ARTS ENGINE
UNIVERSITY OF MICHIGAN

IMPACTS OF THE ARTS: TOPICS & TOOLS FOR ORGANIZATIONAL SENSEMAKING

OVERVIEW

We use the topics that emerge from our interviews as a way to build common understanding, explanations, and refinement of the values and impacts that shape policy and practice in higher education. The Arts Engagement project is a study of ~4000 undergraduate students at the University of Michigan that asked questions about the impacts, precursors, barriers, frequency, and perceptions of co-curricular arts engagement in college. Drawing from a subset of twelve open-ended questions and responses, we created topic models for each question and an interactive decision support tool and browser to facilitate team-based interpretation of the topics.

One example question asked, ***"What role did the arts play in your development as a person, friend, colleague, and student during college?"***

With estimates of prevalence for each of the topics for the question above, we found the arts played a strong social role in creating (10%) and strengthening (8%) social bonds, supporting friendships (15%), and in their personal identity formation (9%). The arts also fostered different perspectives that helped them gain skills (12%), gain a deeper appreciation (11%), and become more open-minded through cultural understanding (12%) and finding a balance in life (14%). Another 7% indicated the arts played other miscellaneous roles, and only 2% of the responses indicated that the arts did not play a significant role.

TOPIC ESTIMATION

Topic prevalence was estimated with the *stm* package in R (Roberts et al. 2014, 2017). Using the topic search technique from Lee and Mimno (2014), this technique consistently overestimated the number of topics by human interpretation of the most representative and comprehensible number of topics for the data. Held-out likelihood was near uniform for topic numbers above seven. The three most frequent number of topics estimated were used as a starting point for manual inspection of the response clustering and adjustment of the topic model to maximize the human interpreted representation, coherence, and exclusivity of underlying topics.

TEAM-BASED INTERPRETATION

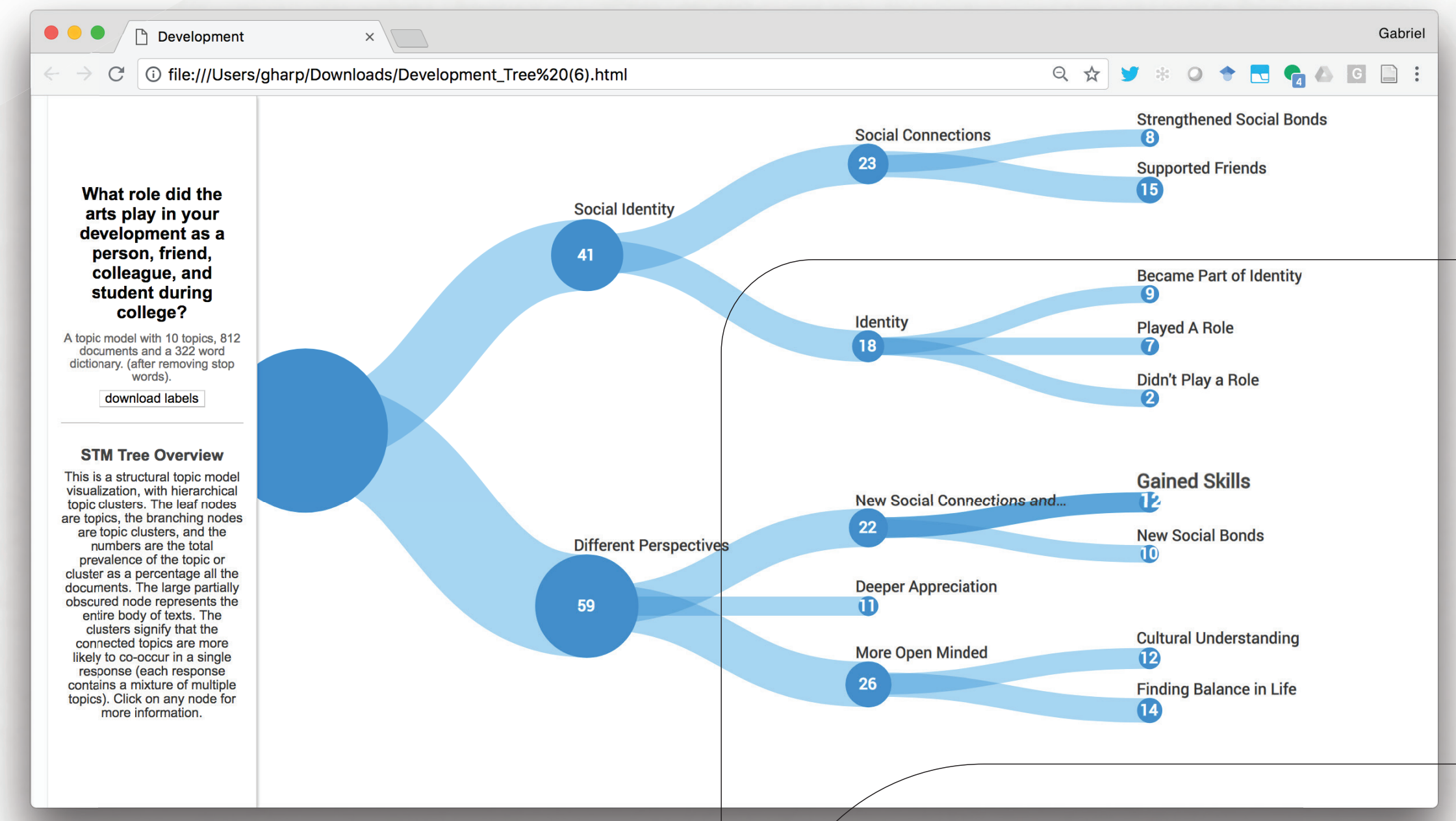
Manual inspection of the results, topic identification, and "tuning" was assisted with a custom made interactive topic browser which provided decision support for team-based interpretation of each child topic and parent node. Starting with the leaf nodes and working back towards the root of each topic tree, topic labeling and description of the model and clusters was completed by the research team which included domain and community experts for the population sampled and the topic areas under study. The group iteratively performed close reading of representative stems and representative responses (10-50 per topic), examined topic relationships and topic proportions, discussed, interpreted, and revised as needed the topic labels — leading to further development of descriptions, contextualization, and policy implications for key audiences.

Research Team:
Gabriel Harp
Deb Mexicotte
Jack Bowman
Mengdan Yuan

artsengine.umich.edu
a2ru.org/insights

a2ru ALLIANCE FOR THE ARTS IN RESEARCH UNIVERSITIES
AWMF
ART WORKS. | National Endowment for the Arts
arts.gov

ArtsEngine and the Alliance for the Arts in Research Universities (A2RU) highlight arts and design-driven teaching and research in order to show the impact on creative practice, knowledge production, and student learning experiences. Our work is supported by an Alliance of over forty research universities, the U-M Office of the Provost, the Andrew W. Mellon Foundation, the National Endowment for the Arts, the U-M College of Engineering, the A. Alfred Taubman College of Architecture and Urban Planning, the Penny W. Stamps School of Art & Design, the U-M School of Music, Theatre & Dance, and the U-M Office of Research.

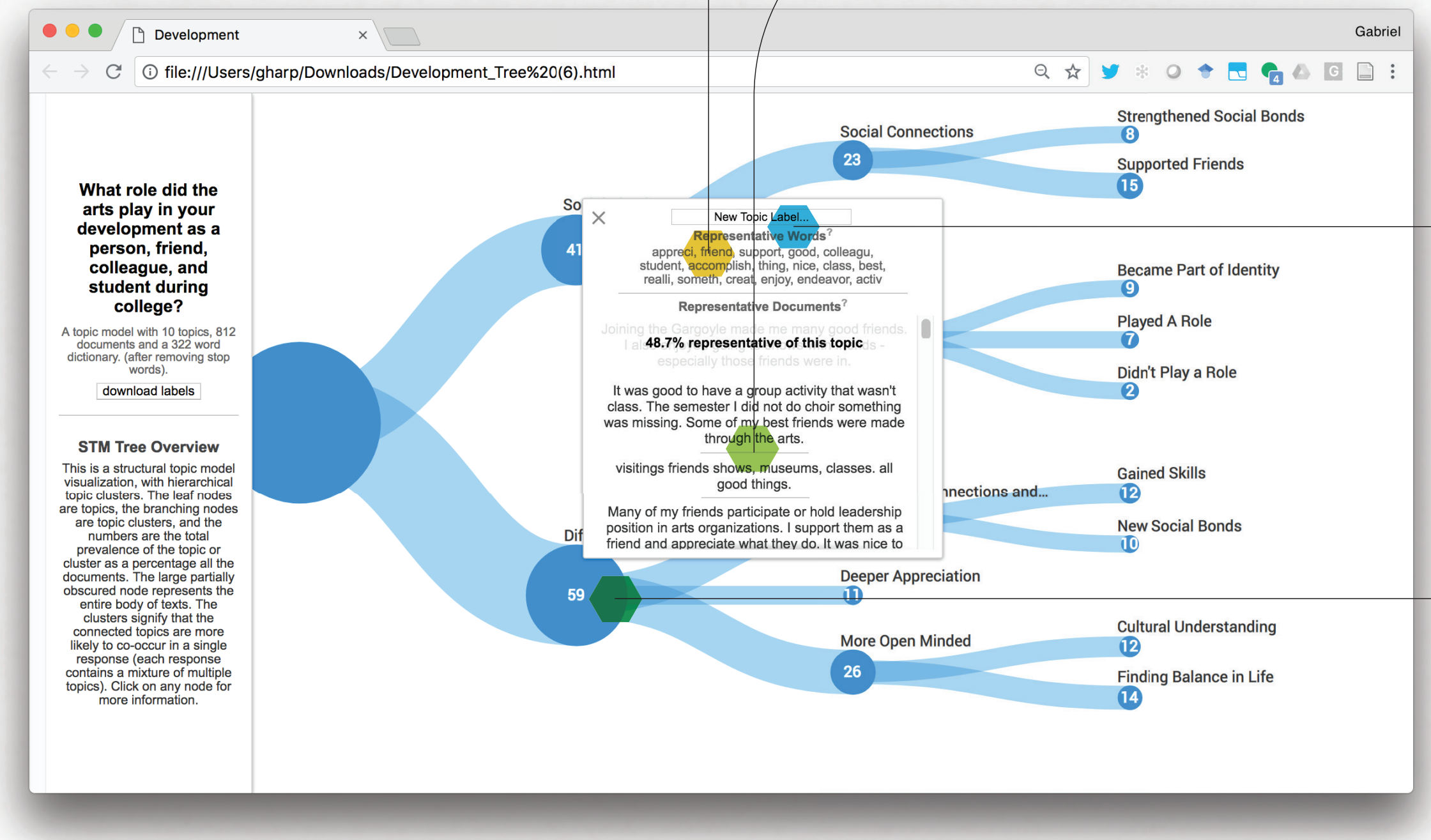


What role did the arts play in your development as a person, friend, colleague, and student during college?

A topic model with 10 topics, 812 documents and a 322 word dictionary. (after removing stop words).

STM Tree Overview

This is a structural topic model visualization, with hierarchical topic clusters. The leaf nodes are topics, the branching nodes are topic clusters, and the numbers are the total prevalence of the topic or cluster as a percentage all the documents. The large partially obscured node represents the entire body of texts. The clusters signify that the connected topics are more likely to co-occur in a single response (each response contains a mixture of multiple topics). Click on any node for more information.



What role did the arts play in your development as a person, friend, colleague, and student during college?

A topic model with 10 topics, 812 documents and a 322 word dictionary. (after removing stop words).

STM Tree Overview

This is a structural topic model visualization, with hierarchical topic clusters. The leaf nodes are topics, the branching nodes are topic clusters, and the numbers are the total prevalence of the topic or cluster as a percentage all the documents. The large partially obscured node represents the entire body of texts. The clusters signify that the connected topics are more likely to co-occur in a single response (each response contains a mixture of multiple topics). Click on any node for more information.

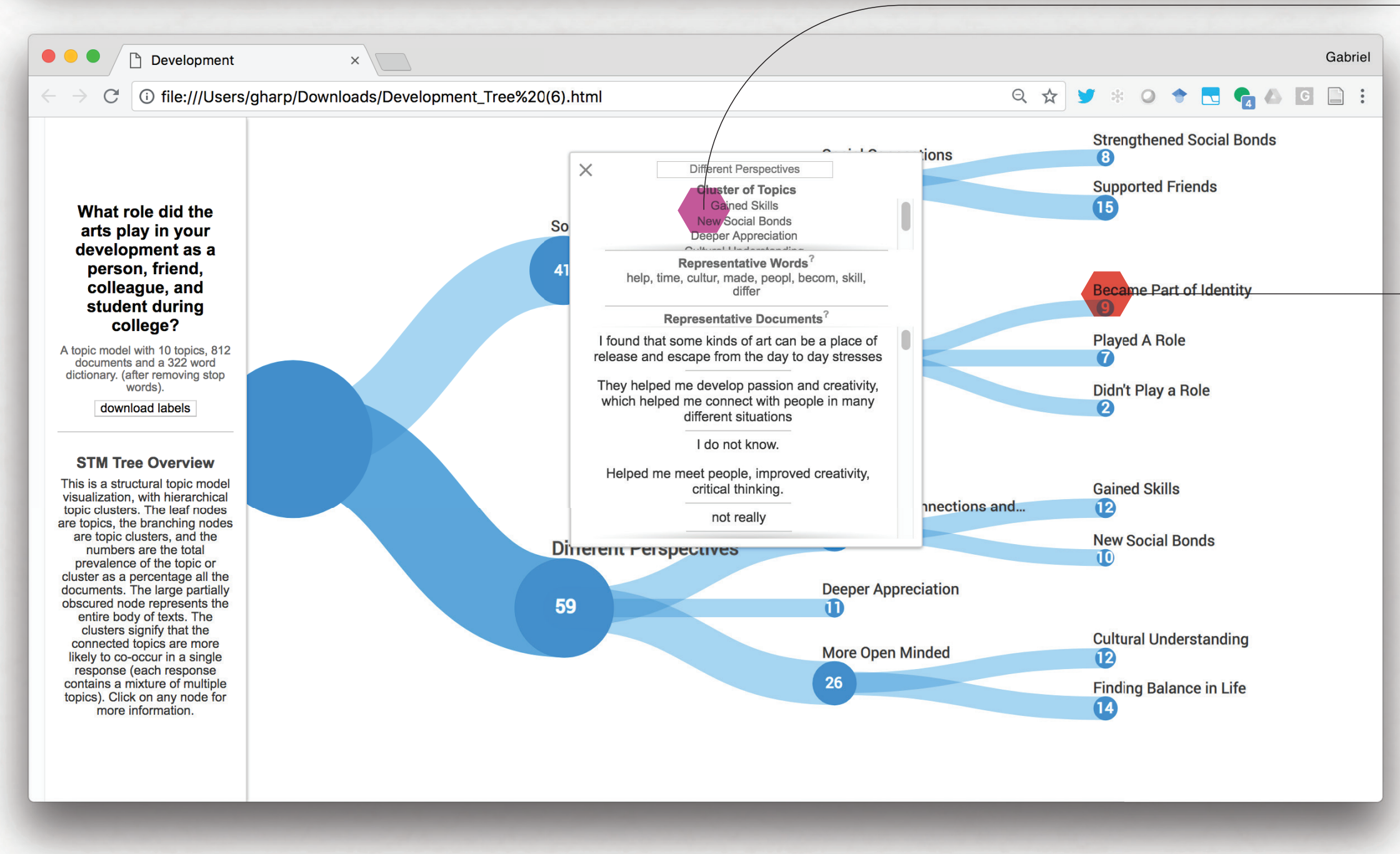
Representative Words
appreci, friend, support, good, colleagu, student, accomplish, thing, nice, class, best, realli, someth, creat, enjoy, endeavor, activ

Representative Documents
Joining the Gargoyle made me many good friends. I al... ds - especially those friends were in.
48.7% representative of this topic

It was good to have a group activity that wasn't class. The semester I did not do choir something was missing. Some of my best friends were made through the arts.

visitings friends shows, museums, classes. all good things.

Many of my friends participate or hold leadership position in arts organizations. I support them as a friend and appreciate what they do. It was nice to



What role did the arts play in your development as a person, friend, colleague, and student during college?

A topic model with 10 topics, 812 documents and a 322 word dictionary. (after removing stop words).

STM Tree Overview

This is a structural topic model visualization, with hierarchical topic clusters. The leaf nodes are topics, the branching nodes are topic clusters, and the numbers are the total prevalence of the topic or cluster as a percentage all the documents. The large partially obscured node represents the entire body of texts. The clusters signify that the connected topics are more likely to co-occur in a single response (each response contains a mixture of multiple topics). Click on any node for more information.

Cluster of Topics
Gained Skills
New Social Bonds
Deeper Appreciation

Representative Words
help, time, cultur, made, peopl, becom, skill, differ

Representative Documents
I found that some kinds of art can be a place of release and escape from the day to day stresses

They helped me develop passion and creativity, which helped me connect with people in many different situations

I do not know.

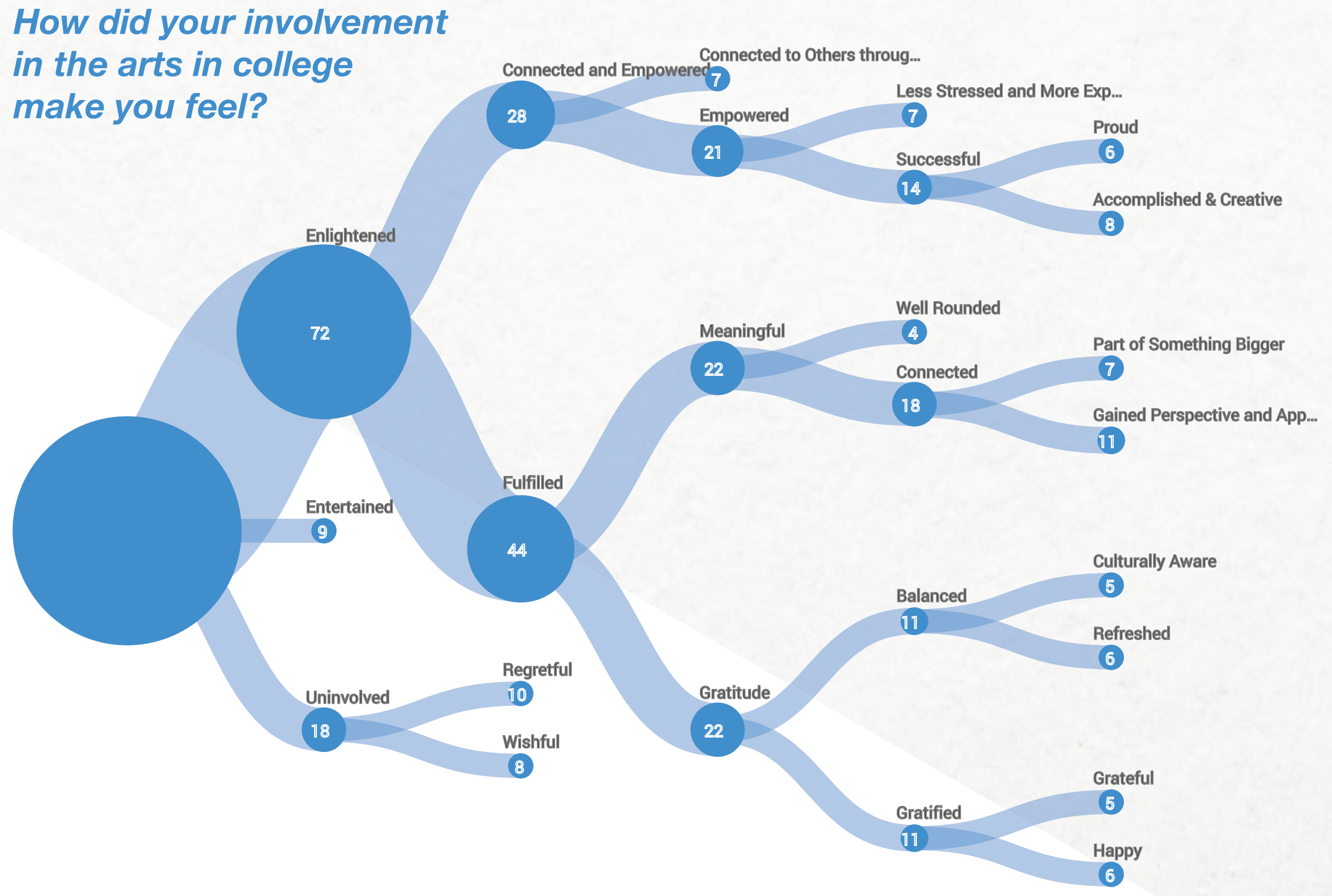
Helped me meet people, improved creativity, critical thinking.

not really

EXAMPLES OF OTHER QUESTIONS ANALYZED

How did your behavior or thinking change (as a result of engaging in the arts)?
In what ways do you think you can grow (through arts engagement)?
Please describe any transformative or meaningful arts experience you had during college.
What do you see as the barriers preventing you from being involved in the arts?
What role did the arts play in your college experience, both positive and negative?
How do you define "the arts"?
What role do you see the arts playing in society?
If the arts had any impact on your career choice, please describe.
In what other ways do you see yourself being involved in the arts after college?

How did your involvement in the arts in college make you feel?



VISUALIZATION FEATURES

Representative Stems. We use these stems to generally identify the topic. The stems are generated from various metrics based on frequency, exclusivity, and log probabilities which are then combined through weighted voting. Stems are used instead of words in order to merge nouns, adjectives, and verb tenses; e.g. appreciation, appreciates, appreciating, etc all become appreci. Certain words such as art, arts (i.e. ones specific to the question phrasing) are removed ahead of time in order to produce more meaningful results.

Representative Documents. Documents contain multiple topics as percentages of the whole response; e.g. the response *"It's too expensive and competitive"* might be considered to 49% relate to the expensive topic, 49% to the competitive topic, and 2% to the other topics in the model. 10 to 50 documents are displayed in order of document topic proportion, with a minimum of 15% representative for the given topic.

Labeling. Topics and clusters can be labeled, saved to the visualization, and downloaded to reintegrate with the source datasheet for subsequent analysis or use elsewhere.

Topic Clusters. The tree visualization intuitively show inter-topic relationships and provides the ability to infer higher level parent topics based on the identity of child nodes. The D3-based visualization is created from hierarchical clustering on the topic proportion dissimilarity matrix using the stm R package, and then further customized for our use-case of team-based interpretation.

Representative Cluster Responses. These responses assist with inference of parent node topics, and are generated from cluster constituent topic responses with a minimum of 10% representation of each topic, and 10% variance between topics.

Node Proportions. Each topic contains a number representing the sum of document topic proportions that correspond to that topic. Topic clusters contain the sum of their constituent topics; e.g. if 10 of 100 total responses were 50% topic A, then topic A would be 5% of all document proportions.

Packaged as an HTML file. One of the nice things about packaging the topic model, labels, tree, representative documents, and topic descriptions together in an HTML file is that it can be easily shared as an all-in-one package for communicating about the subject. This makes for simpler dissemination to generate enthusiasm among different members of the institution — all of which have different levels of familiarity with the topics and techniques used to uncover them.

ANALYTIC RESOURCES AND TECHNIQUES

TOPIC MODELS (stm R package; Roberts et al., 2014)
text-mining for discovery and estimation of hidden semantic structures (topics) in texts

CLOSE READING AND INTERPRETATION (team-based)
visualization-supported reading of individual responses in order to interpret the results and overall relationships — individually and through group discussion and collaboration

DICTIONARIES
organized collections of terms and phrases used to help classify, categorize, and calculate the frequency of words in text

LINGUISTIC INQUIRY AND WORD COUNT (LINGUISTIC AND PSYCHOLOGICAL)
calculates the degree to which linguistic, cognitive, and affective categories of words are used in a text (Tausczik and Pennebaker, 2010; Pennebaker et al., 2015)

DOCUSCOPE LANGUAGE ACTION TYPES (RHETORICAL PATTERNS)
analyzes rhetorical features such as persuasiveness or first-person reporting; covers 40 million linguistic patterns of English classified into over 100 categories of rhetorical effects (Kaufer et al., 2004; Ishizaki and Kaufer, 2012)

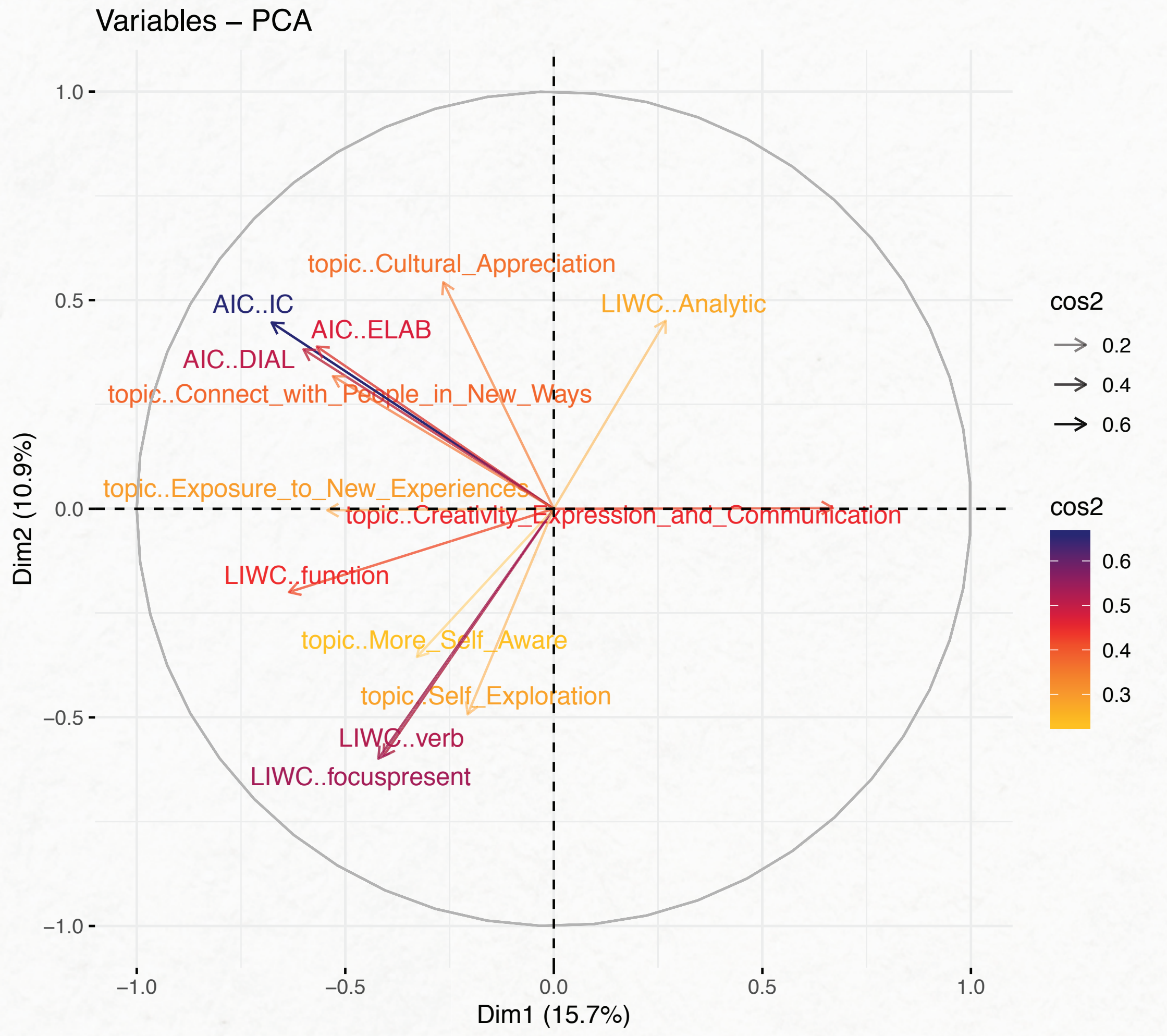
AUTOMATED INTEGRATIVE COMPLEXITY (COGNITIVE STRUCTURE)
measures the cognitive structure implied in a speaker's verbal content including how well they identify different dimensions of an issue and integrate different ideas (Conway et al., 2014; Houck et al., 2014)

DEMOGRAPHICS (survey collected)
characteristics of a population based on factors such as sex, grade point average, high school locale, ethnic group, income, etc.

NEXT STEPS

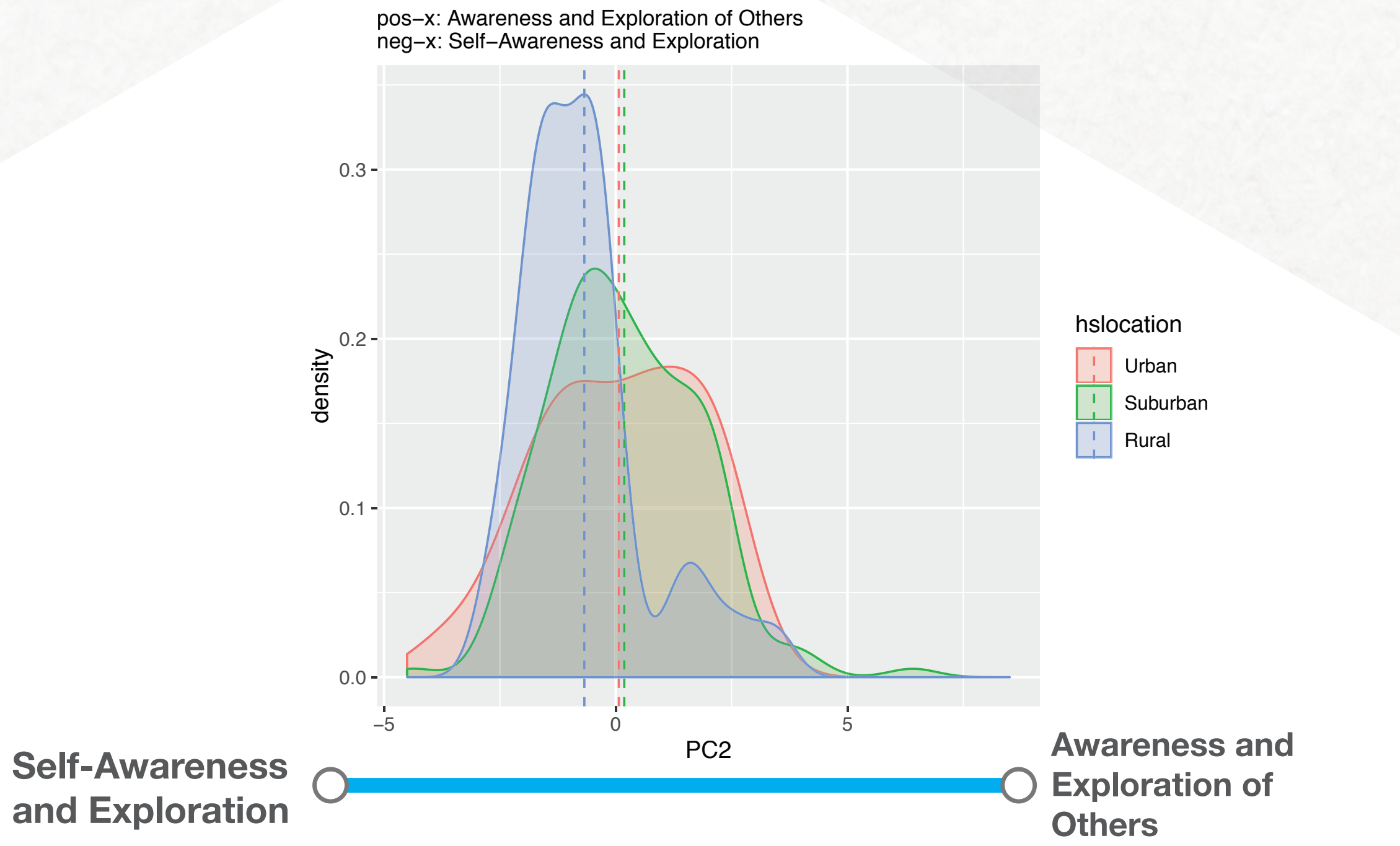
The results of this work are being used to build better models and frameworks for how the arts, design, and interdisciplinary practices create impacts in higher education; to stimulate wider national conversations around the roles and impacts of the arts in student learning and success; and they are expected to support sensemaking around the undergraduate student experience as part of U-M's Presidential Arts Initiative launched in Fall 2019.

Principal Components Analysis (PCA) that includes dictionary-based cognitive, psychographic, and rhetorical measures of the responses as additional variables—along with topic prevalences—has suggested integrative dimensions in the data. Using the component scores from the PCA, we have further analyzed the variation in those scores and their relationship with demographic factors provided by the survey. This work is leading to more refined research questions, improved design of empirical studies, and better insights for policy ad practice. We hope to extend these approaches to help other universities and organizations easily ask and interpret their own students' and stakeholders' experiences.



ABOVE: This example PCA biplot shows correlations between topic prevalence, LIWC, DocuScope LAT, and AIC variables for the question (n responses=971), *"In what ways do you think you can grow (through arts engagement)?"* Four components were extracted that explained 44% of the variation among responses.

BELOW: For the *"ways you can grow"* question above, Principal Component 2 (~11% variation explained) is a dimension that contrasts *self vs. others*. In the example below, An ANOVA with Tukey post-hoc contrast between principal component 2 and high school location indicates a contrast between Rural and Suburban locations (estimate: -0.3874909; tukey adj.p.value: 0.03574811).


pos−x: Awareness and Exploration of Others
neg−x: Self−Awareness and Exploration
density
PC2
hslocation
Urban
Suburban
Rural
Self-Awareness and Exploration
Awareness and Exploration of Others