments in the context of the COVID-19 pandemic
- Sample is divided in a reference sample containing comments from 2019, a COVID-19 sample with comments under COVID-19 related articles and a non-COVID-19 article with comments posted under all other articles during the pandemic
- Populism dictionary [4] used to identify potentially populist comments (populism considered a rare phenomenon)

	Reference	COVID-19	Non-COVID
Comments	9,683,153	14,391,704	16,885,895
Articles	49,421	14,135	68,642
Users	55,780	64,401	82,354

Statistics of the full samples

Score	Number of comments
0	587
1	12
2	551
3	46
4	4

Number of dictionary terms in the annotation sample of size 1,200

DER STANDARD

<https://www.derstandard.at>

Annotation Study

- Five participants with different academic backgrounds
- Guidelines based on ideational definition [8] and prior work by Ernst et al. [1]
- Multiple choice annotation of populist motives: anti-elitism, people-sovereignty, people-centrism, none
- Binary populism label assigned, if one motive is present -> majority vote
- Metric: Krippendorff's alpha for inter-annotator agreement

Group 5

Trugschluß... das hat im Abendland Tradition und wir sind nie davon abgelenkt.
Die Menschheit kann nur zum Fehlschluß, dass, dadurch, dass sich gewisse Imperien durch die faschistischen Achsenmächte eingeschränkt und bedroht fühlen und sie letztlich aufrechterhalten, wie eigentlich den Faschismus bekämpfen.
Dem war nur nie so.
Der Faschismus ist von oben herab immer gerne gesehen, passt er doch perfekt ins Konzept von Kapitalismus UND "divide et impera".

Check all that apply

None
 People-Sovereignty
 Anti-Elitism
 People-Centrism

[Previous](#) [Next](#)

Literatur

Populism:

- Ideational populism coexist with and both
- Other definitions: political communication style [7]
- Growing consensus on ple-centrism and people
- Operationalization of messages conveyed

Citizens as Populist Agents

- Most literature focuses on populist actors [1, 9]
- „Participatory populism“ in the online sphere also

Populism Detection:

- Early studies rely on surveys [6]
- Automated methods: application of machine learning to manifestos or political discourse

Group 5

Trugschluß... das hat im Abendland Tradition und wir sind nie davon abgerückt.

Die Menschheit kam nur zum Fehlschluß, dass, dadurch, dass sich gewisse Imperien durch die faschistischen Achsenmächte eingeschränkt und bedroht fühlten und sie letztlich ausradierten, sie eigentlich den Faschismus bekämpften.

Dem war nur nie so.

Der Faschismus ist von oben herab immer gerne gesehen, passt er doch perfekt ins Konzept von Kapitalismus UND ""divide et impera""

ⓘ Check all that apply

None

People-Sovereignty

Anti-Elitism

People-Centrism

[Previous](#)

[Next](#)

lemic backgrounds
nition [8] and prior

list motives: anti-elitism, none
ne motive is present ->

r-annotator agreement

5

faschistischen Achsenmächte eingeschränkt und bedroht fühlten und sie letztlich aus-

in Kapitalismus UND ""divide et impera""

Next

Examples

Leider hat die Elte die Umverteilung von unten nach oben forciert, wodurch wir heute in diese präkaren Arbeitsverhältnisse getrieben wurden. Die Politiker können dies (1) sofort wieder ändern.

Warum sie das nicht tun? Weil Sie die Befehle und Wünsche dieser Eliten ausführen und nicht mehr dem Volk dienen.

Es ist Zeit für einen Systemwechsel!

Das Volk will nicht impfen, das Volk will lieber Maske tragen und testen ! (3)

Da kann die Regierung am Kopf stehen und mit den Beinen Fliegen fangen, das Volk macht, was das Volk will !

Und wählt Mit freundlichen Grüßen ...

Wenn auch unwahrscheinlich, so hoffe ich doch, dass meine Zeilen den Weg zu Euch, meine geschätzten VOLKsvertreter finden. Wenn auch unwahrscheinlich, so biete (2) ich Euch ein persönliches Treffen an, indem Ihr eine Stimme aus genau diesem VOLK, welches Eure „Hausverstand-Ergüsse“ finanziert, hören könnt. Denn es ist uns allen bewusst, dass die leisen Stimmen Eures Volkes nur im Wahljahr zu hören sind. Wenn auch dieses persönliche Treffen höchst unwahrscheinlich ist, so ließe ich mich gerne belehren, warum EURE Hausverstand besser ist als Unserer;

- (1) Anti-Elitism
- (2) People-Sovereignty
- (3) People-Centrism

Methods: Populism Detection

- Data split into 800 training, 200 validation and 200 test samples
- Baselines: Dictionary (Gründl) [3], dictionary (R&P) [9], elastic net regression (EN), logistic regression (LR), support vector machine (SVM), random forest (RF) [5]
- Transformer-based model: bert-base-uncased, removal of potentially noisy content (HTML-tags, non-ASCII characters, digits, single-letter words, multiple white space)
- Metrics: Accuracy, precision, recall, F1-score (decisive criterion)

Results: Annotated Corpus

	Agreement
Anti-elitism	0.79
People-centrism	0.54
People-sovereignty	0.72
Populism	0.79

Agreement of the annotators reported for each motive

	Full	Training	Validation	Test
Populist	297	198	50	49
None	903	602	350	351

Distribution of binary populism labels in the full, training, validation and test sample

	P0	P1	P2
Anti-elitism	90	102	76
People-centrism	18	13	32
People-sovereignty	6	10	6
None	302	288	308
Populism	98	112	92

	P0	P1	P3
Anti-elitism	95	100	109
People-centrism	19	16	22
People-sovereignty	8	11	7
None	296	290	284
Populism	104	110	116

	P0	P1	P4
Anti-elitism	87	97	104
People-centrism	14	10	14
People-sovereignty	9	8	9
None	304	299	286
Populism	96	101	114

Exact labels given by each participant for the annotation sample (here divided in four batches)

Results: Classification

	Gründl	R&P	EN	LR	SVM	RF	BERT
Accuracy	0.69	0.77	0.77	0.81	0.74	0.77	0.80
Precision	0.44	0.53	0.54	0.92	0.46	0.58	0.57
Recall	0.88	0.51	0.39	0.22	0.33	0.22	0.79
F1	0.59	0.52	0.45	0.36	0.38	0.32	0.66

Result of different methods on 200 training samples

Discussion

Annotation Study

- High inter-annotator agreement on binary populism label
- Anti-elitism the most prevalent motive in the sample
- Indicator for pandemic populist, as political and scientific elites where held accountable for changes in social life
- High class-imbalance: only 24.8% comments labelled populist

Populism Detection:

- BERT-model outperforms precision of dictionary by 12% leading to an 7% increase in F1
- Reduction of false positives
- BERT-model outperforms all other machine-leaning baselines

Examples

Leider hat die Elite die Umverteilung von unten nach oben forciert, wodurch wir heute in diese präkeren Arbeitsverhältnisse getrieben wurden. Die Politiker könnten dies (1) sofort wieder ändern.

Warum sie das nicht tun? Weil Sie die Befehle und Wünsche dieser Eliten ausführen und nicht mehr dem Volk dienen.

Es ist Zeit für einen Systemwechsel!

Das Volk will nicht impfen, das Volk will lieber Maske tragen und testen ! (3)

Da kann die Regierung am Kopf stehen und mit den Beinen Fliegen fangen, das Volk macht, was das Volk will !

Und wählt Mit freundlichen Grüßen ...

Wenn auch unwahrscheinlich, so hoffe ich doch, dass meine Zeilen den Weg zu Euch, meine geschätzten VOLKsvertreter finden. Wenn auch unwahrscheinlich, so biete (2) ich Euch ein persönliches Treffen an, indem Ihr eine Stimme aus genau diesem VOLK, welches Eure „Hausverstand-Ergüsse“ finanziert, hören könnt. Denn es ist uns allen bewusst, dass die leisen Stimmen Eures Volkes nur im Wahljahr zu hören sind. Wenn auch dieses persönliche Treffen höchst unwahrscheinlich ist, so ließe ich mich gerne belehren, warum EURE Hausverstand besser ist als Unserer;

- (1) Anti-Elitism
- (2) People-Sovereignty
- (3) People-Centrism

Leider hat die Elite die Umverteilung von unten nach oben forciert, wodurch wir heute in diese präkeren Arbeitsverhältnisse getrieben wurden. Die Politiker könnten dies sofort wieder ändern.

Warum sie das nicht tun? Weil Sie die Befehle und Wünsche dieser Eliten ausführen und nicht mehr dem Volk dienen.

Es ist Zeit für einen Systemwechsel!

People-sovereignty	0.74
Populism	0.79

Agreement of the annotators reported for each motive

Populist	297	198	50	49
None	903	602	350	351

Distribution of binary populism labels in the full, training, validation and test sample

	P0	P1	P2
Anti-elitism	90	102	76
People-centrism	18	13	32
People-sovereignty	6	10	6
None	302	288	308
Populism	98	112	92

	P0	P1	P3
Anti-elitism	95	100	109
People-centrism	19	16	22
People-sovereignty	8	11	7
None	296	290	284
Populism	104	110	116

	P0	P1	P4
Anti-elitism	87	97	104
People-centrism	14	10	14
People-sovereignty	9	8	9
None	304	299	286
Populism	96	101	114

Exact labels given by each participant for the annotation sample (here divided in four batches)

Methods: Populism Detection

- Data split into 800 training, 200 validation and 200 test samples
- Baselines: Dictionary (Gründl) [3], dictionary (R&P) [9], elastic net regression (EN), logistic regression (LR), support vector machine (SVM), random forest (RF) [5]
- Transformer-based model: bert-base-uncased, removal of potentially noisy content (HTML-tags, non-ASCII characters, digits, single-letter words, multiple white space)
- Metrics: Accuracy, precision, recall, F1-score (decisive criterion)

Result of different methods on 200 training samples

Note:

- Best performance for each metric marked in bold font
- Recall of Gründl dictionary naturally high, as it was used for the sampling
- Methods with low recall hardly detect any populist comments

- Indicator for pandemic populist, as political and scientific elites where held accountable for changes in social life
- High class-imbalance: only 24.8% comments labelled populist

Populism Detection:

- BERT-model outperforms precision of dictionary by 12% leading to an 7% increase in F1
- Reduction of false positives
- BERT-model outperforms all other machine-learning baselines

Examples

Leider hat die Elte die Umverteilung von unten nach oben forciert, wodurch wir heute in diese präkaren Arbeitsverhältnisse getrieben wurden. Die Politiker könnten dies (1)

Wenn auch unwahrscheinlich, so hoffe ich doch, dass meine Zeilen den Weg zu Euch, meine geschätzten VOLKsvertreter finden. Wenn auch unwahrscheinlich, so biete (2) ich Euch ein persönliches Treffen an, indem Ihr eine Stimme aus genau diesem VOLK, welches Eure „Hausverständ-Ergüsse“ finanziert, hören könnt. Denn es ist uns allen bewusst, dass die leisen Stimmen Eures Volkes nur im Wahljahr zu hören sind. Wenn auch dieses persönliche Treffen höchst unwahrscheinlich ist, so ließe ich mich gerne

Warum sie das nicht tun? Weil Sie die Befehle und Wünsche dieser Eltern ausführen und nicht mehr dem Volk dienen.

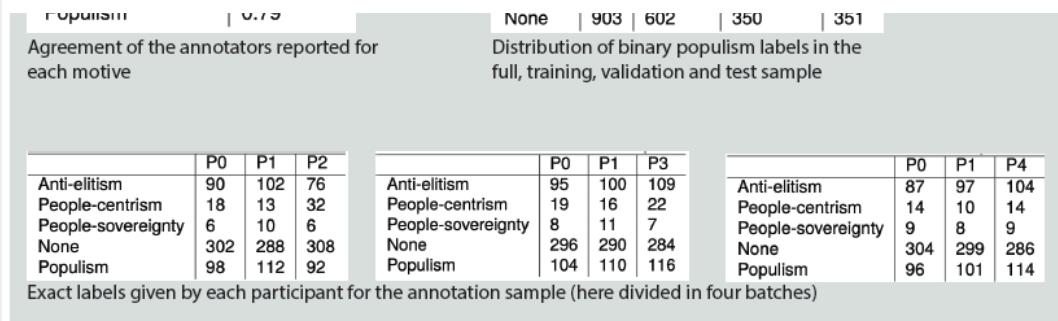
Methods: Populism Detection

- Data split into 800 training, 200 validation and 200 test samples
- Baselines: Dictionary (Gründl) [3], dictionary (R&P) [9], elastic net regression (EN), logistic regression (LR), support vector machine (SVM), random forest (RF) [5]
- Transformer-based model: bert-base-uncased, removal of

Das Volk will nicht impfen, das Volk will lieber Maske tragen und testen !

Da kann die Regierung am Kopf stehen und mit den Beinen Fliegen fangen, das Volk macht, was das Volk will !

Und wählt Mit freundlichen Grüßen ...



Note:

- Best performance for each metric marked in bold font
- Recall of Gründl dictionary naturally high, as it was used for the sampling
- Methods with low recall hardly detect any populist comments

entific elites where held accountable for changes in social life

- High class-imbalance: only 24.8% comments labelled populist

Populism Detection:

- BERT-model outperforms precision of dictionary by 12% leading to an 7% increase in F1
- Reduction of false positives
- BERT-model outperforms all other machine-leaning baselines

Examples

Leider hat die Elte die Umverteilung von unten nach oben forciert, wodurch wir heute in diese präkaren Arbeitsverhältnisse getrieben wurden. Die Politiker können dies (1) sofort wieder ändern.

Warum sie das nicht tun? Weil Sie die Befehle und Wünsche dieser Eliten ausführen und nicht mehr dem Volk dienen.

Es ist Zeit für einen Systemwechsel!

Das Volk will nicht impfen, das Volk will lieber Maske tragen und testen ! (3)

Da kann die Regierung am Kopf stehen und mit den Beinen Fliegen fangen, das Volk macht, was das Volk will !

Und wählt Mit freundlichen Grüßen ...

Wenn auch unwahrscheinlich, so hoffe ich doch, dass meine Zeilen den Weg zu Euch, meine geschätzten VOLKsvertreter finden. Wenn auch unwahrscheinlich, so biete (2) ich Euch ein persönliches Treffen an, indem Ihr eine Stimme aus genau diesem VOLK, welches Eure „Hausverstand-Ergüsse“ finanziert, hören könnt. Denn es ist uns allen bewusst, dass die leisen Stimmen Eures Volkes nur im Wahljahr zu hören sind. Wenn auch dieses persönliche Treffen höchst unwahrscheinlich ist, so ließe ich mich gerne belehren, warum EUER Hausverstand besser ist als Unserer;

- (1) Anti-Elitism
- (2) People-Sovereignty
- (3) People-Centrism

Anti-elitism	0.79
People-centrism	0.54
People-sovereignty	0.72
Populism	0.79

Agreement of the annotators reported for each motive

	Full	Training	Validation	Test
Populist	297	198	50	49
None	903	602	350	351

Distribution of binary populism labels in the full, training, validation and test sample

	P0	P1	P2
Anti-elitism	90	102	76
People-centrism	18	13	32
People-sovereignty	6	10	6
None	302	288	308
Populism	98	112	92

	P0	P1	P3
Anti-elitism	95	100	109
People-centrism	19	16	22
People-sovereignty	8	11	7
None	296	290	284
Populism	104	110	116

	P0	P1	P4
Anti-elitism	87	97	104
People-centrism	14	10	14
People-sovereignty	9	8	9
None	304	299	286
Populism	96	101	114

Exact labels given by each participant for the annotation sample (here divided in four batches)

Methods: Populism Detection

- Data split into 800 training, 200 validation and 200 test samples
- Baselines: Dictionary (Gründl) [3], dictionary (R&P) [9], elastic net regression (EN), logistic regression (LR), support vector machine (SVM), random forest (RF) [5]
- Transformer-based model: bert-base-uncased, removal of potentially noisy content (HTML-tags, non-ASCII characters, digits, single-letter words, multiple white space)
- Metrics: Accuracy, precision, recall, F1-score (decisive criterion)

Recall	0.88	0.51	0.39	0.22	0.33	0.22	0.79
F1	0.59	0.52	0.45	0.36	0.38	0.32	0.66

Result of different methods on 200 training samples

Note:

- Best performance for each metric marked in bold font
- Recall of Gründl dictionary naturally high, as it was used for the sampling
- Methods with low recall hardly detect any populist comments

label

- Anti-elitism the most prevalent motive in the sample
- Indicator for pandemic populist, as political and scientific elites where held accountable for changes in social life
- High class-imbalance: only 24.8% comments labelled populist

Populism Detection:

- BERT-model outperforms precision of dictionary by 12% leading to an 7% increase in F1
- Reduction of false positives
- BERT-model outperforms all other machine-learning baselines

Examples

Leider hat die Elite die Umverteilung von unten nach oben forciert, wodurch wir heute in diese präkaren Arbeitsverhältnisse getrieben wurden. Die Politiker können dies (1) sofort wieder ändern.

Warum sie das nicht tun? Weil Sie die Befehle und Wünsche dieser Eliten ausführen und nicht mehr dem Volk dienen.

Es ist Zeit für einen Systemwechsel!

Das Volk will nicht impfen, das Volk will lieber Maske tragen und testen ! (3)

Da kann die Regierung am Kopf stehen und mit den Beinen Fliegen fangen, das Volk macht, was das Volk will !

Und wählt Mit freundlichen Grüßen ...

Wenn auch unwahrscheinlich, so hoffe ich doch, dass meine Zeilen den Weg zu Euch, meine geschätzten VOLKsvertreter finden. Wenn auch unwahrscheinlich, so biete (2) ich Euch ein persönliches Treffen an, indem Ihr eine Stimme aus genau diesem VOLK, welches Eure „Hausverstand-Ergüsse“ finanziert, hören könnt. Denn es ist uns allen bewusst, dass die leisen Stimmen Eures Volkes nur im Wahljahr zu hören sind. Wenn auch dieses persönliche Treffen höchst unwahrscheinlich ist, so ließe ich mich gerne belehren, warum EURE Hausverstand besser ist als Unserer;

- (1) Anti-Elitism
- (2) People-Sovereignty
- (3) People-Centrism

Methods: Populism Detection

- Data split into 800 training, 200 validation and 200 test samples
- Baselines: Dictionary (Gründl) [3], dictionary (R&P) [9], elastic net regression (EN), logistic regression (LR), support vector machine (SVM), random forest (RF) [5]
- Transformer-based model: bert-base-uncased, removal of potentially noisy content (HTML-tags, non-ASCII characters, digits, single-letter words, multiple white space)
- Metrics: Accuracy, precision, recall, F1-score (decisive criterion)

Results: Annotated Corpus

	Agreement
Anti-elitism	0.79
People-centrism	0.54
People-sovereignty	0.72
Populism	0.79

Agreement of the annotators reported for each motive

	Full	Training	Validation	Test
Populist	297	198	50	49
None	903	602	350	351

Distribution of binary populism labels in the full, training, validation and test sample

	P0	P1	P2
Anti-elitism	90	102	76
People-centrism	18	13	32
People-sovereignty	6	10	6
None	302	288	308
Populism	98	112	92

	P0	P1	P3
Anti-elitism	95	100	109
People-centrism	19	16	22
People-sovereignty	8	11	7
None	296	290	284
Populism	104	110	116

	P0	P1	P4
Anti-elitism	87	97	104
People-centrism	14	10	14
People-sovereignty	9	8	9
None	304	299	286
Populism	96	101	114

Exact labels given by each participant for the annotation sample (here divided in four batches)

Results: Classification

	Gründl	R&P	EN	LR	SVM	RF	BERT
Accuracy	0.69	0.77	0.77	0.81	0.74	0.77	0.80
Precision	0.44	0.53	0.54	0.92	0.46	0.58	0.57
Recall	0.88	0.51	0.39	0.22	0.33	0.22	0.79
F1	0.59	0.52	0.45	0.36	0.38	0.32	0.66

Result of different methods on 200 training samples

Discussion

Annotation Study

- High inter-annotator agreement on binary populism label
- Anti-elitism the most prevalent motive in the sample
- Indicator for pandemic populist, as political and scientific elites where held accountable for changes in social life
- High class-imbalance: only 24.8% comments labelled populist

Populism Detection:

- BERT-model outperforms precision of dictionary by 12% leading to an 7% increase in F1
- Reduction of false positives
- BERT-model outperforms all other machine-learning baselines

Examples

Leider hat die Elte die Umverteilung von unten nach oben forciert, wodurch wir heute in diese präkaren Arbeitsverhältnisse getrieben wurden. Die Politiker können dies (1) sofort wieder ändern.

Warum sie das nicht tun? Weil Sie die Befehle und Wünsche dieser Eliten ausführen und nicht mehr dem Volk dienen.

Es ist Zeit für einen Systemwechsel!

Das Volk will nicht impfen, das Volk will lieber Maske tragen und testen ! (3)

Da kann die Regierung am Kopf stehen und mit den Beinen Fliegen fangen, das Volk macht, was das Volk will !

Und wählt Mit freundlichen Grüßen ...

Wenn auch unwahrscheinlich, so hoffe ich doch, dass meine Zeilen den Weg zu Euch, meine geschätzten VOLKsvertreter finden. Wenn auch unwahrscheinlich, so biete (2) ich Euch ein persönliches Treffen an, indem Ihr eine Stimme aus genau diesem VOLK, welches Eure „Hausverstand-Ergüsse“ finanziert, hören könnt. Denn es ist uns allen bewusst, dass die leisen Stimmen Eures Volkes nur im Wahljahr zu hören sind. Wenn auch dieses persönliche Treffen höchst unwahrscheinlich ist, so ließe ich mich gerne belehren, warum EURE Hausverstand besser ist als Unserer;

- (1) Anti-Elitism
- (2) People-Sovereignty
- (3) People-Centrism

Methods: Populism Detection

- Data split into 800 training, 200 validation and 200 test samples
- Baselines: Dictionary (Gründl) [3], dictionary (R&P) [9], elastic net regression (EN), logistic regression (LR), support vector machine (SVM), random forest (RF) [5]
- Transformer-based model: bert-base-uncased, removal of potentially noisy content (HTML-tags, non-ASCII characters, digits, single-letter words, multiple white space)
- Metrics: Accuracy, precision, recall, F1-score (decisive criterion)

Results: Annotated Corpus

	Agreement
Anti-elitism	0.79
People-centrism	0.54
People-sovereignty	0.72
Populism	0.79

Agreement of the annotators reported for each motive

	Full	Training	Validation	Test
Populist	297	198	50	49
None	903	602	350	351

Distribution of binary populism labels in the full, training, validation and test sample

	P0	P1	P2
Anti-elitism	90	102	76
People-centrism	18	13	32
People-sovereignty	6	10	6
None	302	288	308
Populism	98	112	92

	P0	P1	P3
Anti-elitism	95	100	109
People-centrism	19	16	22
People-sovereignty	8	11	7
None	296	290	284
Populism	104	110	116

	P0	P1	P4
Anti-elitism	87	97	104
People-centrism	14	10	14
People-sovereignty	9	8	9
None	304	299	286
Populism	96	101	114

Exact labels given by each participant for the annotation sample (here divided in four batches)

Results: Classification

	Gründl	R&P	EN	LR	SVM	RF	BERT
Accuracy	0.69	0.77	0.77	0.81	0.74	0.77	0.80
Precision	0.44	0.53	0.54	0.92	0.46	0.58	0.57
Recall	0.88	0.51	0.39	0.22	0.33	0.22	0.79
F1	0.59	0.52	0.45	0.36	0.38	0.32	0.66

Result of different methods on 200 training samples

Note:

- Best performance for each metric marked in bold font
- Recall of Gründl dictionary naturally high, as it was used for the sampling
- Methods with low recall hardly detect any populist comments

Discussion

Annotation Study

- High inter-annotator agreement on binary populism label
- Anti-elitism the most prevalent motive in the sample
- Indicator for pandemic populist, as political and scientific elites where held accountable for changes in social life
- High class-imbalance: only 24.8% comments labelled populist

Populism Detection:

- BERT-model outperforms precision of dictionary by 12% leading to an 7% increase in F1
- Reduction of false positives
- BERT-model outperforms all other machine-leaning baselines

Examples

Leider hat die Elte die Umverteilung von unten nach oben forciert, wodurch wir heute in diese präkaren Arbeitsverhältnisse getrieben wurden. Die Politiker können dies (1) sofort wieder ändern.

Warum sie das nicht tun? Weil Sie die Befehle und Wünsche dieser Eliten ausführen und nicht mehr dem Volk dienen.

Es ist Zeit für einen Systemwechsel!

Das Volk will nicht impfen, das Volk will lieber Maske tragen und testen ! (3)

Da kann die Regierung am Kopf stehen und mit den Beinen Fliegen fangen, das Volk macht, was das Volk will !

Und wählt Mit freundlichen Grüßen ...

Wenn auch unwahrscheinlich, so hoffe ich doch, dass meine Zeilen den Weg zu Euch, meine geschätzten VOLKsvertreter finden. Wenn auch unwahrscheinlich, so biete (2) ich Euch ein persönliches Treffen an, indem Ihr eine Stimme aus genau diesem VOLK, welches Eure „Hausverstand-Ergüsse“ finanziert, hören könnt. Denn es ist uns allen bewusst, dass die leisen Stimmen Eures Volkes nur im Wahljahr zu hören sind. Wenn auch dieses persönliche Treffen höchst unwahrscheinlich ist, so ließe ich mich gerne belehren, warum EURE Hausverstand besser ist als Unserer;

- (1) Anti-Elitism
- (2) People-Sovereignty
- (3) People-Centrism

Methods: Populism Detection

- Data split into 800 training, 200 validation and 200 test samples
- Baselines: Dictionary (Gründl) [3], dictionary (R&P) [9], elastic net regression (EN), logistic regression (LR), support vector machine (SVM), random forest (RF) [5]
- Transformer-based model: bert-base-uncased, removal of potentially noisy content (HTML-tags, non-ASCII characters, digits, single-letter words, multiple white space)
- Metrics: Accuracy, precision, recall, F1-score (decisive criterion)

Results: Annotated Corpus

	Agreement
Anti-elitism	0.79
People-centrism	0.54
People-sovereignty	0.72
Populism	0.79

Agreement of the annotators reported for each motive

	Full	Training	Validation	Test
Populist	297	198	50	49
None	903	602	350	351

Distribution of binary populism labels in the full, training, validation and test sample

	P0	P1	P2
Anti-elitism	90	102	76
People-centrism	18	13	32
People-sovereignty	6	10	6
None	302	288	308
Populism	98	112	92

	P0	P1	P3
Anti-elitism	95	100	109
People-centrism	19	16	22
People-sovereignty	8	11	7
None	296	290	284
Populism	104	110	116

	P0	P1	P4
Anti-elitism	87	97	104
People-centrism	14	10	14
People-sovereignty	9	8	9
None	304	299	286
Populism	96	101	114

Exact labels given by each participant for the annotation sample (here divided in four batches)

Results: Classification

	Gründl	R&P	EN	LR	SVM	RF	BERT
Accuracy	0.69	0.77	0.77	0.81	0.74	0.77	0.80
Precision	0.44	0.53	0.54	0.92	0.46	0.58	0.57
Recall	0.88	0.51	0.39	0.22	0.33	0.22	0.79
F1	0.59	0.52	0.45	0.36	0.38	0.32	0.66

Result of different methods on 200 training samples

Note:

- Best performance for each metric marked in bold font
- Recall of Gründl dictionary naturally high, as it was used for the sampling
- Methods with low recall hardly detect any populist comments

Discussion

Annotation Study

- High inter-annotator agreement on binary populism label
- Anti-elitism the most prevalent motive in the sample
- Indicator for pandemic populist, as political and scientific elites where held accountable for changes in social life
- High class-imbalance: only 24.8% comments labelled populist

Populism Detection:

- BERT-model outperforms precision of dictionary by 12% leading to an 7% increase in F1
- Reduction of false positives
- BERT-model outperforms all other machine-learning baselines

Contribution

- First annotated dataset for populist user comments in the German language
- Proposed model outperforms state-of-the-art of populism detection in Austrian news comments by 0.07 regarding F1-score

Future Work

- Fine-tune SOTA large language models for the task to improve classification performance
- Incorporate comments from a diverse set of news outlets
- Extend the sample with data from Germany e.g.
- More fine-grained annotation entailing mentioned elites

References

- [1] Nicole Ernst, Sina Blassnig, Sven Engesser, Florin Büchel, and Frank Esser. 2019. Populists prefer social media over talk shows: An analysis of populist messages and stylistic elements across six countries. *Social Media + Society*, 5(1):2056305118823358.
- [2] Nicole Ernst, Sven Engesser, Florin Büchel, Sina Blassnig, and Frank Esser. 2017. Extreme parties and populism: an analysis of facebook and twitter across six countries. *Information, Communication & Society*, 20:1–18.
- [3] Charlotte Galpin and Hans-Jörg Trenz. 2019. Participatory populism: Online discussion forums on mainstream news sites during the 2014 european parliament election. *Journalism Practice*, 13(7):781–798.
- [4] Johann Gründl. 2022. Populist ideas on social media: A dictionary-based measurement of populist communication. *New Media & Society*, 24(6):1481–1499.
- [5] Kirk A. Hawkins and Bruno Castanho Silva. 2016. An head-to-head comparison of human-based and automated text analysis for measuring populism in 27 countries.
- [6] Jan Jagers and Stefaan Walgrave. 2007. Populism as political communication style: An empirical study of political parties' discourse in belgium. *European Journal of Political Research*, 46(3):319–345.
- [7] Benjamin Krämer. 2014. Media Populism: A Conceptual Clarification and Some Theses on its Effects. *Communication Theory*, 24(1):42–60.
- [8] Cas Mudde. 2004. The populist zeitgeist. *Govern-ment and Opposition*, 39(4):541–563.
- [9] Matthijs Rodouijin and Teun Pauwels. 2011. Measuring populism: Comparing two methods of content analysis. *West European Politics*, 34(6):1272–1283.
- [10] Anne Schulz, Philipp Müller, Christian Schemer, Dominique Wirs, Martin Wettstein, and Werner Wirth. 2018. Measuring populist attitudes on three dimensions. *International Journal of Public Opinion Research*, 30:316–326.
- [11] Jigile Ulinskaite and Lukas Pukels. 2021. Identifying populist paragraphs in text: A machine-learning approach. *CoRR*, abs/2106.03161.
- [12] Kurt Weyland. 2001. Clarifying a contested concept: Populism in the study of latin american politics. *Comparative Politics*, 34(1):1–22.
- [13] Dominique Stefanie Wirs. 2018. Persuasion through emotion? an experimental test of the emotion-eliciting nature of populist communica-tion. *International Journal of Communication*, 12:25.

Contact

1) Faculty of Informatics, TU Wien

ahmadou.wagne@tuwien.ac.at, thomas.kolb@tuwien.ac.at, julia.neidhardt@tuwien.ac.at

2) <https://recsys-lab.at>

Data

<https://doi.org/10.48436/vbkwj-b8t85>

Available upon request



funded by:



Contact



Ahmadou Wagne
Research Unit of Data Science
ahmadou.wagne@tuwien.ac.at
<https://recsys-lab.at/team/ahmadou-wagne/>