# AI Policies: a quantitiative document analysis

**Companion to a dashboard presentation**

John Little

2024-01-22

Ithaka S+ R convened a two-year research project in March of 2023.(Cooper, Ruediger, and Schonfeld 2023) Yakut Gazi (PI, DLI) and Joe Salem (Library) are chairing a local cohort[1] charged with conducting the survey centered around a *qualitative inquiry* protocol. An early phase of year-one will conduct interviews with a subset of local university personnel involved in research, AI, and policy at Duke.

An additional phase of year-one, before on-site interviews are conducted, includes the qualitative analysis of the following policy documents with the possibility of a larger corpus.

1. *DKU Guide for Teaching and Generative AI*.(Duke Kunshan University 2023)
2. *Artificial Intelligence Policies: Guidelines and Considerations.*(Duke and Innovation 2023)
3. Guidance for the use of Artificial Intelligence Tools for Academic Assignments in MD Program.(Bulletin and Duke University School of Medicine, n.d.)

The documents were qualitatively assessed and classified per an a priori taxonomy. Additionally, the documents were quantitatively assessed vis-a-via the standard text-mining algorithms: *word frequency*, and *Term-Frequency-Inverse Document Frequency* (TF-IDF) of single-words and bi-grams. The visualizations of quantitative text-mining analysis are included below. The code for the analysis can be found on GitHub.(Little 2024)

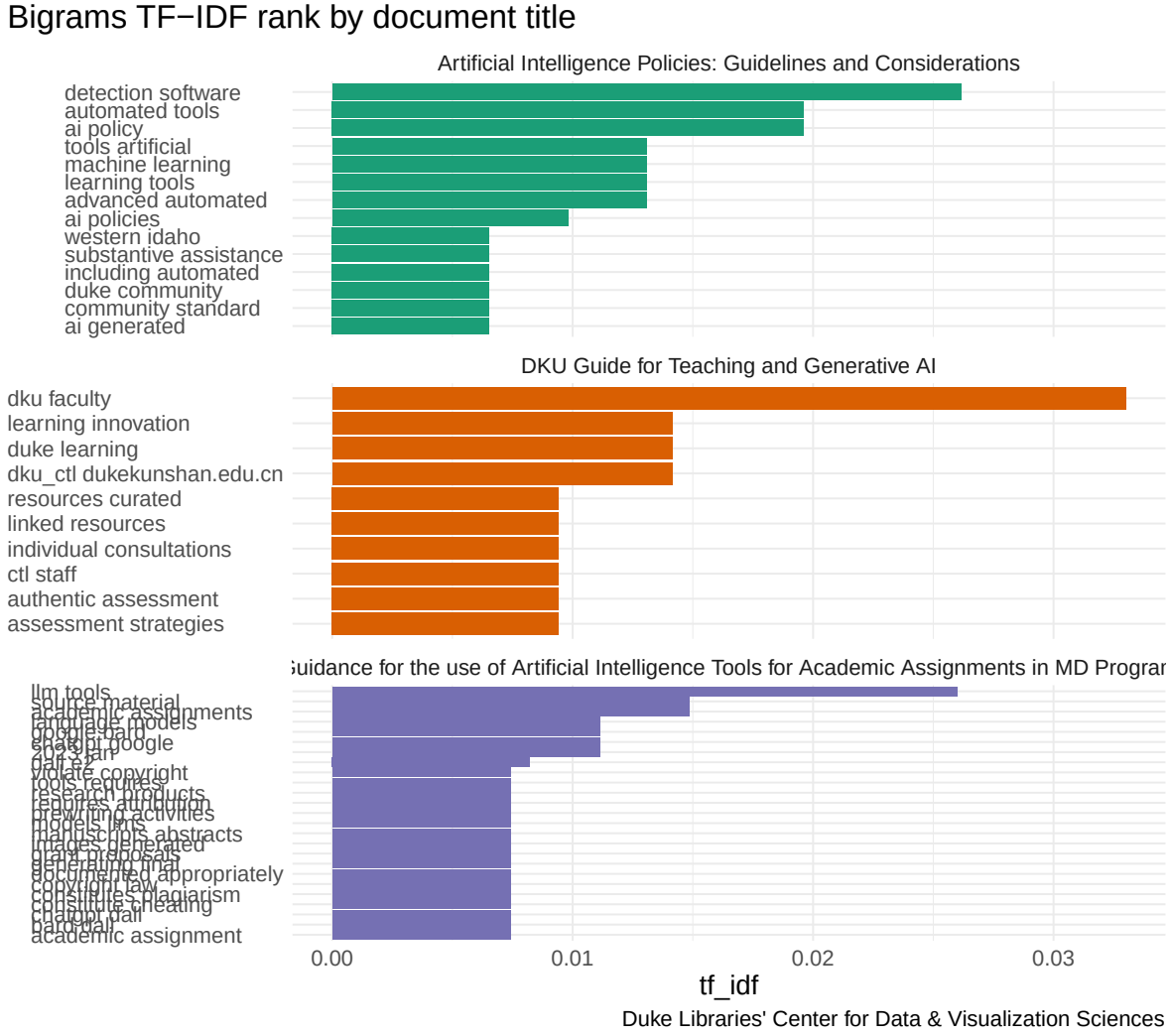
## Definitions

**N-grams** A contiguous sequence of n items from a given sample of text or speech. The items can be phonemes, syllables, letters, words, or base pairs according to the application.

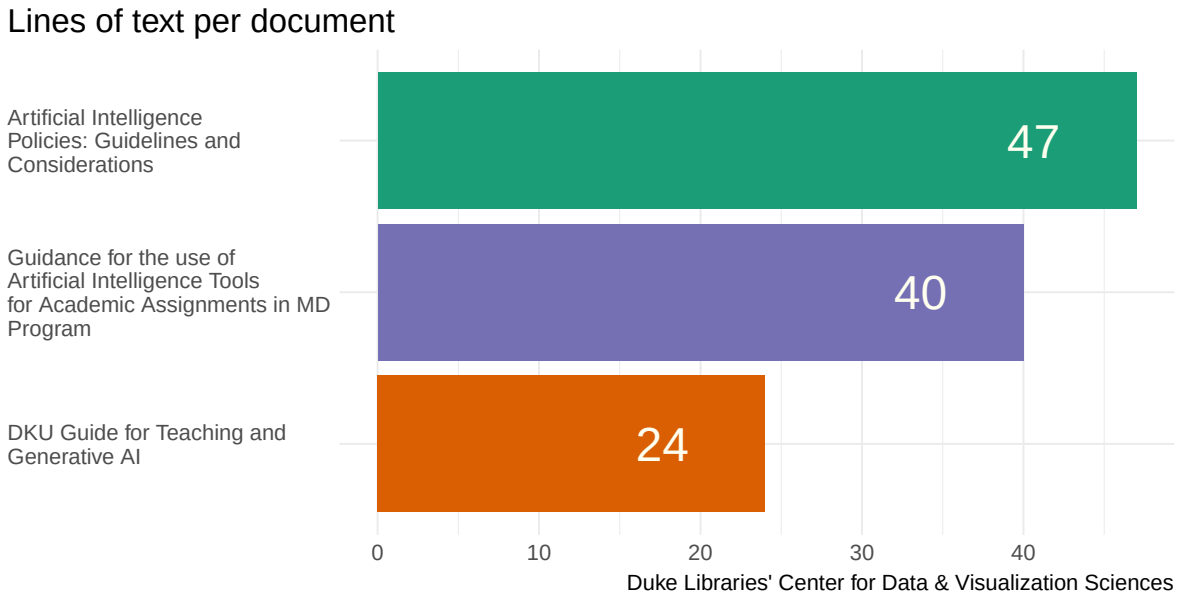
[1] Linda Daniel; John Little; Greay Reavis; Xinzhu Wang

**TF-IDF** A numerical statistic intended to reflect how important a word is to a document in a collection or corpus. It is often used as a weighting factor in searches for information retrieval, text mining, and user modeling. The TF-IDF value increases proportionally to the number of times a word appears in the document and is offset by the number of documents in the corpus that contain the word, which helps to adjust for the fact that some words appear more frequently in general.

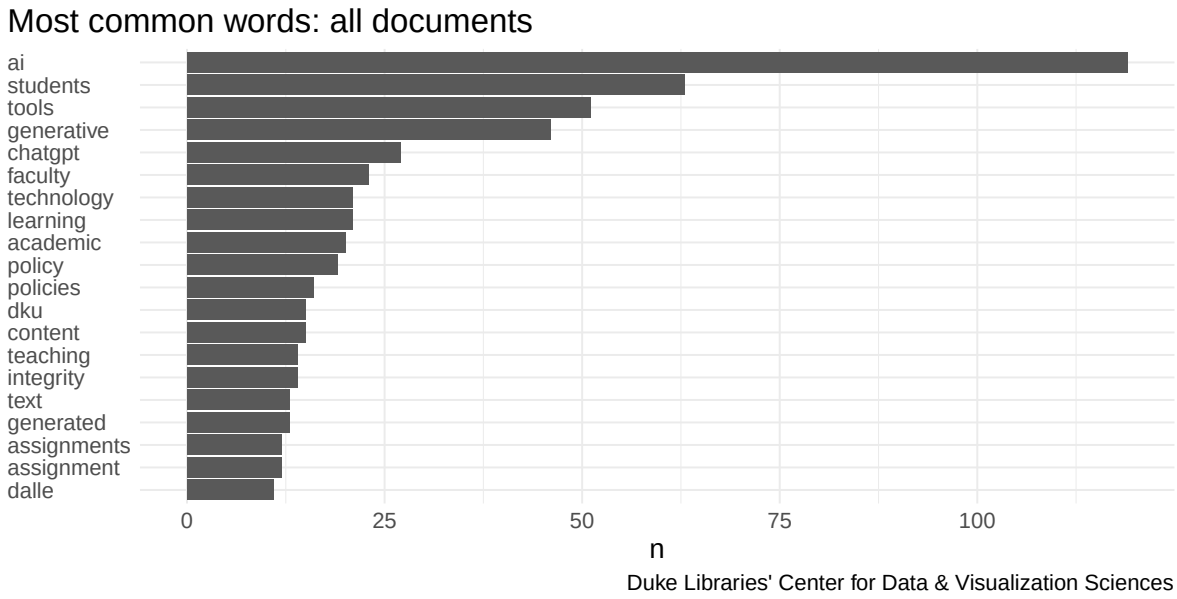
## Bi-grams TF-IDF rank by document title



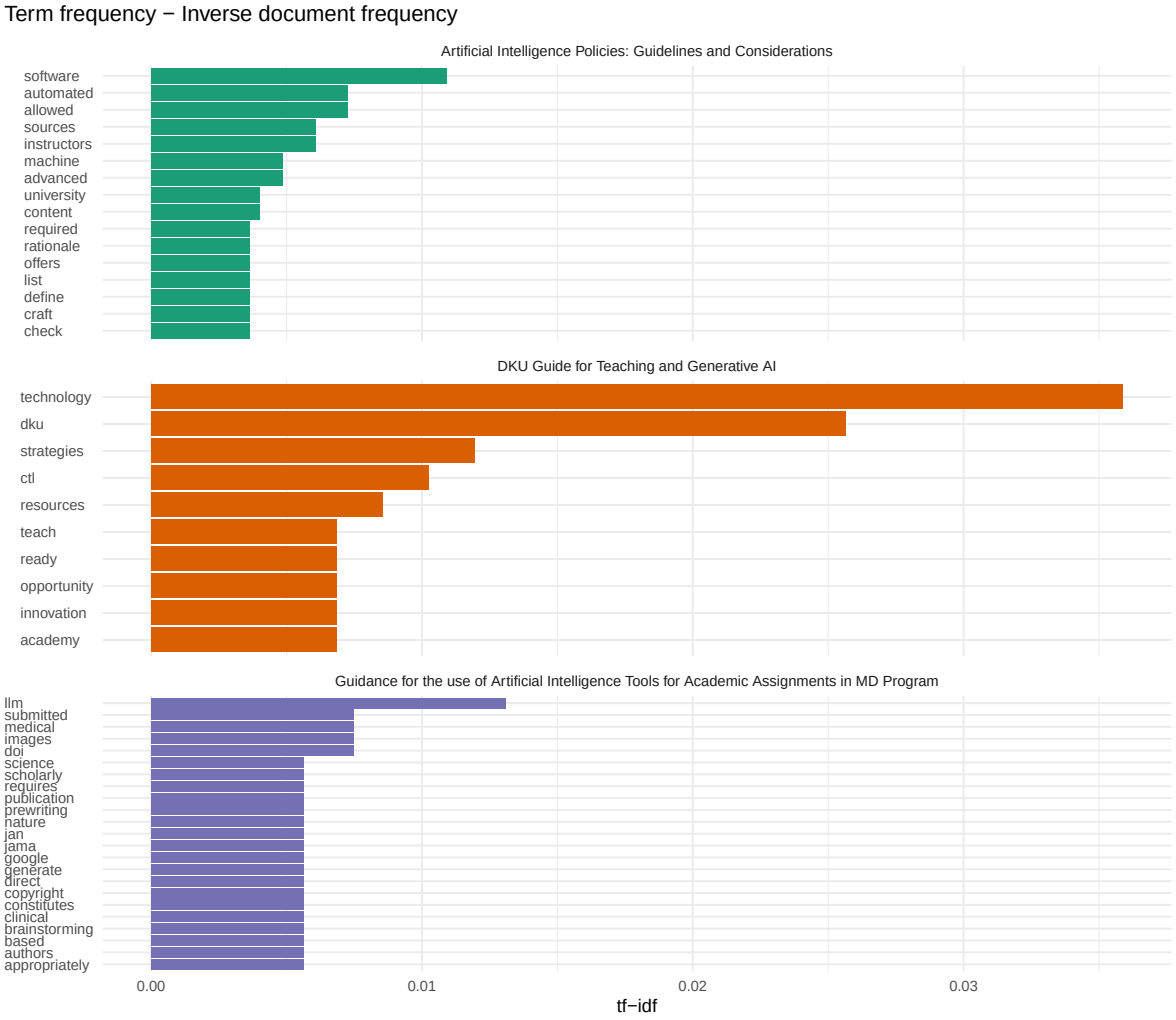
## Lines of text per document




## Word frequency of all words across all documents in the corpus

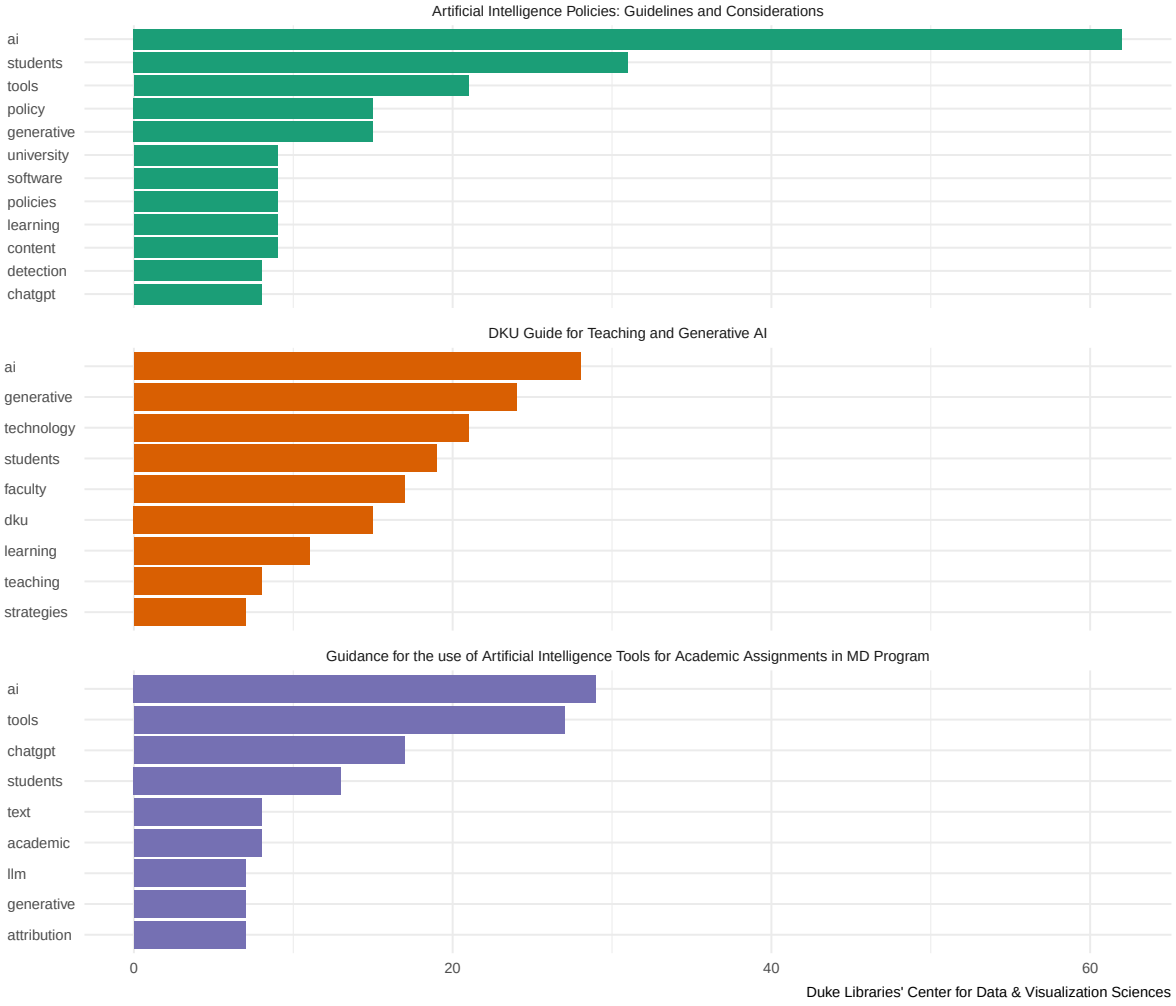


## The most common words in each document



## Term-Frequency - Inverse Document Frequency (TF-IDF) of each document

Most common words



**Word Cloud**


specific
sources
permitted
including
models
limitations
include
knowledge
resources
material
plagiarism
automated
writing
information
critical
strategies
encourage
generated
intelligence
education
academic
software
attribution
generative
text
results
e2
2
tools
teaching
llm
ctl
chatgpt
ai
learning
dalle
policy
faculty
2023
duke
students
policies
allowed
tool
technology
dku
detection
artificial
content
integrity
citation
scientific
support
assignment
university
guidance
assignments
thinking
courses
assessment
research
instructors
syllabus
explore
source
language
student
understand
generating

Bulletin, 23-24 School of Medicine, and Duke University School of Medicine. n.d. "Guidance for the Use of Artificial Intelligence Tools for Academic Assignments in MD Program." https://medicine.bulletins.duke.edu/allprograms/dr/duke_peoplesoft-catalog.coursedog.com.

Cooper, Danielle, Dylan Ruediger, and Roger Schonfeld. 2023. "Making AI Generative for Higher Education." https://sr.ithaka.org/blog/making-ai-generative-for-higher-education/.

Duke Kunshan University. 2023. "DKU Guide for Teaching and Generative AI - ." https://www.dukekunshan.edu.cn/center-for-teaching-and-learning/faculty-resource-guide-teaching-with-ai/.

Duke, Learning, and Innovation. 2023. "Artificial Intelligence Policies: Guidelines and Considerations." https://learninginnovation.duke.edu/ai-and-teaching-at-duke-2/artificial-intelligence-policies-in-syllabi-guidelines-and-considerations/.

Little, John. 2024. *Libjohn/Analysis_ai_ithaka*. https://github.com/libjohn/analysis_Ai_ithaka.