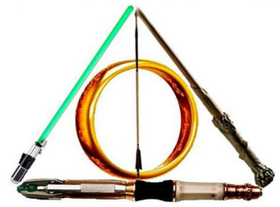
Novelty and Success of Fanfictions in Fandoms

Elise Jing[1], Yong-Yeol Ahn[1], and Simon Dedeo[2]*

1 Center for Complex Networks and Systems Research, Indiana University
2 Social and Decision Sciences, Carnegie Mellon University

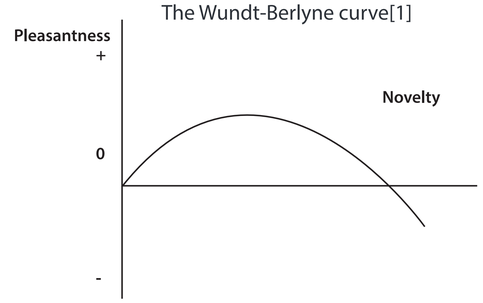
*simon.dedeo@gmail.com



Motivation

The creative industries want to know: What is the connection between novelty and success?

The Wundt-Berlyne curve[1]



• Classical theory: the Wundt-Berlyne curve. The reversed U-shaped curve suggests that increasing novelty first brings up pleasantness; after reaching a certain threshold, more novelty will result in a decline of pleasantness.

• Fandoms are (online) communities where people create and consume fan works, with fanfictions being a popular form.

• In fandoms, authors write on shared topics and compete for success. Does the same pattern of novelty and success hold?

Acknowlegement

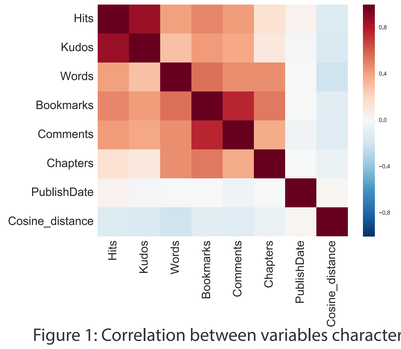
The authors thank the AO3 staff for helping with collecting information from the site.

References

[1] Sluckin, W., Hargreaves, D. J., & Colman, A. M. (1983). Novelty and human aesthetic preferences. Exploration in animals and humans, 245-269.

Results

• We developed measurements for a fiction's novelty, using the fiction's distance to the "average" fiction of its fandom as an indicator.
• A Negative relationship is identified between a fiction's novelty and the Kudos it receives by examining correlations and fitting a Negative Binomial regression model.



Negative Binomial Regression Results	
	Kudos
Words	1.685e-05 ***
Chapters	-4.427e-05 *
PublishDate	-7.503e-05 ***
Cosine distance	-2.912e+00 ***
Constant	4.751e+00 ***
Observations	22044
Signif. codes: 0 '***' 0.001 '**' 0.01 '*'	

Figure 1: Correlation between variables characterizing a fiction's length ("Words" and "Chapters"), publish time ("PublishDate") , novelty ("Cosine_distance"), and success ("Hits", "Kudos", "Bookmarks" and "Comments").

• A consistent pattern across fandoms is observed, where the increasing of novelty is generally associated with the decreasing of Kudos.

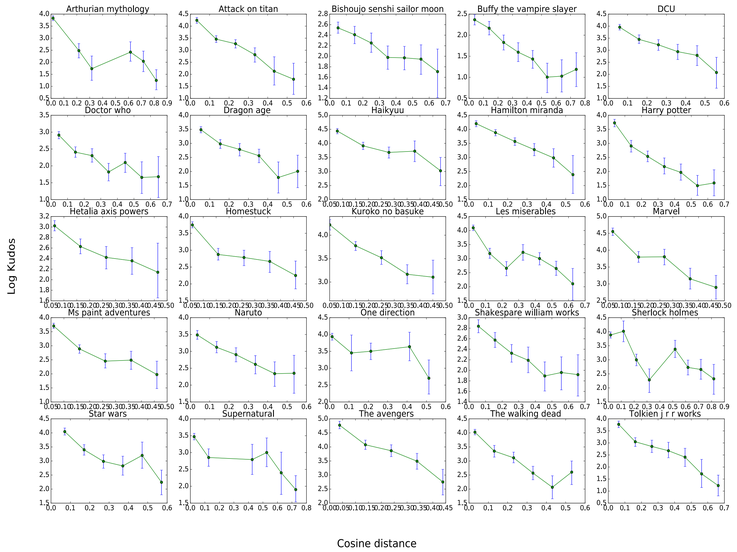


Figure 2: Relationship between novelty and Kudos in 25 fandoms. x-axis: the cosine distance between fictions and the "average" fictions of their fandoms, respectively, ranging between 0-1 and with a bin width of 0.1. y-axis: the average number of Kudos that fictions in this interval receive, on a log-scale.

Methods

Dataset: Archive of Our Own beta

• 25 most popular fandoms in AO3's 10 categories of creative works.
• 904,760 fictions from 132,162 authors.

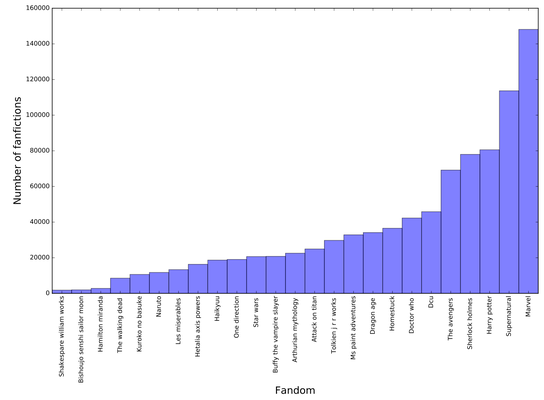


Figure 3: The distribution of the amount of fictions in the AO3 dataset.

Modeling & analysis

• Data collection: we used a Python web scraper to collect data from the AO3 website. Information collected includes the fictions' texts, and metadata about the author, publication, and reception from their readers.
• Pre-processing: only fictions written in English with >500 words were analysed. We also removed stopwords and tokens with low frequency.
• Modeling the fictions: we created a unigram language model with Simple Good-Turing smoothing to represent each fiction as a vector in the vector space of all fictions.
• Measuring novelty: an "average" fiction is defined as the center point in the vector space. Novelty of a fiction is measured by the cosine distance between the vector representations of the "average" fiction and itself.

Discussion

• Using computational linguistics tools, we analysed a large dataset of fanfictions, and found that in fandoms, fictions with high novelty tend to be less successful.
• The classical Wundt-Berlyne curve posits that increasing novelty will first lead to an increase, then a decline of success. Our findings show a different pattern, suggesting that traditional theories may not fully capture the behavior of creators and consumers in fandoms.