



Topic Modeling and/as Genre Study

Moriah L. Kirdy | University of Pittsburgh

moriah.kirdy@pitt.edu | @moriahlkirdy

Slides and talk transcript available on my website:

<https://moriahlkirdy.com/talks>

May, 25

- * • Call landlord
- ✕ Plan vacation
- > Task List Item
- > Task List Item
- < Dinner with FM next month
- Take in laundry
- ! — The dress is blue and black
- ☹ — The gym is closed saturday



Weekly spread, Bullet Journal example from <https://zenofplanning.com/bullet-journal-weekly-spread> (original source, @Jenna Meyers on Instagram)

“Say we’ve got a collection of documents, and we want to identify underlying ‘topics’ that organize the collection. Assume that each document contains a mixture of different topics. Let’s also assume that a ‘topic’ can be understood as a collection of words that can have different probabilities of appearance in passages discussing the topic....”

“...Of course, we can’t directly observe the topics; in reality all we have are documents. Topic modeling is a way of extrapolating backward from a collection of documents to infer the discourses (‘topics’) that could have generated them.”

~ Ted Underwood, “Topic Modeling Made Just Simple Enough” <https://tedunderwood.com>

“We formally define a *topic* to be a distribution over a fixed vocabulary... We assume that these topics are specified before the data has been generated” [with a footnote that reads: “Technically, the model assumes that the topics are generated first, before the documents.”]

~ David Blei, “Probabilistic Topic Models”

“[O]ne never writes or speaks in a void. What fills that void is not only cultural context (ideological and material baggage surrounding our every action) and situational context (the people, languages, and purposes involved in every action) but also generic context, the existing genres we have read or written or that others say we should read or write...”


**“...genres are always already existing ...
As opposed to an abstract concept of
genre, the context of genres is the
existence of particular genres, the
already existing textual classifications
and forms already established and being
established within a given culture, the
set of typified rhetorical actions already
constructed by participants in a society”**

~ Amy Devitt, *Writing Genres*



A smattering of topic-modeling studies with text collections that cohere as genres

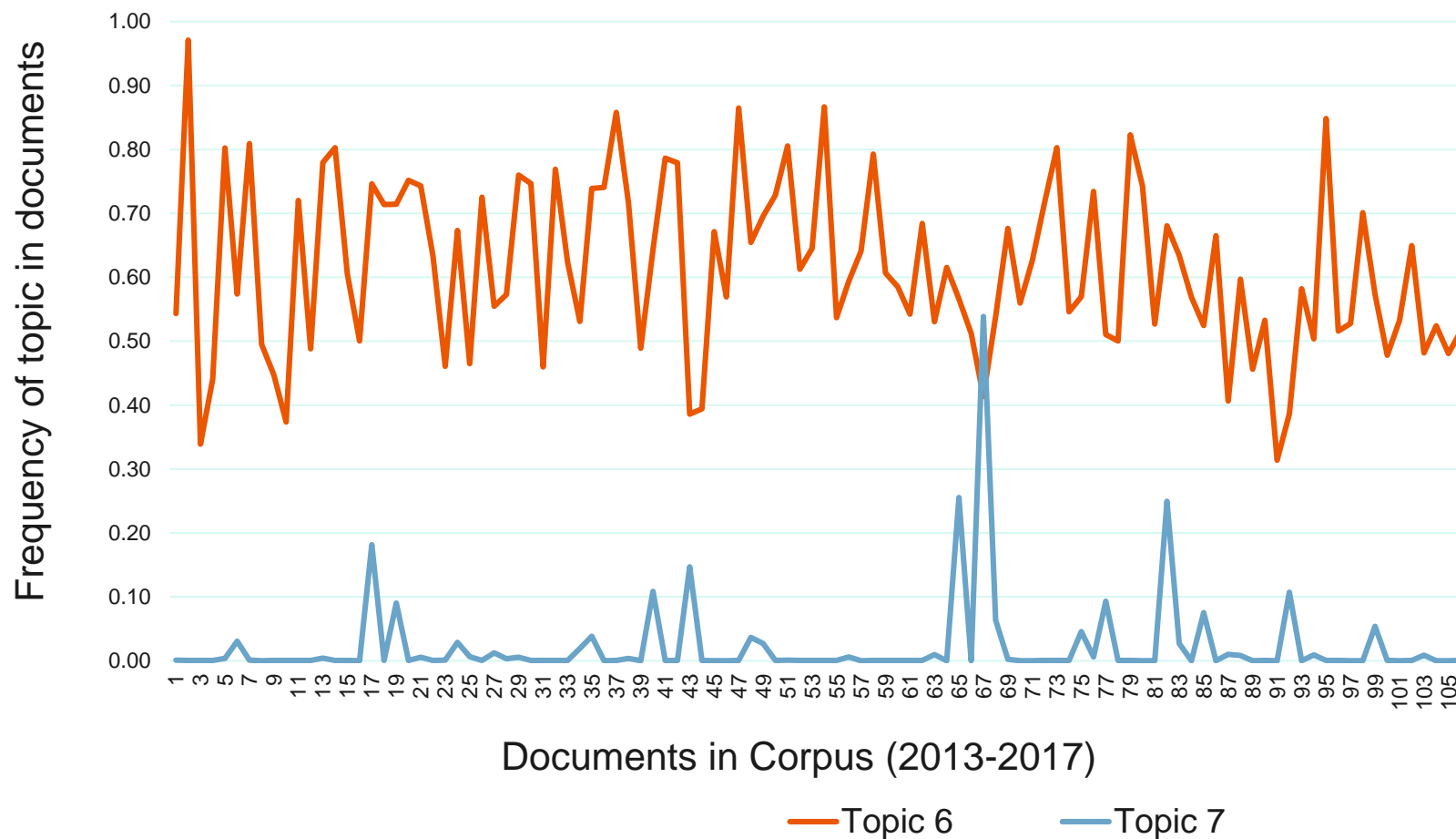
- **Ships' logs:** Schmidt, Benjamin "Words Alone: Dismantling Topic Models in the Humanities" *Journal of Digital Humanities*
- **Literary Scholarship:** Underwood, Ted & Andrew Goldstone, "What can topic models of PMLA teach us about the history of literary scholarship?" www.tedunderwood.com
- **Newspaper Articles from the *Pennsylvania Gazette* from 1728-1800:** Block, Sharon "Doing More with Digitization" *Common-Place*
- **Comp/Rhet Dissertation Abstracts:** Miller, Benjamin "Mapping the Methods of Composition/Rhetoric Dissertations: A "Landscape Plotted and Pieced" *College Composition and Communication*
- ... so many of the studies!



Topic 6	2.92662	Bullet_journal page pages notebook list day journal things time month system make love work planner daily monthly find post write
Topic 7	0.09472	Collections pen daily_log signifiers task entries future_log collection add index simply monthly_log bullets migrate leuchtturm entry log write migration fountain

Topic 6 (Bullet Journal, page, pages, notebook...)

Topic 7 (collections, pens, daily_log, signifiers...)





Topic Modeling and/as Genre Study

Moriah L. Kirdy | University of Pittsburgh

moriah.kirdy@pitt.edu | @moriahlkirdy

Slides and talk transcript available on my website:

<https://moriahlkirdy.com/talks>