

# Topic Modeling for Makerspace Artifact Analysis

Dominique Exley,<sup>1</sup> Johanna Okerlund,<sup>2</sup> and David Wilson<sup>2</sup>

<sup>1</sup>Winthrop University

<sup>2</sup>University of North Carolina at Charlotte

dominique.exley@gmail.com, jokerlun@uncc.edu, davils@uncc.edu

## Abstract

As the making phenomenon becomes more prevalent, diverse, and vast, it becomes increasingly challenging to identify general temporal or spatial trends in types of making endeavors. Identifying trends in what participants are making is important to makerspace leaders who seek to understand the impact of the making phenomenon on the world or who are interested in broadening participation within their own maker contexts. This paper shows how topic modeling by means of LDA can be used to analyze maker artifacts, and illustrates how these types of insights can be used to make inferences about the making phenomenon, as well as to inform efforts to broaden participation.

## Introduction

The making phenomenon encompasses a recent wave of emphasis on Makerspaces, Fab Labs, Hackerspaces, and digital fabrication technology, which provide opportunities for people to explore their creativity by making new tangible artifacts. Makerspaces have become increasingly popular in schools, universities, libraries, and communities around the world. Communities of makers can consist of all age groups, backgrounds, and disciplines. One commonality that defines all makerspaces is the act of making; the process of manifesting an idea into a digital or physical product. This process almost always involves the use of digital fabrication machines such as 3-D printers and laser cutters. Other tools and materials for making include micro-controllers, sewing machines, conductive thread, and woodworking tools.

Champions of these spaces and enthusiasts of the making phenomenon cite the promise of making to promote learning and personal agency, spark collaborations, enable entrepreneurs to develop ideas, and support social goals such as sustainability. (Blikstein 2013; Bull et al. 2010; Gershenfeld 2008; Hui and Gerber 2017; Hsu, Baldwin, and Ching 2017; Martin 2015). However, there has also been some critique about how the making phenomenon has not yet lived up to its promises, such as falling short on being truly open for all (Bardzell et al. 2017; Roedl, Bardzell, and Bardzell 2015). In order to help makerspace leaders, researchers, and designers better understand how to guide their

communities, it is important to have an understanding of the maker communities themselves.

Research in Human-Computer Interaction (HCI) has made some steps toward understanding maker-related people, practices, spaces, and communities. In particular, this research often focuses on helping to design tools and environments for makers and to spread maker empowerment (Bardzell et al. 2017). Complementary research also tries to understand the implications of makerspaces relative to community, identity, education, and entrepreneurship (Sheridan et al. 2014). Methodologies for studying making are typically ethnographic, local, and aim to uncover the nuances of making, rather than generalize across the entire phenomenon. For example, interviews and case studies have been used to gain an understanding of the inner-workings of maker communities (Carucci and Toyama 2019; Toombs, Bardzell, and Bardzell 2015). Kuznetsov and Paulos conducted one of the broadest studies of the maker phenomenon by surveying 2600 users of online maker platforms such as Instructables, Etsy, and Craftster (Kuznetsov and Paulos 2010). However, this massive survey barely scratches the surface of the vast and dynamic making phenomenon and surveys fall short of ever being able to capture how values, conversations, and interests in the making phenomenon change over time.

Additionally, while the social aspects of the making phenomenon have been studied in depth, comparatively little research has addressed the study of material artifacts that come out of these spaces either at a small scale or broad scale. Thus, the aim of our research is to explore how Artificial Intelligence techniques can be used to understand trends in the artifacts that are created by people in the making phenomenon as a step towards understanding their impact on the world beyond. To expand the reach of existing makerspace research, we ask the following questions:

- How can artifact classification lead to a deeper understanding of what is happening in the making phenomenon?
- In what ways can the insights gained from artifact classification expand positive maker activities that ripple outside of internal communities?

## Methodology

To perform the artifact analysis, we conducted data collection and topic modeling, as detailed below. In the topic modeling stage we focused on finding thematic structures within the set of tags associated with each artifact. We considered these tags to be the best overall indicators of an artifact's essence.

### Data Collection

In order to create a topic model of maker artifacts, a representative data set needed to be acquired. The makerbot Thingiverse website claims to be the “world’s largest 3-D printing community”<sup>1</sup> and contains over a million 3-D models; the models and the information attached to them are referred to as ‘things’. Users of the site upload their designs and share them with others. ‘Things’ can be downloaded, liked, modified, collected for later use, or commented on. Studying artifacts from this site allows us to get an idea of what kind of artifacts the general population of makers is creating. The artifacts used in our study’s analysis come exclusively from the site. We drew from the site’s categories of Hobby, Art, and Household.

Thingiverse employs a REST API to manage the site’s vast collection of ‘things’. This API is used by developers to create tools that can be deployed as plugins hosted on the website, standalone web apps, or desktop applications. Typically, calls are made to a distinct API URL and JSON or XML responses are returned. For our project, informational details for set of ‘things’ were gathered and then analyzed. Particular data points of interest on each ‘thing’ included: ID, name, date added, date modified, like count, collect count, download count, view count description, instructions, and tags. We focused primarily on tags, which are keywords that are defined by the user when they upload an artifact.

The data set we collected contains 105,688 unique artifacts from three categories with a total of 33,961 unique tags. Individually, the Hobby, Art, and Household categories contained 44,872, 44,870, and 15,948 unique artifacts respectively. Furthermore the categories contained 21,809, 8,879, and 13,743 unique tag descriptors.

### Topic Modeling

We utilized topic modeling, which involves the use of algorithms to uncover hidden thematic structures within collections of documents (Blei 2012). Specifically, we used Latent Dirichlet Allocation (LDA), where each document is represented as a probability distribution over a set of topics and each topic is represented as a probability distribution over a set of keywords. LDA has been used to answer scientific and sociological questions such as how scientific communities have changed over time (Padilla et al. 2014; Hall, Jurafsky, and Manning 2008) and trends in newspaper coverage of arts funding (DiMaggio, Nag, and Blei 2013). To our knowledge, neither LDA, nor topic modeling in general, have been applied in the study of broad trends in the

making phenomenon. Li et al. discuss the challenges of using topic modeling on documents that include both text and tags due the questions of how the tags should be weighted compared to the document text and to each other (Li, Li, and Pan 2013). We mitigate those concerns by focusing only on the tags and considering all tags for a given artifact to be of equal weight.

We used the Python package `gensim`’s implementation of LDA to determine what topics existed within our data set. After pre-processing our data set, the two key parameters we input to the model were a dictionary of terms and the corpus. The dictionary consists of a unique id mapped to each individual term. The corpus refers to a given word’s id and it’s respective frequency within the set of documents. Lastly, we input a number of topics that resulted in a distinct enough separation among keywords. The resulting output reflects a keyword’s importance within a given topic.

## Results

For our initial study, we conducted an analysis with 2 topics in order to get a baseline understanding of whether and how this kind of topic modeling can provide insight into maker communities.

The results of the topic model on the Hobby category with 2 topics are shown in Figure 1. The first topic appears to refer to electronics projects, with keywords such as `arduino`, `homebrew`, and `sensor`. The second topic can be summed up simply as drones. Notable keywords here include `quad`, `quadcopter`, and `fpv`.

Similarly, we performed a topic model with two topics for the Art category (Figure 2). The first topic exposed in this category is print artwork. Keywords of importance in are `logo`, `sign`, and `design`. The second topic is likely to be sculpture, with keywords such as `figure` and `statue`.

We also performed a topic model with two topics for the Household category (Figure 3). The first topic, which can be roughly described as utensils, includes keywords of importance such as `ikea`, `wall mount`, and `water pipe`. The next topic describes holiday decorations. Keywords highlighted in this topic are `christmas`, `decoration`, and `ornament`.

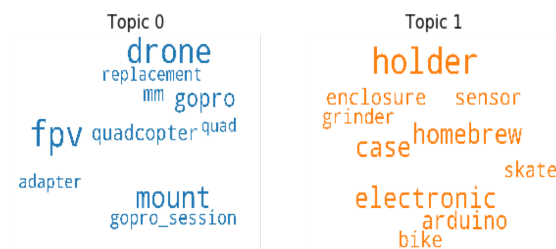


Figure 1: Word cloud visualization for Hobby category.

## Discussion and Future Work

The artifact analysis we conducted helped shed light on what types of making endeavors are happening in the making phenomenon. This knowledge can be used to expand on these

<sup>1</sup>[www.makerbot.com/thingiverse/](http://www.makerbot.com/thingiverse/)



Figure 2: Word cloud visualization for Art category.



Figure 3: Word cloud visualization for Household category.

activities and emphasize positive benefits to society as a whole. For example, the household category of useful objects resonates with the touted sustainable maker ideal of locally producing needed artifacts rather than purchasing them from a global mass production market that is harming the planet (Gershenfeld 2008).

We can also see what types of making endeavors seem to be more dominant, which could inform ways to target makers or advertise makerspaces. For example, the Hobby category shows that drones and electronics were the main categories. Makerspace leaders might tap into the popularity of such projects and use them to attract new makers to the space. However, we might also recognize that there are many other types of hobbies mentioned in makerspace literature, such as crafting or sewing (Buechley and Hill 2010; Fox, Ulgado, and Rosner 2015), that were not uncovered by our topic model. Perhaps generating a model with a larger number of topics would uncover such hobbies. Or perhaps our topic modeling analysis has shed light on how the Thingiverse website does not afford experiences for such hobbyists or does not provide ample invitation for them to participate.

The resulting topics for the Hobby category were not surprising in light of previous research that identified the Arduino micro-controller as popular for small electronics projects amongst hobbyist makers (Buechley and Hill 2010). However, some of the topics in the other categories were more novel. Both the Art and Household categories had one topic that centered around functional artifacts (print artwork, utensils) and one that centered around aesthetic or expressive artifacts (sculpture, holiday decorations). The fact that these distinctions between topics exist in multiple categories of artifacts suggests that perhaps maker endeavors

could be broadly grouped into functional and expressive projects, though more AI-driven and qualitative investigations are needed to explore this idea.

Our initial study shows promise for adopting topic models to help understand the making phenomenon. Going forward, our research will look at additional topic perspectives and topic model parameterizations across different contexts. One main challenge is the limited size of our data set compared to the total number of created artifacts. Other platforms, such as Fablabs.io, exist for makers to share artifacts, and future studies may also investigate artifacts shared in these communities. Additionally, future studies could compare different large scale communities, focus on the differences between these communities that the topic model analysis sheds light on, and use these results to inform methods that inspire co-operation between makers with the aim of contributing to a growing body of maker-created artifacts that has a positive impact on the world beyond.

However, none of the online platforms offers a complete representation of the incredibly vast and diverse making phenomenon and should always be considered in relation to insights from other types of investigations. Understanding the partial perspective each maker platform offers is still useful as it insights into how newcomers' perceptions of what the making phenomenon is for may be influenced. For example, many novice makers turn to Thingiverse for 3D printing inspiration. Knowing the landscape of projects newcomers are confronted with can help makerspace leaders know if they need to deliberately showcase other interesting but less viral applications of 3D printing.

Overall, these results show promise for artifact classification to lead to a deeper understanding of what is happening in the making phenomenon, especially when considered in relation to qualitative small-scale studies. Insights from this type of analysis help makerspace leaders discover types of making that are not well-supported on various platforms, discover ways to attract new makers, and broaden participation in their local making contexts.

## Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1723744 and Grant No. 1757844, as well as the UNCC CCI GAANN fellowship.

## References

- Bardzell, J.; Bardzell, S.; Lin, C.; Lindtner, S.; Toombs, A.; et al. 2017. HCI's making agendas. *Foundations and Trends® in Human-Computer Interaction* 11(3):126–200.
- Blei, D. M. 2012. Probabilistic topic models. *Communications of the ACM* 55(4):77–84.
- Blikstein, P. 2013. Digital fabrication and 'making' in education: The democratization of invention. *FabLabs: Of machines, makers and inventors* 4:1–21.
- Buechley, L., and Hill, B. M. 2010. Lilypad in the wild: How hardware's long tail is supporting new engineering and design communities. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, DIS '10, 199–207.

- New York, NY, USA: Association for Computing Machinery.
- Bull, G.; Maddox, C.; Marks, G.; McAnear, A.; Schmidt, D.; Schrum, L.; Smaldino, S.; Spector, M.; Sprague, D.; and Thompson, A. 2010. Educational implications of the digital fabrication revolution. *Journal of Research on Technology in Education* 42(4):331–338.
- Carucci, K., and Toyama, K. 2019. Making well-being: Exploring the role of makerspaces in long term care facilities. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, 1–12. New York, NY, USA: Association for Computing Machinery.
- DiMaggio, P.; Nag, M.; and Blei, D. 2013. Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of us government arts funding. *Poetics* 41(6):570–606.
- Fox, S.; Ulgado, R. R.; and Rosner, D. 2015. Hacking culture, not devices: Access and recognition in feminist hackerspaces. In *Proceedings of the 18th ACM conference on Computer supported cooperative work & social computing*, 56–68.
- Gershenfeld, N. 2008. *Fab: the coming revolution on your desktop—from personal computers to personal fabrication*. Basic Books.
- Hall, D.; Jurafsky, D.; and Manning, C. D. 2008. Studying the history of ideas using topic models. In *Proceedings of the 2008 conference on empirical methods in natural language processing*, 363–371.
- Hsu, Y.-C.; Baldwin, S.; and Ching, Y.-H. 2017. Learning through making and maker education. *TechTrends* 61(6):589–594.
- Hui, J. S., and Gerber, E. M. 2017. Developing makerspaces as sites of entrepreneurship. In *CSCW*, 2023–2038.
- Kuznetsov, S., and Paulos, E. 2010. Rise of the expert amateur: Diy projects, communities, and cultures. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, 295–304.
- Li, S.; Li, J.; and Pan, R. 2013. Tag-weighted topic model for mining semi-structured documents. In *Twenty-Third International Joint Conference on Artificial Intelligence*. Cite-seer.
- Martin, L. 2015. The promise of the maker movement for education. *Journal of Pre-College Engineering Education Research (J-PEER)* 5(1):4.
- Padilla, S.; Methven, T. S.; Corne, D. W.; and Chantler, M. J. 2014. Hot topics in chi: trend maps for visualising research. In *CHI'14 extended abstracts on human factors in computing systems*. 815–824.
- Roedl, D.; Bardzell, S.; and Bardzell, J. 2015. Sustainable making? balancing optimism and criticism in HCI discourse. *ACM Transactions on Computer-Human Interaction (TOCHI)* 22(3):15.
- Sheridan, K.; Halverson, E.; Litts, B.; Brahms, L.; Jacobs-Priebe, L.; and Owens, T. 2014. Learning in the making: A comparative case study of three makerspaces. *Harvard Educational Review* 84(4).
- Toombs, A. L.; Bardzell, S.; and Bardzell, J. 2015. The proper care and feeding of hackerspaces: Care ethics and cultures of making. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 629–638.