Visualizing Knowledge Networks in Online Courses
Marni Baker-Stein, Sean York & Brian Dashew

As networking platforms have become more ubiquitous in the personal consumer space, data derived from social interaction is increasingly being used in the commercial space to analyze markets, make decisions, and develop new, personalized tools. However, even as social tools and design develop a presence in the learning space, research using social data to develop new understandings about knowledge production, teaching, and learning in online social learning spaces is fairly limited. This article is a practitioners’ progress report on a research collaboration between Columbia University School of Continuing Education and Pearson Higher Education Technology, established with the goal of developing a framework and methodology for studying how social interactions and knowledge construction unfold in online courses that employ both formal and informal social learning activities. The work describes an emergent methodology for analyzing data produced by social and conversational interactions in online learning environments, using threaded discussion data from a group of students and faculty at Columbia University School of Continuing Education. It overviews the graph database schema and technologies employed, and describes examples of how the data is used to describe, differentiate among, and visualize individuals, conversations, and patterns of concept connectedness. Finally, it discusses relative strengths and weaknesses of the approach, suggesting ways it might evolve to improve our understanding of social networking and engagement in online learning environments, and how it can optimally impact student learning.

Keywords: Social, analytics, knowledge, networks, visualization

Note: The figures labeled as Interactive may be viewed by downloading the Internet Learning Journal app from the iOS App Store.

I - Introduction

With the rise of consumer-facing networking platforms like Facebook, Twitter, LinkedIn and Instagram, “social” has become a dynamic engine of commercial enterprise powered by huge amounts of data. This data is structured and presented in ways that drive the continuous development of real time, highly personalized tools for social and professional networking. And, perhaps even more critically, it has the potential to give us unprecedented insights into the social mechanisms that underpin cultural practices of learning and knowledge production.

However in research efforts targeted at understanding student success and learning in higher education and specifically in online courses and programs, we have only recently begun to explore the potential uses and impacts of “social”. Learning management systems that support online instruction increasingly provide (or integrate
with) social networking tools to facilitate community building and social knowledge networking. Yet related research efforts that seek to understand student behavior in online courses have focused primarily on attendance patterns and wayfinding behaviors, content engagement and assessment outcomes, leaving the social dimensions of these environments relatively unexplored.

It is often said that we value what we can measure, and we measure what we value. A review of the technology impacting the state of higher education instruction and research indicates both value and measurability may be shifting towards an examination of the social space as a powerful means of surfacing knowledge construction activity. The 2014 Horizons Report (New Media Consortium, 2014) lists the growing ubiquity of social media as among the drivers of change likely to impact education within the next two years. The report also lists two trends as three to five years away from having a significant impact on the state of higher education: the rise of data-driven learning and assessment, and a shift towards viewing students as creators of content. We believe these and other trends listed in the Horizons Report indicate the time is now to gain insights into the conditions that promote social knowledge networking in online courses and to identify practical methods to measure its impacts.

With these goals in mind in 2011 the researchers launched a collaborative research effort between the Columbia University School of Continuing Education program development and instructional design team, and Pearson Higher Education Technology. Together, we defined an exploratory methodology and an initial set of logical questions to guide research-engaging data produced from the social networking environment of an online master’s degree program offered at Columbia University. Our goal was to develop a framework and methodology aimed broadly at allowing us to better understand social interactions and knowledge construction in online courses that employ both formal and informal social and cooperative learning activities.

We will first elaborate our definition of Social Knowledge Networking (SKN) and the logic we applied in structuring our data and identifying the initial questions that grounded our research. Next, we provide a generic description of our emergent methodology for analyzing the data produced by social and conversational interactions in online learning environments. Then we present an overview of the graph schema and technologies we used, followed by results for each of our three research questions. Finally, we discuss relative strengths and weaknesses of the method, suggesting ways it might evolve to improve our understanding of how social networking and engagement work in online learning environments and how it can optimally impact student learning.

II - Analytical Framework

Our initial analytical framework incorporated relevant concepts from content analysis, knowledge network analysis, and conversational analysis into a custom model, represented in Figure 1.

A. The Knowledge Map

Foundational to this framework is the recognition that each course contains an underlying knowledge map. The map represents the conceptual skeleton of the course, including those concepts provided by the instructor via course resources, lectures, or activity prompts, and those introduced via discussion in the course. Part
of our aim is to be able to understand and visualize the topic spread of student and instructor-generated content from the course discussions across (and beyond) this map. In this paper we deal mostly with conversational concepts. Work on course concept structure and use of ontologies is ongoing.

Knowledge Activity

To characterize the focus of student and instructor-generated content in the context of each course discussion, we further wanted to be able to identify the level of knowledge activity that resulted from participant engagement. To this end we developed a custom rubric to align with the types of knowledge activities prompted by collaborative and discussion assignments included in the program under study.

Levels of Knowledge Activity:
1. Absorb: Determining the meaning of instructional messages and course concepts.
2. Transfer: Transferring an understanding between contexts or disciplinary environments.
3. Apply: Carrying out or using a theory, concept or procedure in a given situation.
4. Innovate: Putting elements together to form a novel or coherent whole or to identify an original product or solution.

Not surprisingly, knowledge activity is one of the most difficult attributes to code consistently and our understanding continues to develop as we dive deeper into our data and can see more clearly the types and nuances of knowledge activity occurring in our context.
C. Conversational Influences

To understand conversational influences on the topicSpread of course conversations, we implemented an approach suggested by a prior research partnership between Pearson Learning Solutions and Texas Christian University (Zelenka, 2012). This approach measures the conversational force of individual contributions to course discussions by extracting topics that appear in each response, and then considers the relationship of these contributions to topics already introduced in the conversation. The TCU/Pearson research team proposed that there are four levels of discussion thread contribution that impact topicSpread.

Levels of Topic Spread:
1. Participation: A student or instructor response does not cover topics relevant to the discussion but merely states agreement or disagreement or offers social conversation.
2. Explanation: A response covers topics that have already been introduced in a thread.
3. Elaboration: A response provides additional closely-related topics about topics already introduced in a particular top-level threaded response.
4. Expansion: A response connects topics already introduced in the discussion to more distantly-related topics.

These codes are assumed to form an ordered hierarchy, with expansion representing the highest level of topicSpread.

As we read the data more deeply, we noted a number of common speech acts that seemed to be impacting the levels of knowledgeActivity and topicSpread across conversations in the learning community.

One of these was a Topic Spread Request, in which a discussant would ask another person to Explain, Elaborate, or Expand upon some topic. If the response to such a spread-Request was coded for topicSpread at the same level, we would consider the request to have been met.

Levels of Topic Spread Request:
1. Explanation: A discussant requests an explanation of topics that have already been introduced in a thread.
2. Elaboration: A discussant requests a response containing additional closely-related topics about topics already introduced in a particular top-level threaded response.
3. Expansion: A discussant requests a response that will connect topics already introduced in the discussion to more distantly-related topics.

The Columbia team had designed and written the courses and read the assigned readings. They therefore acted as our content experts when it came to applying topicSpread scores for our entire response set.

It should be noted that topicSpread is not intended as a way of valuing contributions, beyond the observation of whether new concepts are introduced, and how closely or distantly related they are to assigned content resources and prior discussion. For example, a topicSpread of Level 4/Expand might correspond with the use of an analogy that clarifies a course concept, or it could signal a distracting departure from relevant topics. Determining concept relevance is a significant area for ongoing research.

We added the following conversational moves in addition to topicSpread and spreadRequest:
• Question: Does a discussant ask a question?
• Personal Story: Does a discussant tell a story from personal experience?
• Citation: Does a discussant make reference to a book, article, or other work (citation)?
• Challenge: Does a discussant challenge another discussant?

D. Task Target and On-Targetness

To understand conversations in our formal learning environment, we also felt it important to consider the targeted behavior of the collaborative activity or discussion prompt. Activities (discussion prompts) were coded using the knowledge-Activity and topicSpread categories. For example, tasks might ask students to Transfer and Elaborate (knowledgeActivity=2/Transfer, topicSpread=3/Elaborate). General topical alignment was also considered.

Each discussant’s comment, as well as the entire thread, was coded for whether or not it was on target in relation to the original task prompt. These binary attributes are called onTargetPost, and onTargetThread.

E. Metadata Attributes

Finally, we identified a set of quantitative attributes that provide more information about individual participants as well as the shape and structure of conversations themselves. These included:

• word count (of participants, conversations, and individual responses)
• number of posts (for each participant and conversation)
• number of unique participants (in each conversation)
• time stamp (of each participant’s posts and the conversation as a whole)
• proximity of posts in time (of each participant’s posts and for the conversation as a whole)
• level of the response tree at which a response is posted (responseLevel)

F. Intersectionality

We believed that our richest insights from this type of exploratory study would spring from our ability to identify and visualize the intersection of individual, conversational and content characteristics. For example, do certain combinations of individual students generate more ‘productive’ or ‘successful’ conversations? Are student and instructor questions treated differently? What kinds of instructor strategies might be effective in various kinds of conversations? How does the introduction of certain concepts or resources impact the depth or number of participants in a conversation? See Figure 2 for some examples of these intersectionalities.

With this emergent framework as our guide, we manually coded a data set of 948 threaded discussion posts for targeted attributes; designed a graph schema and graph database to aid in describing and analyzing the problem space; and began the project of designing queries and visualizations to facilitate analysis of the threaded discussion data from graph computing and Natural-Language Processing (NLP) perspectives.

G. Tools Development and Scalability

We decided to employ or build technology solutions where feasible, but to not limit our questions to what was possible with current technologies. We favored a data design that would speak well to our questions, even if at first it would require significant labor to op-
In order to conduct sophisticated analyses of social interaction in online learning, we determined that we must first be able to identify, count, qualify and visualize individual behaviors and interactions among the network of participating faculty and students. We also wanted to visualize the traverse of anonymized faculty and student conversations across the content map of the course and program.

To this end, we formulated the following high-level research questions:

- **RQ1:** Can we identify, differentiate, and visualize individual characteristics and behaviors in an online discussion or course?
- **RQ2:** Can we identify, differentiate, and visualize conversation characteristics and behaviors in an online discussion or course?
- **RQ3:** Can we identify and visualize content focus over time in an online discussion or course?
Our research questions address fundamental challenges of doing sophisticated analyses of online discussions. Conversations have structural and other non-content attributes, but are also contexts where unique individuals come together and co-create a body of content. The problem of identifying and quantifying individual influence on conversational content and structure is a complex one, as is the problem of identifying how conversational structure and content might arise as a combined expression of the attributes and behaviors of multiple individuals. In the following sections we will describe our approach to each question, and discuss our findings.

IV - Methodology

In this paper, we present the current state of the qualitative, quantitative, and visual research methodology that has emerged over the past three years of collaborative work. The Columbia and Pearson teams adopted an iterative, grounded approach to data gathering and analysis, beginning with a thick, digital ethnography of discussions in several online courses. Methods included close readings of discussion texts, analysis of conversational moves and strategies, and detailed analysis of engagement with assigned and unassigned resources. We identified quantitative and qualitative attributes as described above in the Analytical Framework, which we then applied to the data on successive passes over a period of several months. The result was a set of rich, augmented discussion data containing both automated and hand-coded attributes for each discussion response, along with detailed digital-ethnographic field notes.

Then, in order to analyze the data from network and visualization perspectives, we employed a variety of software tools and techniques. These approaches included creation of a graph database with a custom schema designed to model threaded discussion data, a domain specific language (DSL) for exploring that data, and use of Natural-Language Processing (NLP) tools, network visualization tools (such as Gephi), graphic design software (such as Adobe Illustrator), and spreadsheets. We relied heavily on open source software, and wrote our own code as well. We were able to automate some tasks with custom scripts and parsers, while others required hours of painstaking, repetitive work. Thus, the present work is presented as a practitioners’ progress report on the project of defining a set of Social Knowledge Networking attributes relevant to emergent digital pedagogies, and of devising ways to measure and reason about them. Our examples are intended to be illustrative rather than definitive. Our methodology is presented as one of exploratory inquiry, rather than as a proven, streamlined approach to answering the kinds of questions we engage here.

We draw our data examples from a single week of anonymized, small-group, threaded discussion data, consisting of one instructor prompt, seven individual thread response trees, and a total of 64 comments, over a period of four days. All names are invented code names, applied without regard to gender or course role. The seven students are Alakel, Danen, Fesler, Loret, Viska, Renlit, and Kerrad. Naya is the lead instructor, and Jakata is a TA. Radsel, a participant from another group, cross posts one comment in Fesler’s thread.

For each research question, we provide a brief conceptual overview of our approach; a technical summary describing the processes and technologies involved; a situated example to illustrate an application of the model to real data; and a discussion where we explore Instructional Design insights and implications for future work.
Because network thinking is fundamental to our approach, we will preface our data analysis with a conceptual overview of our graph database schema, and a technical summary of the graph technologies we used. We will reference this schema in our discussion of each research question.

A. Conceptual Overview: Graph Database Schema

We engaged with the applied graph science experts at the Aurelius consulting group, creators of the open-source TinkerPop graph computing stack, to model the conversational data as a network schema (a ‘directed property graph’), build a graph database against that schema, and design a domain specific language (DSL) for traversing and interrogating the threaded discussion graph. We found several benefits to modeling the data as a graph, as shown in Figure 3.

First, as a data structure, the graph allows us to pose many questions in an exploratory and intuitive manner. Second, the familiar concept map construct eased discussion and reasoning about the data among more- and less-technical researchers. This was particularly important given that we expected to discover new and important questions over the course of the study. Finally, the graph-structured data is easily exported in forms that can be used with existing network visualization tools. This allowed us to use visualization as a first-class investigative tool over the course of the study, as well as a post-hoc story-telling tool.

Figure 3. A Graph Database Schema for Threaded Discussion Data.
B. Technical Summary: Graph Database Technology

For the analysis presented in this article, we populated a Neo4J graph database with our research data according to the schema described above. This research graph comprised roughly 46,000 vertices and 144,000 edges. We currently use the distributed graph database Titan to maintain our production dataset, consisting of approximately 400 million vertices and 1.2 billion edges. Because TinkerPop is graph vendor agnostic, we are able to use the same tools to manipulate both our Titan production graph and our Neo4J research graph. We built a custom DSL using Gremlin, the graph traversal language built into TinkerPop. The DSL composes custom graph traversals, queries, and calculations that can be executed in various contexts in the graph, such as for a whole course, a whole discussion, a single thread, or a group or individual over time. Queries can generate sub-graphs that can be used for visualization, or to test traversals, statistical methods, machine learning techniques, or other approaches. While we will provide limited examples to illustrate our approach, an in-depth discussion of these technologies is beyond the scope of this paper. You can learn more about them at http://tinkerpop.com, and https://github.com/tinkerpop.

Gremlin enables the flexible construction of traversals for exploratory data analysis in the graph. For example, where ‘g’ is the graph, and ‘V’ is the set of all vertices in the graph, the following Gremlin query would generate a list of all concepts mentioned by a person named Renlit over the history of all of Renlit’s responses:

\[
g.V().has('personName', 'Renlit').out('wrote').out('mentions')\]

In this manner, we can construct complex and unanticipated queries to explore and interrogate the data, and evolve new queries based on emergent understandings of the data. Query results are themselves graphs, which can be used for visualization and other analytical work. If an interesting metric is discovered, it can be codified as an algorithm, expressed as a ‘step’ and used inline with other Gremlin commands. For example, imagine we have created a method for determining whether or not a person is a ‘Thought Leader’ in a course, based on some graph traversal. We could express that algorithm in a Gremlin step called isThoughtLeader, and use that step to discover all concepts discussed by thought leaders as follows:

\[
g.V().has('type', 'person').isThoughtLeader().out('wrote').out('mentions')\]

The output of such algorithms can be tested and used to inform learning environment design, or studied in conjunction with other factors in the course of ongoing research.

VI - RQ1 Findings: Can we identify, differentiate and visualize individual attributes and behaviors in an online discussion or course?

A. RQ1 Conceptual Overview

There are many kinds of learner data available, depending on the environment, activity, platform, or product under study. In a general conversational context, much of what we can know about a person is derived from:

- What they contribute: Number, size, content, and attributes of individual comments
• How they interact: Timing and response tree depth of contributions, behavioral patterns, conversational moves and strategies, and individual influence
• With whom they interact: Which threads they contribute to, to whom they respond, who responds to them, and the identity, number, and variety of their co-discussants

Not surprisingly, multiple passes through the data revealed many insights and avenues for exploration that were not apparent during earlier readings. While it may seem straightforward to see how conversational and participatory elements manifest at an individual comment or thread level, it is much more difficult to understand the historical context of a contribution, or to consistently apply a discussion rubric over a large amount of conversational data.

We approached the problem of modeling and differentiating individuals using a construct we term a ‘comparative corpus diagram,’ an example of which is shown in Figure 4.

An individual’s corpus is a collection of all responses they have authored in some context or time period. A comparative corpus diagram is a graphical and statistical representation of multiple individual response corpora, with responses sized and colored for various attributes and arranged for easy comparison among individuals. When we analyzed corpora coded for attributes from our SKN model, we found them to be a compelling supplement to the digital-ethnographic narratives of individuals and conversations in our data set.
B. RQ1 Technical Summary

A corpus diagram requires a graph containing a person, and all associated responses. We collected those responses by following all outgoing ‘wrote’ edges from a given person, as follows:

g.V.has('personName','Renlit').out('wrote')

We wrote the results to an in-memory Tinkergraph, exported the data as GraphML, and imported to Gephi for further modeling. We applied a consistent set of visualization rules, such as node sizing based on wordCount, and color mappings for the values of various attributes. Finally, we applied a force-directed graph layout algorithm to the model to obtain a readable presentation. Based on that model, we used Gephi to export a separate SVG vector graphics file for each attribute's color scheme, and overlaid them using Adobe Illustrator. As a final step, we exported to PDF format while preserving top-level Illustrator layers, resulting in a layered PDF. We used these PDFs as data analysis tools, and to generate the comparative corpus diagrams presented in this paper.

C. RQ1 Example

The following figures use comparative corpus diagrams, coded for a handful of attributes, to illustrate a few similarities and distinctions among three participants: Renlit and Loret, who are students, and Naya, who is a lead course instructor. Each corpus diagram represents the entire history of each discussant's contributions over multiple weeks and courses, and is accompanied by a brief description of the participant based on our digital-ethnographic observation data. A brief comparison will illustrate how elements of these participants' digital-ethnographic descriptions can be detected using comparative corpus diagrams, and the potential of the approach to support identification and differentiation of individuals based on their patterns of discussion participation.

Figure 5 compares corpora for Renlit, Loret, and Naya, coded for usage of personalStories. Renlit’s diagram shows the highest level of story usage across the entire data set, and reflects the digital-ethnographic description of Renlit’s tendency to answer questions using personalStories rooted in a professional context. Loret shows story usage at a significantly lower level than Renlit, but more in line with typical student numbers. Naya, on the other hand, uses only one personalStory in a corpus of 91 responses, the largest corpus in the data set. Naya’s responses are significantly shorter than most student responses, with an average wordCount of 61. We can’t infer that all instructors in all situations will show such a marked difference from students in this regard, but in combination with other data points, these provide a promising starting point for differentiating participants.

Figure 6, coded for questions, reveals a striking correlation between Naya’s corpus diagram, and the digital-ethnographic description of Naya as favoring short, probing questions as a participation strategy. A comparison of Naya with Loret and Renlit is also revealing. For stories, Renlit was prolific and Naya barely registered, with a gap of about 70%. For questions, the situation is flipped, with Naya asking many questions and Renlit asking relatively few, with a gap of approximately 50%. And in both cases, Loret is in between, in some cases appearing more like the other student, and in some appearing more instructor-like, as reflected in the digital-ethnographic description.
Figures 5, 6 and 7 are presented sequentially for comparative purposes.

**Figure 5.** Corpus Comparison – story.

Renlit tends to answer questions using personal stories rooted in a professional context (wine distribution). This obtains in Renlit’s own lead posts, and in responses to others. Given the focused nature of these stories, Renlit’s posts generally do not include questions for others.

Loret’s posts often begin with praise for the prior author, then use questions to encourage expansion on the original post. In some ways, Loret acts the part of teacher or facilitator, but also occasionally probes beyond the scope of the conversation.

Naya is the lead instructor. Naya’s posts are generally short, posing probing questions prompting students to go deeper. When a student post includes multiple parts (or responses to multiple questions), Naya selects one part and encourages development in that area.

**Figure 6.** Corpus Comparison – question.

Renlit tends to answer questions using personal stories rooted in a professional context (wine distribution). This obtains in Renlit’s own lead posts, and in responses to others. Given the focused nature of these stories, Renlit’s posts generally do not include questions for others.

Loret’s posts often begin with praise for the prior author, then use questions to encourage expansion on the original post. In some ways, Loret acts the part of teacher or facilitator, but also occasionally probes beyond the scope of the conversation.

Naya is the lead instructor. Naya’s posts are generally short, posing probing questions prompting students to go deeper. When a student post includes multiple parts (or responses to multiple questions), Naya selects one part and encourages development in that area.

**Figure 7.** Corpus Comparison – responseLevel.

Renlit tends to answer questions using personal stories rooted in a professional context (wine distribution). This obtains in Renlit’s own lead posts, and in responses to others. Given the focused nature of these stories, Renlit’s posts generally do not include questions for others.

Loret’s posts often begin with praise for the prior author, then use questions to encourage expansion on the original post. In some ways, Loret acts the part of teacher or facilitator, but also occasionally probes beyond the scope of the conversation.

Naya is the lead instructor. Naya’s posts are generally short, posing probing questions prompting students to go deeper. When a student post includes multiple parts (or responses to multiple questions), Naya selects one part and encourages development in that area.

84
Finally, in Figure 7, we show comparative corpora for responseLevel – the level of the conversation tree at which the discussant contributes each response. The slightly cooler cast to Renlit’s corpus indicates that Renlit tends to participate somewhat later in a thread than Loret – compare, for example, at L2 (orange) and L4 (light green). Note also that Loret and Renlit each have four responses at L1, indicating that they have each initiated four threads. Naya, on the other hand, has no posts at L1 because instructors do not typically respond directly to their own discussion prompts. This may seem self-evident, but it is encouraging to see an intuitive result illustrated so plainly in the data. Finally, to the question of Loret as a student who presents as instructor-like in certain ways, what happens if we disregard the L1 responses in the Loret and Renlit diagrams? The remainder of Loret’s corpus falls somewhat between Naya’s and Renlit’s for wordCountAvg, as well as for the distribution of responseLevels. For example, Loret’s proportion of L2 to L4 posts is much more similar to Naya’s than it is to Renlit’s.

D. RQ1 Discussion

A simple graph traversal, derived from the schema shown in Figure 3, can yield a participant corpus data structure that is amenable to visual and statistical analysis. The examples above show that comparative corpus diagrams can be used as exploratory tools for generating rough insights about individual differences among discussants, and as useful models to support reasoning about individuals. They
can provide valuable data and insights that instructors can use to help students, and that students and instructors can use to help themselves. For example, longitudinal analysis could show changes in the character of a student’s corpus over time, or reveal instructor strategies and interventions that work more or less well for individual students. An instructor might realize she tends to interact more with advanced students, even though they are not the ones who need the most support. It is also important to note that automated metrics could be based on the structural and mathematical properties of the schema, so that even if the metrics are imperfect or approximate, they can provide a consistent yardstick against which to better understand, measure, and improve social environments for learning. Two instructors may come to different conclusions about a student based on their expertise and course requirements, but they would have the same tools and evidence available to support their decision-making process. A wide variety of education research studies could conceivably be conducted using a consistent descriptive baseline of participation metrics, conceptual content, social learner models, and comparative conversation analysis tools.

Access to such tools could also have powerful implications for instructional design and teaching practice. One instructor, upon viewing SKN data for a course, realized that although challenges are a desirable behavior for the course, they were seldom being used by students. The instructor subsequently added an activity that explicitly required challenges as an output of student work.

Figure 21. Instructor Participation Across Seven Discussion Threads. Instructor posts (Jakata and Naya) are highlighted, showing a distinctive pattern of posting across threads within a narrow time window.
Visually identifying difference might also allow instructors to more easily target messaging and feedback to individual students. Figure 21 illustrates the relatively regimented participation patterns of the instructors in our data set, as compared with the more free-flowing timing of student contributions. Instructor corpora are also strikingly similar to each other, as compared with the diversity of student corpora. Though we cannot be sure of the reason for this regimented behavior, it is safe to suggest that as class sizes increase, it becomes difficult simply to read the massive volume of student contributions, much less to fairly assess contributions or craft individualized responses. Corpus diagrams could help instructors in large online courses by presenting high-level summaries and signifiers to help them target attention, participate more effectively, and perhaps gauge the effectiveness of various response interventions over time and at scale. These visualizations can not only provide instructors with a better understanding of student contributions, but also perhaps provide students and instructors with tools for perceiving, assessing, and focusing their own behaviors and interaction strategies.

Although there is not enough space to discuss it here, we have also experimented with creating a ‘concept corpus’ for each participant. This model connects a person directly to the concepts mentioned throughout their response corpus, producing a concept graph of that person’s favored discussion topics over time, which could be used to recommend content, connect with peer tutors, or form effective work groups. 8.3. RQ3 Example describes the construction of a concept graph for a discussion thread, and Interactive 4 allows basic exploration of that concept graph. This example can be used to imagine how an individual concept corpus could be utilized.

**VII - RQ2 Findings:** Can we identify, differentiate, and visualize conversation attributes and behaviors in an online discussion or course?

**A. RQ2 Conceptual Overview**

In 6. RQ1 FINDINGS, we considered a collection of hand-coded response attributes across a discussant corpus as a means of representing, differentiating, and reasoning about individual discussants, using digital-ethnographic readings as an analytical anchor. The individual corpus, as the unit of analysis, was constructed based on the relations between a person and their associated response nodes in the graph. What makes that analysis possible is consistent and replicable corpus generation based on the underlying structure of the graph. Individual corpora may vary, but their underlying structural properties are the same.

Now, in 7. RQ2 FINDINGS, we investigate the interactional, influential, temporal, and co-creational aspects of individuals participating in discussion threads. We approach this problem using the same SKN attributes and digital-ethnographic descriptions, mapped onto the somewhat more complex graph structures of threaded discussion response trees. We also describe a graph-structural influence metric, called DiscussionRank, that can be used to gauge the impact of a response, response author, particular speech act, or other event on the evolution of a discussion thread.

**B. RQ2 Technical Summary**

For conversation modeling, we can use the Discussion--contains-->Response and Response--hasResponse-->Response relations in our schema to extract a basic subgraph of the desired discussion.
Figure 8. Response Threads for 24 Discussion Prompts.
That subgraph serves as the foundation for further exploration, analysis, and visualization. While the graph structure of an individual author’s response corpus is a simple, hub-and-spoke structure, the recursive, branching structure of a discussion thread requires a more complex traversal that, when expanded to return all threads associated with a collection of discussion prompts, yields a visualization like that shown in Figure 8.

While this early test visualization reveals notable structural differences across threads, and maps an intriguing geography of citation usage, the elements of time, authorship, and conversational content are notably absent. We will address time and authorship here, and explore content more closely in 8. RQ3 FINDINGS.

We had neither the resources nor the inclination to approach the problem of time-based graph visualization programmatically in the early phases of our research, and existing tools were too constraining. We therefore took an exploratory, design-based approach to modeling conversational graph structures in time, and performed it on a small data set to help us begin thinking about the problem. This visualization approach maps conversational terrains in a way that attempts to capture network structure, individual response attributes, attribute trends, and individual participation patterns over time.

Figure 9. Graph timeline representation of a single week’s discussion over a period of four days, in ten-minute intervals, coded for use of citations.
First, we used Gephi to size, color, lay out, and export the threads to SVG format, and imported to Illustrator layers, as described in 6. RQ1 FINDINGS. We arranged response nodes along a horizontal timeline of the week's discussion based on timestamp data, and labeled nodes with key data points such as author name and wordCount. We could then turn on or off the layers containing color-coded versions of the nodes, to reveal how the values of automated and hand-coded response attributes relate to the combined temporal and graph-structural model of a conversation.

Figure 9 shows the resulting visualization for one week of discussion. Next, we wanted to consider each participant's corpus as a context for their contributions to specific conversations. We added corpus diagrams around the timeline, and connected each timeline response to its position in the author's corpus, as shown in Figure 10.

In this visualization, each thread is assigned a color. Each response in the thread is circled in the same color, and a line of that color connects the response to its position in the author’s corpus. This allows us to see, for example, how typical a comment is for that author with respect to size, quality, use of questions, depth in the discussion tree, etc. In addition, the colored lines emanating from a corpus diagram provide a quick view of the extent to which that author is participating in each of the week's threads. For example, Danen contributes two comments to Danen's own thread (orange), and one comment each to Alakel's, Viska's, and Loret's threads (brown, green, pink). These
connections provide entry points for analysis of the structure and evolution of conversations across the data set. Interactive 2 enables you to browse a week of discussion data interactively, creating custom views like that shown in Figure 10. For example, you can turn on and off various SKN attributes, overlay connector lines, and explore participation patterns, structural elements, corpus statistics, and response typicality with respect to author corpora.

This is a good example of a tool that could be automated to give instructors quick and insightful views into ongoing conversations, enabling them to choose where and when to interact to best effect. Such a tool would provide an intuitive, visual way to explore and compare response attribute distributions, temporal patterns of interaction, conversation structural properties, an individual’s influence on a conversation, or the possible impact of conversational features or events on subsequent discourse.

The DiscussionRank Metric For Conversational Influence

Graph traversals for reasoning about recursively-branching response trees will necessarily be more complex than those we used to investigate the simple hub-and-spoke structure of the response corpus. One excellent example of a graph structural property for measuring participation and influence in a threaded discussion is the DiscussionRank measure devised by Marko Rodriguez of Aurelius, as part of Aurelius’ engagement with Pearson on this study. Roughly inspired by Google’s PageRank algorithm, DiscussionRank measures author influence on a discussion thread based on a count of responses generated,
with a diminished weight for the author's own posts. At each level of the tree at which an author posts, they are assigned one point for their own post, and one point for each subsequent response that is not their own. The process is repeated until there are no more posts by that author in the tree. For a given thread, a person with a higher DiscussionRank score can be said to have generated more discussion than a person with a lower score. This metric does not claim to evaluate the quality or relevance of discussion. Just as PageRank (Page et al., 1999) considers a link to a web page as a vote of importance without otherwise judging the quality of the page, DiscussionRank considers a response to be a vote of importance for a conversation. The resulting metric serves as a consistent, replicable yardstick for investigating what happens when multiple individuals enter into conversation together, and against which we can compare other quantitative and qualitative measures. Figure 11 describes the DiscussionRank counting method. In the example shown, DiscussionRank flips the ‘scoreboard’ upside-down as compared with a basic measure based on the raw number of posts an author contributes.

In its basic application, DiscussionRank is assigned to a person: the author of the initial post. Thus we could compare Renlit’s thread to Kerrad’s thread (as we will do in 7.3. RQ2 Example), to assess which author’s thread produced the most discussion activity. However, this metric can be extended and repurposed in interesting ways. First, the initial response node need not be the lead post. Imagine that both a student and a teacher post questions at the third level of a thread. One could measure DiscussionRank from each point to determine which person’s post generated the most subsequent discussion. One could also analyze data over longer periods of time, in various situations, and under different activity structures, to see which individuals are more or less highly ranked under specific conditions. Secondly, DiscussionRank can measure not only the generative

Figure 10. A sample view of the discussion shown in Figure 9, with author corpora added.
influence of an author on a conversation's structure, but also the influence of particular conversational features on a discussion, with or without regard to the author. For example, imagine that instead of counting DiscussionRank from the first post in a thread and comparing thread scores, we count from each post that includes a citation, and compare cited resource scores. If the DiscussionRank score after a citation to Resource A is calculated at 3, and the score subsequent to Resource B is 5, then Resource B could be said to have generated more discussion than Resource A. This method could be applied to other conversational features as well to help investigate their impact on conversation.

C. RQ2 Example: Chronology of a thread with author corpora

Our sample discussion data is from Group 4, Week 1, in response to a discussion prompt that assigns learners to discuss depictions of predictive analytics in the media, and to describe how predictive analytics could be or are used in their workplace. We will examine the thread for which Renlit is the lead author and posts seven times, with three other students and two instructors posting once each.

This thread is characterized by a strong primary line of discussion between the lead author and two instructors, in a question-answer-question-answer structure.
We will explore the main line of conversation here. There is another question-answer exchange with a fellow student, two less impactful side interactions, and a summative post by the lead author. These can be explored in detail using Interactive 3.

The following figures describe and illustrate structural, temporal, and SKN attributes of responses in the example thread. For context, we also provide some minimal narrative on the content of the conversation, based on our digital-ethnographic analysis. Rather than move through the entire thread chronologically, we will cover the main body of the conversation here, and leave other responses for discovery in Interactive 3. In the section on 8. RQ3 FINDINGS, we will use a concept graph to examine the actual content of the thread in more detail.

Renlit opens this thread with a detailed description of a media piece focused on analytics, as well as an in-depth description of analytics in the wine industry. The wine industry example contains anecdotes from Renlit's own experience, so this response was coded as containing a personalStory (Figure 12). Renlit's corpus diagram indicates high usage of personalStories across all threads – the highest proportion in the data set – and we can see that Renlit tells stories in most of the posts on this thread.
Figures 14, 15, 16, and 17 are presented sequentially for comparative purposes.

**Figure 14.** Thread Chronology. Response 5.

**Figure 15.** Thread Chronology. Response 6.

**Figure 16.** Thread Chronology. Response 3.

**Figure 17.** Thread Chronology. Response 8.
In Response 2, Renlit responds to Renlit’s own lead post with another media example. In the corpus diagram for responseLevel (Figure 13), we can see that Renlit’s three longest posts are first-level responses, in keeping with the first-level post in the current example (Renlit.582). Response 2 (Renlit.112) is in the lower tier of Renlit’s second-level responses by wordCount.

Renlit responds to Naya’s question/spreadRequest nudge with another detailed explanation of analytics applications in the wine industry. Figure 17 illustrates that while most student posts in this thread are coded at topicSpread=Level 3/Elaborate, Renlit’s final response in the question-and-answer chain with Jakata and Naya increases to topicSpread=Level 4/Expand. By this point, the conversation has become a technical and specific discussion between the lead author and the two instructors. It is interesting to note that both instructors have nudged the lead author deeper into material from the lead post, but neither has explicitly attempted to open the discussion to other participants.

**Comparative Thread Analysis**

We can compare and differentiate individual thread graph timeline diagrams just as we can individual corpus diagrams. Figure 18, Figure 19, and Figure 20 compare the now-familiar Renlit thread to the Kerrad thread, which took place simultaneously in the same discussion group and is pictured at the bottom of Figure 10. Figure 18 shows instructors Jakata and Naya each asking short, prompting questions (spreadRequest=Level 3/Elaborate) of both lead
Figure 18. Comparison of Renlit and Kerrad Threads – question.

Figure 19. Comparison of Renlit and Kerrad Threads – knowledge.
authors, at the same time in each thread (points B and E for Jakata, and points C and G for Naya). Despite the similarity of the interventions, the subsequent values for topicSpread, knowledgeActivity, and DiscussionRank are distinct for each thread.

Figure 19 shows an increase in knowledgeActivity subsequent to Jakata’s question at B, with no change after the partner post at E.

Figure 20 shows topicSpread increasing to Level 4/Expand after Naya’s question at C, but no change after the partner post at G.

As a final point of comparison, we can use discussionRank to assess the generative influence of individual questions on subsequent discussion (see Figure 11 for an explanation of how to calculate discussionRank). For example, Jakata’s discussionRank score is 7 at point B, and 3 at point E. The differences in knowledgeActivity, topicSpread, and discussionRank values for Jakata’s questions at B and E signal some variation in influence, even given the similar instructional questioning strategy. There could be many reasons that similar interventions in similar contexts would produce varying results. In the case of the Kerrad thread, Kerrad expresses initial apprehensions about statistics and analytics. As a result, the responses from the rest of the group are focused on helping Kerrad to understand analytics in the context in which they were presented. By contrast, the Renlit thread is more focused and technical in nature. The Kerrad conversation remains more static at a level of explanation, whereas Renlit’s thread shows more change. The ability to perceive such trends and distinctions in conversations using a set of familiar metrics could help instructors more effectively engage with, assess, and support learners in online social spaces.
D. RQ2 Discussion

The thread graph timeline visualization allowed us to see the corpus data in context, revealing both how individual attributes are expressed in a conversational context, and how others responded to these behaviors.

The timeline also helped us to see phenomena that were clear in neither the corpus visualizations nor the LMS discussion board display. For example, we have commented above on the influence of Jakata and Naya’s successive questioning on the evolution of the thread. But also note the pattern of instructor participation across all threads in this week of discussion, highlighted in Figure 21.

Jakata’s question is one of several of similar format posted across multiple threads within a 20-minute period on the evening of Day 2. Based on the data for the entire week, and taking into account Jakata’s corpus diagram, this short, targeted nudge for elaboration appears to be a templated strategy for engaging in and promoting discussion. Naya appears to employ a similar approach, only later in the week. Note that the Alakel and Fesler threads at the top of Figure 21 are not yet extant during the time Jakata is posting, and Jakata never returns to post in those threads. The ability to identify this pattern does not necessarily invalidate the approach. Indeed, it appears to work fairly well for Renlit in this case, as a validation of Renlit’s examples and as encouragement to use personal and professional experience as tools with which to engage with course concepts. But does the strategy work consistently in varying contexts, and for students who post at different times? And what of the other student par-
participants? What strategies might an instructor employ to bring others into a discussion that centers around a participant’s particular area of expertise? What more might an instructor be able to do with tools that support the ability to navigate, understand, and participate effectively in an unfolding discussion? We hope future research in this area will begin to address these and other questions, in service of improving effectiveness, efficiency, and engagement around social and cooperative learning activity in online environments.

Recall from our discussion of corpus data that we noted the consistency of instructor responses. The timeline data provides some insight into the impact of this consistent behavior. We used the binary attribute onTargetPost, for example, to search for instances where an instructor response to an off-target post led to a subsequent on-target post. In the case of the avowedly small data set we queried, this event took place only twice over three weeks of discussion. This points to a need for more effective instructor responses—assuming that onTargetPost is a valued attribute for a given context. Assessing the best response type for given post characteristics is another layer of future research that could emerge from this approach.

The timeline visualizations also helped us to recognize flaws in the structure of discussion activities. For example, a typical assignment asks students to respond to an initial prompt and then to post responses to a set number of other students. Yet the data suggest this type of activity structure leads to sprawl. For the week visualized and discussed in 7. RQ2 FINDINGS, a single prompt leads to 24 unique endpoints. This highlights the fact that ‘social’ learning assignments should be clear about the goals of conversation—converging, diverging, problem-solving, etc.—and specify writing activities that guide students towards these behaviors. We might even come to recognize particular data fingerprints associated with different social and cooperative activities, and distinguish between their more and less successful forms.

VIII - RQ3 Findings: Can we identify and visualize content focus over time in an online discussion or course?

A. RQ3 Conceptual Overview

We felt it was critical for our model to surface important concepts in a conversation, how the concepts are related to each other, and how they change over time. The topicSpread score provides one method of tracking changes in content over time: a rising or falling trend in the topicSpread scores for successive discussion responses can provide a sense of the degree of topical expansion or stasis in the discussion. However, topicSpread remains a numerical score, yielding no information about the actual topics under discussion. It is also a subjective, manually-applied score at present, and could be difficult or computationally expensive to replicate automatically. Below, we describe our initial efforts to understand the topical evolution of a conversation over time, including an examination of the discussion concepts themselves, as extracted using NLP and situated in our graph schema.

A. RQ3 Conceptual Overview

We began our investigation of topical focus using exploratory visualizations. We used our Gremlin DSL to extract discussion graphs that contained response nodes and concept nodes, and their connecting edges (re-
response--mentions-->concept). For some visualizations, we also added person, resource, and citation nodes. The addition of resource nodes, for example, allowed us to see the overlap between student-mentioned concepts, and concepts in assigned reading material, as shown in Figure 26. We applied force-directed layouts, with concept labels sized by the number of responses mentioning them (concept InDegree). Based on a close reading of the Renlit thread, we determined that the four major concept categories under discussion were Media, Analytics, General Business, and Wine. We then assigned each concept to one of those four categories, or left it unlabeled. The categories are color-coded, so that the resulting visualization (Figure 23 and Interactive 4) provides a rough understanding of the mixture of topic areas covered in each response. Unlabeled concepts were omitted from this visualization for simplicity.

While this approach to understanding topical focus admittedly has its attendant flaws and assumptions, we believe this kind of diagram can provide some insight into how we might gauge the prominence of individual concepts in a conversation; the categories of concepts under discussion; the emergence, progression, and disappearance of concepts over time; and the degree to which each participant contributes to discussion around a particular concept. The visual elements are underpinned by real graph relations, amenable to counting and interpretation by algorithms. One improvement we intend to make in ongoing work is to relate the discussion concepts to an ontology of the course domain, with the goal of understanding conversational content against the conceptual structure of course content.

Figure 22. Concept Overlap Between One Post and One Assigned Text.
Figure 23. Categorized Concept Graph for Renlit Thread.

Interactive 4. Categorized Concept Graph showing Response Node 5.
C. RQ3 Example: Concept Progression and Concept Overlap

Below, we illustrate two approaches we used to explore topic focus over time: 1) Categorized concept progression; and 2) Concept overlap. We revisit Renlit’s thread from 7.3 RQ2 Example, in which Jakata and Naya ask successive questions that lead Renlit to delve deeply into the technical applications of analytics in the wine industry. The discussion prompt assigned students to discuss media depictions of predictive analytics, and to describe how analytics are used or might be used in their own work or industry. Figure 22 shows the graph-based chronology for the Renlit thread from 7. RQ2 FINDINGS, for reference. Each response node is numbered chronologically, for easy comparison with the categorized concept graph in Figure 23.

Concept Progression

Figure 23 illustrates a categorized concept graph for a single discussion thread, with Renlit as the lead author. The twelve responses are arranged in a circle, each labeled with its chronological order in the discussion, and the author’s name, ascending clockwise. 01 RENLIT is the first post, and 12 RENLIT is the last. The grey arrows describe the response tree structure, and indicate where questions are present. Edges are drawn between responses and the concepts they mention. If a concept is only mentioned in a single post, it floats to the outside of that post. If a concept is mentioned in multiple posts, it floats to the middle and is sized according to the number of posts that mention it (concept InDegree). We will call these multiple-connected concepts the ‘central’ concepts, and take them as a high-level representation of discussion content for purposes of analysis. You can explore the concept graph diagram interactively in Interactive 4. Select response nodes and central concepts in succession to get an idea of who is talking about what, and how much.

In 01 RENLIT, Renlit opens the conversation with a broad post covering all four main concept categories, including some media depictions of analytics, and a detailed example of analytics in the wine industry. The post is judged onTargetPost=true. After Renlit quickly follows up with another media example in 02 RENLIT, we are presented with three question-and-answer pairs, as shown in Figure 24. Renlit responds individually to questions from Jakata (03), Naya (06), and Loret (04).

In 7.3. RQ2 Example where we color-coded the timeline diagram for questions, spreadRequests, topicSpread, and other attributes, we pieced together the influence of Jakata and Naya’s questions on the evolution of the thread. Now that we are able to view the categorized concept graph of the thread, we can see lexical clues to the content of these questions and their responses. For example, the digital ethnography indicates that 03 JAKATA poses a question about the use of indices in the wine industry. Note that the dominant Wine concept category (red) in Jakata’s question appears to carry over into 05 RENLIT, where Renlit answers Jakata’s question. We see a large cloud of new wine-related concepts connected to 05 RENLIT, including particular wines, vintages, stock bottles, rainfall data, neighborhood shops, Liv-Ex’s fine wine indices, and Wine Spectator ratings, mixed in with some business-related concepts such as business decisions, investors, profit, dollars, and retail. The post also connects to several central concepts, including wine, wine business, bottle, data, and retailer. Analytics and Media concepts are absent. When we look at the distribution of concept categories over
time, we can easily imagine that by focusing on Renlit's professional domain of wine distribution, rather than on the media depictions, Jakata's question has acted as a sort of conceptual filter that helps to shape and guide the subsequent conversation. Naya's subsequent question in 06 NAYA also focuses on wine, seeming to reinforce the effect in 08 RENLIT, where we again see Wine concepts dominate. We see a similar effect, in different concept categories, between Loret's question in 04 LORET, and Renlit's response in 07 RENLIT. Both responses are dominated by the blue and yellow Analytics and Business categories, with no reference to Media, and with some reference to Wine by Renlit.

Now, we can combine our insights from the timeline diagrams in 7.3. RQ2 Example and from the categorized concept diagram in Figure 23 and Interactive 4. 03 JAKATA, 04 RENLIT, and 06 NAYA all contain questions, and all issue a spreadRequest of Level 3/Elaborate. In each case Renlit's response contains a personalStory, and the topicSpread score for the response either meets or exceeds the requested topicSpread level (see Figure 17). And in each case, the dominant concept categories in Renlit's response match the dominant categories in the question. This appears to support a hypothesis that questions can act as important inflection points in a discussion, and indicates potentially interesting ways of finding and surfacing those inflection points to enhance learning and instructional effectiveness.

To round out our description of this thread with respect to topical focus, let's look at responses 09-12, in Figure 23 or in Interactive 4. In 09 VISKA, Viska contributes a short post describing the Lion Nathan Group's QR codes, in case Renlit is not aware of this. Note that this post, and Renlit's thank you note in post 11, are the only posts in the thread that do not mention any of the larger, central concepts. Alakel's post 10 ALAKEL is a short comment on analytics, which connects to only one central concept – data – and receives no response. Another interesting feature of this thread is Renlit's summative post, 12 RENLIT. Here, Renlit retreats from the increasingly technical narrative, and resets the topical focus to the Media, Business, and Analytics concept categories. The digital-ethnographic data indicates that Renlit is returning to the discussion prompt to ensure the thread overall has met the assigned goals. It is interesting to note that although Media concepts feature prominently in three of Renlit's seven posts,
Visualizing Knowledge Networks in Online Courses

no other participants mention Media concepts, and none of the thread's central concepts are Media-related. Nonetheless this thread was judged as onTargetThread=true, perhaps due to the deep dive into Renlit's professional experience, which was also part of the assignment.

D. RQ3 Discussion

The early topical focus visualization shown in Figure 23 was time-intensive and involved a number of manual steps, but it can serve as a roadmap for automated approaches. The underlying graph structure lends itself to automated data extraction and visualization methods, and can be used as input to statistical, algorithmic, machine learning, mathematical, and other modes of analysis. To illustrate this point, we will outline a simple, example metric for calculating individual concept overlap scores in a discussion.

The conceptOverlap metric emerged from our desire to somehow quantify the ways in which participants are connecting with each other against the backdrop of the discussion's concept graph. It is important to note that in this example we calculate concept mentions by post, not by author. If Renlit is the only participant to mention a concept, but mentions the same concept in multiple posts, the concept score will still increment. Properly weighting and interpreting such factors is an important area for future work.

The basic formula produces the ratio of the number of central (multi-connected) concepts mentioned by a person, to the total number of concepts mentioned in the conversation. To state it in graph terms: for a given author, count the number of concepts the author mentions where the concept InDegree > 1, then divide by the total number of concepts regardless of InDegree. We can also produce these ratios with respect to each concept category, to see how individuals are contributing to the relative prominence of central concepts. ConceptOverlap values for the Renlit thread are provided in Figure 25.

Upon further testing a score like conceptOverlap can be adjusted, weighted, and modified. For example, overlap values could be weighted depending on the number of participants mentioning each central concept, the associated level of topicSpread or knowledgeActivity, or concept relevance as determined by comparison with an ontology of course content.

Also note that conceptOverlap need not only be measured between individual posts. For example, it could also be measured between two individuals over multiple conversations, between an individual and the resources they cite, among members of

<table>
<thead>
<tr>
<th></th>
<th>Total Central Concepts Mentioned</th>
<th>Ratio to Total Concepts</th>
<th>Ratio to Central Concepts Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Renlit</td>
<td>18</td>
<td>0.110</td>
<td>1.000</td>
</tr>
<tr>
<td>Jakarta</td>
<td>1</td>
<td>0.006</td>
<td>0.056</td>
</tr>
<tr>
<td>Naya</td>
<td>3</td>
<td>0.018</td>
<td>0.167</td>
</tr>
<tr>
<td>Viska</td>
<td>0</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Loret</td>
<td>4</td>
<td>0.024</td>
<td>0.222</td>
</tr>
<tr>
<td>Alakel</td>
<td>1</td>
<td>0.006</td>
<td>0.056</td>
</tr>
</tbody>
</table>

Figure 25. Question and Answer Pairs in Renlit Thread.
a group, or among multiple conversations or courses. As mentioned earlier, such metrics could serve as a foundation for content, peer tutor, or study group recommendations. They could also serve to support instructor facilitation, student awareness and engagement, dimensions of assessment, comparative analysis for research, suggested conversational entry points based on personal interests, and more. As one example, Figure 26 illustrates concept overlap between the text of Renlit’s lead post, and the text of How to Lie with Statistics (Huff, 1954), an assigned reading cited in the post.

Ongoing work in this area includes automated concept categorization, automated approaches to scoring topicSpread, mapping concepts to an ontology, and linking topicSpread scores to the actual concepts under discussion.

IX - Discussion and Implications for Future Research

The emergence of social tools in educational settings combined with a developing awareness of big data and visualization techniques mark a critical opportunity to develop techniques for collecting meaningful data that enable us to better assess social behaviors in online courses. This area has been previously under-represented in research, and conditions are favorable for us to develop a deeper understanding of the tools and pedagogies that support learning in social and cooperative online learning spaces.

Our research to date details a methodology for capturing individual and conversational patterns present in online Social Knowledge Networks. And although we are encouraged by the findings so far, we have gone deep but not broad. A more rigorous examination is required to draw clear conclusions about this work.

A. Learning Activity Design

We suggest that the most effective approach for assessing the productivity of a discussion is not a standardized “counting mechanism,” but a tailored approach more dependent on activity type. A discussion in which students share their own experiences and engage in interviewing activities should have a different fingerprint than one in which students are working to develop a single solution to a problem. Identifying the anticipated data fingerprints associated with a library of activity types, and their variations, will be a critical step to defining student and instructional strategies for success.

B. Learner and Instructor Strategies

Similarly, whether learner and instructor strategy is effective depends at least in part on our expectations for the discussion.

We can also ask questions about how instructor strategies might vary depending on the students to whom they are responding. This connection, however, relies on us knowing more about the nature of corpora. In particular, does the character of a corpus stay the same across a student’s academic career? Or does it change based on the composition of their cohort, their development through a program, or other factors? These questions may lead us to identify new metrics for predicting and supporting team and cohort success, and the ways in which individuals may influence one another over the course of their interactions. If we can begin to measure these influences, we might be able to establish and support successful cooperative and collaborative teams, learning communities, peer tutoring relationships, and more.
C. Tools and Platform

Another area of future research and development concerns Learning Management Systems and other platforms in which learning-focused discussions are hosted. The traditional linear, threaded discussion forum might make the effective facilitation of discussion difficult. Consider the case of Jakata’s entry into the week 1 discussion: Jakata responds to all visible posts in a brief timespan but receives no notification of new posts after two new students respond. Further, these new posts are pushed to the bottom of a chronological display, meaning that when Jakata logs in, these responses may not be immediately visible. Rich opportunity lies in investigating the kinds of layouts, signals, entry points, notifications, and recommendations that give rise to more expressive and efficacious social learning environments.

D. Data Science, Automation, and Algorithms

The numerical, categorical, text, and other attributes of each response in a corpus or a discussion are available within the native graph structure of the data for detailed statistical, graph-structural, and other analyses, as well as for visualization. This enables a combination of high-level visual survey and detailed data analysis that we hope can help speed the research-into-practice cycle for online social and cooperative learning environments.

Of course this does not mean we have discovered how to reverse-engineer deep, digital-ethnographic descriptions from course or discussion data. Most attributes for this study were manually coded by human experts. However, if over time we can develop the capabilities to automatically apply some or all of these, or other, codes, we believe it will lead to valuable new ways of designing, describing, navigating, supporting, and evaluating social and cooperative learning activity in online courses at scale. Therefore, the Pearson team continues to evolve, scale, and automate this research-based graph database system for social and cooperative learning and discourse. For example, we have implemented experimental versions of: NLP-based question and citation identification; a preliminary topicSpread metric; a conversation influence metric; an ontology comparison model for understanding conversation concept structures; a measure of response reciprocity among a community of learners and instructors; and visualization components for viewing participant conversations and corpora in ways similar to those presented in this paper. Some of these features are currently available in experimental alpha release form to individual students and instructors using the OpenClass LMS platform, on the Learner Intelligence alpha page.

E. Closing Thoughts

We have suggested here that the confluence of data-driven decisions in education and the proliferation of social media tools make the time right for a deep exploration of how knowledge is constructed in online social learning spaces. Our goal, in particular, was to define a set of individual, conversational, and content-based attributes and behaviors that might support the formation of thriving social knowledge networks.

We have accomplished something of our goal, in that we have been able to identify and visualize trends and behavior in those three areas. We recognize, however, that the work is far from complete, and we hope that this paper serves as a catalyst for additional research into this important, emerging field.
We have described above some major themes and opportunities to guide our—and others’—future research. In summary, we believe the ability to answer these questions may have a transformational impact on institutions’ fundamental approaches to teaching and learning. For example:

- Learning to measure and value individual patterns of behavior in the context of discussion and collaborative activities in online courses allows for more holistic assessment of student performance and potentially more proactive and actionable interventions to identify and assist at-risk students.
- Learning to identify, measure, and value conversation patterns in the context of discussion and collaborative activities in online courses, will assist in the development of new pedagogies, course and activity management strategies, and technologies aimed at increasing the productivity and positive impact of these activities in online courses.
- Learning to visualize topic spread and conversation swell around particular topic areas, and to evaluate them against structured concept graphs, will assist in the development of program, course, and activity design, adaptively matching students with helpful content, promoting lifelong learning behaviors, and more.
- Concerns that models and typologies may originate in this kind of research and spur action on measures of student performance that are not yet well understood and that may change across contexts and across time.
- Concerns over reporting (to students, faculty, administrators, and systems) and the creation of records of fine-grained student performance that persist over time, as well as a multitude of other ethics and data privacy issues.
- Recognition that findings will precede mechanisms for implementation, and a commitment to supporting teacher educators, faculty and instructional designers in effectively and responsibly adopting new methods.
- Considering to what extent, when students become co-creators of course content, there should be oversight/monitoring/policing of the flow of information to assure that it is helping students rather than confusing or overwhelming them.

We hope this report will contribute to a responsible evolution of online and blended teaching and learning, through an increased awareness and understanding of the social spaces in which these increasingly occur.

References


About the Authors

Marni Baker-Stein, is the Chief Innovation Officer for the University of Texas System's Institute for Transformational Learning, Dr. Stein is an authority on online program and curricular development, delivery, and assessment; next-generation strategic enrollment and student lifecycle management infrastructure implementation; and student-centered, outcomes-focused, competency-based instructional design. Prior to joining the University of Texas System, she was Senior Associate Dean of Columbia University's School of Continuing Education, where she was responsible for the development, design, and evaluation of all online programming initiatives, and Director of Program Development for the University of Pennsylvania's College of Liberal and Professional Studies. At Penn, she developed the university's Open Learning Commons, an innovative social networking platform to host online courses, communities, and open educational resources, and designed pioneering online programs in positive psychology and environmental sustainability, which enrolled thousands of students from 63 countries. Frequently invited to speak on technology-enhanced curricular and pedagogical innovation, Dr. Stein has a Ph.D. in Teaching, Learning and Curriculum from Penn. An accomplished educational researcher, her scholarship focuses on social and knowledge networking behaviors in online courses and the impact of design, instructional strategies, and platform technology upon student engagement in e-learning.

Sean York is the Chief Architect, Innovation & Advanced Development for Pearson Global Higher Education. He has worked for 14 years in educational technology, instructional design, and software development, and holds an M.S.Ed. in Instructional Systems Technology from Indiana University. He leads research that moves theory into practice in large-scale software development for education. He has been with Pearson since 2005 in Instructional Design and Software Architecture roles, and works on the technical and theoretical foundations for Pearson’s social and cooperative learning infrastructure.

Brian Dashew is the Director of Instructional Design at Columbia University’s School of Continuing Education, where he works with program faculty in a number of professional certificate and Master’s degree programs. Brian is also a doctoral student in Adult Learning and Leadership at Teachers College-Columbia University. Prior to his time at Columbia, Brian was an instructional designer at Marist College and eCornell.

About the Authors

Some Pearson platforms and technologies referenced in the paper are Patent Pending, U.S. Publication No. US-2014-0272911-A1, and U.S. Provisional Application No. 62/072,932. Patent applications cover underlying methods and technologies relating to surfacing and implementing educational interventions at scale using network-based methods and analytics, not the specific analytical framework described in this paper. The goal is to support a diverse body of research and collaboratively produce work to improve the quality and effectiveness of social learning tools, platforms, and emergent pedagogies.