1.3 Applied Uniqueness, Naturality & Readability Measures

Without loss of generality, the quality score we present in the article could be adapted to incorporate other linguistic components. The software tool employed in our empirical application studies implements the components content uniqueness (s_d), naturality similarity (s_n) and readability similarity (s_r) as follows:

Uniqueness measurement (s_d). For our quality score (qs_g), we derive a uniqueness measure (s_d) to assess if the content is sufficiently unique for the search engine. In addition to the definitions around formula (3) in the main manuscript, we apply a critical value (s_{cv}) to ensure that the generated content is sufficiently unique based on the length of the keyword (kw) and parameter b.

$$s_{cv} = (100 - (100/(kw + 1)^b))/100 \quad \text{(W5)}$$

By implementing this non-compensatory filtering rule we ascertain that content that fails to achieve this minimum level of uniqueness is discarded from further content selection. The value b determines the factor of increasing conservativeness the larger the n-gram size (kw+1), as repeating small sized n-grams is less of a concern than repeating large sized n-grams (W5). In our setup, we set b to 1.1 after an evaluation phase in which we look at a) the machine output, b) acceptable duplicate rates in human content impressions, and c) content retaining rates for the whole range of common n-gram sizes. For example, that means that with an n-gram size of 3, s_{cv} ~.70 (i.e., 70% unique), an n-gram size of 5, s_{cv} ~.82 (i.e., 82% unique), and an n-gram-size of 7, s_{cv} ~.88 (i.e., 88% unique).

Naturality similarity measures (s_n). To quantify the naturality similarity between the generated content and the top ranked search results, we applied 12 linguistic measures which assess the lexical richness and composition of a text using the R package languageR.

Specifically, we use the following measures: tokens, types, hapax legomena, dis legomena, tris legomena, Yule's K, Zipf's R, Type-Token-Ratio, Herdan's C, Guiraud's R, Sichel's S, Lognormal. More information on the precise meaning, practical examples and literature sources can be found in Baayen and Shafaei-Bajestan (2019).

Readability similarity measures (s_r). For the readability similarity measure, we applied 46 pre-existing measures of readability contained in the R package quanteda (see Benoit et al. 2018). We make use of the following measures: ARI, Bormuth.MC, Bormuth.GP, Coleman, Coleman.C2, Coleman.Liau.ECP, Dale.Chall, Dale.Chall.PSK, Danielson.Bryan, Dickes.Steiwer, DRP, ELF, Farr.Jenkins.Paterson, Flesch, Flesch.PSK, Flesch.Kincaid, FOG, FOG.PSK, FOG.NRI, FORCAST, FORCAST.RGL, Fucks, Linsear.Write, nWS, nWS.2, nWS.3, nWS.4, RIX, Scrabble, SMOG, SMOG.C, Spache, Spache.old, Strain, Traenkle.Bailer, W, St, C, Sy, W3Sy, W2Sy, W_1Sy, W6C, W7C, Wlt3Sy, W_wl.Dale.Chall. More information on the precise meaning, calculation and literature sources can be found in Benoit et al. (2018).

Appendix References

Baayen RH, Shafaei-Bajestan E (2019) Analyzing linguistic data: A practical introduction to statistics. Package 'languageR'. Version 1.5.0. *CRAN*. Accessed May 20, 2019, https://cran.r-project.org/web/packages/languageR/languageR.pdf

Benoit K, Watanabe K, Wang H, Nulty P, Obeng A, Müller S, Matsuo A, (2018) "quanteda: An R package for the quantitative analysis of textual data." *Journal of Open Source Software*. 3(30). https://doi.org/10.21105/joss.00774

Berger J, Sherman G, Ungar L (2020b) TextAnalyzer. Accessed November 11, 2020, http://textanalyzer.org

Bronnenberg BJ, Kim JB, Mela CF (2016) Zooming in on choice: How do consumers search for cameras online? *Marketing Science*. 35(5):693-712.

Danaher PJ, Mullarkey GW, Essegaier S (2006) Factors affecting website visit duration: A cross-domain analysis. *Journal of Marketing Research*. 43(2):182-194.

Edelman B, Zhenyu L (2016) Design of search engine services: Channel interdependence in search engine results. *Journal of Marketing Research*. 53(6):881-900.

Flanigan, AJ, Metzger, MJ (2007) The role of site features, user attribtues, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*. 9(2):319-342. https://doi.org/10.1177/1461444807075015

Jerath K, Ma L, Park YH (2014) Consumer click behavior at a search engine: The role of keyword popularity. *Journal of Marketing Research*. 51(4):480-486.

Kamoen N, Holleman B, Bergh H (2013) Positive, negative, and bipolar questions: The effect of question polarity on ratings of text readability. *Survey Research Methods*. 7(3):181-189.

Liu J, Toubia O (2018) A semantic approach for estimating consumer content preferences from online search queries. *Marketing Science*. 37(6):930-952.

Maechler M, Rousseeuw P, Croux C, Todorov V, Ruckstuhl A, Salibian-Barrera M, Verbeke T, Koller M, Conceicao ELT, Palma MA (2020) Basic robust statistics. Package 'robustbase'. Version 0.93-6. *CRAN*. Accessed May 20, 2020, https://cran.r-project.org/web/packages/robustbase/robustbase.pdf

Pennebaker JW, Booth RJ, Boyd RL, Francis ME (2015) Linguistic inquiry and word count: LIWC2015. Austin, TX: Pennebaker Conglomerates. Accessed November 1, 2020, www.LIWC.net.

Pitler E, Nenkova A (2008) Revisiting Readability: A unified framework for predicting text quality. *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*. 186-195.

Radford A, Narasimhan K, Salimans T, Sutskever I (2018) Improving language understanding by generative pre-training. OpenAI.

Roberts C (2010) Correlations among variables in message and messenger credibility scales. *American Behavioral Scientist*. 54(1):43-56.

Rocklage MD, Rucker DD, Nordgren LF (2018) Persuasion, emotion and language: the intent to persuade transforms language via emotionality. *Psychological Science*. 29(5):749-760.

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomze AN, Kaiser L, Polosukhin I (2017) Attention is all you need. *31st Conference on Neural Information Processing Systems (NIPS 2017)*. 1-15.