Chen, W. et al. (Eds.) (2017). Proceedings of the 25th International Conference on Computers in Education. New Zealand: Asia-Pacific Society for Computers in Education

# Discovering Dynamics of an Idea Pipeline: Understanding Idea Development within a Knowledge Building Discourse

**Alwyn Vwen Yen LEE & Seng Chee TAN***
*Nanyang Technological University, Singapore*
*sengchee.tan@ntu.edu.sg

**Abstract:** Idea development is an important process within a knowledge-building discourse and it is crucial to understand the dynamics of idea development throughout the discourse, such as the growth, flourishing or fading of ideas. This study proposes a framework called Idea Pipeline that explores and tracks the dynamics of idea development within a knowledge-building discourse. This pipeline consists of three phases: discovery, identification and analysis, and 'rise above' of ideas. Each phase of the pipeline will be illustrated using findings from a comparison study of two online knowledge building discourses. During the discovery phase, a text miner is used to identify groups of related keywords from the discourse; this is represented as keyword graphs with weighted frequencies to show the diversity of ideas that were embedded within the knowledge-building discourse. In the idea identification and analysis phase, network analysis was conducted to help label key ideas that were promising to the discourse community; this would provide the community with information to decide which ideas to pursue so that advancement of communal knowledge could be achieved leading to the 'rise above' phase. This Idea Pipeline framework can be an additional method for the temporal analysis of a computer-supported collaborative learning discourse over a longer duration of weeks or even months.

**Keywords:** Online discourse analysis, knowledge building discourse, idea pipeline, idea development, text mining, network analysis

## 1. Introduction

In the field of computer-supported collaborative learning (CSCL), researchers have access to processual data over a duration of weeks, months, or even years. Yet, many CSCL researchers ignore the temporal information and choose not to analyze changes over time. Reimann (2009) argued that experimental studies that have strict control for variables may not work well for CSCL studies, not just because of the practical difficulties in controlling the variables, but also the fundamental assumption that the factors exerting consistent and continuous influence on the learning events over the entire period may not be valid. Few CSCL researchers have studied processual changes over time, and among the few (e.g., uptake of ideas by Suthers, Dwyer, Medina, & Vatrapu, 2010), the current methods rely extensively on labor-intensive manual coding and interpretations. In addition, methods such as a conversational analysis that examines the temporal development of ideas are not applicable in explaining longer term unfolding of events.

This paper intends to address this research gap in CSCL by proposing a framework called Idea Pipeline that explores and tracks the dynamics of idea development within a knowledge-building discourse over time. We piloted and illustrated the workings of this framework using a comparison study of existing data, which were obtained from two online knowledge building (KB) discourses conducted on Knowledge Forum (Scardamalia, 2004). We showed how technologies, such as text mining and network analysis tools, can be leveraged to provide insights into learning behaviours in online discourse and to automate part of the analyses. The analyses in this study were carried out after the completion of a course, but the methods can be applied to facilitate knowledge building among learners at various junctures of their discourse.

## 2. Background

As knowledge is constructed through social interactions such as discourse (Stahl, 2004), researchers have been harnessing technological tools to investigate discourse, such as converting knowledge and ideas into innovative product (Hansen & Birkinshaw, 2007), or assisting students in making promisingness judgements to achieve better knowledge advances that benefit their community (Chen, Scardamalia & Bereiter, 2015). This study is situated in knowledge building classrooms (Scardamalia & Bereiter, 2015), where collaborative improvement of ideas forms the central tenet of students' learning. Online discourse is conducted on Knowledge Forum, a CSCL platform that is specifically developed to support knowledge building using discourse scaffolds and analytics. Knowledge building engages students in discursive activities intended to enhance collective understanding (Bereiter, 2002). Students in a knowledge building community work on authentic knowledge problems that arise from their attempts to understand the world, and the knowledge problems are often identified and articulated as ideas. Students construct and progressively improve on ideas represented as knowledge artifacts (such as online notes containing their ideas) and this process often leads to a diversity of competing ideas. This provides the natural impetus for discussions that lead to collective idea improvement. By improvement of ideas, we refer to the enhancement of coherence, quality, and utility of ideas. Continual engagement in idea improvement process may lead to 'rise above', which refers to the higher-level outcome of integrating several ideas or framing the inquiry using a higher-level principle or theory.

The process where students deliberate, work on and improve their shared ideas collaboratively is similar to the innovation processes of improving the design of an object, such as an improved version of a mobile phone. The classical view of the innovation process consists of four phases: idea generation, idea formulation, problem-solving, and utilization (Myers & Marquis, 1969). In the innovation process, a large number of ideas are generated in the early stages, which need to be evaluated and selected, so that resources can be devoted to more valuable and promising ideas, which could culminate in innovative products or solutions that are critical to the organizations. This process has been largely implemented in production lines and large organizations such as Microsoft Corporation (Bailey & Horvitz, 2010). Nevertheless, similar processes of discovery, identification, analysis and application of ideas can be adapted for a learning community to engage learners in deep learning and meaning making, while steering the students away from focusing solely on knowledge acquisition.

The introduction of the Idea Pipeline framework is significant in several ways. First, in terms of methods, digital data associated with the online discourse allows for some degree of automation, such as the use of analytical tools to calculate and measure learning outcomes, and to identify and analyze ideas within discourse (Lee, Tan & Chee, 2016). In other words, technologies such as text mining and learning analytics can be leveraged to automate parts of the analysis of student interactions and discourse. Second, the proposed Idea Pipeline can track and monitor the development and movement of ideas within the learning community, while seeking to understand the dynamics that stimulate the growth, flourishing or fading of ideas within discourse. This framework can be useful to the CSCL community, as the process of analyzing ideas in discourse over time is complex and still requires substantial research in determining development trajectories and interactions of ideas among different agents in the community. The Idea Pipeline allows us to form a deeper understanding and comprehends idea developments visually within a knowledge building discourse in a more succinct manner. The methods proposed in this study are also different from *Big Ideas Tool* (Chen, Chuy, Resendes & Scardamalia, 2010) that requires the students to identify ideas they feel are promising.

## 3. Idea Pipeline – A Framework for Tracking and Analyzing online Discourse

The proposed Idea Pipeline framework consists of three sequential phases (Figure 1), with emphasis placed on the processes of evolution of ideas through phases of (a) discovery, (b) identification and analysis, and (c) 'rise above' of ideas. First, a pool of ideas is discovered from a communal discourse using a text mining tool to determine related keywords from the discourse with weighted frequencies to show the diversity of ideas and relationship among ideas embedded within the knowledge-building discourse; this can be visualized as keyword graphs (see Figure 2). As most ideas within the initial

phase are competing and often in preliminary forms with uncertain prospects, identifying *promising ideas* is the key focus in the second phase of Identification and Analysis. Ideas that can bring the community forward to help achieve a higher level of understanding are considered promising ideas (Chen et al., 2015). *Promisingness* can be considered from either the student's or the teacher's perspectives. For example, to a teacher, a promising idea is aligned to a learning objective. To a group of students, a promising idea has potential to lead to their shared learning goal. This study analyzes promisingness of ideas from the students' perspectives based on three criteria: (a) the relevancy of ideas to the context; (b) the level of interests of the community in pursuing ideas deserving of further development; and (c) the degree of impact that ideas have on communal discourse through synthesis and improvement of ideas. Finally, the 'rise above' phase in this study consists of integrating diverse ideas to achieve a higher level of understanding, often manifested as a deeper conceptual representation, or higher level of abstraction by using a theory or principle in the explanation for an inquiry. For students who were not able to achieve 'rise above' through communal discourse, results from the first two phases of the pipeline can inform the students about other promising ideas present in the discourse, and they can choose to reestablish their interests and continue contributing to the discussion.

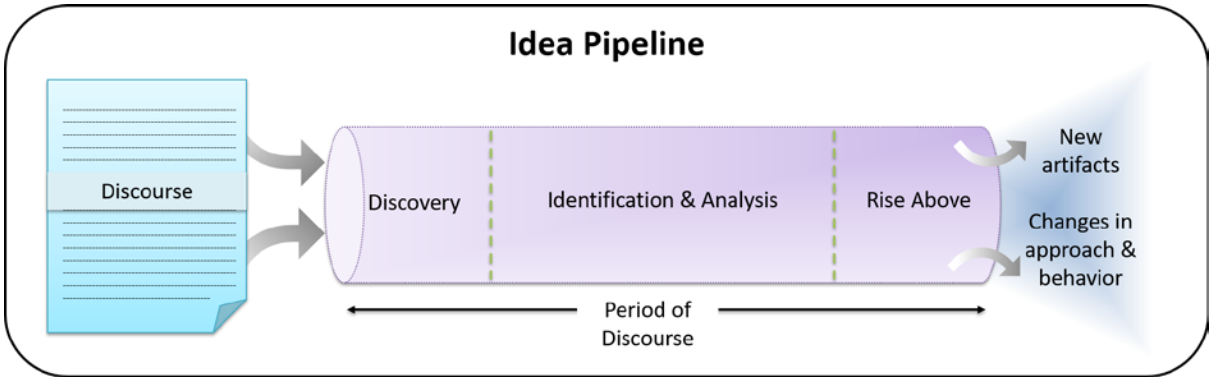


Figure 1. Idea Pipeline Framework

In terms of technology and methods, a text mining tool called SOBEK (Reategui, Epstein, Lorenzatti & Klemann, 2011) is used in the first phase of idea discovery. SOBEK can identify the ideas (as keywords) in a text and construct a relational keyword graphs. Second, building on a previous study (Lee et al., 2016), the Idea Identification and Analysis methodology ($I^2A$) is applied in the second phase of the Idea Pipeline. In essence, the $I^2A$ methodology identifies and evaluates a pool of ideas within a discourse by using a combination of network analysis and temporal analytics to identify promising ideas that relate to context and are significantly interesting to the community. The usage of network analysis helps to assess the interactivity between different agents and components of the discourse network, while temporal analytics provide insights into the often-overlooked temporal dimension of discourse data. The third phase of the pipeline describes the 'rise above' of ideas through qualitative textual analysis. This 'rise above' phase is complex, usually covers a wide range of developments, and is represented mostly through artifact developments, changes in learning approaches or behavioral changes. It can include a translation of verbalized intentions into actual action or practice that are often not captured through discourse. As automation of the 'rise above' phase was not a goal for this study; instead, a qualitative approach was adopted for the analysis of the 'rise above' phase.

## 4. Knowledge Building Classes – Context of the Study

The two knowledge-building classes in this study were two cohorts of educators enrolled in a graduate-level course in two different years, but the courses were conducted using similar content and resources. The course focused on understanding CSCL and knowledge building by encouraging the participants to act as knowledge builders. Community A consisted of 15 participants, including two lecturers and 13 graduate students. Community B consisted of 11 participants, including two lecturers and nine graduate students. This study analyzed the textual discourses (Discourse A and Discourse B) and the dynamics related to the idea development process. Both communities underwent 13 weeks of instruction and collaborative discourse on topics of CSCL and knowledge building as part of their

master's in the education program. Knowledge Forum was used as the main online discourse platform for students to communicate and share ideas with one another. The teaching staff helped to facilitate learning and co-creation of knowledge within the knowledge building community. The notes on Knowledge Forum, referred to as 'discourse units,' were periodically archived for reference and was subsequently anonymized for further analysis at the end of the course.

Over 13 weeks of 3-hour sessions, the participants discussed about CSCL and the core principles of knowledge building, the ways of designing and facilitating collaborative improvement of ideas, and how they could apply what they have learned to design learning environments for their students. This study focuses on investigating how the dynamics of idea development – including processes such as generation and discovery of ideas, sharing of diverse ideas and improvement of such ideas – are manifested in a knowledge building discourse (Scardamalia, 2002). As we illustrate the working of the Idea Pipeline framework, we seek to answer the following research questions: (1) What are the key ideas pursued by the two classes of graduate students in their respective discourses? (2) What are the promising and potential ideas that can be identified from their discourses? (3) Was there any evidence of higher level of understanding ('rise above') found in their discourses?

## 5. Methods

### *5.1 Text Miner SOBEK*

In this study, the key ideas in the discourses were harvested from the participants' notes in the Knowledge Forum. SOBEK (Reategui et al., 2011) was used to identify relevant keywords and related concepts from unstructured text data, using frequency analysis of textual material and filtering of important keywords. SOBEK generates relational graphs of keywords – a graphical representation of nodes of keywords with relative sizes to indicate their frequencies and connections to indicate relationships among the keywords. An in-built thesaurus within SOBEK helps to minimize unnecessary repetition of keywords by filtering out common words such as related prepositions and adverbs, and further helps in aggregating words belonging to similar concepts and meaning.

### *5.2 Idea Identification and Analysis ($I^2A$) methodology*

$I^2A$ (Lee et al., 2016) is a mixed method approach that uses network analysis with keywords and temporal analysis of textual discourse, to identify and trace the development of ideas that become promising talking points within a discourse. We first used a network measure, the betweenness centrality (BC) coefficient, to assess meaningful connections between keywords within discourse units (notes in Knowledge Forum). The presence of keywords in a discourse unit represents the key theme or meaning of ideas within the discourse unit. When the BC values of discourse units are calculated over time, the BC value of each discourse unit will vary according to the network changes and the presence of keywords in the discourse unit; the plot of BC values over time forms a BC temporal trend for each discourse unit. The pattern of the BC trends indicates the nature of the ideas contained in a discourse unit. For example, a discourse unit showing a trend of sustained BC values over time is likely to contain communally interesting and promising ideas. Qualitative analysis is used to validate this classification result. In the previous study (Lee et al., 2016), keywords were provided by the teacher, and these keywords were used for identifying promising ideas to the teachers. In this study, $I^2A$ was deployed similarly, but the keywords were extracted from the student discourses, and therefore, it was used to determine ideas promising to the students.

The promisingness of an idea, when assessed from the student's perspective, is based on three criteria: (a) the relevancy of ideas to context; (b) the level of interests of the community to pursue ideas deserving of further development; and (c) the degree of impact that ideas have on communal discourse through synthesis and improvement of ideas. When ideas in the discourse satisfy these three criteria, the ideas are considered as *promising ideas*. These are ideas that the community feel are relevant and are interesting to continue pursuing after some initial interest. More time can be allocated for discussion, such that interests in these promising ideas can be sustained within the knowledge building discourse. The community can then build on each other's ideas, achieve 'rise above' as a community and

eventually achieve a higher level of understanding and advancement of communal knowledge. However, the proportion of promising ideas in discourse is often far and few. It would be useful to apply less strict criteria to uncover ideas with a lower degree of relevancy, communal interests and impact. These ideas are labelled as *potential ideas* that may contain less promising but yet important elements that help advance discussions.

### *5.3 Finding 'Rise Above' from Discourse*

The method to identify 'rise above' of ideas in a discourse is difficult due to its wide range of developments that can occur simultaneously throughout the whole discourse. Although 'rise above' represented through artifact developments are easier to track using quantitative methods (e.g., finding specific phrases or usage of scaffolds), the changes in learning approaches and behavioral changes are, however, harder to find. Therefore, a qualitative approach was adopted in this study to find discourse units with 'rise above.' We first used the $I^2A$ methodology to search for discourse units that contain promising ideas, and scrutinized the contents for any 'rise above.' We subsequently broadened our search to look for 'rise above' in discourse units that are related to the identified discourse units. Notably, 'rise above' notes contained shared experiences from participants and reflection of their follow-up actions towards some group activities or individual tasks. The experiences and actions are often influenced by feedback from the community or are drawn from inspiration that originated from promising ideas found within the discourse. 'Rise above' notes can also be authored by multiple authors and be the result of group discussions conducted over several group meetings.

## 6. Findings and Discussions

### *6.1 Discovery Phase*

We ran the two knowledge building discourses (281 notes in Discourse A; 204 notes in Discourse B) through SOBEK and obtained a table of weighted frequencies and two relational keyword graphs that contain relevant and commonly used keywords within the respective discourse. Results in Table 1 show that Discourse B contained fewer unique keywords as compared to Discourse A. The frequencies of keywords in Discourse B were adjusted with a proportional weighting to account for the different number of notes, and to further provide a fairer comparison of keyword frequencies between the two discourses. The weighting shown in Table 1 is based on the difference in note count (281/204=1.38x), with the assumption that all notes have equal possibility of contributing promising content to the discourse.

Table 1: Weighted frequency (original frequency) of keywords at the end of Discourses A and B.

| Keywords between the two discourses | Discourse A, 1x weight (Original frequency) | Discourse B, 1.38x weight (Original frequency) |
|---|---|---|
| Based | 63 (63) | No significant presence |
| Community | 79 (79) | No significant presence |
| Discourse | 82 (82) | No significant presence |
| Idea | 67 (67) | 118.68 (86) |
| Information | 55 (55) | No significant presence |
| Knowledge | 325 (325) | 248.4 (180) |
| knowledge building / kb | 424 (424) | 313.26 (227) |
| Learning | 172 (172) | 238.74 (173) |
| Process | No significant presence | 75.9 (55) |
| Students | 151 (151) | 193.2 (140) |
| understanding | 114 (114) | 81.42 (59) |

In this study, both communities of students underwent similar knowledge building activities. The ideas discovered within the discourses are a reflection of authentic thoughts, opinions, and actions that participants contemplated. This situation was observed to be consistent throughout both discourses

as both communities sought to build new knowledge and apply their knowledge on top of their current teaching roles in schools. The results show that Discourses A and B had 10 and 7 identified keywords respectively, with Discourse A showing a more diverse range of ideas. The larger number of keywords and connections between keywords in Discourse A showed that Community A, with the relatively larger number of participants and notes, generated a more diverse list of ideas.

Further, even though both communities contain six common keywords (see Table 1), the differences in the discussion foci and content for both communities are revealed in the relational keyword graphs. The relational keyword graphs were generated shortly after the discourses started (Figure 2) and after the discourses ended (Figure 3). We collapsed the term 'kb' with the keyword 'knowledge building' since both keywords refer to the same concept. The keyword graphs show a stark contrast in the development of ideas in the respective discourses. One significant difference is in the starting points of discourse for both communities. Community A began the discussion on knowledge building, whereas Community B started to discuss on the topic of CSCL, even though both communities were provided with the same teaching materials and instructed by the same teaching staff. Over the period of discourse, the relational keyword graphs for both communities expanded but Community A was more concerned with exploring the concepts of 'knowledge building,' 'community knowledge' and 'understanding,' while Community B was more interested in discussing 'student learning' and the 'learning process.'

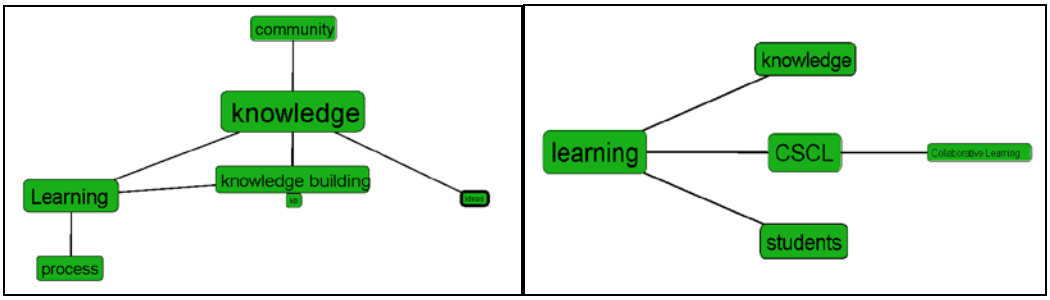


Figure 2. Different discussion foci for Discourse A (left) and Discourse B (right) at the initial part of the discourse

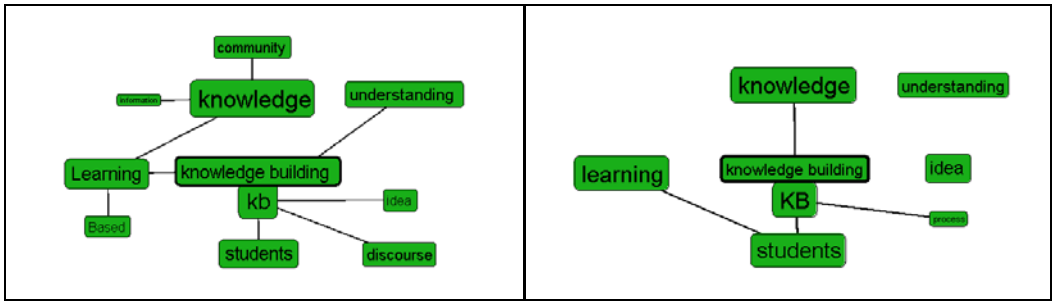


Figure 3. A more diverse spread of keywords in Discourse A (left) than in Discourse B (right) at the end of the discourse

As knowledge building is the foci in both courses, both keyword graphs show that ideas related to the conceptions of 'knowledge building' have been communicated and discussed between students in their respective discourses. The keyword network of Discourse A has noticeably higher connectivity, repetition, and a variety of keywords that were used to represent a wider spread of ideas, opinions, and thoughts. Burchfield (1996) proposed that increasing repetition of keywords that are important to the reader's comprehension allows discourse participants to readily connect ideas and increase coherency. One can also argue that this is plausible at the superficial level, but the issue of how knowledge coherence is achieved is not trivial. This study does not depend on keyword frequency alone to assess cohesiveness, but rather, combines the visualizations of keyword frequencies with inter-keyword connections to show the intensity of keyword usage and strength of connections between keywords in the discourse. Thus, Discourse A is likely to contain arguments that are more cohesive. There were fewer edges in Discourse B (connections between different nodes in the network) observed in the keyword graph, which is a reflection that the ideas were not properly connected in the discourse. The

presence of isolated keywords or keywords not well connected to other related keywords shows the lack of coherence among the ideas in Discourse B. This might explain the lack of activity and conversations in Discourse B, as compared to the wider range of activities and discussions found in Discourse A.

In relation to the content of discourse, a more detailed scrutiny of the textual discourse provided insights into the intentions and learning objectives set out by both communities. Despite the fact that both communities were using the same materials taught by the same teaching staff, with the same instructional objectives, other than the list of six common keywords (Table 1) that are related to the central theme of knowledge building, the overall foci for both communities were found to be quite different. Community A placed greater emphasis on understanding the concept of using knowledge building as a form of enhancing communal understanding, whereas Community B gravitated towards application – using knowledge building and technological tools to help students improve their individual learning processes. In summary, from the analysis of Discourse A, we discovered that knowledge building discourse among the participants did help in the contribution of diverse ideas. At the same time, even though the presence of diverse ideas and unique viewpoints serve as inspiration for other community participants, there remains a need for the community to eventually come together and take responsibility for advancing group knowledge. With regard to Discourse B, although the initial topic of CSCL dominated early parts of the discourse, the lecturer guided the community back to discussions regarding knowledge building by the end of discourse. Discourse B was able to reach a consensus in due course by the end of discourse, albeit with a lower diversity of ideas.

### *6.2 Identification and Analysis ($I^2A$) Phase*

In essence, $I^2A$ methodology can be used to identify promising ideas through network analysis and measures such as betweenness centrality. Groups of keywords discovered using SOBEK in the discovery phase were used to calculate BC values, which are indicative of the importance of keywords and centrality within the discourse network. When BC is calculated over time, BC temporal trends are formed, and these trends constitute patterns that can be used for classifying ideas within discourse into different idea types, such as promising and potential ideas. For discourse units with BC trends that exceeded the BC thresholds, the discourse units are considered to possibly contain promising ideas. In this study, the BC values that act as thresholds for promising ideas are calculated to be close to 0.026 and 0.049 for Discourse A and Discourse B respectively, and the thresholds for potential ideas are calculated to be slightly lesser (<10% of BC threshold for promising ideas) for potential ideas. The final number of promising and potential ideas for each discourse are shown in the results (Figure 4).

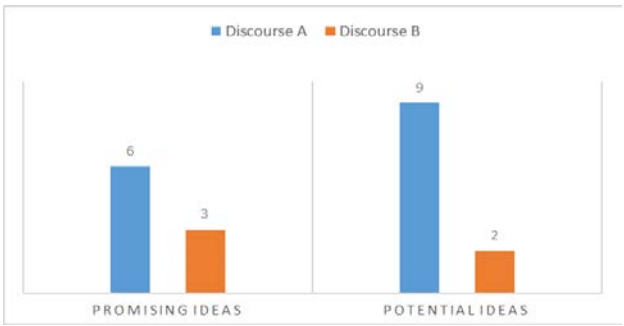


Figure 4. Frequency of promising and potential ideas for Discourse A and Discourse B

In a straight-up comparison between Discourse A and B, Discourse A contains more promising and potential ideas, indicating high relevancy of ideas to context and extensive interests from the community. The results show that there were six promising ideas in discourse A and nine potential ideas, which can be further improved to become promising talking points. The analysis in the earlier idea discovery phase has shown that there is a larger diversity of keywords and ideas in Discourse A than Discourse B. The results obtained from the $I^2A$ analysis provide corroborative evidence that cohesive arguments and diversity of ideas in Discourse A have provided more interesting talking points that the community can participate in. Eventually, the increased interests and participation in discourse allowed the community to attain a higher level of understanding through discussions. Discourse B contained fewer keywords and diversity of ideas, and there was a corresponding fewer promising and

potential ideas. This suggests that it is less likely to find significant interests or participation from participants in Community B, and ideas in Discourse B are likely to have a muted impact on subsequent parts of the discourse.

Since each discourse unit is an individual node connected with others in the discourse network, each discourse unit has its BC trend, and there would be the same number of BC trends as the total number of discourse units in a discourse. Overall, the plots of most BC trends tend to increase to a peak, then decreases over time, and is consistent with the temporal analysis of BC trends using network analysis. However, a graph containing BC trends of all discourse units can appear disorienting and mask the key trends of Discourse A and B. Further, we are more interested in analyzing the unique behaviors of BC trends that vary over the period of discourse, such as the significant or abrupt changes in BC values at specific temporal positions of the discourse. To better reflect the major difference between Discourse A and B, in this study, we condensed and averaged the majority of significant BC trends and their values within a particular discourse into a single BC trend. This BC trend is labelled as a generalized BC trend that is representative of most of the BC trends in either Discourse A or Discourse B. The overall general shape of BC trends for both discourses were plotted (Figure 5) and critical differences can be identified from the two trends. The generalized BC trend of Discourse B exhibited an abnormally high peak early in the discourse, but it dropped to insignificant levels soon after. Discourse A had a gradually increasing generalized BC trend with local peaks towards the middle of the discourse, before decreasing and plateauing at a slightly lower than average BC value towards the end of discourse. Both BC trends were truncated at discourse frame 70, as there were no significant BC trend movements observed after that.

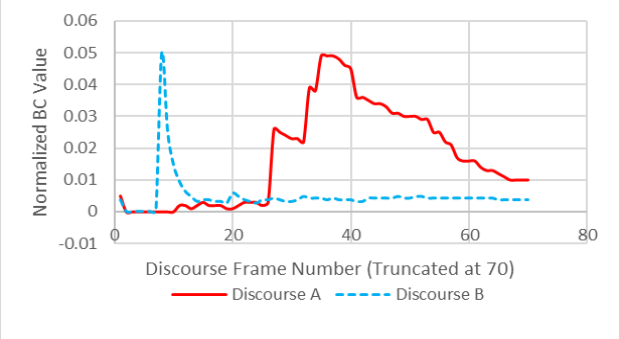


Figure 5. Generalized BC trends for Discourse A (solid line) and Discourse B (dotted line)

Further examination of the textual data provided corroborative evidence consistent with the earlier observations and analysis from the discovery phase of the Idea Pipeline. Participants in Community B initially embarked on discussions on CSCL, but they were eventually guided back to the topic on KB, and the cooling of interests in discussions after the switch is reflected by the significant decrease in BC trend. On the contrary, Community A (solid line) embraced the topic of KB right from the start of discourse and subsequently engaged in discussions to further improve ideas and 'rise above' through collaborative knowledge building discourse. There was heightened interest among the participants, reflected by the peak in the BC trend, before communal consensus brought the discussions to a close with lower activity towards the end of Discourse A. Overall, Discourse A contains sufficiently interesting promising ideas that became talking points, which sustained the interests of the community to continuously build on and improve each other's ideas in order to achieve better understanding. Discourse B was not able to generate the same amount of communal interests, and results indicated that discussions were also not as broad and extensive. These results show that the usage of $I^2A$ within the Idea Pipeline allows us to identify and analyze the flow of ideas and interests within the community, in a more visible and scalable manner.

### *6.3 'Rise Above' Phase*

This phase was difficult to analyze using quantitative measures or with the aid of analytical tools, as the extent of 'rise above' of ideas by participants cannot be accurately measured or tracked when actions and consequences occur outside of online discourse. However, we were able to use the $I^2A$

methodology to identify discourse units that contain promising ideas, and the methodology helped us save time in searching for 'rise above' in the discourse. The search was expanded to look for a qualitative interpretation of instances, where shared experiences from participants and consensus from the community helped sustained interest in discussions that eventually improved and advanced communal knowledge. These instances of discourse provided insights into the extent and success of discussing and integrating ideas that were deemed promising by the community, and are quite frequently spotted within Discourse A. We extracted two examples from Discourse A to illustrate how participants built on promising ideas, improved the ideas throughout the course, and integrated peers' ideas to achieve deeper understanding of conception of knowledge building.

> Let me quote my example: Initially, I have the problem of understanding whether kb is a theory, pedagogy or technology. Throughout this course, I have read our classmates' different perspectives, ideas pertaining to this problem of understanding and I tried my very hard to synthesize these ideas and arrive at my theory that kb is a theory … This synthesis of idea leads me to ask more questions such as is kb related to the theory of connectivism? Can other alternatives technologies be used to support a principle-based pedagogy? Do you think I am achieving rise above? or which dimension of rise above do you think I belong? (DU268 by student S6)

> I came into this course seeking a validation - that I am already practising kb and that kb is pbl (problem-based learning), pbl is kb. As I read more on kb and as I prepared the discussion on pbl & kb, I realised I have obtained my rise above. I now have a much clearer idea to this problem of understanding on 'what is the difference between kb and pbl?'. I have learned that pbl and kb while highly similar has distinct differences, and I should not be attempting to merge the pedagogies as one, though they are very similar … Understanding a pedagogy would be trying to understand its paradigm of education and understanding its origin. (DU269 by student S11)

Both students S6 and S11 mentioned that they achieved a better understanding of their own inquiries and problems through discussions and self-directed learning by the end of the discourse. They discovered and identified conflicting ideas that were viewed as a problem, and analyzed it through discourse with the community, in the form of sharing sessions and presentations. The students eventually worked on their ideas with further intention to improve ideas, either through posing of questions by student S6 to invoke responses or through self-reflection evident in student S11's case, in the hope of ascertaining that his/her statements and contributions are accepted by the community. By undergoing such a process of idea development through the Idea Pipeline, novel ideas can be discovered with a new perspective and be improved through sharing over multiple idea development cycles, such that eventually, the community as a whole can achieve a better level of understanding through communal discourse.

There was, however, no significant 'rise above' of ideas in Discourse B and it could be due to a few reasons. First, the initial deviation of discourse topic required a concerted effort from both the community and teaching staff to get the community back on-track, and it could have detracted the community from focusing on topics that they have interests in. From another perspective, the sentiment within Community B is often muted and less enthusiastic as compared to Community A, therefore pointing to a possibility that Community B might have little genuine interest in discussing the topics.

## 7. Conclusions and Future Directions

In summary, we proposed the usage of the Idea Pipeline framework to discover the dynamics of the idea development process, by tracking and monitoring ideas, through the phases of discovery, identification and analysis, and 'rise above' of ideas in a discourse. By using a comparison study of two knowledge building discourses, we illustrated how promising ideas can be identified and how these ideas improve and evolve in communities of different learning approaches and engagement. The comparison studies demonstrated how the Idea Pipeline and associated methods can be used, with different types of communities and discourse, to analyze textual data for visualization of idea diversity and identification

of promising ideas in knowledge building discourse using network analysis. The identification of promising ideas in discourse is critical to knowledge building, as communal interests need to be sustained for the process of continuous improvement of ideas, so that 'rise above' can occur to benefit the learning community. The Idea Pipeline framework and proposed methods provide a way to uncover this critical information to the community, through discourse analysis, network analysis and usage of temporal analytics. In our pilot study, by observing information such as the visualization of scarce keywords in the keyword network and the lackluster BC trend during mid-discourse, teacher interventions could be introduced to help Community B eventually achieve more 'rise above' by the end of the knowledge building discourse. Moving forward, we intend to study how a knowledge building community can use the Idea Pipeline framework and methods for near real-time monitoring of idea creation and development processes so that the participants could be guided towards a more productive knowledge-building discourse. We have also proposed the Idea Pipeline framework to schools and teachers with the aim of studying larger population to validate and enhance the Idea Pipeline framework. In addition, our future work include development of a systematic method to identify and characterize instances of 'rise above' of ideas in knowledge building discourse.

## Acknowledgements

The research reported here was supported by the Centre for Research and Development in Learning, Nanyang Technological University (CRADLE@NTU). We would also like to thank participating teachers and students in this study.

## References

Bailey, B. P., & Horvitz, E. (2010). What's your idea?: A case study of a grassroots innovation pipeline within a large software company. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2065-2074). ACM.

Bereiter, C. (2002). Education and mind in the knowledge society. Mahwah, NJ: Erlbaum.

Burchfield, R. W. (1996). *Modern English Usage*. Clarendon Press.

Chen, B., Chuy, M., Resendes, M., & Scardamalia, M. (2010). "Big Ideas Tool" as a new feature of knowledge forum. *In 2010 Knowledge Building Summer Institute (KBSI).* Toronto, Ontario, Canada.

Chen, B., Scardamalia, M., & Bereiter, C. (2015). Advancing knowledge-building discourse through judgments of promising ideas. *International Journal of Computer-Supported Collaborative Learning*, *10*(4), 345-366.

Hansen, M. T., & Birkinshaw, J. (2007). The innovation value chain. *Harvard business review*, *85*(6), 121.

Lee, A. V. Y., Tan, S. C., & Chee, K. J. K. (2016). Idea Identification and Analysis (I2A): A search for sustainable promising ideas within knowledge-building discourse. In *Proceedings of the 12th International Conference of the Learning Sciences* (pp. 90-97). Singapore: National Institute of Education.

Myers, S., & Marquis, D. G. (1969). *Successful industrial innovations: A study of factors underlying innovation in selected firms.* National Science Foundation NSF: Washington DC.

Reategui, E., Epstein, D., Lorenzatti, A., & Klemann, M. (2011). Sobek: A text mining tool for educational applications. In *International Conference on Data Mining*, 59-64.

Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning, 4*(3), 239-257.

Scardamalia, M. (2002) Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal education in a knowledge society* (pp.67-98). Chicago: Open Court.

Scardamalia, M. (2004). CSILE/Knowledge Forum. *Education and technology: An encyclopedia*, 183-192.

Scardamalia, M., & Bereiter, C. (2015). Knowledge building: Theory, pedagogy, and technology. In Sawyer, R. K. (Ed.), *The Cambridge Handbook of the Learning Sciences* (2nd Ed., pp. 397-417-118). NY: Cambridge University Press.

Stahl, G. (2004). Building collaborative knowing. In *What we know about CSCL* (pp. 53-85). Springer Netherlands.

Suthers, D. D., Dwyer, N., Medina, R., & Vatrapu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction. *International Journal of Computer-Supported Collaborative Learning, 5*(1), 5-42.