UNIVERSITY OF MANNHEIM

„Glass-box“ Machine Learning in Text Cassification: The Case of Integrative Complexity in Social Media Comments

Hartmut Wessler

Institute for Media and Communication Studies
University of Mannheim

wessler@uni-mannheim.de

Lecture Series „Data Science in Action“
September 22, 2022

Hartmut Wessler

- Professor of Media and Communication Studies, U Mannheim
- Disclaimer: I don’t usually self-identify as a data scientist ☺
- I’m a communication scholar who uses methods from Natural Language Processing (NLP) to study the democratic qualities of political communication online and offline

The team

Dr. Chung-hong Chan

Julia Jakob

Timo Dobbrick

Dr. Rainer Freudenthaler

The papers

The International Journal of Press/Politics
1–21
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/19401612211044018
journals.sagepub.com/home/hij
SAGE

The Integrative Complexity of Online User Comments Across Different Types of Democracy and Discussion Arenas

Julia Jakob[1], Timo Dobbrick[1], and Hartmut Wessler[1,2]

Abstract

This study is the first to compare the integrative complexity o
across distinct democratic political systems and in discussion a
mary use functions. Integrative complexity is a psycho-ling
increasingly used by communication scholars to study the a
political debate contributions. It captures the sophistication o
in terms of differentiation and integration, mapping whether

COMMUNICATION METHODS AND MEASURES
https://doi.org/10.1080/19312458.2021.1999913

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Enhancing Theory-Informed Dictionary Approaches with "Glass-box" Machine Learning: The Case of Integrative Complexity in Social Media Comments

Timo Dobbrick[a], Julia Jakob[a], Chung-Hong Chan[a], and Hartmut Wessler[b]

[a]Mannheim Centre for European Social Research, University of Mannheim; [b]Institute for Media and Communication Studies, University of Mannheim

ABSTRACT

Dictionary-based approaches to computational text analysis have been shown to perform relatively poorly, particularly when the dictionaries rely on simple bags of words, are not specified for the domain under study, and add word scores without weighting. While machine learning approaches usually perform better, they offer little insight into (a) which of the assumptions underlying dictionary approaches (bag-of-words, domain transferability, or additivity) impedes performance most, and (b) which language

What I want to talk about today

Desirable synergies	• Substantive research & methods development
The research problem	• Measuring integrative complexity in text
The traditional approach	• Manual coding
The computational approach	• Dictionary + Machine learning (ML)
The advantages of „glass-box“ ML	

Desirable synergies

Opposing perspectives

Communication Science

Data Science (NLP)

Core concern

Fascinating new methods options

Research problem: Integrative complexity

Theory-building

Method: Computational content analysis (Text-as-data)

Just another application

Core concern

Desirable synergies

- Computational content analysis methods help solve research problems in communication science
  - On a larger scale (Big Data)
  - More economically (less work for similar insights)

- They do not make research more more valid or better *per se*
  - „Validate, validate, validate“ (Grimmer & Stewart 2013)
  - Outputs always need to be validated by human researchers

- Special problem for Machine Learning, especially Deep Learning:
  - Black box: Parameters don’t mean anything to humans
  - → “Glass-box” ML: Algorithms that produce interpretable intermediate results that can aid in theory building

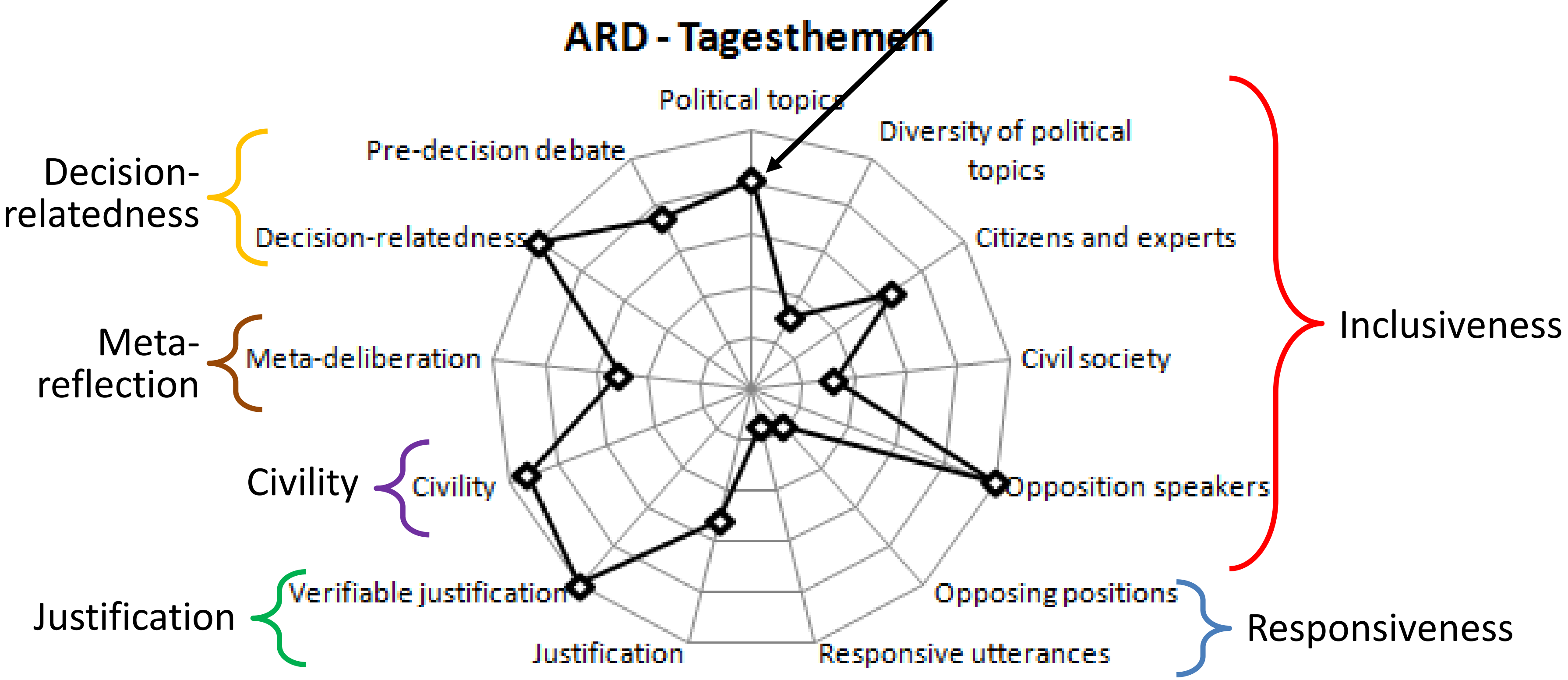
The research problem

The theoretical context

- Qualities of mediated discourse
  - Inclusiveness (of actors and ideas)
  - Civility (versus hate speech, impoliteness and intolerance)
  - Justification/Reason-giving
  - Reciprocity (actors referring to each other)
  - Decision-relatedness
  - Meta-reflection (of rules and conduct of public debate itself)

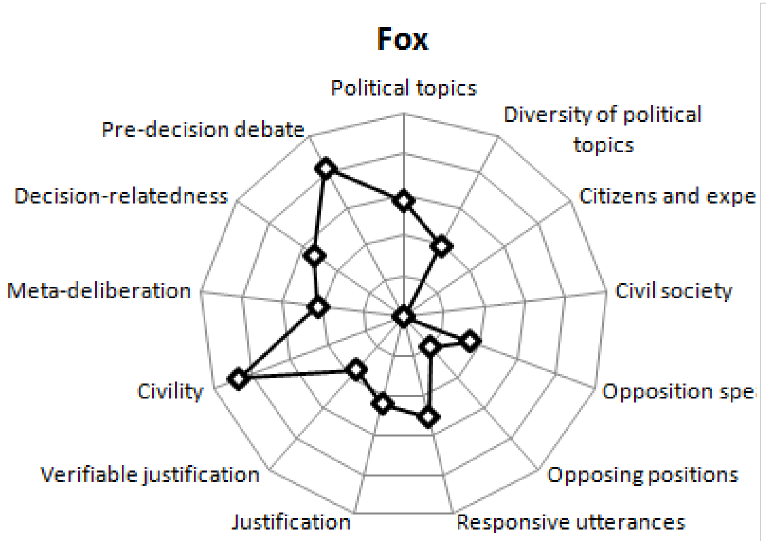
Example: Qualities of TV news

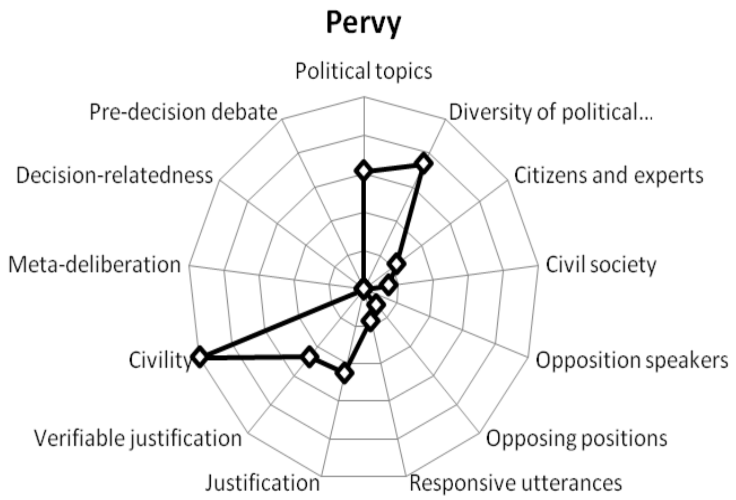
„The close to the outer bound, the better“



Source: Wessler & Rinke, 2014

Example: Qualities of TV news



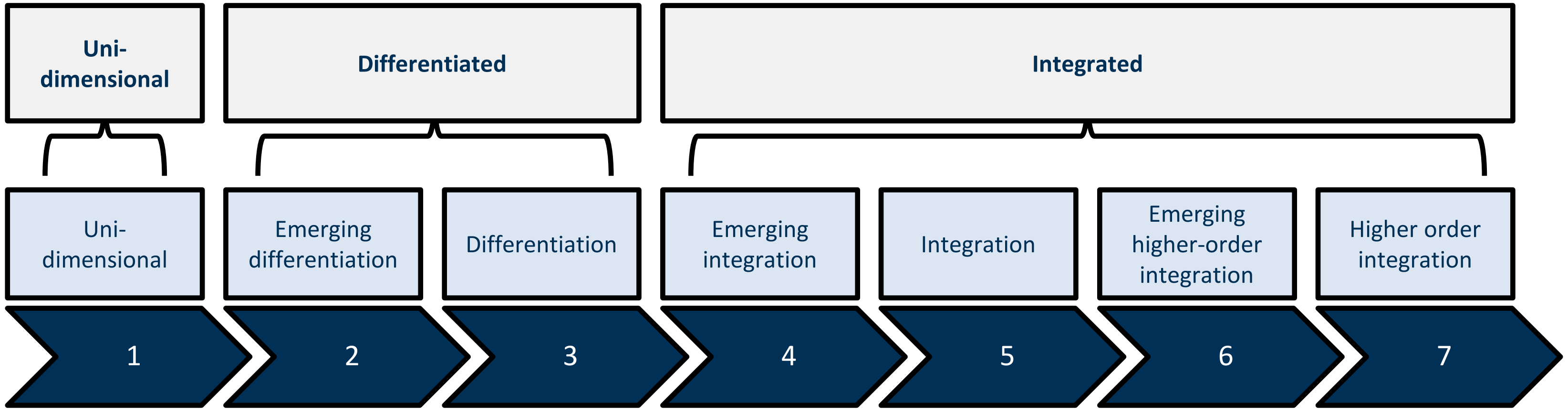


Integrative complexity

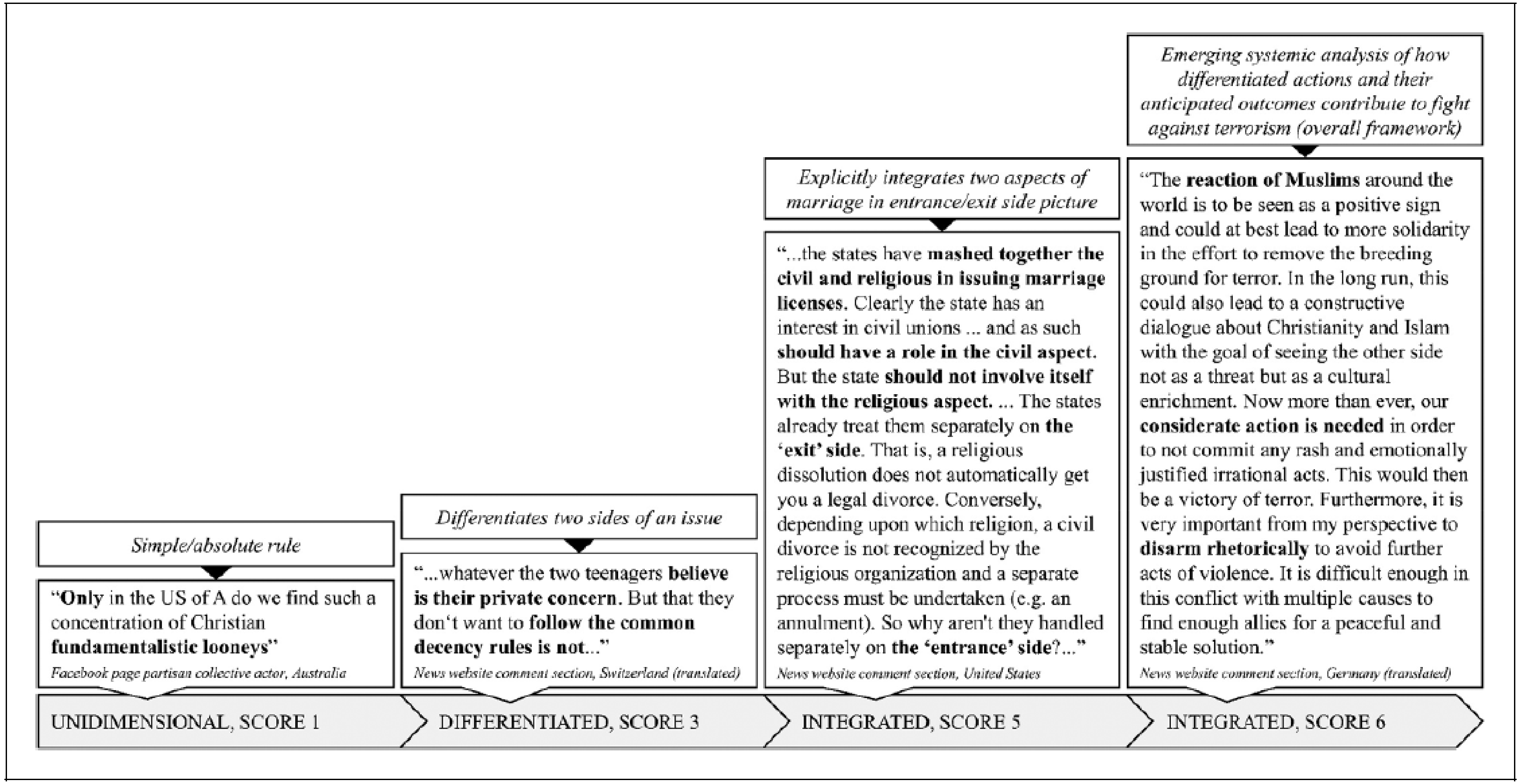
- Deliberative and communitarian discussion norms require statements to be justified soundly with reasons (Freelon 2015)
- Traditionally, justification was operationalized as:
  - A reason is stated in addition to a claim, making the statement more complex
- New operationalization of justification:
  - Aspects of or perspectives on the topic are differentiated and then related to each other (integrated) in a statement → Integrative complexity (Suedfeld et al. 1992)

Integrative complexity

- Maps the range from simple to complex argumentation in debate contributions (Beste & Wyss, 2014)
- Captures the sophistication of statements by their degree of *differentiation* and *integration* (Suedfeld et al., 1992)



Examples for simple and integratively complex user comments



Types of democracy

- *Consensus democracies* strive for argumentatively sustained compromise, actors in *majoritarian democracies* tend to clearly dissociate from each other (Ljiphart 2012; Steiner et al. 2004)
  - **H1:** The integrative complexity of online user comments is higher in consensus-oriented than in majoritarian democracies.

Consensus-oriented democracies

- Germany
- Switzerland

Majoritarian democracies

- Australia
- United States

Primary use function

- Opinion diversity and disagreement foster well-reasoned public statements (Zhang et al. 2013; Maia et al. 2020)
  - ➢ **H2:** The integrative complexity of online user comments is higher in arenas that are *used primarily for issue-driven debates* with plural opinions than in forums that are *rather used for preference-driven, like-minded discussions*.

Issue-driven discussions

- Website comment sections of mainstream news media
- Facebook pages of mainstream news media

Preference-driven discussions

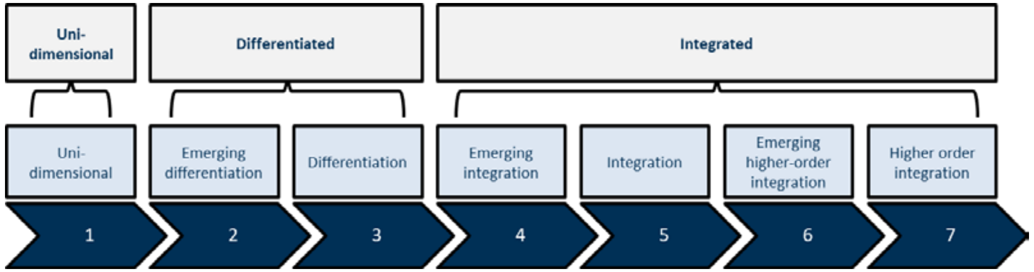
- Facebook pages of partisan collective actors and alternative media
- Twitter

The two papers compared

	Analytical paper (Jakob et al., 2021)	Methods paper (Dobbrick et al., 2021)
Countries studied	CH/DE – USA/AUS	
Media arenas studied	Users comments from: • News websites (legacy news media) • Facebook pages of legacy news media • Facebook pages of alternative media and partisan actors • Twitter	
Topic studied	Public role of religion	
Period studied	Aug 2015 – July 2016	
Method used	Manual content analysis N = 4,800 user comments (300 randomly selected per country and arena out of a total N=1,236,551 comments)	Automated content analysis (dictionary + machine learning) N = 4,800 comments as gold standard (available here: https://osf.io/z4an2/)

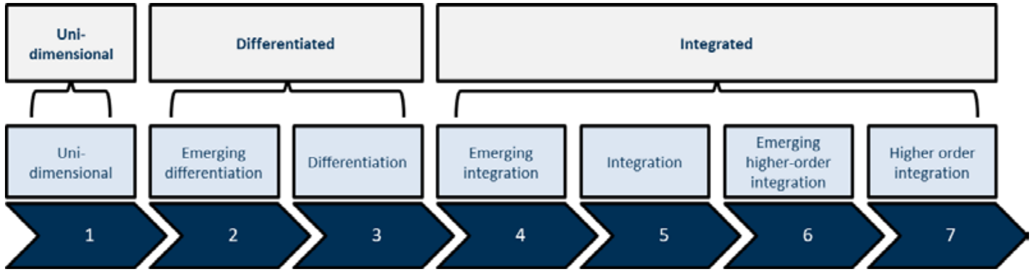
The traditional approach: Manual content analysis

Manual content analysis



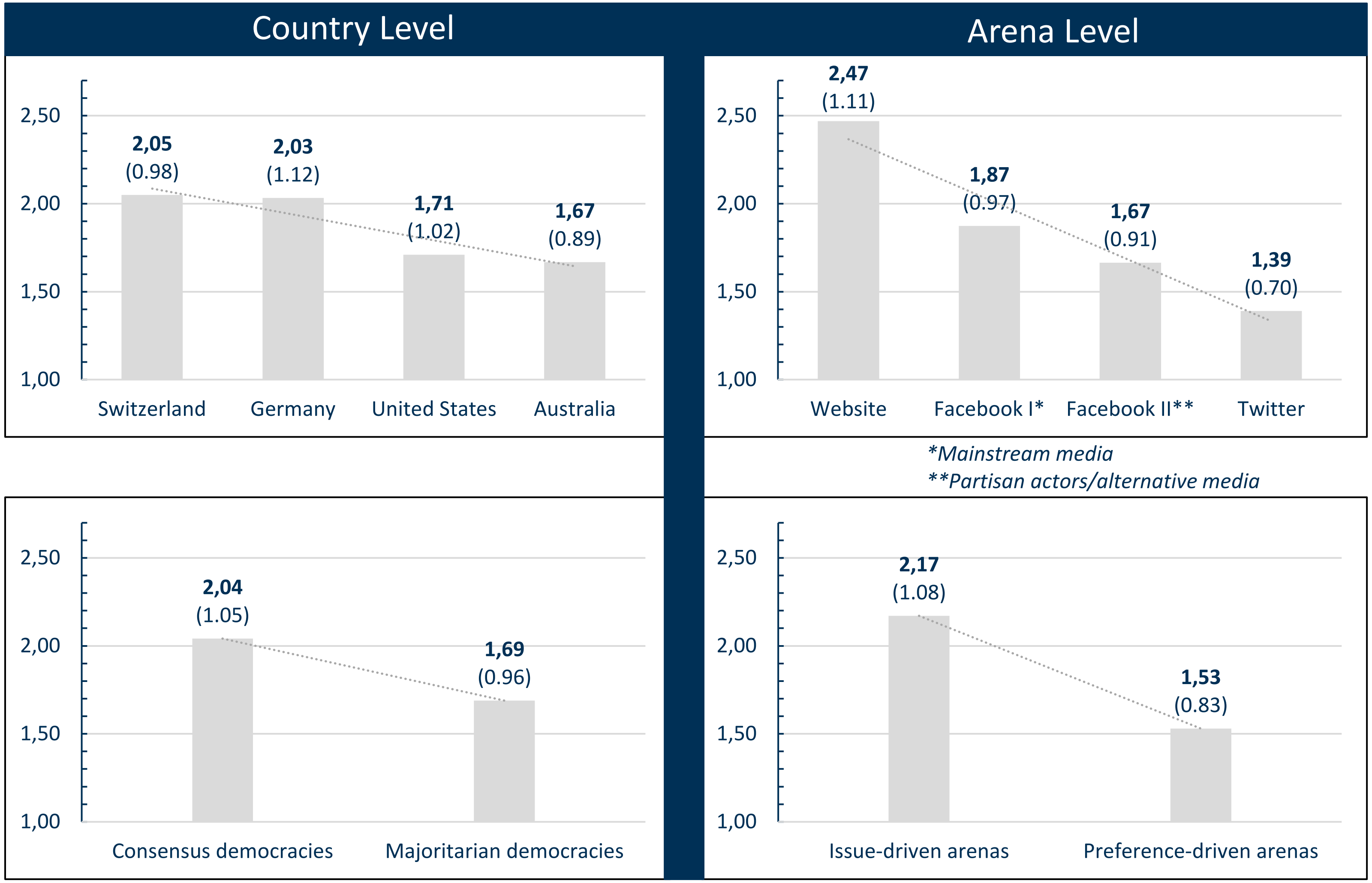
- *N = 4,563* user contributions for statistical analysis (randomly sampled from a larger data set of *N* = 1,236,551 contributions)
- Integrative Complexity scored on ordinal scale
  - 1 = one aspect or perspective only
  - 3 = at least 2 aspects or perspectives on the topic, but no integration
  - 5 = connection in form of superordinate category, mutual influence or synthesis
  - 7 = connection drawn as part of systemic conceptual framework
- Three coders: Krippendorff's alpha .85 (.88 and .86 for the tandems)

Descriptive results



- Half of the user comments were unidimensional
- 44% scored 2 or 3, i.e. they differentiated (at least) two aspects or perspectives
- Only 7 % had higher scores, i.e. they drew connections between aspects of or perspectives on the topic

Mean integrative complexity



N = 4,563, *SD* in brackets

Hypothesis tests*

*controlled for the number of words in a user comment

✓ H1: The integrative complexity of online user comments is *higher in consensus-oriented* than in majoritarian democracies.

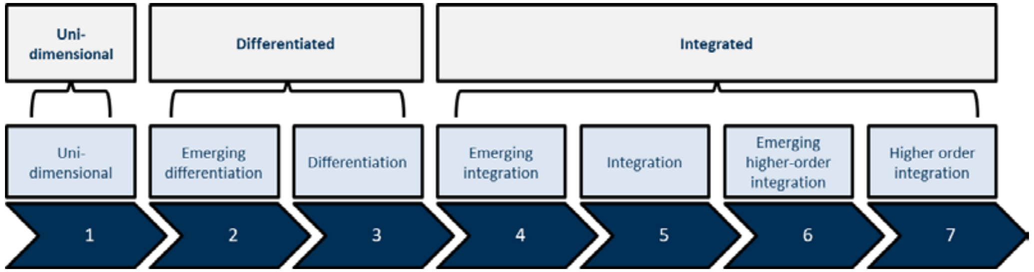
✓ H2: The integrative complexity of online user comments is *higher in arenas that are used primarily for issue-driven debates with plural opinions* than in forums that are rather used for preference-driven, like-minded discussions.

Integrative complexity in context

- The sophistication of online user comments is comparable with that of statements in **U.S. congressional speeches** (Tetlock 1983), **presidential primary debates** (Conway et al. 2012) or **State of the Union addresses** (Thoemmes and Conway 2007)
- Much less refined than for example after **participation in deliberative mini-publics** (Jennstål 2019)
- Findings confirm that the **“spirit of accommodation”** (Lijphart 1975: 103) in consensus-oriented democracies can improve the quality of political debates (Steiner et al. 2004; Wyss et al. 2015)
- Study highlights the **value of arenas used primarily for issue-driven discussions** for democratic discourse (Schudson 1997)

The computational approach

Computational content analysis



- Automatically classifying user contributions into the seven categories based on their content
- Content captured through the LIWC dictionary (Pennebaker et al., 2015) – Linguistic Inquiry and Word Count
- Integrative Complexity is defined by a theoretical selection of ten features from LIWC (Owens & Wedeking, 2011)

Which LIWC features theoretically define integrative complexity?

Sixl	Percentage of words with six or more letters
+Discr	Discrepancy: should, would, could, etc.
+Tent	Tentativeness: maybe, perhaps, guess, etc.
+Incl	Inclusiveness: and, with, include, etc.
+Cause	Causation: because, effect, hence, etc.
+Insig	Insight: think, know, consider, etc.
+Inhib	Inhibition: block, constrain, stop, etc.
-Cert	Certainty: always, never, etc.
-Negate	Negations: no, not, never
-Excl	Exclusiveness: but, without, exclude, etc.

Which method works best for classifying integrative complexity?

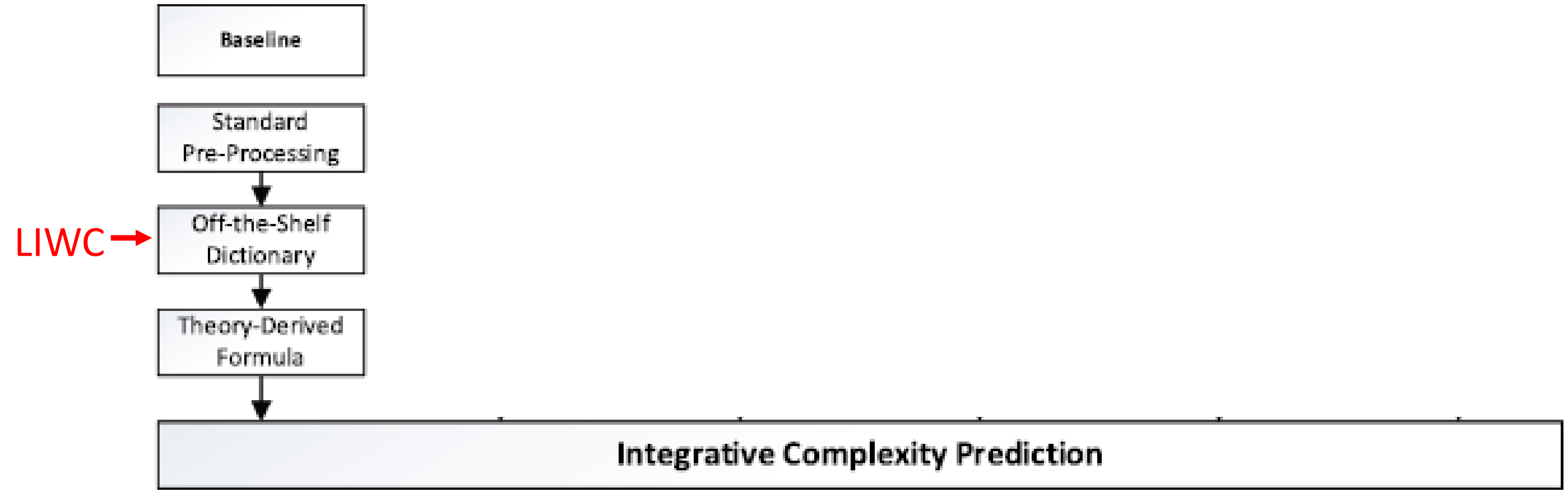


Figure 1. Overview of the Baseline, Assumption-Based Framework, and Shotgun Approaches. *Note*: Boxes highlighted in gray show the process steps that differ from the baseline process.

Which method works best for classifying integrative complexity?

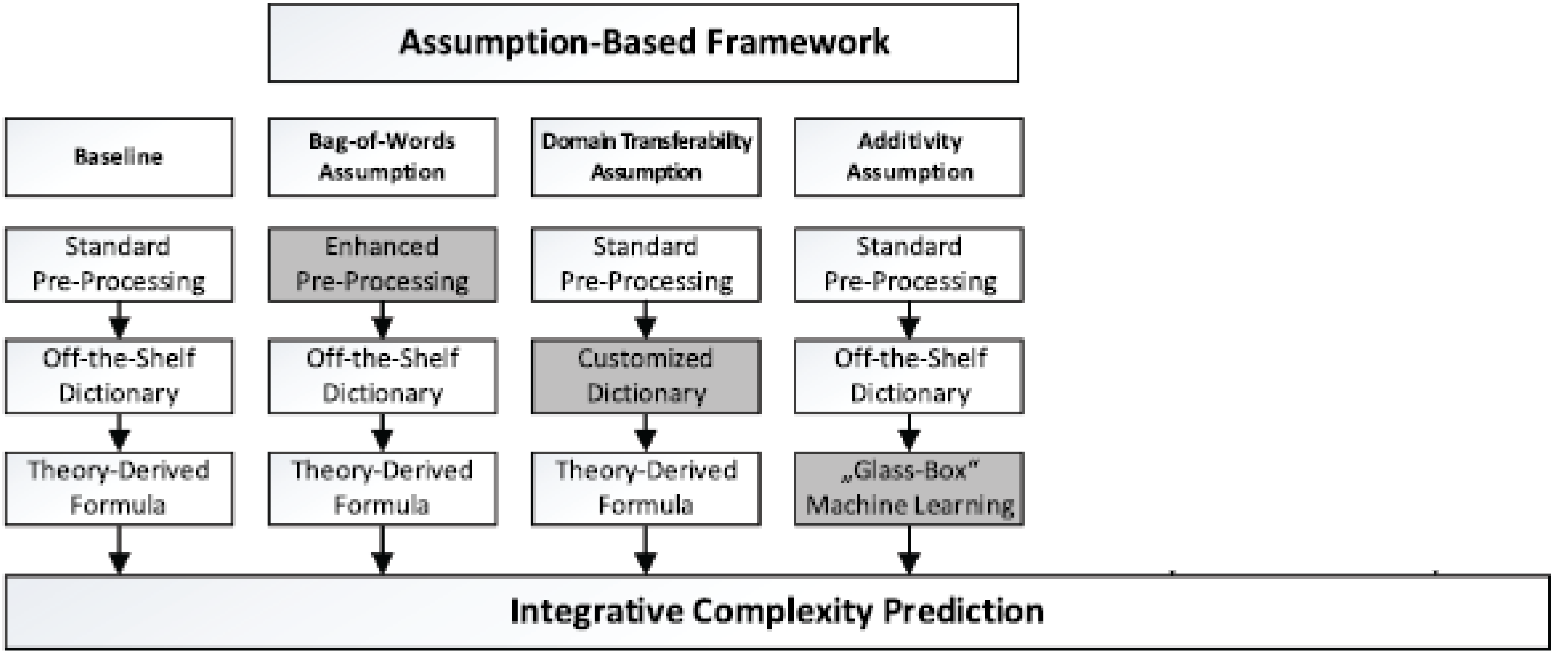


Figure 1. Overview of the Baseline, Assumption-Based Framework, and Shotgun Approaches. *Note*: Boxes highlighted in gray show the process steps that differ from the baseline process.

Three (problematic!) assumptions in using off-the-shelf dictionaries

1. Bag-of-words assumption:
   – Word order and grammatical functions are deemed irrelevant for capturing meaning

2. Domain transferability assumption:
   – The off-the-shelf dictionary is deemed equally applicable to all knowledge domains

3. Additivity assumption:
   – Each word has equal (or a predefined) weight and contributes accordingly to the classification

Which method works best for classifying integrative complexity?

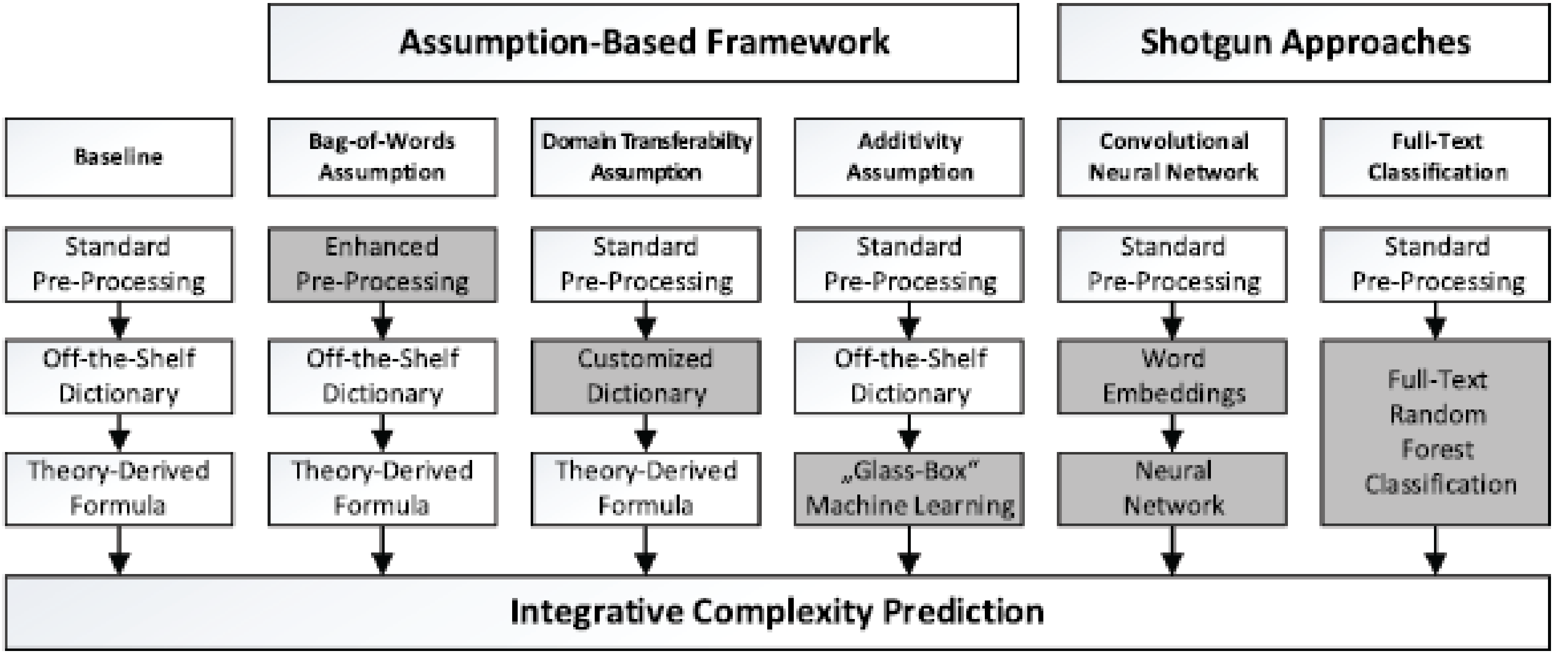


Figure 1. Overview of the Baseline, Assumption-Based Framework, and Shotgun Approaches. *Note*: Boxes highlighted in gray show the process steps that differ from the baseline process.

Which method works best for classifying integrative complexity?

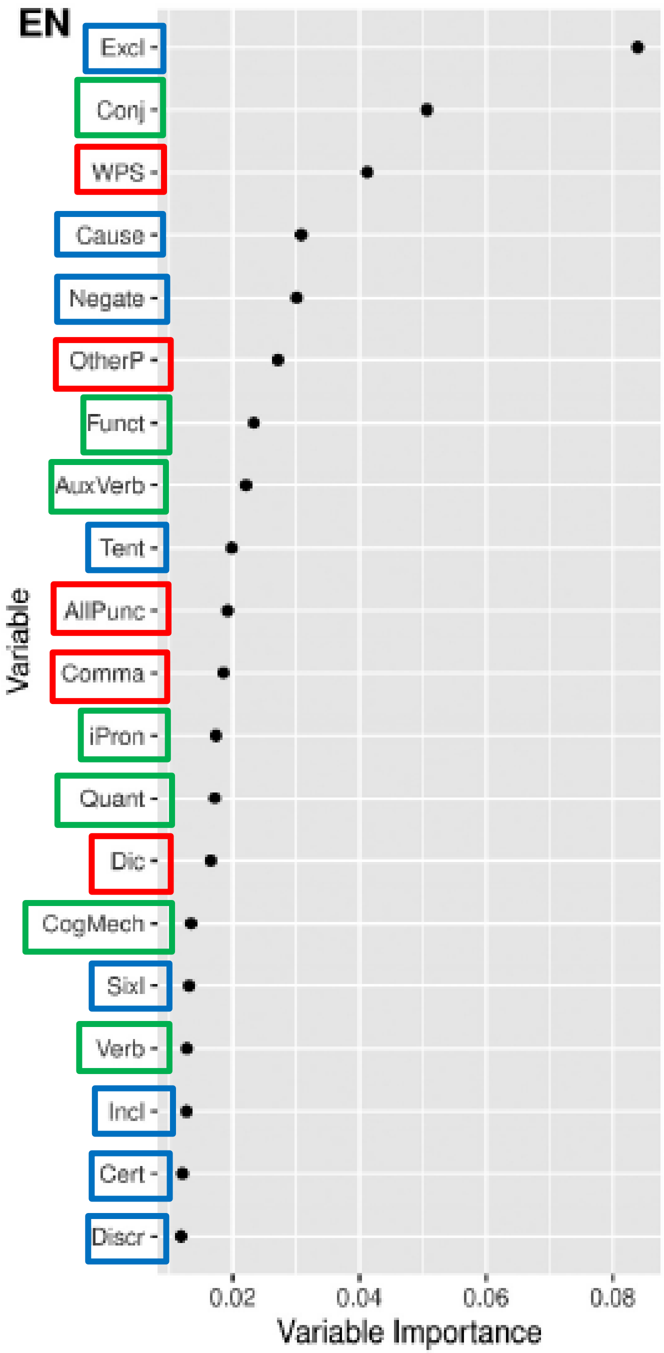
Table 1. Results of the Cross-Validation for the Different Approaches.

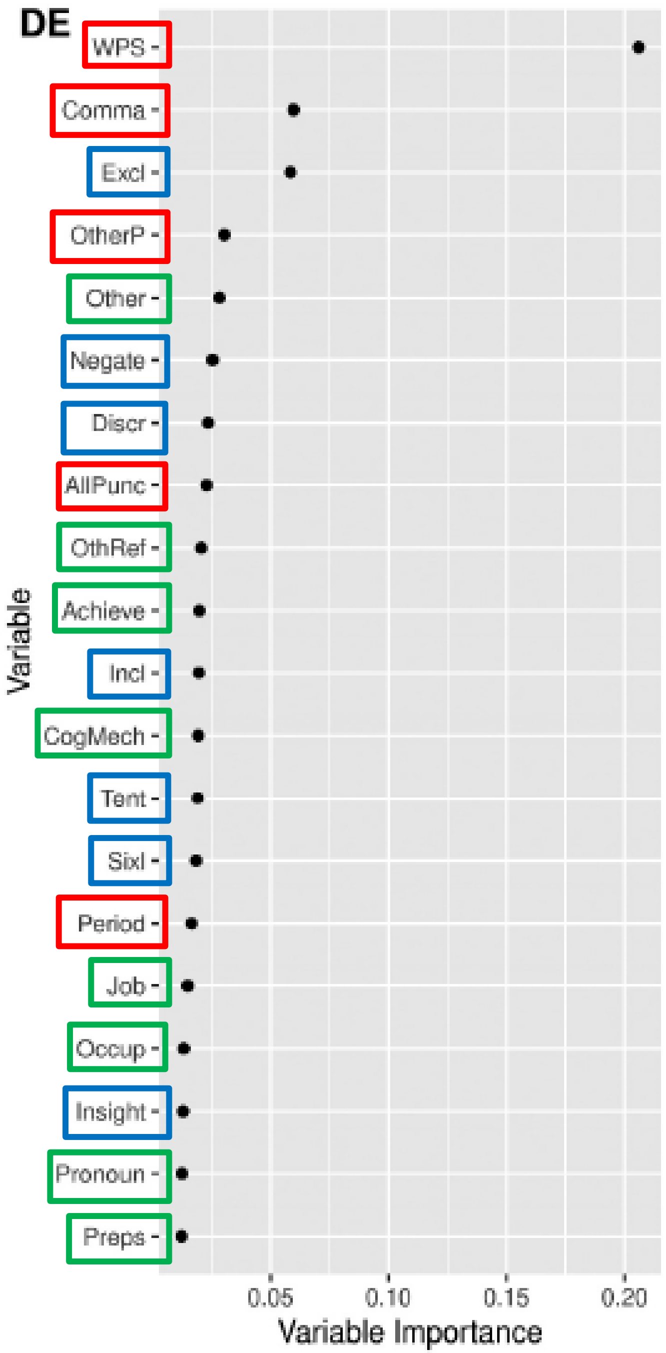
	English		German	
Approach	RMSE	Corr	RMSE	Corr
Baseline (Wyss et al., 2015)	30.30	−.07	32.90	.01
Assumption: Bag-of-words				
Negation (Young & Soroka, 2012)	31.40	−.04	34.40	.02
POS-tagging (Jacobi et al., 2015)	51.30	.11	62.60	.16
POS-tagging (Benamara et al., 2007)	48.60	.15	58.90	.08
Lemmatization (Haselmayer & Jenny, 2017)	22.50	−.06	29.30	.04
Assumption: Domain transferability				
Adj. word choices (Diesner & Evans, 2015) – 5%	25.80	.04	30.10	.10
Adj. word choices (Diesner & Evans, 2015) – 10%	24.60	.06	29.80	.11
Assumption: Additivity				
10 features				
Linear regression [1]	0.93	.30	1.04	.22
Lasso regression	0.93	.30	1.04	.22
M5P	0.75	.64	0.84	.62
Random forest regression	0.72	.68	0.78	.67
All features [2]				
Linear regression [1]	0.83	.52	0.84	.60
Lasso regression	0.83	.52	0.84	.61
M5P	0.76	.62	0.81	.65
Random forest regression	0.70	.70	0.75	.71
Shotgun full-text machine learning				
CNN (fastText Word Embeddings)	0.75	.71	0.84	.69
Random forest	0.76	.73	0.85	.72

Note: RMSE = root mean squared error. POS = part-of-speech. CNN = convolutional neural network.
[1] Please refer to Online Appendix II for the regression coefficients.
[2] The table reports the performance for machine learning models trained without word count as a feature (see section "Additivity assumption").

Results

- Theory-informed dictionary + supervised ML performs as well as shotgun approaches (i.e., full-text classification, deep learning)
- The additivity assummption is attacked best by our approach
  - Automatically assigning weights to words rather than equal or predefined weights
  - Randomly choosing the number of features to be included in the ensemble of decision trees
- In addition, Variable Importance analysis yields insights on the components of integrative complexity important for theory building





Which features drive IC classification?

Features contained in original theory

Additional formal features

Additional content features

Theoretical insights on IC derived from „glass-box“ ML

- Most important features driving IC classification:
  - Exclusiveness: but, except, without, etc.
  - Words per sentence
  - Conjunction: and, but, whereas, etc.
  - Punctuation (Comma, etc.)
- Only theoretically derived feature not showing up in TOP 20
  - Inhibition: block, constrain, stop, etc.
- Theoretically most interesting new content features:
  - Achievement: earn, hero, win, etc.
  - Cognitive Processes: cause, ought, etc.
  - Quantifiers: few, many, much, etc.

The advantages of „glass-box“ ML

Advantages (and limitations) of „glass-box“ ML

Advantages

- „Glass-box“ ML can yield not only good predictions but interpretable results on top
- It can validate and/or expand the theoretical base of the concepts underlying the classification
- Even if classification performance were below that of full-text deep learning approaches, this additional benefit might be worth a little trade-off
- If supervised machine-learning is combined with theory-driven word lists (dictionaries), this might answer the call for more theory-driven computational research to some degree
- Size of training data set can be smaller for dictionary-based ML classification than for full-text classification

Advantages (and limitations) of „glass-box“ ML

Limitations and open questions

- Good dictionaries do not exist for all languages
- It remains open whether dictionary translation works well for more distant languages (e.g., Chinese, Japanese, Arabic, Hebrew)
- It is unclear whether the combination of dictionaries + „glass-box“ ML works equally well for other constructs and in other topical domains
- To produce the desirable synergies mentioned earlier ML methods should be made more accessible and methods training in communication science should be expanded to include ML

Thank you for your attention!

wessler@uni-mannheim.de

Sources

- Beste, S., & Wyss, D. (2014, September 3–6). *Cognitive complexity as a proxy for high quality deliberation? A theoretical and empirical exploration of cognitive complexity and deliberative quality in the EuroPolis discussions* [Conference Presentation]. European Consortium for Political Research General Conference, Glasgow, United Kingdom.
- Conway III, L. G., Gornick, L. J., Burfeind, C., Mandella, P., Kuenzli, A., Houck, S. C., & Fullerton, D. T. (2012). Does complex or simple rhetoric win elections? An integrative complexity analysis of U.S. presidential campaigns. *Political Psychology, 33*(5), 599–618. https://doi.org/10.1111/j.1467-9221.2012.00910.x
- Dobbrick, T., Jakob, J., Chan, C., & Wessler, H. (2021). Enhancing theory-informed dictionary approaches with "glass-box" machine learning: The case of integrative complexity in social media comments. *Communication Methods and Measures.* https://doi.org/10.1080/19312458.2021.1999913
- Freelon, D. (2015). Discourse architecture, ideology, and democratic norms in online political discussion. *New Media & Society, 17*(5), 772–791. https://doi.org/10.1177/1461444813513259
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis, 21*(3), 267–297. https://doi.org/10.1093/pan/mps028
- Jakob, J., Dobbrick, T., & Wessler, H. (2021). The integrative complexity of online user comments across different types of democracy and discussion arenas. *The International Journal of Press/Politics.* https://doi.org/10.1177/19401612211044018

Sources

- Jennstål, J. (2019). Deliberation and complexity of thinking. Using the integrative complexity scale to assess the deliberative quality of minipublics. *Swiss Political Science Review, 25*(1), 64–83. https://doi.org/10.1111/spsr.12343
- Lijphart, A. (1975). *The politics of accommodation: Pluralism and democracy in the Netherlands*. (2nd ed.). University of California Press.
- Lijphart, A. (2012). *Patterns of democracy: Government forms and performance in thirty-six countries* (2nd ed.). Yale University Press.
- Maia, R. C. M., Hauber, G., Choucair, T., & Crepalde, N. J. B. (2020). What kind of disagreement favors reason-giving? Analyzing online political discussions across the broader public sphere. *Political Studies, 69*(1), 108–128. https://doi.org/10.1177/0032321719894708
- Owens, R. J., & Wedeking, J. P. (2011). Justices and legal clarity: Analyzing the complexity of U.S. Supreme Court opinions. *Law & Society Review, 45*(4), 1027–1061. https://doi.org/10.1111/j.1540-5893.2011.00464.x
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015*). The development and psychometric properties of LIWC2015*. University of Texas at Austin. https://repositories.lib.utexas.edu/handle/2152/31333
- Schudson, M. (1997). Why conversation is not the soul of democracy. *Critical Studies in Mass Communication, 14*(4), 297–309. https://doi.org/10.1080/15295039709367020
- Steiner, J., Bächtiger, A., Spörndli, M., & Steenbergen, M. R. (2004). *Deliberative politics in action: Analyzing parliamentary discourse* (1st ed.). Cambridge University Press.

Sources

- Suedfeld, P., Tetlock, P. E., & Streufert, S. (1992). Conceptual/integrative complexity. In C. P. Smith (Ed.), *Motivation and Personality: Handbook of Thematic Content Analysis* (1st ed., pp. 393–400). Cambridge University Press. https://doi.org/10.1017/CBO9780511527937.028
- Tetlock, P. E. (1983). Cognitive style and political ideology. *Journal of Personality and Social Psychology, 45*(1), 118–126. https://doi.org/10.1037/0022-3514.45.1.118
- Thoemmes, F. J., & Conway III, L. G. (2007). Integrative complexity of 41 U.S. presidents. *Political Psychology, 28*(2), 193–226. https://www.jstor.org/stable/20447033
- Wessler, H., & Rinke, E. M. (2014). Deliberative performance of television news in three types of democracy: Insights from the United States, Germany, and Russia. *Journal of Communication, 64*(5), 827–851. https://doi.org/10.1111/jcom.12115
- Wyss, D., Beste, S., & Bächtiger, A. (2015). A decline in the quality of debate? The evolution of cognitive complexity in swiss parliamentary debates on immigration (1968–2014). *Swiss Political Science Review, 21*(4), 636–653. https://doi.org/10.1111/spsr.12179
- Zhang, W., Cao, X., & Tran, M. N. (2013). The structural features and the deliberative quality of online discussions. *Telematics and Informatics, 30*(2), 74–86. https://doi.org/10.1016/j.tele.2012.06.001