

# OnDiscuss: An Epistemic Network Analysis Learning Analytics Visualization Tool for Evaluating Asynchronous Online Discussions

Yanye Luther<sup>1</sup>[0000-0002-0995-3356], Marcia Moraes<sup>1</sup>[0000-0002-9652-3011],  
Sudipto Ghosh<sup>1</sup>[0000-0001-6000-9646], and James Folkestad<sup>2</sup>[0000-0003-0301-8364]

{yanye.luther, marcia.moraes,  
sudipto.ghosh, james.folkestad}@colostate.edu

Department of Computer Science<sup>1</sup>, School of Education<sup>2</sup>, Colorado State University,  
Fort Collins CO 80523, USA

**Abstract.** Asynchronous online discussions are common assignments in both hybrid and online courses to promote critical thinking and collaboration among students. However, the evaluation of these assignments can require considerable time and effort from instructors. We created OnDiscuss, a learning analytics visualization tool for instructors that utilizes text mining algorithms and Epistemic Network Analysis (ENA) to generate visualizations of student discussion data. Text mining is used to generate an initial codebook for the instructor as well as automatically code the data. This tool allows instructors to edit their codebook and then dynamically view the resulting ENA networks for the entire class and individual students. Through empirical investigation, we assess this tool's effectiveness to help instructors in analyzing asynchronous online discussion assignments.

**Keywords:** Epistemic Network Analysis · Learning Analytics · Asynchronous Online Discussion

## 1 Introduction

Social interactions play an important role in learning process. According to the social learning theory developed by Vygotsky [25], learning takes place primarily in social and cultural settings where students interact with their peers, teachers, and parents, as active participants in the creation of their own knowledge. Interpersonal interactions and discussions are a fundamental part of teaching and learning in this context [17, 6]. Therefore, asynchronous online discussions (AOD) are widespread in hybrid and online courses [3, 8]. Some of the benefits of AOD are deeper understanding of course material [7], communicating more effectively with group members [7], improvements in critical thinking and writing skills [3], and increasing student performance in meeting learning outcomes [4]. Despite of these benefits, instructors reported struggling to assess students' contributions in forum activities due to difficulties in following the discussions, the

lack of specific reports related to the subjects discussed, the students' contributions to those subjects, and the lack of visualizations to convey messages in a graphical format [13].

In this paper, we present OnDiscuss, a learning analytics (LA) visualization tool to support instructors in evaluating asynchronous online discussions. The tool utilizes both text mining and Epistemic Network Analysis (ENA). Latent Dirichlet Allocation (LDA) [5] is used to perform automatic topic extraction from the discussion data to create an initial codebook [20]. The use of LDA with the intervention and addition of the instructor's own keywords demonstrated significant potential to assist instructors in evaluating discussion based assignments [16]. However, in [16] the ENA visualizations were only demonstrated to a single instructor, requiring manual presentation. This current study investigates the potential for broader application by exploring how different instructors utilize a tool that enables them to modify their own codebook and observe immediate changes in ENA models.

This paper aims to answer the following research questions:

- **RQ1.** Are ENA visualizations of asynchronous online discussion data helpful for instructors who are novices in ENA?
- **RQ2.** Do the ENA visualizations reduce the time and effort spent assessing asynchronous online discussions for instructors of any experience level with ENA?

## 2 Related Works

ENA has been used in several works related with collaborative learning in asynchronous online discussions [19, 11, 23, 21]. Rolim et al. [19] used ENA to provide insights on the relationship between social and cognitive presence in asynchronous online discussions. Gašević et al. [11] proposed the use of social epistemic network signature (SENS) which combines ENA and Social Network Analysis (SNA) to analyze collaborative learning. Swiecki and Shaffer [23] extended SENS and proposed the integrated social-epistemic network signature (iSENS), an approach that provides the simultaneous investigation of cognitive and social connections in collaborative learning. Scianna and Kaliisa [21] proposed an analysis workflow and visualization method called Social Sentiment Embedded Epistemic Networks (SSEEN) to consider how sentiment manifests and explore its usefulness in understanding student interactions in asynchronous online discussions.

Besides its use in collaborative learning analysis, ENA has been used to visualize many aspects of learning and education [12, 9]. Fougård analyzed instructors' ability to assess student papers and considered a different range of number of topics in order to capture different levels of complexity of the ENA models. Visually ENA had the potential to indicate the quality of the student assignment. They reported it can be a struggle to choose the correct number of codes and keywords and that it should be left to the instructor to make that informed decision [10]. Vega introduces participatory quantitative ethnography (QE) which

includes participants in co-construction and co-interpretation of ENA models. This work demonstrates the deeper analysis that arises from interpretations of data by modifying codes, adding connections, and reacting to codes [24].

Unlike previous works that used ENA in collaborative learning analysis, this work combines LDA and ENA in a tool that: builds an initial codebook for instructor, automatically codes the data of asynchronous online discussions using that initial codebook, and generates individual ENA visualizations of student's content connections. In addition, past works do not have the ability to modify the codebook and thus instantly update the ENA model on their own without any researcher intervention.

### 3 Methodology

#### 3.1 Participant Selection

In order to answer the research questions, the researchers selected one instructor that did not have previous experience with ENA and one instructor with previous experience with ENA to evaluate their own asynchronous online discussion with the aid of OnDiscuss. Both instructors are from the same Research 1 land-grant university. The instructor without experience in ENA will be referred as instructor A, the novice, and the instructor with experience will be referred as instructor B, the expert. Instructor A taught a graduate level Computer Science course with discussion discourse amongst 5-6 students and instructor B taught a graduate level Education course with discourse amongst about 20 students.

The courses that were chosen for the instructors to analyze had to meet the following criteria. The course must have asynchronous online discussion assignments. Those assignments must require students to assimilate knowledge and synthesize concepts into a coherent discussion post rather. Additionally, these discussions needed to be entirely raw text since OnDiscuss can't parse potential keywords in images, videos, links to websites, etc.

#### 3.2 OnDiscuss Description and Functionality

OnDiscuss is a learning analytics visualization tool for instructors that utilizes text mining algorithms and Epistemic Network Analysis (ENA) to generate visualizations of student discussion data. As highlighted by reviewers in [15], the process of establishing a codebook from scratch may present challenges for instructors new to ENA. Our tool uses the same process described in [16] to generate five topics with ten keywords each derived from one discussion data to be used as an initial codebook for the instructor. The rationale about the initial number of topics and keywords can be found in [20]. After the initial codebook is built, the tool uses text mining algorithms to automatically code the discussion for each post. That coded data is then represented in a resulting ENA network. The tool uses rENA [2] package and the following settings to build the ENA model: unit is the student ID, the conversation is the utterance from the student in their discussion post, and an infinite stanza window. That configuration

enabled us to generate for each student their connections among the codes that they were supposed to be discussing.

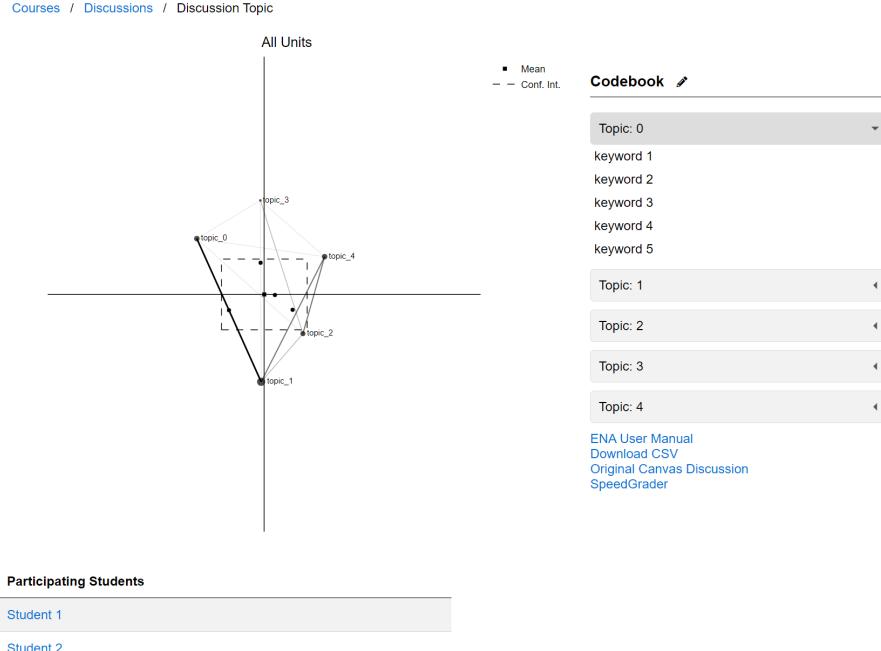


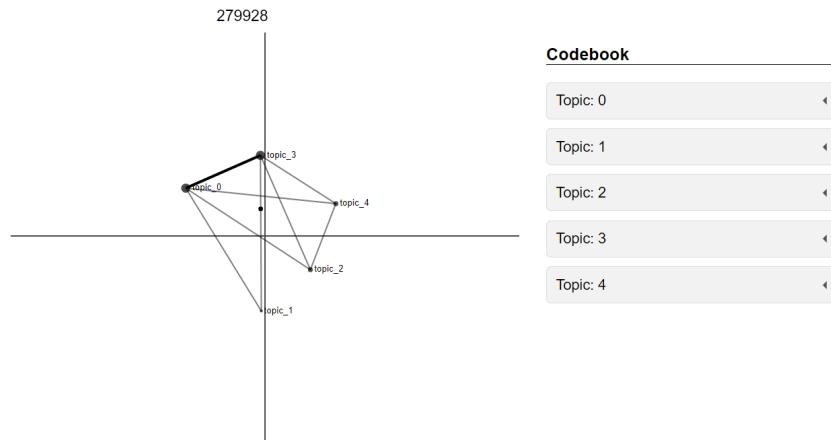
Fig. 1: Example Class/Group ENA View for a Discussion Topic

The tool is integrated with Canvas Learning Management System (LMS) which allowed instructors to view a list of all discussions published to Canvas within a course. Clicking on a discussion would display the group ENA model and the associated codebook as shown in Figure 1. Instructors can edit their codebook by adding, removing, and editing keywords. They cannot edit number of topics since [20] found 5 to be the optimal number of topics for grouping keywords. However, they can edit the names of the topics to be more meaningful since LDA will just assign the topics as numbers 0-4. Once the edits are completed, the ENA models will automatically be updated to account for the codebook changes. Instructors are also able to view individual students' networks as well as the discussion posts that contributed to the network as shown in Figure 2.

Also shown in Figure 1, OnDiscuss provided a link to the discussion in Canvas as well as a link to the SpeedGrader for easy access to grade and view the original discussion from Canvas itself. There was also a link to the "ENA User Manual" which simply provided reinforcement of what ENA is and how to interpret an

### Student 2

Courses / Discussions / Discussion Topic / Student 2



### Submission

Applying input space partitioning to test object oriented software can be a valuable approach but it also comes with its challenges. Here are some challenges and potential solutions based on practical experience and insights Complex Object Structures Object oriented software often involves complex and deeply nested object structures which can lead to a large and intricate input space. It is challenging to generate comprehensive test cases that cover all possible interactions. Solution Utilize modeling techniques such as UML diagrams to visualize and understand the object relationships. This can help identify critical paths and objects for testing.

Fig. 2: Example Individual ENA View for a Discussion Topic

example network. A link to download the comma separated value (CSV) file was provided in the correct format for the ENA Web Tool [1].

### 3.3 Experimental Procedure

The experiment was conducted as an empirical observational study [18], where the researchers observed the instructor's interaction with OnDiscuss during semi-structured interview [14] sessions that last around one hour. Researchers recorded the interview sessions as well as took notes about instructor's interactions and questions. In response to any queries raised by the instructor during the session, the researchers offered assistance as needed. Because discussion topics were pulled from past semesters of the course, prior to the interview session the instructors were asked to review the selected discussion. This was done to have the instructors reacquaint themselves with the discussion topic and the students' postings and discourse of the prompt.

To begin the session, the researchers delivered a presentation that went over the basics of what ENA is, a case study illustrating ENA interpretation, and the impact of discussion data and codebook modifications on visualizations. The

presentation started by explaining that ENA identifies the co-occurrences in segments of discourse data and modeling the weighted structure of co-occurrences as a dynamic network model [22]. The researchers then presented a case study on one discussion topic from a single semester of a course. An initial codebook was generated using LDA which was used to create the group class and individual student models where instructors learned the nodes represented topics, the thickness of the lines represented the strength of connections between topics, and the other points in the group network are the centroids of the individual students' networks. The presentation gave examples of student posts and how keywords co-occurred in the discourse data to create the ENA model. Because this case study was performed on a single discussion topic from a single semester, additional networks of the same discussion topic from a different semester were shown to demonstrate the impact of the discussion data using the same codebook. Finally, the researchers made their own edits to the initial codebook and demonstrated how such modifications could influence the resulting networks. This presentation was delivered to both instructors, irrespective of their familiarity with ENA.

After the presentation the instructors navigated to the chosen discussion assignment in OnDiscuss. The initial codebook was completely generated by LDA from previous semesters of the same discussion topic [20]. The instructors were allowed to make as many edits and iterations of the codebook while inspecting the models produced by each iteration. Once the instructor was satisfied with the codebook and/or the group and individual ENA models, a closing semi-structured interview was conducted by the researchers.

## 4 Results

### 4.1 Instructor A

The novice instructor with ENA asked many questions about the basics of ENA during the presentation. For example, *What is a codebook?*, *How are the topics positioned in 2D space?*, *What should I be expecting to understand from these visualizations?*. After responding to the instructor's questions and completing the presentation, the instructor navigated to their chosen discussion on OnDiscuss. At this point, they had access to the initial codebook created from running LDA on all the previous semesters of the same discussion topic (Table 1) and the resulting group network (Figure 4a).

The instructor started their interaction with the tool by removing all the initial LDA generated keywords from each topic and added their own keywords. They said that half of all the automatically generated keywords were useful individually but the groupings of the keywords into topics was not helpful. Consequently, they removed all keywords and added those they thought would be present. The first iteration of their codebook didn't have keywords that occurred in the discourse data so the network only consisted of a single line.

Researchers intervened and suggested that this could've been due to multiple reasons such as not including the stems of keywords, not choosing enough

Table 1: Instructor A Initial Codebook

Topic	Keywords
0	devic, interfac, child, applic, potenti, post, input_paramet, parent, behavior, run
1	write, want, team, choic, custom, field, look, product, interfac, array
2	boundari, import, select, api, handl, rest_api, rest, encapsul, sure, partit_test
3	partit_method, categori_partit, leak, determin, memori_leak, lan- guag, applic, system, databas, partit_test
4	subclass, tester, abstract, output, group, model, oop, overlap, detect_memori, disjoint

keywords to represent a topic, and choosing keywords not present in the discussion. Taking this advice, the instructor included the proper stems, added more keywords to existing topics, and completely reworked a topic that only had 2 keywords to a new topic with 12 keywords. The codebook that they were most pleased with is shown in Table 2 and the resulting group network is shown in Figure 4b.

Table 2: Instructor A Best Codebook

Code	Keywords
Observability	observability, visible, get, state, visibility, observable, getter, access, field, accessor
Controllability	configure, object, control, modify, state, mutate, mutator, update
inheritance	inheritance, child class, parent class, overriding, depth, sub-class, superclass, sub class, super class, inherit, interface, abstract class
testing	black box, black-box, white-box, white box, automate, automation, industry, difficult, easy
object oriented programming	public, private, protected, package, simple, complex, abstraction, specialization, data, encapsulation, method, field

Analyzing the networks from the codebook shown in Table 2, they stated that the connection strength between topics in the group ENA shown in Figure 4b did indeed reinforce what they had gathered from rereading the discussion. Also the individual ENA graphs were representative of both the "strong" students and "weak" students contributions. When analyzing a strong student's network

shown in Figure 3a, the instructor was shocked to not see any connections to testing but stated *This student made good strong connections to controllability and observability*. When analyzing a weak student's network shown in Figure 3b, the instructor stated *I was expecting weak connections to controllability and observability because they were initially answering a different question from the prompt*.

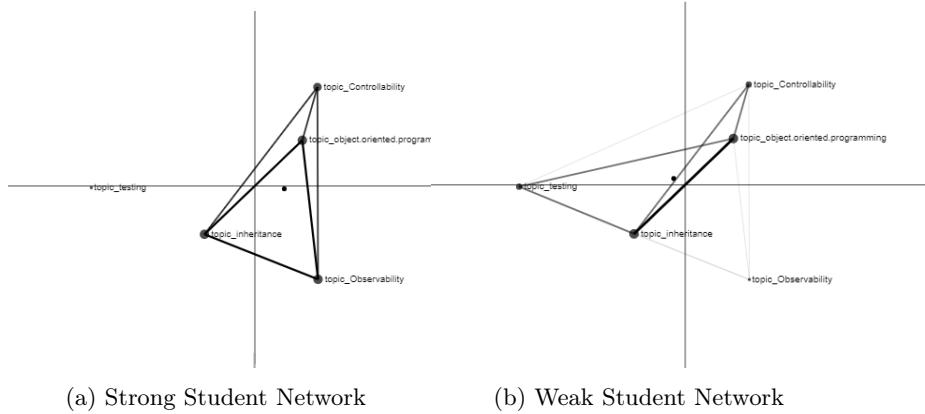


Fig. 3: Instructor A Individual ENA Visualizations

When asked if having these visualizations would impact how they reframe future discussion prompts, instructor A stated they wouldn't change the prompt but rather better prepare the students in class for the discussion. Instructor A explained that they'd display the group network to the entire class to discuss the topics without connections. The instructor saw this as a helpful tool to then reinforce concepts that were missed without needing to read the entire discussion thus saving time and energy. They also noted that these ENA models would not only be helpful to the instructor but also to the students.

#### 4.2 Instructor B

During the presentation, the expert instructor with ENA asked questions such as *What is LDA?*, *What stanza is being used?*, *How are the student points being placed?*. When instructor B began exploring OnDiscuss they were pleased with the initial codebook provided shown in Table 3. This initial codebook was created with more discussion data from various semesters so LDA was able to perform better topic modeling with more data. Since this codebook was used in previous works the topics had meaningful names instead of the LDA assigned numbers and it had some added keywords from the same instructor [16].

This instructor made very minimal edits to the initial codebook. They only removed "fall" and "desire difficulty" from the "effortful learning" topic. The

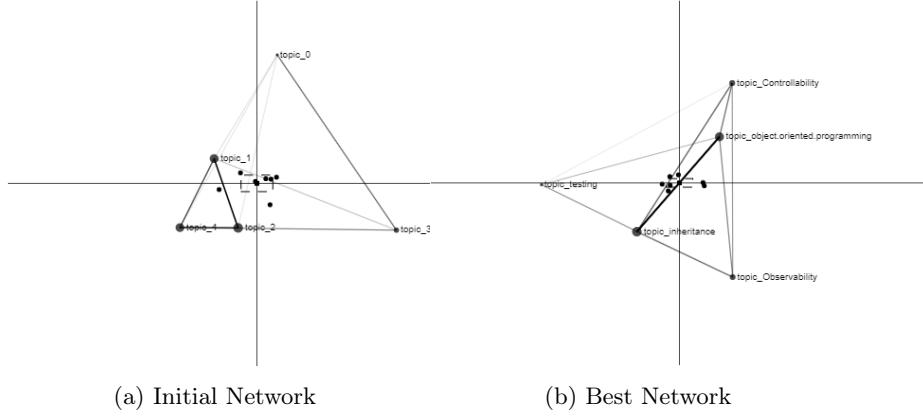


Fig. 4: Instructor A Group ENA Visualizations

network model changed as the instructor expected; however, they didn't think much could be concluded from the change between Figure 5a and Figure 5b. Figure 5b had a new connection between "beyond learning styles" and "illusion of mastery" that did not appear in Figure 5a. Comparing networks between codebook edits wasn't meaningful to Instructor B. Instead they wanted a baseline model to compare the resulting network to in order to make more informed edits to the codebook. They suggested creating a baseline model from a textbook or an asynchronous online discussion amongst the instructors responding to the same discussion prompt as the students.

While examining individual student interactions, the instructor expressed the need for two separate networks: one focusing solely on each student's initial post and another incorporating both their original post and subsequent replies. This would provide more isolated insight into their contribution and engagement with other students as well as their initial contribution. Additionally, the instructor suggested that a clearer indication of occurring keywords would be helpful when reading the student's discussion posts in the tool rather than just the plain text.

## 5 Discussion

### 5.1 RQ1. Are ENA visualizations of asynchronous online discussion data helpful for instructors who are novices in ENA?

The novice instructor started with no understanding of ENA. Throughout the presentation and the time exploring the tool, this instructor was unsure what made more a "good" codebook and ENA network. Since this instructor was in the field of computer science, they seemed more inclined to understand the math behind the ENA visualizations. Future works should be done with novice instructors in other fields to determine if they also have the same inclination.

Table 3: Instructor B Initial Codebook

Topic	Keywords
effortful learning	desire, plf, resonate, parachute, land, jump, commun, parachute land, land fall, difficult, difficulties, mistakes, failure, effortful learning, desirable difficulty, desirable, effortful
beyond learning styles	dylexia, learn style, individual, learn differ, disable, intelligent, prefer, support, dyslex, focus, instructional style, learning styles
illusion of mastery	confidence, feedback, calibration, confidence memory, accuracy, peer, answer, event, state, calibration learn, illusion of mastery, illusions of mastery, misunderstanding, illusion of knowing, illusions of knowing, illusion of learning, illusions of learning, re read, cram
retrieval practice spaced out practice interleaving	mass, mass practice, interleaving practice, space retrieval, tend, day, long term, week, myth, practice space, retrieval practice, retrieval process, testing effect, test effect, recall knowledge, retrieval, actively retrieving, periodically testing, retrieval activity, retrieval activities, low stakes, effective retrieval must be repeated, flash cards, quizzing, practice and retrieval, quiz over time, continually retrieve the information, frequently quizzing, retrieval practice activity, retrieval practice activities, testing efforts, active retrieval, practice, testing for its benefit in the learning process, short quiz, active recall, process of retrieval, practice sessions, self testing, recall the information, RPA, RPAs, spacing out, spacing out practice, spaced practice, spacing practice, spaced out practice, spaced out, spaced retrieval, space retrieval, space practice, retrieval spaced, retrieve spaced, spaced application, spaced knowledge, space knowledge, spaced retrieval, retrieval practice is spaced, interleaving, interleaved practice, interleave, interleaved

This could inform future revisions to the presentation to emphasize and include more of the mathematical background behind ENA.

Despite this once the instructor created a codebook with keywords they were pleased with, they were able to successfully draw conclusions from the networks. The placement of the points in the ENA network were still confusing to this instructor; however, they were able to extract lots of meaning from the thickness of the edges of the both the individual and group networks. They were able to make interpretations such as *The entire class isn't making many connections to these topics so I should emphasize them more in lectures and I thought this student made a strong contribution to the discussion and it's nice to see that visually they made lots of strong connections to all of the topics.* Instructor A

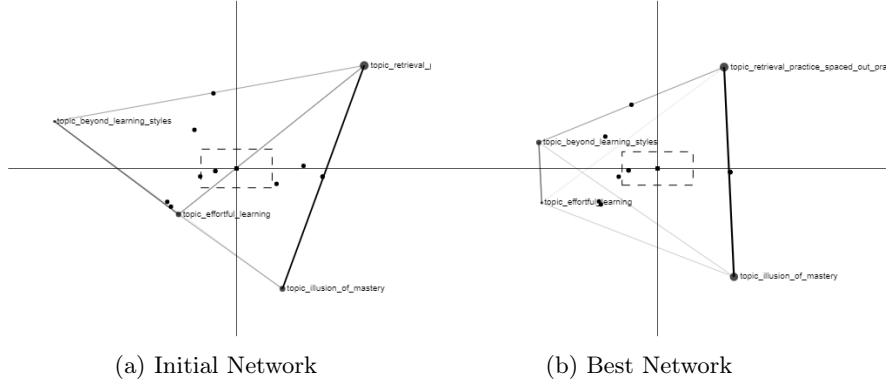


Fig. 5: Instructor B Group ENA Visualizations

stated once they understood how to create a better codebook they'd feel even more familiar and inclined to use ENA.

The ENA presentation likely played a vital role in the novice instructor's basic understanding of ENA which manifested into their own ENA interpretations. We hypothesize that additional exposure to modifying codebooks and interpreting ENA networks from other discussions would be beneficial for their continued development. However, through just a single session this instructor demonstrated promising proficiency in ENA, indicating the potential for other instructors unfamiliar with ENA.

## 5.2 RQ2. Do the ENA visualizations reduce the time and effort spent assessing asynchronous online discussions for instructors of any experience level with ENA?

Both Instructor A and B reported that having these ENA networks and a modifiable codebook was a helpful supplemental tool for evaluating their asynchronous online discussions. Both said in the semi-structured interview that these supplemental networks would make grading and analyzing the discussions faster. This tool is not meant to fully replace reading through the students' discussion posts but is instead a supplemental tool. The instructors stated it was faster for them to assess the discussions with the visual aid provided by the ENA networks. The instructors also stated by utilizing the individual networks they'd be able to provide more personalized feedback to individual students.

Instructor A found the group and individual networks helpful standalone while Instructor B thought they were somewhat useful and in order to be even more useful they needed more context such as a baseline network. Despite their differing perspectives and experiences with ENA, both instructors were able to draw their own conclusions from the networks based on their own codebooks. This diversity in interpretations highlights the power and flexibility of ENA and

the tool, as instructors not only have the ability to handcraft their codebooks but also to derive nuanced insights tailored to their specific teaching contexts.

### 5.3 Limitations

The ENA networks only represent the raw text data within the discussions. Consequently, this may overlook other forms of media, such as images, videos, and links to external websites, which could provide valuable insights or perspectives that are not captured through text alone.

Another potential limitation concerns the study population. The instructors were purposefully chosen by the researchers to participate in the research project. While this approach was necessary to ensure access to relevant discussion data and expertise with ENA, it introduced a selection bias. Also although the participants in our study are drawn from the same university, it is important to note that they exclusively represent graduate level courses with relatively small class sizes.

Instructor A had considerably less available discussion data than Instructor B to create the codebook. Instructor A had 363 posts with 37,254 total words while Instructor B had 2,648 posts with 444,364 total words. Because Instructor B had a lot more data, LDA was able to provide more accurate topic modeling [5]. Moreover, Instructor B's dataset was used in prior works so the codebook could potentially be more refined than Instructor A's.

## 6 Conclusion and Future Works

In this study, we explored the potential of utilizing a tool that enables instructors to edit their own codebook and visualize the resulting ENA visualizations in real time. By providing instructors with the ability to directly interact and customize, this tool opens up new opportunities to popularize ENA and make it more accessible to educators who may or may not be familiar with the intricacies of ENA.

Our findings highlight several key insights regarding the implications of this tool for enhancing the accessibility and usability of ENA as a learning analytics visualization tool. The tool is helpful to those unfamiliar with ENA and future refinements like the addition of a baseline ENA model can make it more helpful to those familiar with ENA. Despite the tool's automated keyword generation capabilities, it is clear that instructor intervention remains crucial for refining the codebook. Therefore, while automated techniques like Latent Dirichlet Allocation (LDA) provide valuable insights given a large amount of data, they must be complemented by expert guidance.

Many potential future works could stem from this study in terms of both adding new features to the tool and exploring different populations. Adding a baseline ENA model from either the textbook or discussion posts of the instructor could have potential to assist the instructor even further by allowing them to make comparisons between the class and base model. Further research should

also explore applying OnDiscuss to different subjects other than Computer Science and Education, to undergraduate courses, and to even larger class sizes since it would be even more time consuming to assess the discussion assignments.

## References

1. Ena web tool (2024), <https://app.epistemicnetwork.org/login.html>
2. rena package (2024), <https://cran.r-project.org/web/packages/rENA/index.html>
3. Aloni, M., Harrington, C.: Research based practices for improving the effectiveness of asynchronous online discussion boards. *Scholarship of Teaching and Learning in Psychology* **4**(4), 271 (2018)
4. Alzahrani, M.G.: The effect of using online discussion forums on students' learning. *Turkish Online Journal of Educational Technology-TOJET* **16**(1), 164–176 (2017)
5. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. *Journal of machine Learning research* **3**(Jan), 993–1022 (2003)
6. Davis, M.L., Witcraft, S.M., Baird, S.O., Smits, J.A.: Learning principles in cbt. In: *The science of cognitive behavioral therapy*, pp. 51–76. Elsevier (2017)
7. Decker, J., Beltran, V.: Graduate students' perceptions of the benefits and drawbacks of online discussion tools. *International Journal of Online Pedagogy and Course Design (IJOPCD)* **6**(1), 1–12 (2016)
8. Fehrman, S., Watson, S.L.: A systematic review of asynchronous online discussions in online higher education. *American Journal of Distance Education* **35**(3), 200–213 (2021)
9. Fernandez-Nieto, G.M., Martinez-Maldonado, R., Kitto, K., Buckingham Shum, S.: Modelling spatial behaviours in clinical team simulations using epistemic network analysis: Methodology and teacher evaluation. In: *LAK21: 11th International Learning Analytics and Knowledge Conference*. p. 386–396. LAK21, Association for Computing Machinery, New York, NY, USA (2021). <https://doi.org/10.1145/3448139.3448176>
10. Foug, S.S., Siebert-Evstone, A., Eagan, B., Tabatabai, S., Misfeldt, M.: Epistemic network analysis of students' longer written assignments as formative/summative evaluation. p. 126–130. LAK '18, Association for Computing Machinery, New York, NY, USA (2018). <https://doi.org/10.1145/3170358.3170414>
11. Gašević, D., Joksimović, S., Eagan, B.R., Shaffer, D.W.: Sens: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior* **92**, 562–577 (2019)
12. Herder, T., Siewecki, Z., Foug, S.S., Tamborg, A.L., Allsopp, B.B., Shaffer, D.W., Misfeldt, M.: Supporting teachers' intervention in students' virtual collaboration using a network based model. In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. p. 21–25. LAK '18, Association for Computing Machinery, New York, NY, USA (2018). <https://doi.org/10.1145/3170358.3170394>
13. de Lima, D.P., Gerosa, M.A., Conte, T.U., de M. Netto, J.F.: What to expect, and how to improve online discussion forums: the instructors' perspective. *Journal of Internet Services and Applications* **10**, 1–15 (2019)

14. Magaldi, D., Berler, M.: Semi-structured interviews. Encyclopedia of personality and individual differences pp. 4825–4830 (2020)
15. Moraes, M., Folkestad, J., McKenna, K.: Using epistemic network analysis to help instructors evaluate asynchronous online discussions. In: Second International Conference on Quantitative Ethnography: Conference Proceedings Supplement. pp. 19–22 (2021)
16. Moraes, M., Ghaffari, S., Luther, Y., Folkesdtad, J.: Combining automatic coding and instructor input to generate ena visualizations for asynchronous online discussion. In: Arastoopour Irgens, G., Knight, S. (eds.) *Advances in Quantitative Ethnography*. pp. 381–394. Springer Nature Switzerland, Cham (2023)
17. Prawat, R.S.: Teachers' beliefs about teaching and learning: A constructivist perspective. *American journal of education* **100**(3), 354–395 (1992)
18. Pyrczak, F., M., D.: The empirical approach to knowledge. In: *Making Sense of Statistics*, pp. 1–5. Routledge (2018)
19. Rolim, V., Ferreira, R., Lins, R.D., Găsević, D.: A network-based analytic approach to uncovering the relationship between social and cognitive presences in communities of inquiry. *The Internet and Higher Education* **42**, 53–65 (2019)
20. Saravani, S.M., Ghaffari, S., Luther, Y., Folkestad, J., Moraes, M.: Automated code extraction from discussion board text dataset. In: *Advances in Quantitative Ethnography: 4th International Conference, ICQE 2022, Copenhagen, Denmark, October 15–19, 2022, Proceedings*. pp. 227–238. Springer (2023)
21. Scianna, J., Kaliisa, R.: 'sseen': a networked approach to uncover connections between sentiment, social, and epistemic elements of student online forum discourse. *Educational technology research and development* pp. 1–23 (2023)
22. Shaffer, D.: *Quantitative Ethnography*. Cathcart Press, Madison, WI (2017)
23. Swiecki, Z., Shaffer, D.W.: isens: an integrated approach to combining epistemic and social network analyses. In: *Proceedings of the tenth international conference on learning analytics & knowledge*. pp. 305–313 (2020)
24. Vega, H., Irgens, G.A.: Constructing interpretations with participants through epistemic network analysis: Towards participatory approaches in quantitative ethnography. In: Wasson, B., Zörgő, S. (eds.) *Advances in Quantitative Ethnography*. pp. 3–16. Springer International Publishing, Cham (2022)
25. Vygotsky, L.S.: *Mind in society*. Harvard University Press, Cambridge, MA (1978)