COVER SHEET FOR PROPOSAL TO THE NATIONAL SCIENCE FOUNDATION

PROGRAM ANNOUNCEMENT/SOLICITATION NO./DUE DATE
NSF 18-1

☐ Special Exception to Deadline Date Policy

FOR NSF USE ONLY

NSF PROPOSAL NUMBER

FOR CONSIDERATION BY NSF ORGANIZATION UNIT(S) (Indicate the most specific unit known, i.e. program, division, etc.)
DRL - Default

DATE RECEIVED	NUMBER OF COPIES	DIVISION ASSIGNED	FUND CODE	DUNS# (Data Universal Numbering System)	FILE LOCATION
				006046700	

EMPLOYER IDENTIFICATION NUMBER (EIN) OR TAXPAYER IDENTIFICATION NUMBER (TIN)
356001673

SHOW PREVIOUS AWARD NO. IF THIS IS
☐ A RENEWAL
☐ AN ACCOMPLISHMENT-BASED RENEWAL

IS THIS PROPOSAL BEING SUBMITTED TO ANOTHER FEDERAL AGENCY? YES ☐ NO ☒ IF YES, LIST ACRONYM(S)

NAME OF ORGANIZATION TO WHICH AWARD SHOULD BE MADE
Indiana University

AWARDEE ORGANIZATION CODE (IF KNOWN)
0018093000

ADDRESS OF AWARDEE ORGANIZATION, INCLUDING 9 DIGIT ZIP CODE
Indiana University
509 E 3rd St
Bloomington, IN. 474013654

NAME OF PRIMARY PLACE OF PERF
Indiana University

ADDRESS OF PRIMARY PLACE OF PERF, INCLUDING 9 DIGIT ZIP CODE
Indiana University
509 E 3RD ST
Bloomington ,IN ,474013654 ,US.

IS AWARDEE ORGANIZATION (Check All That Apply) ☐ SMALL BUSINESS ☐ MINORITY BUSINESS ☐ FOR-PROFIT ORGANIZATION ☐ WOMAN-OWNED BUSINESS

☐ IF THIS IS A PRELIMINARY PROPOSAL THEN CHECK HERE

TITLE OF PROPOSED PROJECT **Net.Create: Using Network Analysis to Support Digital Humanities in Large History Classrooms**

REQUESTED AMOUNT	PROPOSED DURATION (1-60 MONTHS)	REQUESTED STARTING DATE	SHOW RELATED PRELIMINARY PROPOSAL NO. IF APPLICABLE
$ **299,661**	**24** months	**08/15/18**	

THIS PROPOSAL INCLUDES ANY OF THE ITEMS LISTED BELOW
☐ BEGINNING INVESTIGATOR
☐ DISCLOSURE OF LOBBYING ACTIVITIES
☐ PROPRIETARY & PRIVILEGED INFORMATION
☐ HISTORIC PLACES
☐ VERTEBRATE ANIMALS IACUC App. Date _____
PHS Animal Welfare Assurance Number _____
☒ TYPE OF PROPOSAL **EAGER**

☐ HUMAN SUBJECTS Human Subjects Assurance Number _____
Exemption Subsection _____ or IRB App. Date _____
☒ INTERNATIONAL ACTIVITIES: COUNTRY/COUNTRIES INVOLVED **CI GM NL FR**
☒ COLLABORATIVE STATUS **Not a collaborative proposal**

PI/PD DEPARTMENT
History

PI/PD FAX NUMBER
812-855-3378

PI/PD POSTAL ADDRESS
509 E 3RD ST
Bloomington, IN 474013654
United States

NAMES (TYPED)	High Degree	Yr of Degree	Telephone Number	Email Address
PI/PD NAME **Kalani Craig**	**PhD**	**2013**	**812-855-7581**	**craigkl@indiana.edu**
CO-PI/PD **Joshua A Danish**	**DPhil**	**2009**	**812-856-8330**	**jdanish@indiana.edu**
CO-PI/PD **Cindy E Hmelo-Silver**	**PhD**	**1994**	**812-856-2363**	**chmelosi@indiana.edu**
CO-PI/PD **Ben T Loh**	**PhD**	**2003**	**773-220-8889**	**ben@inquirium.net**
CO-PI/PD				

PROJECT SUMMARY

Overview:

This Net.Create: Using Network Analysis to Support Digital Humanities Learning in Large History Classrooms is an innovative approach to network analysis that brings simultaneous multi-user predictive network data entry and live visualization into university history classrooms to support pedagogy and learning. Net.Create will aim to bridge gaps in technical expertise, visualization integration and the use of network analysis in history classrooms by providing a scaffolded, easy-to use interface and curricular materials that are designed from the ground up to be successfully implemented in a lecture classroom. The Net.Create platform explicitly supports collaborative generation of data as a way of helping students to see the challenges of interpreting evidence from a historical text and transforming it into a digital format, as well as the value that network analysis as a digital format has in understanding the complexity of historical interactions. We will then use Net.Create to systematically test the potential of network analysis as a new genre for learning in the humanities and to better understand how complex network visualization tools mediate learning and engagement in undergraduate humanities classrooms. We will explicitly contrast students' learning processes and learning outcomes when they either create and refine networks, or work with prebuilt networks, allowing us to develop theory around how network analysis can be most effectively harnessed in both large university history classrooms and smaller, advanced history seminars. Our findings will generate theoretical frameworks for using technology to support student engagement and the development of historical-thinking practices. Our findings will also provide practical guidance for how instructors in a variety of humanities disciplines can adapt valuable network visualization and analysis skills, typically the focus of STEM disciplines, to support students as they engage with humanities learning.

Intellectual Merit:

Network analysis is an increasingly popular and powerful computational tool for the analysis of large data sets. Digital historians have used these tools to represent and analyze historical contexts because they support scholars in looking at a broad range of connections between people, places, and events. While humanities pedagogues are optimistic that these affordances also provide unique opportunities for supporting students in learning history, available tools do not support easy integration into humanities classrooms, and there is not yet empirical or theoretical support for how this might be accomplished effectively. As part of this integration, Net.Create aims to support novice history learners in recognizing how historical practices are grounded in argumentation rather than in single authoritative accounts by scaffolding students in creating and refining visualizations of historical corpora, allowing them to see these rich contexts, and then challenging them to develop and defend historical argument using these visualizations. Analysis will explore how the Net.Create tool and curricular activities contribute to students' historical practices and to student understanding of network analysis approaches, and whether these network-analysis practices result in new historical learning, retention, and understanding.

Broader Impacts:

There is a demonstrated need for a low-tech-barrier tool and accompanying activities that can be used to support network analysis in a variety of environments. Social network analysis in particular is increasingly useful in 21st-century workplaces that rely on employees who can effectively connect computational and social practices, as well as understand the statistical data that underpins visualizations. The design of Net.Create is intended to support history students in a wide range of contexts in engaging in challenging historical practices, and in appreciating how these practices are grounded in constructing arguments using evidence from the past. These approaches should then generalize beyond the 200 students of this initial pilot to students across the country via our freely available tools, and across disciplines through our proposed easy customization process. Net.Create and the associated curriculum also provide a unique opportunity to teach these increasingly valuable techniques, and for helping students to use humanities classrooms to learn cutting-edge digital approaches to scholarship which can apply across disciplinary boundaries.

TABLE OF CONTENTS

For font size and page formatting specifications, see PAPPG section II.B.2.

	Total No. of Pages	Page No.* (Optional)*
Cover Sheet for Proposal to the National Science Foundation		
Project Summary (not to exceed 1 page)	1	
Table of Contents	1	
Project Description (Including Results from Prior NSF Support) (not to exceed 15 pages) **(Exceed only if allowed by a specific program announcement/solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)**	8	
References Cited	2	
Biographical Sketches (Not to exceed 2 pages each)	7	
Budget (Plus up to 3 pages of budget justification)	10	
Current and Pending Support	7	
Facilities, Equipment and Other Resources	2	
Special Information/Supplementary Documents (Data Management Plan, Mentoring Plan and Other Supplementary Documents)	1	
Appendix (List below.) **(Include only if allowed by a specific program announcement/ solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)**		
Appendix Items:		

*Proposers may select any numbering mechanism for the proposal. The entire proposal however, must be paginated. Complete both columns only if the proposal is numbered consecutively.

Net.Create: Using Network Analysis to Support Digital Humanities Learning in Large History Classrooms

1 Project Description

1.1 Introduction

Net.Create: Using Network Analysis to Support Digital Humanities Learning in Large History Classrooms is being submitted as an NSF EAGER proposal. Its innovative approach to network analysis brings simultaneous multi-user predictive network data entry and live visualization into university history classrooms to support pedagogy and learning. These network tools scaffold students in creating and refining network data and visualizations that compartmentalize and then re-aggregate the details, significance and interactions of people in historical events, allowing students to not just see the rich contexts in which historical interactions are placed, but to build that context from scratch.

Network analysis is increasingly seen as a powerful computational tool for making sense of big data sets that seek to describe the large network of interdependencies created by the overlap of many individual relationships between pairs of entities. Humanities scholars have used networks to identify broad-reaching patterns that are otherwise challenging to see in large corpora of data which can span both time and space (Carrington, Scott, & Wasserman, 2005; Scott, 2013; Wasserman & Faust, 1994; Wasserman & Galaskiewicz, 1994). When used in classrooms, these digital-humanities experiences explicitly prepare students for a twenty-first century workplace in which these computational tools and analytical techniques will play a key role. Teaching history students using network analysis as a tool for understanding history requires interpretation of data as students transform unstructured data into structured data according to the disciplinary norms of history. Specifically, novice history learners struggle to cope with the overwhelming volume of "facts" of history, which are more easily remembered and, more importantly, more clearly understood when they are seen not as facts but as interpretive detail in the context of a historical argument about the cause and meaning of historical interactions between people, places, and things. The interpretive data gathering process necessary for humanities network analysis, which Johana Drucker calls *capta* (2011), is relevant for what many scholars consider the interpretive nature of "objective" datasets common to data science (boyd & Crawford, 2012). It also emphasizes the interpretive nature of what novice historians generally consider "objective" historical facts. Network analysis approaches also take advantage of students' experience with social media in making patterns in historical relationships more visible, meaningful, and memorable for students as they build comprehension of historical events and then use that comprehension to develop historical argumentation (Craig, 2017b; Drucker, 2013; Gochenour, 2011).

For students to benefit in these ways from generating data networks in a typical college classroom, they need to be able to work concurrently on a single large network. Small contributions from each student are feasible, but then those small contributions are aggregated into a larger network context, providing a more complete historical framework for each students' individual contributions. However, there are no existing network analysis tools that allow simultaneous multi-user network data entry, and thus no empirical studies of network analysis in humanities classrooms. Given the potential of network analysis tools to support students in simultaneously learning valuable computational tools and historical concepts, there is a need to explore network analysis in history classrooms in a systematic manner to provide more robust theoretical and design guidelines for how this interdisciplinary genre of learning might be supported.

Our proposal aims to develop an innovative network analysis and visualization tool intended specifically to integrate predictive text and network analysis of historical events to support pedagogy within a wide-range of historical contexts. We will then use this new tool to systematically test the potential of network analysis as a new genre for learning in the humanities and to better understand how complex network visualization tools mediate learning and engagement in undergraduate humanities classrooms. Our planned data analysis will provide insights into how students in humanities classrooms engage with and learn these new computational tools, and how they learn the history content as a consequence. Our findings will enable us to develop theory around how these network visualizations can

be most effectively used in both large university history classrooms and smaller, advanced history seminars; provide practical guidance for how history instructors across contexts and specializations might instruct their students in the use of new computational tools; incorporate the Net.Create tool with minimal need for professional development or advanced technical expertise; and to develop theoretical frameworks for using technology to support student engagement and historical-thinking practices.

1.2 Background & Related Research

Network analysis tools are fundamentally intended to support users in making sense of complex data. In the humanities, students are expected to consume texts comprised of complex interwoven details, which requires reading comprehension, attention to detail and use of that comprehension, and detail recall to craft an analysis of historical trends drawn from the data in that text (Theibault, 2013). While there are famous exceptions like Charles Minard's map of German troops marching toward Russia in World War II (Minard, 1869), historians depend on oral or written forms of argument, rarely generating visualizations or thinking about representations as a mechanism for analysis beyond maps and timelines that convey basic temporal and geographic information (Prangsma, 2007). As students replicate these disciplinary practices in history classrooms, they are asked to read primary and secondary sources from the historical time period being studied (Carr 2008; Wineburg 2001), and then use the information from those texts to interpret the significance of historical events and persons in historical context, mostly in textual form (Barton & Levstik 2004). Students run into a number of common challenges as they read and discuss these texts: they may not understand the language used in the past, the context in which events occurred, the source of the text itself, or how these frameworks might influence their reading and recall of detail (Craig, 2017c; Shopkow, Díaz, Middendorf, & Pace, 2012; Sipress & Voelker, 2009). For example, when studying the 1348 Black Death outbreak of bubonic plague, we have seen that students are quick to judge early responses to disease through the lens of current medical knowledge, leading them to assume that historical figures were foolish, rather than recalling specific details about medical training and putting those details in context of medieval medicine (Craig, Mahoney, & Danish, 2017).

It is particularly challenging to reconstruct the relationships between the many participants in a historical text. Additionally, interactions that do appear in the historical record are rarely spelled out for students to see in an explicit manner (Wineburg, 1991). The specific nature of interactions between people during historical events are thus subject to interpretation, and therefore could appear in a network analysis in a variety of different forms. Network analysis can support history students as they reflect on, debate, and then explicitly represent their interpretation of the connections within these historical texts by making careful decisions about how to reduce information in a network to single nodes (an item in the network) and edges (the connections between items in a network) and then using aggregate visualization comprised of many node-and-edge pairings in the network to rebuild their understanding of the larger historical context (Carrington, Scott & Wasserman, 2005; Scott, 2013; Wasserman & Faust, 1994; Wasserman & Galaskiewicz, 1994). The visual representations of each node and edge in the rebuilt network connect visual cues to rough statistical models of centrality—how many connections a node has to other nodes—in a way that will make otherwise challenging connections more open to analysis. For instance, in a network about a plague outbreak in Honolulu, Hawaii, in 1899-1900, Hawaiians are only barely present in the narrative. Students often fail to notice their absence, and the significance of that absence in the Hawaiian Republic's response to plague, until the network graph makes it visible.

The simultaneous view of individual interactions between two nodes and the aggregate interactions of the whole network imparts a rich contextual depth ideal for the study of history. History's disciplinary focus is on understanding specific historical agents and their actions and then placing an individual historical agent's conceptual framework of the world in which they live into a larger context made up of many historical agents and many world views (Carr, 1961/1987; Holt, 1990; Wineburg, 1991). The increasing use of social network analysis (SNA) to understand the role of the individual as one of a series of vast global connections in the networked age—in social media platforms, to name one very obvious example (Adamic, Buyukkokten, & Adar, 2003; Heer & boyd, 2005)—has made network analysis a

particularly attractive tool for historians seeking to employ modern technological methods in service of traditional historical epistemic practices (Meeks & Grossner, 2012; Uboldi & Caviglia, 2015).

Engaging students in disciplinary practices, including discussions about how the current tools of the discipline may influence key contemporary debates, are an important goal for pedagogues across disciplines (Engle & Conant, 2002). The major barrier in using network-analysis as a pedagogical tool in humanities classrooms is the fact that existing tools for visualization are highly specialized, and thus require a great deal of expertise on the part of both students and instructors to implement effectively (Graham, Milligan, & Weingart, 2016). It is thus challenging to bring these activities into a history classroom without additional scaffolding and translation. Furthermore, these materials are often intended for small group workshops, and therefore further work is needed to implement network-analysis tools successfully in traditional history lectures which are typically much larger at the introductory level.

In addition to the design challenges that instructors may face with existing network analysis tools, these tools are built on the assumption that data already exists in electronic form, and thus can be easily imported (Koeser & Croxall, 2013). However, the vast majority of archival data used in humanities research and teaching remains undigitized and therefore not subject to automated importation (Putnam, 2016). Additionally, historical documents do not lend themselves to the simplicity of an easily reducible node-edge relationship. Finding nodes and edges often means reading closely in a single document for historical connections—a practice that is already central to history research and pedagogy (Ad Hoc Committee on the Evaluation of Digital Scholarship by Historians, 2015; Liu, 2013). Thus, the process of examining a historical text for potential data that can be used in a network analysis tool is more than simply mining a text for fact. Students need to reflect deeply on the nature of historical significance for each detail they encounter and recreate the interactions drawn from historical records that they are exploring in their classrooms. It is particularly important that students come to appreciate that the same historical texts might be interpreted to create different network visualizations, and thus they need to treat these not as "definitive" but as another form of historical source that requires deep interrogation. More broadly, the process of creating a network from the ground-up will give students a deeper appreciation for how network visualization and analysis tools support the interpretation of all forms of data cleaning.

1.2.1 Initial Pilot and Previous Work

PI Craig has already developed and implemented a prototype for the network tool and used that to support a pilot test of network analysis in an undergraduate history classroom at IU: The Black Death, a course on outbreaks of bubonic plague from the sixth century Mediterranean to twentieth-century Hawaii (Craig 2017a; Craig, Mahoney, & Danish, 2017). Because of its subject matter at the nexus of history, medicine, and science, the Black Death class attracts a large number of science and pre-med students as well as students majoring in journalism, sociology, criminology, and history. Prior to the first network-analysis class session, students were asked to read an introduction to a 250-page monograph on an outbreak of bubonic plague Honolulu in 1899. When they came to class, they were then briefly introduced to the idea of network visualization, and to the process of creating network nodes and edges using a simple web form provided by Craig. The students worked together in groups of 5-6, to create nodes and edges to track the people, groups, places, events, and relationships between them within a 10-page segment of the monograph. The result was a network that documented the historical interactions in the entire monograph, which students used as a focal point for their discussions monograph in the following 3 class sessions. Between the first and second network-analysis sessions, students read the monograph. They were then asked in the second class-session to revisit 10 pages of the monograph, document the nodes and edges they saw on paper, and then revise or add to the network entries from the first session.

Analysis of the students' networks show that their recognition of historical figures and the relationships between figures improved dramatically. In their first attempt to craft a network visualization, 93 students in 16 groups entered 74 individual nodes—42 people, 18 groups of people, 7 locations, 4 things (newspapers, navy ships) and 3 events—and tracked 80 connections between these in a single 75-minute class session (Craig, 2017a). This was a low-density network (a low node-to-edge ratio in which

some nodes had very few edges), which reflects the fact that this was the students' first encounter with the 10 new pages they were assigned, and that they lacked the context a full reading would provide.

In the second network encounter, students found both factual and interpretive errors in 30 of the 74 original entries and entered an additional 68 new entries, in addition to creating 100 new edges. The final network included 96 individual people, 47 groups of people, 19 places, 10 things and 3 events, with 187 edges. The network density in the revised network was much higher, suggesting that students noticed more detail. The class discussion that followed suggests that the additional engagement with detail and the *capta* process of reducing historical interaction to network data also supported higher-order historical thinking. Students pointed to a number of contentious moments that required debate about significance, context and agency in the relationships between native Hawaiians, the mostly-native Hawaiian National Guard, and the Chinese and Japanese residents of Chinatown. Reconstructing the relationship between the recently deposed Queen, Liliuokalani, and the National Guard corrected their previous assumption that the Hawaiian National Guard (HNG) was mostly made up of white settlers but also raised questions about how the HNG balanced enforcement that often put Hawaiians under the control of the very group of white settlers that overthrew their native government.

Craig's observation of the discussions that followed the network analysis matched unsolicited student comments, which noted that engaging with a network analysis task prior to a full reading of the historical account supported better detail recall and easier understanding of human response to the events the monograph described. This is consistent with research on preparation for future learning in which engaging with a problem prior to encountering new information supports better learning and transfer (Schwartz & Bransford, 1998). More importantly, the initial encounter with 10 pages of unfamiliar text plus the network analysis primed students to comprehend the contents of the remainder of the monograph more completely than a similarly timed introductory exercise in the previous year's version of the class. In that previous exercise, student groups presented a short bullet-point summary of the people and events in 10 new pages to the rest of the class but struggled to see how their summary connected to other summaries from the monograph. In particular, students in previous years had primarily attributed the controversy in early 1900 over fire as an effective containment method to a difference between the interests of doctors, politicians, and business owners. Students who engaged in the network analysis exercise attributed the controversy to a much more nuanced combination of previous personal relationships between historical figures and the interests that differed between these professional groups (as the monograph's author does).

2 Research Methods & Plan

Our goal with the proposed EAGER project is to demonstrate the transformative potential of network input and analysis tools such as Net.Create for learning history content, and also engaging with network analysis tools. First, we want to show how entering and revising network data can support historical thinking, interpretation, and learning. This is an important contribution to the literature because network analysis tools have previously focused on pre-populated data, and thus demonstrating the value of entering data will help researchers to understand the possible value of this kind of tool. Second, we wish to explore how students use the unique features of a network visualization to explore historical concepts. This will have the benefit of both demonstrating the value of applying network visualizations to the humanities (an approach not yet explored for undergraduate students) and of demonstrating how students in the humanities might come to learn and appreciate these increasingly important digital sense-making tools. To accomplish these goals, we are proposing a 2-year study with 2 rounds of data collection, an initial implementation in Fall of 2018 to help us continue to refine our prototype activities and design, and a second implementation in Fall of 2019 intended to allow us to produce experimental contrasts to more concretely demonstrate the impact of using NetCreate in an undergraduate history classroom.

2.1 Research Questions

1. **How does network analysis support students appropriation of historical norms:** How do the practices of digital history affect learning and engagement with the epistemic foundations and disciplinary practices of historians?

2. **How do network-analysis tools help students understand the interactions between people and events that they encounter in history reading:** How does student retention and application of information about people and events in a network, and therefore the evidence available to use for crafting a history argument, change with experience in network creation, visualization, and analysis?
3. **How do students appropriate network analysis in service of history argumentation:** How do different modes of interacting with and generating network diagrams affect student retention and application of information in crafting a historical argument?
4. **How do the features of Net.Create support basic understanding of network analysis:** How do students appropriate the vocabulary of network analysis and connect underlying statistical data to network visualizations?

2.2 Research Design

2.2.1 Curricula and Activity Design

Our design approach builds on activity theory as a way of simultaneously articulating the features of our activity, and the software tools needed to help students engage in those activities so that they can transform their engagement with the content (Danish, 2014). This approach, coupled with analysis of our early pilot explorations (Craig et al., under review) has led us to focus on *capta*'s data-gathering, -entry and revision cycles as core activities in helping students to engage in sense-making with a historical monograph. Specifically, we have seen that students struggle with understanding the significance of key actors (people, organizations, artifacts) within a historical context, and the significance of their relationships with each other when attempting to engage a large corpus of data such as a lengthy historical text. Asking students to enter the actors into the network visualization tool as nodes, and then determine the nature of the relationships between them (the edges) is a first step in challenging students to make sense of the historical text in interaction with their peers.

Those unfamiliar with historical analysis might assume that this kind of data entry is uncontroversial and that there is only one "right" network. However, that belief would ignore the fact that there are, of course, many potentially valid answers regarding who is important within a historical text and which relationships are important to highlight and document. Furthermore, as our students come to realize in class: just because one historical source indicates a specific relationship does not mean that other sources agree, nor that there is consensus, which means that deeper study and inquiry is required to develop a supportable argument about what happened in a specific context. In fact, our early exploratory Net.Create work (Craig et al., under review) demonstrated that the most robust historical discussions emerge between students as they argue about what the relationship between historical actors might be, and how to translate their readings into specific data points for entry into the computer. Thus, our focus in designing Net.Create is on scaffolding data entry and revision so that students are engaged in interpretation of the text they are visualizing and have ample opportunity to discuss and debate their interpretations, those of their peers, and those of disciplinary experts, as they work through a historical event.

2.2.2 Net.Create Software Tool

In order to produce a network visualization of a historical text, students need to first review the text and transform the information in it for input into a database. As students draw information from historical texts, transform that information for input into a database, and see that text visualized live in a network that the entire class collaborated on, they engage in a humanities modeling exercise that simultaneously supports alternative pathways for learning the traditional content of history and provides new opportunities for learning the epistemic foundations and disciplinary practices of historians. While the contextualization in our activity design is more common to history classrooms, the practice of analyzing relationships between people, things, places, and even concepts is generalizable to many humanities disciplines, and Net.Create's tool design will therefore support broader use outside of history classrooms. To support productive engagement with historical argumentation, the Net.Create tool consists of 2 core components:

2.2.2.1 Net.Create Editor: A scaffolded node/edge input and editing form

In one screen, students will enter nodes (items, people, etc.) with additional notes that help students avoid duplicate entries and disambiguate very similar entries (dates, nicknames, citations, etc.). In the second, students will use drop-down menus to create edges (relationships of various types) between nodes they have previously entered, allowing for the input of multiple edge entries between the same nodes provided those edges happen at different points in time. Note that the types of nodes and classes of edges will be initially provided by the instructor to help scaffold students' engagement with the content, although students will be encouraged request additional node/edge types.

The node entry form will scaffold students' recording of key historical events using disciplinary norms of citation. In addition, the node entry form will incorporate predictive text and error checking to help encourage students as they encounter some common historical issues—for instance, a historical agent with a regularly-used nickname, or two people with the same last name and similar first names but different birth and death dates. The edge entry form will extend this scaffolding by providing a list of nodes previously entered by peers, from which students can choose to form edge pairs, as well as a way to add missing nodes. These predictions and limitations in the ways in which students can enter nodes and edges provide scaffolding that highlights common misunderstandings students have about historical figures and their relationships in the historical text while still encouraging students to form their own interpretations of those historical relationships independently of the instructor.

2.2.2.2 Net.Create Visualizer: An open-source network visualization interface

Net.Create will simultaneously create live network visualizations based on the most recent student input. Most network visualization tools require two separate database tables, one for nodes and one for edges, each with a differing set of variables customized to each type of data. Additionally, the tool will calculate the number of times two nodes appear in an edge relationship in order to create a "weight" column that tells the network-visualization tool how strong the edge between two nodes should be; this requirement mitigates an intentional behavior in Gephi, a very common network-visualization tool, which assumes that duplicate edges are errors and stops importing when the first duplicate is found.

As students input a new node or edge, the network visualization will adjust in order to provide students with a clear view of the current network. The immediate effect in network-analysis terms is that students can slowly add to their understanding of the network as a whole rather than trying to absorb the entire network in a single glance. For instance, from a historical-thinking perspective, few historical agents in a historical context have no connections. Therefore, as students add new nodes, we can prompt students to look immediately at the network visualizations—and therefore more closely at the historical context of the node they have just entered—to determine whether they need to also add more historical context for that node.

2.3 Research Plan

This proposal embodies a program of iterative design-based research (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003; McKenney & Reeves, 2012) that is appropriate for a design and development study (Institute of Education Sciences, 2013) where our goal is to understand the value of a tool (Net.Create) while continuing to refine it.. Our schedule is partially shaped by the currently approved IU History Department schedule, which allows us to incorporate our designs into classes taught by the PI.

Year 1 will focus on tool refinement and activity development and then collect data focused on RQ2 and RQ4 based on the initial design with a small number of students in a large history class comprised of approximately 100 students. This is similar to the context in which the pilot tool was previously used and will therefore provide us with an opportunity to explore how Net.Create is used in a traditional, introductory historical survey (The Black Death). This will also allow us to continue to refine the tool and our understanding of how a tool such as this can support history learning and retention.

During year 2, we will test all four research questions at scale with a refined tool by looking at how the tool supports engagement, history-content retention, and the disciplinary practices of history associated with argumentation using primary sources with 6-8 groups of 5-6 students each in a large (100 students) classroom. While it is challenging to do a true experimental contrast within this kind of classroom context because so much of our design is tied to the in-class interactions, we propose to use our

planned division of labor to better understand how the engagement in creating or revising a network can support students in learning the content that they are engaging with. That is, students will typically be divided into groups who develop networks for specific chapters (e.g., group 1creates the portion of the network tied to chapter 1, group 2 focuses on chapter 2, etc.). Thus, we can randomly assign groups of students into conditions where they are assessed based on either a) the network they created, b) a network they revised, or c) a network they have only briefly seen. Our intention is that this kind of assignment will help us to continue to understand how entering, refining, or simply viewing a network supports different levels of engagement with the content and thus learning.

2.3.1 Measures Used

We will use the following data sources for addressing the research questions in the two implementations.

Video data, artifacts, self-report questionnaire	We will develop new coding schemes, grounded in students' comments to explore the following issues (Chi, 1997): **1.** Students' historical reasoning. Our prior analysis indicate that it is valuable to distinguish between basic identification of historical actors (basic historical reasoning) and more advanced historical reasoning such as interpretation, contextualization, and historical empathy. **2.** How the tool supports and encourages the two kinds of reasoning indicates above.
Pre/Post history-knowledge measures	Improvement in short-answer questions focused on identification of significant historical figures and their relationship to each other. We will look to see how students identify and describe the key figures, how they articulate their relationships, and how these are justified using citations to the text.
Pre/Post network-analysis understanding	Improvement in short-answer questions focused on information drawn from viewing and analyzing pre-drawn network visualizations
Hand-drawn network diagrams	Complexity and accuracy of representation based on comparison to computer-generated network visualization for the same data
Self-report questionnaire	Likert-scale measures of confusion surrounding, interest in, and self-perception of improvement in: tool use, activity, network-analysis methods, history knowledge, and history methods

3 Why an EAGER?

3.1 Intellectual Merits

Network analysis is an increasingly popular and powerful computational tool for understanding relationships within large data sets. Digital historians have used these tools to represent and analyze historical contexts because they support scholars in exploring a broad range of connections between people, places, and events. While humanities pedagogues are optimistic that these affordances also provide unique opportunities for supporting students in learning history, available tools do not support easy integration into humanities classrooms, and there is not yet empirical or theoretical support for how this might be accomplished effectively. Net.Create aims to simultaneously open up this new genre of scholarship to undergraduate history students and provide a research base for activities that effectively integrate network analysis and history learning. As part of this integration, Net.Create also aims to demonstrate how a tool of this kind might overcome traditional barriers to recall and argumentation in history learning.

3.2 Broader Impacts

Social network analysis in particular is increasingly useful in 21st-century workplaces that rely on employees who can effectively connect computational and social practices, as well as understand the process of data collection and analysis that underpins visualizations. The design of Net.Create is intended to support history students in a wide-range of contexts in engaging in challenging historical practices, and

in appreciating how these practices are grounded in constructing arguments using evidence from the past. We will demonstrate the utility of this approach in two such contexts while developing theoretical insights into how these tools support history learning and engagement. These approaches should then generalize beyond the 200 students of this initial pilot to inform future work in network analysis that stretches across disciplinary lines. Net.Create and the associated curriculum also provide a unique opportunity to explore how students can learn these valuable network visualization and analysis skills, typically the focus of STEM disciplines, within a history classroom in service of history learning. If successful, this may provide an alternative context for teaching these valuable techniques, and for helping students to use history classrooms to learn cutting-edge digital approaches to scholarship across disciplinary boundaries.

4 Personnel & Relevant Prior Support

The project will be led by PI Kalani Craig, Ph.D., Clinical Assistant Professor of History at Indiana University and Co-Director of IU's Institute for Digital Arts and Humanities. Dr. Craig's 10 years of experience in web management and software design support her digital-history research, which uses text mining and network analysis to understand conflict in early medieval Europe, and her pedagogical research, which has included studies on network analysis and text mining (Craig, 2017), problem-based learning in history pedagogy with Dr. Hmelo-Silver (Bae et al., 2018), and spatial history and epidemiological simulations in history learning with Dr. Danish (Craig et al., 2017). Craig will lead Net.Create history activities, instructional strategies and assessments, and integration of Net.Create into those activities. Craig will also provide management for hourly staff. *Craig has not received prior support from the NSF.*

Co-PI Joshua A. Danish, Ph.D., Associate Professor of the Learning Sciences at Indiana University. Danish has over 15 years of experience designing and developing educational technology, including studies on visual representations as a tool for students of all ages learn about complex concepts. As an award-winning instructor in the IU School of Education, Danish has also spent a great deal of time exploring undergraduate pedagogy. Danish will lead the representational-practices activities and assessments, including the user interface of the Net.Create tool, and integrate them into the history activities. Danish will also oversee graduate student researchers and research data management. *Danish's prior support includes Science through Technology Enhanced Play (STEP) grant (2013-2016, IIS-1323767, $348,056), which has successfully shown how students can create productive models of the particulate nature of matter (Danish et al., 2015) and been disseminated via international and national conferences such as ICSCL and ICLS and journal articles including IJCSCL.*

Co-PI Cindy E. Hmelo-Silver, Ph.D. is Director of the Center for Research on Learning and Technology, Barbara B. Jacobs Chair in Education and Technology and Professor of Learning Sciences at Indiana University. Her expertise includes using technology to scaffold learning about complex systems, design-based classroom research methods to test learning technologies, and design of problem-based inquiry learning environments. For this project, Dr. Hmelo-Silver will lead the iterative research-design revisions between enactments. *Hmelo-Silver has received funding on 13 NSF grants including REESE DRL DRL # 1439227: Effects of Computer Support for Collaborative Learning in STEM Education: A Research Synthesis (2013-2017, $346, 892,) of computer-supported collaborative learning literature from 2005-2014, for which results have been disseminated at conferences and peer reviewed journals.*

Senior Personnel Dr. Ann McCranie, Assistant Director of Research Administration at Indiana University Network Science Institute, will serve as a bridge between the Net.Create team and the network-science community. Dr. McCranie will provide two weeks of consulting centered on the effective communication of network-analysis concepts—including betweenness and Eigenvector centrality measures, network-neighborhood and modularity measures, and density measures—in both the development of the Net.Create tool and the classroom activities that make use of the Net.Create tool, Dr. McCranie will also disseminate the learning gains of Net.Create's innovative predictive-text and predictive-connection tools in humanities classrooms to the network-science community.

Net.Create: Using Network Analysis to Support Digital Humanities Learning in Large History Classrooms

References

Ad Hoc Committee on the Evaluation of Digital Scholarship by Historians, American Historical Association (2015). *Guidelines for the Professional Evaluation of Digital Scholarship by Historians*. Retrieved from https://www.historians.org/Documents/Teaching and Learning/Current Projects/Digital Scholarship Evaluation/Guidelines Dig. Scholarship 9.30.2015(0).pdf

Adamic, L., Buyukkokten, O., & Adar, E. (2003). *A social network caught in the Web*.

Bae, H., Xia, F., Chen, Y., Craig, K., Hmelo-Silver, C. (2018) Developing Historical Thinking in PBL CLass Supported with Synergistic Scaffolding. 13th International Conference of the Learning Sciences (ICLS).

Barton, K. C., & Levstik, L. S. (2004). *Teaching history for the common good*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Boyd, d. & Crawford, K (2012) Critical Questions for Big Data, *Information, Communication & Society*, 15:5, 662-679, DOI: 10.1080/1369118X.2012.678878

Carr, E. H. (1961; 1987). *What is History?* (Second Edition ed.). London: Penguin.

Carrington, P. J., Scott, J., & Wasserman, S. (Eds.). (2005). *Models and methods in social network analysis*: Cambridge University Press.

Chi, M. T. H. (1997). Quantifying Qualitative Analyses of Verbal Data: A Practical Guide. *Journal of the Learning Sciences*, *6*(3), 271-315. doi:10.1207/s15327809jls0603_1.

Cobb, P., Confrey, J., diSessa, A. A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. *Educational Researcher, 32*(1), 9-13.

Craig, K. (2017a) Analog Tools in Digital History Classrooms: An Activity-Theory Case Study of Learning Opportunities in Digital Humanities. *International Journal for the Scholarship of Teaching and Learning* 11(1).

Craig, K. (2017b) Analog Tools in Digital History Classrooms: An Activity-Theory Case Study of Learning Opportunities in Digital Humanities. *International Journal for the Scholarship of Teaching and Learning* 11(1).

Craig, K. (2017c) History in 140 characters: Twitter to Support Reading Comprehension and Argumentation in Digital-Humanities Pedagogy, in *The Emerging Learning Design Journal* , 5.1 (Special Issue: Digital Humanities; June 2017).

Craig, K., Mahoney, C., & Danish, J. A. (2017). *Correcting for Presentism in Student Reading of Historical Accounts Through Digital-History Methodologies*. Paper presented at the AERA 2017 Annual Meeting, San Antonio, TX.

Craig, K., Szostalo, M., Uttamchandani, S., Danish, J. (under review). *Correcting for Presentism in Student Reading of Historical Accounts Through Digital-History Methodologies*. Paper presented at the AERA 2017 Annual Meeting, San Antonio, TX.

Danish, J. A. (2014). Applying an activity theory lens to designing instruction for learning about the structure, behavior, and function of a honeybee system. *Journal of the Learning Sciences, 23*(2), 100-148.

Danish, J. A., Enyedy, N., Saleh, A., Lee, C., & Andrade, A. (2015). Science Through Technology Enhanced Play: Designing to Support Reflection Through Play and Embodiment. In O. Lindwall, Häkkinen, P., Koschman, T. Tchounikine, P. & Ludvigsen, S. (Ed.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference* (Vol. 1). Gothenburg, Sweden: The International Society of the Learning Sciences.

Danish, J. A., & Saleh, A. (2014). Examining How Activity Shapes Students' Interactions While Creating Representations in Early Elementary Science. *International Journal of Science Education*, 1-21. doi:10.1080/09500693.2014.923127

Drucker, J. (2011) Humanities Approaches to Graphical Display. *Digital Humanities Quarterly* 5(1). Retrieved from http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html.

Drucker, J. (2013). 7A Network Analysis. *Introduction to Digital Humanities: Concepts, Methods, and Tutorials for Students and Instrucgtors.* Retrieved from http://dh101.humanities.ucla.edu/?page_id=64
Engle, R. A., & Conant, F. R. (2002). Guiding Principles for Fostering Productive Disciplinary Engagement: Explaining an Emergent Argument in a Community of Learners Classroom. *Cognition and Instruction, 20*(4), 399-483. doi:http://dx.doi.org/10.1207/S1532690XCI2004_1
doi:http://dx.doi.org/10.1559/1523040042742402
Gochenour, P. (2011). Nodalism. *Digital Humanities Quarterly, 5*(3).
Graham, S., Milligan, I., & Weingart, S. (2016). *Exploring Big Historical Data: The Historian's Macroscope*: Imperial College Press.
Heer, J., & boyd, d. (2005). *Vizster: Visualizing Online Social Networks.* Paper presented at the IEEE Information Visualization (InfoVis).
Holt, T. (1990) Thinking historically: Narrative, imagination, and understanding. College Entrance Examination Board.
Institute of Education Sciences (2013). Common Guidelines for Education Research and Development: A Report from the Institute of Education Sciences, U.S. Department of Education.
Koeser, R. S., & Croxall, B. (2013). Networking the Belfast Group through the Automated Semantic Enhancement of Existing Digital Content. *Journal of Digital Humanities, 2*(3).
Liu, A. (2013). The meaning of the digital humanities. *PMLA, 128*(2), 409-423. doi:https://doi.org/10.1632/pmla.2013.128.2.409
McKenney, S. E., & Reeves, T. C. (2012). *Conducting educational design research*. New York: Routledge.
Meeks, E., & Grossner, K. (2012). Modeling Networks and Scholarship with ORBIS. *Journal of Digital Humanities, 1*(3).
Minard, C. J. (1869). Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812–1813. Retrieved from https://en.wikipedia.org/wiki/Charles_Joseph_Minard - /media/File:Minard.png
Mohr, J. C. (2006). *Plague and Fire: Battling Black Death and the 1900 Burning of Honolulu's Chinatown*. New York, NY: Oxford University Press.
Prangsma, M. E. (2007). *Multimodal Representations in Collaborative History Learning*. Utrecht University Repository.
Putnam, L. (2016). The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast. *The American Historical Review, 121*(2), 377-402. doi:http://dx.doi.org/10.1093/ahr/121.2.377
Bransford, J., & Schwartz, D. (1999). Rethinking Transfer: A Simple Proposal with Multiple Implications. *Review of Research in Education*, 24, 61-100.
Scott, J. (2013). *Social Network Analysis* (third edition. ed.). Los Angeles: SAGE.
Shopkow, L., Díaz, A., Middendorf, J., & Pace, D. (2012). The History Learning Project "Decodes" a Discipline: The Union of Teaching and Epistemology. In K. McKinney & M. T. Huber (Eds.), *The Scholarship of Teaching and Learning In and Across the Disciplines* (pp. 93-113). Bloomington, IN: Indiana University Press.
Sipress, J. M., & Voelker, D. J. (2009). From Learning HIstory to Doing History: Beyond the Coverage Model. In R. A. R. Gurung, N. L. Chick, & A. Haynie (Eds.), *Exploring Signature Pedagogies: Approaches to Teaching Disciplinary Habits of Mind* (1st ed., pp. 19-34). Sterling, Va.: Stylus Pub.
Theibault, J. (2013). Part 5. See What I Mean? Visual, Spatial, and Game-Based History. In K. Nawrotzki & J. Dougherty (Eds.), Writing History in the Digital Age. doi:http://dx.doi.org/10.3998/dh.12230987.0001.001
Uboldi, G., & Caviglia, G. (2015). Information Visualizations and Interfaces in the Humanities. In D. Bihanic (Ed.), *New Challenges for Data Design* (pp. 207-218). London: Springer London.
Wasserman, S., & Faust, K. (1994). *Social network analysis: methods and applications*. Cambridge; New York: Cambridge University Press.
Wasserman, S., & Galaskiewicz, J. (1994). *Advances in social network analysis: research in the social and behavioral sciences*. Thousand Oaks, Calif.: Sage Publications.
Wineburg, S. S. (2001). *Historical thinking and other unnatural acts: Charting the future of teaching the past*: Temple University Press.

Indiana University – Center for Research on Learning and Technology
DATA MANAGEMENT PLAN

Data Management

1.1. Backup and replication within IU Storage Systems

The backup and/or data replication procedures for IU storage systems are as follows:

- *The Research File System.* RFS is backed up nightly to the SDA and saves versions for at least the previous seven days, seven weeks, and two months. While users must request a restore of one of these versions, the previous day's version of each of the user's files is immediately accessible in the one-day backup directory in each user's account.
- *The IU Data Capacitor.* Data stored on the Data Capacitor system are not backed up automatically. The Data Capacitor was primarily designed for the short-term storage of data. However, data from the Data Capacitor can easily be transferred to the SDA from any of IU's compute resources and thus replica copies may easily be maintained.
- *IU's Scholarly Data Archive (SDA)*. By default, data stored within the IU Scholarly Data Archive are stored in duplicate copies – one in the tape silo located at IU Bloomington, one in the tape silo located at IUPUI in Indianapolis. The HPSS metadata specifying which tapes contain any given file is backed up continuously, with multiple copies existing both in Indianapolis and Bloomington.
- *IUScholarWorks Repository*. Data deposited within the IUScholarWorks Repository are stored on the Scholarly Data Archive system, so these data are by default stored in duplicate in the SDA tape archives in Bloomington and Indianapolis.

The system security and documentation of system security is in compliance with NIST 800 Security Standards.

1.2. Facilities for handling sensitive data

IU has put in place appropriate administrative, technical, and physical controls to protect data in accordance with the HIPAA security rule. Electronic Personal Health Information may be stored on all of the HPC and storage facilities described in this document with the exception of the FutureGrid systems. Additionally, IU offers an enterprise-level cloud-based storage service (BOX) that has been approved for data categorized as restricted by FERPA (including identifiable student data in video form), which will house the bulk of our collected data.

1.3. Services lists and disaster recovery planning

IU maintains a list of information technology services it offers (http://www.indiana.edu/~uitsfin/ua_report_on_cost_and_quality_of_services0809.html) and there is a written disaster recovery plan for every service and every system described in this facilities statement except the FutureGrid systems. IU has an agreement for use of an off-site disaster recovery facility in case of a disaster affecting one or more of IU's campuses. For most research systems, the disaster recovery plan calls for restoring services at the remaining of the two main campuses (IUPUI or IUB) if one of IU's core campuses suffers a disaster.

1.4 Code distribution

In addition to IU open-source repositories, we will release the beta and subsequent versions of our code and network database structures via Github as open-source resources.