On Editorial Trust in Recommendation of Learning Resources

Taha Hassan

Preliminary Exam

Doctor of Philosophy

in

Computer Science

D. Scott McCrickard, Chair

Kenneth R. Edmison, Jr.

Bart Knijnenburg

Sang Won Lee

Mike Horning

Dec 5, 2021

Blacksburg, Virginia

Keywords: Adoption, Trust, Power, Recommendation

On Editorial Trust in Recommendation of Learning Resources

Taha Hassan

(ABSTRACT)

Evaluating the trustworthiness of a recommender system (RS) of learning resources poses three unique challenges. The first challenge is a lack of actionable frameworks to allocate context to the use of its host learning management system (LMS). The second challenge is a multitude of stakeholders in higher education (faculty, students, teaching assistants, instructional designers, LMS administrators, and department leadership) with unique job roles and prerogatives. The third challenge is the elusive nature of trust as a construct, in scope (belief, intention, behavior), and in source and target (stakeholder or algorithm, self or other). To address these challenges, this three-phase dissertation project proposes a novel UX framework for educational recommender systems, which accounts for LMS platform contexts, multiple stakeholders, and editorial trust relationships. In its first phase, the project proposes *Depth of Use* (DOU): a first-principles framework of frequent LMS use-contexts. DOU is found to highlight low-adoption course cohorts, evaluate course design interventions, and improve departmental resource allocation. The second phase of this project proposes a novel model of recommendation trustworthiness based in stakeholder allocation of RS editorial tasks. The study discovers a spectrum of faculty intentions about editorial division-of-labor and its frequent rationales, including student expertise, professional curriculum needs, authorship burdens at scale, and learner disengagement. In its third phase, the project investigates how editorial trust might be enhanced by transparency cues (guarantees, social proof, content tags). The dissertation concludes with a set of design guidelines to aid HCI practitioners in identifying editorial consensus, enhancing editorial transparency and algorithmic explainability, and facilitating reflection.

Contents

List of Abbreviations

DOU	Depth of Use
HYP	Hypothesis
LMS	Learning Management System
RQ	Research Question
RS	Recommender System

Chapter 1

Introduction

Introducing new technologies in the higher education domain at scale is a formidable challenge on the platform, stakeholder, and administrative fronts. On the platform side, service-based (SaaS) learning management systems (LMSs) have demonstrated have enabled rapid dissemination of information between stakeholders but limited suitability in supporting specific pedagogies and tools. For stakeholders, prior research frequently observes obstacles to adoption, such as cognitive burdens of discovery and maintenance, lack of technical support, and financial constraints. On the administrative front, the North American higher education software market has a reputation for its high barrier-to-entry, with risk-averse data governance and time-consuming approval processes. These challenges underscore the need to incorporate platform contexts and needs of multiple stakeholders in assessing the usability of new educational tools. Measuring platform use and allocating its contexts is, however, uniquely difficult because a multitude of native LMS applications generate huge volumes of data and there is little standardization of performance metrics for adoption, student learning or administrative support. Native LMS course analytics, recommendation and student feedback capabilities are presently quite limited. Furthermore, inspiring a change in choice technology for stakeholders in higher education is often a function of department initiatives, professional development programs, and sustained institutional evangelism. This points to a non-trivial challenge of trust creation and sustenance in educating stakeholders about new use-cases of technology and having them assess their fit in existing workflows.

To assist in navigating these challenges, this dissertation project proposes a platform-aware, multistakeholder UX evaluation framework of an educational recommender system. Three constituent studies are proposed. The first study proposes *Depth of Use* (DOU): an intuitive, first-principles framework to assess LMS utilization. DOU is found to highlight low-adoption course cohorts, evaluate course design interventions, and improve departmental resource allocation (study details appear in our ITiCSE'20 paper [32]). The second study proposes a novel model of recommendation trustworthiness based in stakeholder allocation of RS editorial tasks. The study discovers a spectrum of faculty intentions about editorial division-of-labor and its frequent rationales, including student expertise, professional curriculum needs, authorship burdens at scale, and learner disengagement (study details appear in our UMAP'21 paper [33]). The third, and final, study investigates how editorial trust might be enhanced by transparency cues (guarantees, social proof, content tags). The dissertation concludes with a set of design guidelines to aid HCI practitioners in identifying editorial consensus, enhancing editorial transparency and algorithmic explainability, and facilitating reflection.

1.1 Background and Motivation

Engineering trustworthy recommender systems (RS) of learning resources *at scale* is critically contingent on faculty and staff buy-in, trust in the recommendation algorithm, and reusability across disciplines. Service-based learning management systems are, increasingly, the choice platform for productivity, communication and assessment at institutions of higher learning [17], and they are frequently employed to scale pedagogies across disciplines [55]. However, despite significant previous work in learning analytics [4, 35, 42] over the past decade, prior research notes the limited capacity of LMSs to evaluate the impact of student engagement, instructional support, and faculty development on student outcomes [22]. There are several reasons. One, the existing adoption studies rely near-exclusively on self-reported LMS use, and there is no consensus on how to model **frequent use-contexts** of LMS-hosted native and third-party educational apps. Two, **multiple stakeholders** [1] in the higher education domain (faculty, students, instructional designers, LMS administrators, department leadership) have overlapping, but by no means identical, requirements for data aggregation and reporting. Three, a multitude of data sources (app metadata, course site content, team drives, social media) provide challenging, ever-expanding volumes of data to analyze, with complex and varied institutional, and governmental rules for access, analysis, and reporting. This identifies the first of our primary research questions (1.2.1): how can we identify the frequent use-contexts of LMS apps, and their impact on learning outcomes?

Educational recommender systems are important facilitators of teaching and learning practices [43]. Recommendation use-cases in prior research range from suggesting content relevant to a given topic, to suggesting learning pathways for a given learning objective, to suggesting next-term coursework and peers for group work. A recommender system aboard an LMS capable of servicing a broad range of use-cases across departments and user-roles needs to reckon with both platform-level drivers of adoption and initial stakeholder perceptions of relevance, trust and interpretation.

For instance, figure 1.1 illustrates how different stakeholders allocate recommendation authoring (seed) and editing (veto) powers. Almost all faculty (96%) allocate RS seed and veto powers for faculty, while only 52% and 44% allocate the same for teaching assistants and students, respectively. Almost all faculty (96%) do not support article vetoing by students, and all stakeholders agree on decreasing overall editorial power for faculty, teaching assistants and students, in that order. On the one hand, this underscores the importance of stakeholder privacy in the domain as course staff contemplates higher algorithmic agency. On the other

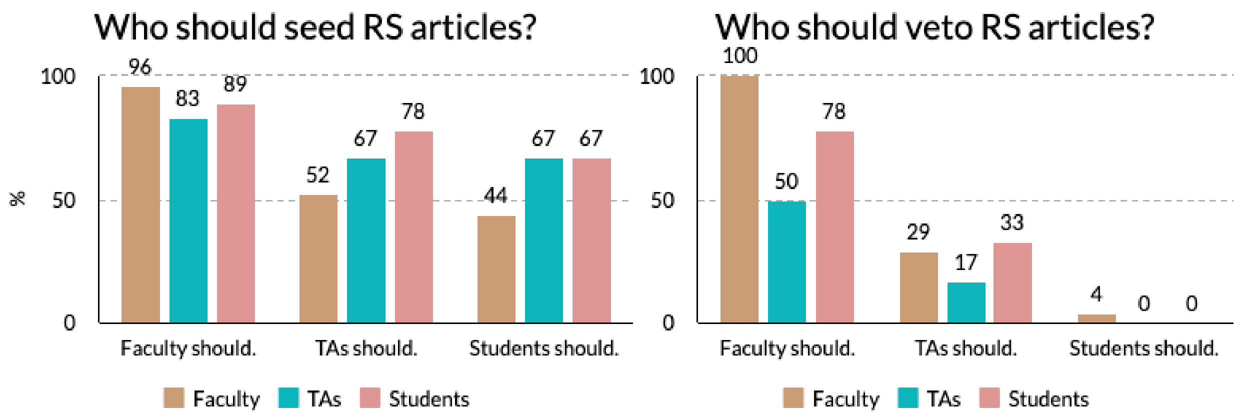

Figure 1.1: Editorial authority: Virginia Tech faculty, TA and students' opinions on who should be allowed to seed (left) and veto (right) articles for a 'Suggested Readings' recommender system (RS). 44% of surveyed faculty allocated the article sourcing task to students, but only 4% preferred students participate in article vetoing.

hand, it points to the variety of stakeholder tasks, duties, roles, prerogatives and potential **biases** we need to account for in designing a reusable educational RS. Existing RS work on trust or group dynamics generally does not grapple with differences in tasks and prerogatives. Herein, we are motivated to ask our second and third primary research questions: how do we interpret stakeholder standards for, and inform the design of, trustworthy educational recommendation, given a potentially **uneven distribution of editorial authority** such as between teachers and learners?

1.2 Research Questions

In the previous section, we outlined three broad research questions (RQs) that inform three constituent inquiries (**Studies I**, **II and III**) of the proposed dissertation project, respectively. We describe these research questions, and their corresponding sub-RQs and hypotheses, in detail as follows. Figure 1.2 outlines the overall structure of the dissertation project, and design parameters for its constituent studies.

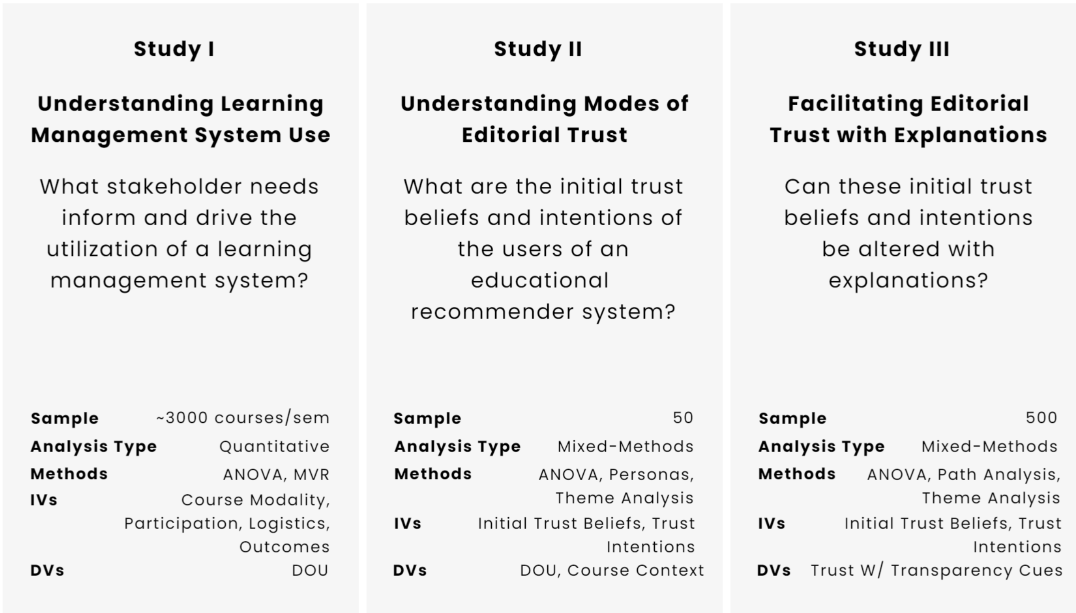

Figure 1.2: Dissertation project structure and constituent study parameters.

1.2.1 RQI: Understanding Learning Management System Use and Contexts

What is the utilization of a learning management system (LMS) for a course and an application native to the LMS, especially as a function of course context?

$$DOU = f(\text{course context}) \tag{1.1}$$

1. **RQI-a**: $DOU = f(\mathbf{modality})$
2. **RQI-b**: $DOU = f(\mathbf{participation})$
3. **RQI-c**: $DOU = f(\mathbf{logistics})$
4. **RQI-d**: $DOU = f(\mathbf{outcomes})$

The hypotheses corresponding to **RQI** are as follows.

(**H1**) Undergraduate courses have higher DOUs relative to graduate DOUs,

(**H2**) STEM courses have higher DOUs relative to non-STEM courses,

(**H3**) Online-only courses have higher DOUs relative to face-to-face courses,

(**H4**) Third-party app use significantly affects DOU,

(**H5**) Course DOU is linked to the #students enrolled full-time in the course,

(**H6**) Course DOU is linked to pageviews for the LMS course website,

(**H7**) Course DOU is linked to the average GPA awarded in that course, and

(**H8**) Course DOU is linked to the DFW rate of that course.

(**H9**) Course DOU is significantly linked to the #teaching staff members for the course, and

(**H10**) Course DOU is significantly linked to the instructor's prior enrollment in on-demand coursework for professional development.

1.2.2 RQII: Identifying Modes of Editorial Trust in Recommendation

How do the stakeholder editorial preferences (trust intentions) vary as a function of trust beliefs (stakeholder trust, trust in automation)? Specifically, how does editorial authority (**E-Auth**) vary as a function of trust in algorithms, delegation, and automation?

$$\text{Recommendation preferences} = f(DOU, \text{course contexts}) \tag{1.2}$$

1. **RQII-a**: E-Auth $= f(Stakeholdertrust)$
2. **RQII-b**: E-Auth $= f(\text{DOU})$
3. **RQII-c**: E-Auth $= f(\text{Course contexts})$

1.2.3 RQIII: Facilitating Editorial Trust with Explanations

How do the trust beliefs and intentions of the stakeholders of an educational recommender system vary with the use of explanations, and with course and platform contexts?

$$\text{RS trust beliefs and intentions} = f(\text{DOU, explanations, course contexts}) \tag{1.3}$$

1. **RQIII-a**: RS trust beliefs and intentions = f(explanations)
2. **RQIII-b**: RS trust beliefs and intentions = f(DOU)
3. **RQIII-c**: RS trust beliefs and intentions = f(course contexts)

1.3 Proposal Overview

The rest of this proposal is organized as follows. Chapter 2 provides the thesis statement for this dissertation project. Chapter 3 talks about understanding the frequent use-contexts of a learning management system. Chapter 4 details an inquiry into recommendation preferences, especially those for editorial trust for the stakeholders of an educational recommendation system. Chapter 5 examines the mechanisms of enhancing editorial trust for an educational recommender system. Chapter 6 describes the anticipated timeline and contributions made thus far in the project. Chapter 7 concludes the proposal document.

Chapter 2

Thesis Statement

Recommender systems that operate in a domain with formal or informal editorial processes require us to model the trust relationships between stakeholders. Multitarget trust can help highlight the change in trust intentions facilitated by recommendation transparency cues, in particular explanations incorporating social proof.

Chapter 3

Literature Review

3.1 Learning Management System Use and Contexts

We review the present state of related work in learning management system evaluation as follows. This includes previous research in human factors of LMS adoption (subsection 3.1.1), LMS success as an information system (subsection 3.1.2), and capacity to support data analytics (subsection 3.1.3).

3.1.1 Human Factors

There is considerable prior work on qualitative grounds for LMS adoption, like teaching and learning efficiency, generational student expectations, and institutional expansion and consolidation [11, 61]. For course instructors, the basic predictors of the pace of LMS adoption are departmental affiliation (STEM vs. non-STEM, say) and course modality (online vs. face-to-face, say). West et al. [61] conducted semi-structured interviews with 30 college instructors over two semesters, about primary use cases, teaching efficacy and efficiency, and overall satisfaction with Blackboard LMS. The study identified so-called 'integration challenges': course instructors finding it difficult to integrate LMS services into their teaching practices. This notion of 'integration' was echoed by McGill and Klobas [48] for the case of student adoption of WebCT, whereby students with a more favorable view of the 'task-technology fit' of LMS services were more likely to have higher LMS utilization. The authors also noted that instructor norms (instructor's view of LMS usability, support staff availability, and access to training resources) affected student utilization of LMS services favorably. Following an institution-wide transition to Canvas LMS, Wilcox et al. [62] surveyed user perceptions on frequent modes of use and platform limitations for Canvas LMS. They identified a generation gap in expectations between students and course instructors, wherein the pervasive student use of the mobile LMS app rendered a subset of Canvas sites - designed by faculty members for the desktop - ineffective in navigation, flow and content organization. We move this research forward by conducting a large-scale study of staff needs (scale, interoperability, ubiquitous access) that have informed Canvas LMS adoption across Virginia Tech.

3.1.2 Information Systems

Likewise, an information systems (IS) perspective on LMS adoption has been thoroughly explored over the years [2, 3, 52]. A bulk of these studies apply and evaluate a canonical model of IS success first discussed by DeLone and McLean [15]. The model factorizes the individual and organizational success of an IS into quality (system, information, and service), use (utilization, intention of use) and net benefits (impact on overall satisfaction, and intention of use) [16]. Adeyinka and Mutula [3] conducted a university-wide study of IS success factors underlying WebCT adoption and operationalized LMS utilization using nature of use (mandatory or optional), frequency of use, access and availability. They found use and intention of use both to be strong correlates of WebCT success. Fathema et al. [20] evaluated TAM using survey data on faculty and student attitudes about Canvas LMS at two public universities. They discovered that system quality and user self-efficacy were strongly linked to system use and perceived usefulness. They also noted that system quality is a multi-faceted notion that incorporates issues like design aesthetics, flexibility of access, degree of customization, and multimedia support. Ngai et al. [51] reported a stronger effect of the perceived usefulness and ease-of-use on system use relative to that of attitude (interest expressed towards adopting a new system). These studies largely employ user-reported system use in their analyses. Nonetheless, there are some early instances of LMS use modeling such as Ozkan and Koseler [52], where study participants reported system use as the number of hours spent daily, on course-related activities with U-Link using a desktop or web application. We contribute a vendor-agnostic resource-specific metric of overall and service-level LMS use.

3.1.3 Learning Analytics and Educational Data Mining

A discussion of the key drivers of learning analytics research in Ferguson [22] and Dawson [14] notes how native LMS data analysis, visualization and recommendation capabilities are presently non-existent or quite limited, even with standard tracking software features. A lot of student activity is external to the LMS, the data volume is huge and ever-expanding, and there is little standardization of the data aggregation and reporting methods, viz-a-viz critical use-cases for all stakeholders involved (faculty, students, instructional designers, LMS administrators, department leadership). These problems persist even as in the past two decades, inroads in educational data mining [19, 53, 54] have helped advance the state of the art in predictive modeling of student engagement, learning and achievement [6, 12, 13, 35]. Simultaneously, LMS log data analyses have been used extensively to model student and faculty use-contexts [9, 47], and to improve LMS features [21], often for specific disciplines and pedagogies [32]. Improving existing pedagogies, assessing learning outcomes and risk-of-failure for students [18, 34], and recommending interventions are all important use-cases that call for a convergence of data sources and a synthesis of approaches [29]. One of the early instances of this approach is Course Signals at Purdue [4]. Course Signals uses students' course outcomes, frequency of interaction with the LMS (Blackboard Vista), prior academic history and demographic information to ascertain a failure-risk measurement. In [63], a short-term warning system for ailing students models the early-term drop in clickthrough rates for modules of an online course. [42] describe a similar early-warning system which identifies isolated students using an analysis of ego networks and micro-communities of high-ability students on an online course forum. We build on these works to describe a cohort analysis and policy claim testing approach. Instructional designers and department leadership can use DOU to test the efficacy of faculty development initiatives, course redesigns, teaching support, and LMS evangelism.

3.2 Modes of Editorial Trust in Recommendation

Trust is an umbrella term for several distinct if connected problems across fields as diverse as recommender systems [59], adaptive web systems [5], user modeling [28], computer networks [37], and game theory [7].

3.2.1 Trust Domains

In a cross-domain review, Artz and Gil [5] describe four broad paradigms of trust analyses (policy, reputation, information web, and general). The first three deal with algorithmic, heuristic, or policy-based views of trust inference for systems with human and software agents. For instance, Massa and Avesani [45][46] develop an influential trust-aware recommendation framework which devises a *local* trust propagation strategy: predicting the trustworthiness of the RS user's neighbors based on proximity to user-labeled trusted 'peers'. A large body of work has been devoted to trust-aware collaborative filtering (CF), trust propagation, aggregation and user-similarity assessment using contextual, personal and network traits ever since [57]. Ma et al. [41], for instance, propose an ensemble fusion of the ratings and web-of-trust similarity matrices. User experiments with recommender systems [38] often identify trust as one of the personal attributes of users or user groups. For instance, Knijnenburg et al. [39] compare five interaction methods (topN, hybrid, explicit, implicit and hybrid) with a recommender system for energy saving measures. They discover that a user's propensity to trust is linked more to their choice satisfaction and trust in the overall recommendation system, rather than their preference for additional control or customization. A parallel body of work on recommendation for groups [25] has grappled with the problem of evaluating group consensus when its various subgroups exhibit independent and overlapping preferences and roles. A subset of this research relies on social choice theory to find broad types of consensus (most popular items, least misery for group members, etc.) [8]. Seko et al. [56] combine the notions of overlapping behavioral tendencies and power-balance in the group to assign consensus. Many open questions remain, including how to evaluate power asymmetry and evaluate the trustworthiness of the algorithmic judgments of consensus.

The final paradigm, in comparison, consists of sociological, psychological, and game-theoretic factorizations of trust [31]. McKnight et al. [49] and Gefen [27] study the qualitative properties of trust perceptions (like competence, benevolence, and integrity) for users of online legal advice and e-commerce websites. Mui et al. [50] propose a computational trust model using the notion of 'reciprocity' (the bi-directional exchange of favors or revenge) to infer social reputation and trust. Braynov and Sandholm [7] discover that a misrepresentation of agents' trustworthiness or distrust can results in sub-optimal degrees of social welfare, profits and cooperation in a bilateral negotiation game. They capture trust exchange between stakeholders, enhancing or undermining their *reputation*, and by implication, trustworthiness of their social neighbors in the process.

3.2.2 Trust Beliefs, Intentions and Behaviors

An important qualitative factorization of trust is contributed by McKnight [49]. McKnight's meta-model posits that **trust beliefs** inform **trust intentions**, which in turn inform **trust behaviors**. This meta-model borrows its canonical notions (beliefs, intentions and behaviors) from the larger social-psychological framework of Theory of Reasoned Action (TRA) [23]. It further argues that **disposition** and **institution** are antecedents to trust beliefs. We describe these notions in a recommendation context as follows.

Trust beliefs Trust beliefs refer to the user-attributed values or standards for a recommender system. McKnight [49] observes that prior models of trust in information systems can be grouped along three key dimensions (**CBI**): competence, benevolence, and integrity.

Trust intentions Trust intentions signal a user's willingness to depend on the recommender system, for instance, by providing personal information, following advice, or making a purchase.

Trust behaviors Trust behaviors are user tasks performed (i.e. adoption or engagement) or risks taken (e.g, consumption and purchase behaviors) in response to information attained via the recommender system.

Disposition and Institution 'Disposition' and 'Institution' describe two antecedents of trust beliefs, according to McKnight [49]. These identify an individual's general tendency to trust people and institutions governing the recommender system (e.g., the web or an LMS), respectively.

We propose a multi-target view of an individual's trust perceptions towards domain stakeholders. In a recommendation context, this requires modeling of (a) the degree to which this individual believes each stakeholder possesses competence, benevolence and integrity overall (trust *beliefs*), and (b) the RS editorial tasks of sourcing, vetoing, rating, and commenting this individual assigns to all stakeholders (trust *intentions*). This model designates the recommendation algorithm as one of the stakeholders. Prior literature on *anthropomorphism* in computing observes that users of a technology artifact form beliefs about trusting the artifact that are similar in language and composition to their beliefs about trust of other humans [36]. Study II survey and interviews suggest the same for stakeholders in higher education (faculty, teaching assistants, students). Analyzing the trust perceptions of human and machine stakeholders together can help assess how both reinforce or diminish each other for different course and editorial contexts.

Why editorial authority? One, we find that there are distinct job-related *roles*, and an informal editorial process of *task allocation* in the domain of higher education. In our survey and interviews for Study II, we observe that stakeholders frequently refer to instructors' *prerogative* in authoring and maintaining the recommended readings. It appears that all stakeholders perceive distinct editorial roles and tasks implied by their job title, so it is useful that a conceptual model of trust should contend with these roles and how tasks are negotiated among them. We discover that these allocations are linked to trust relationships between the stakeholders.

Two, editorial intentions express the intent to trust. This is because trust intentions are typically understood as the willingness to assume some risk in interacting with a digital artifact. McKnight et al. note that the customer of an online retail system indicates their trust by volunteering personal information or making a purchase. In higher education, a faculty member can delegate recommendation authoring or editing tasks, with risks like mishandling of student feedback, misinformation and spam, or absenteeism.

3.3 Trust and Explanations

Explanations are a transparency cue used to communicate the *intent* and *process* of the underlying recommendation algorithm [24], to *persuade* the user to participate (e.g. interact, share information, make a purchase) and increase their confidence in the recommender system [58], or to explain the *tradeoffs* worthy of consideration in decision-making [60]. Explanations are routinely used to enhance the perceived explainability and user trust of knowledge-based systems and decision support systems. Previous research on explanations has evaluated their utility for myriad recommendation use-cases (including music, books, fashion, finance, health) and UX evaluation goals (transparency, trust, satisfaction, persuasion) [26]. Several taxonomies of RS explanations have been researched over the years [24][10]. However, we have found a limited treatment of trust in this literature, owed in part to lack of domain-specific trust models and the difficulty of ensuring construct purity in defining trust metrics [58].

Wang and Benbasat [60] examine common obstacles to trust in recommendation agents, for instance, *agency relationship* and *high choice discretion*. Agency relationships are characterized by *asymmetry of information* and *goal incongruence* between users and the algorithm (machine stakeholder). We observe that in an editorial context for education, this agency relationship is not always perceived as a trust obstacle by students, but a necessary feature of the prerogative of faculty as in-charge of all course content. In Study II, we report obstacles to editorial trust expressed by faculty (authorship burdens, risks of misinformation, student disengagement and disproportionate attention to outcomes). We hypothesize that these obstacles point to underlying deficits in perceptions of competence, benevolence and integrity among stakeholders, and explanations are a potential remedy for these deficits.

Guarantees offer little to no interpretation to RS users about the rationale behind each recommended item. They instead focus on trust engendered by stakeholders. In the case of higher education, course contexts in which students are focused on maximizing their available time for exam prep and assignment completion can potentially favor this mode, and guarantees can reinforce the efficiency of *a priori* editorial authority relationships understood by staff and students.

Social proof explanations can potentially provide material evidence of student participation, and by association, their competence and integrity in editorial matters, to course staff. A reference to RS consumption patterns of high-scoring subset of students in the social proof explanation is one way of harnessing the outcome bias obstacle to trust. Content-based explanations can help reduce information asymmetry between stakeholders and the RS algorithm by enhancing their knowledge about the sources and subject-matter relevance of recommended items. Explanations about the RS algorithm's capability to automatically flag and remove malicious content can potentially alleviate the perceived misinformation risk in adopting a RS for high enrollment courses.

Chapter 4

Understanding Learning Management System Use and Contexts

Online learning management system (LMS) tools are, increasingly, the primary infrastructure for productivity, communication and assessment at institutions of higher learning. Measuring their utilization and impact is thus, critical to assessing the efficacy of online teaching and learning at scale. Existing models of LMS utilization, however, are largely qualitative, opaque to individual LMS tools, and difficult to generalize to the needs of multiple domain stakeholders. Study I proposes **DOU** ('depth of use'), a novel method for computation of intuitive ordinal rankings (low, medium, high) of LMS utilization using instructional design-driven taxonomies. The study proceeds to test hypotheses about the relationship between DOU rankings and course meta-attributes (modality, participation, logistics, outcomes), followed by expert reviews. On the whole, Study I seeks evidence for the efficacy of DOU in *interpreting* and *supporting* stakeholder decisions about adopting new LMS tools (faculty), allocating institutional support (IT administrators), informing LMS evangelism (instructional designers) and assessing administrative policy impact (department leadership).

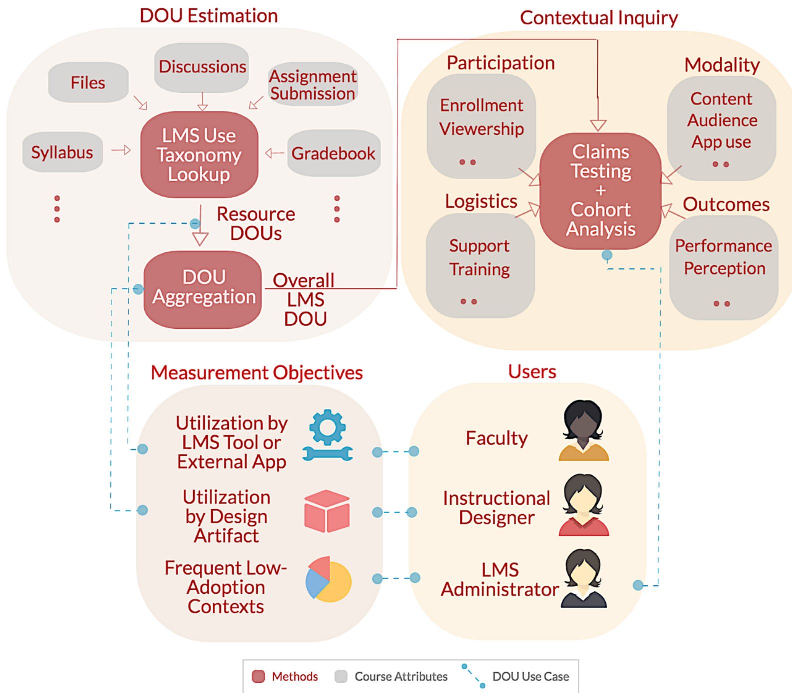

Figure 4.1: Study I overview (clockwise from top left): DOU Estimation (data sources and methods), Contextual Inquiry (hypothesis variables and methods), key user groups, and measurement objectives fulfilled by DOU. (Top-left) We estimate a DOU score of overall LMS use - low, medium or high - for each course in our analysis. We then test claims (top-right) about how DOU is linked to course modality, participation, logistics, and outcomes. Finally, we investigate DOU use-cases (bottom-left and bottom-right) for faculty, instructional designers, and LMS administrators.

4.1 Study Objectives

To answer its primary research question (section 1.2.1), Study I proposes a novel measurement model called 'Depth of Use' (DOU) [32]. This framework assigns an ordinal DOU score of LMS use (low, medium or high) to each of forty-thousand college courses offered between 2017 and 2019 at Virginia Tech. DOU uses a *vendor-agnostic* taxonomy of LMS use developed in collaboration with Virginia Tech instructional designers. Figure 4.1 describes the overall approach. We then hypothesis-test (ANOVA, multivariate regression analysis) these scores against course attributes like modality, participation, logistics and outcomes.

This helps determine the frequent contexts where faculty and staff might deem a subset of LMS services effective. For instance, the study discovers a consistent impact of overall LMS use on learning outcomes, and increasing reliance by faculty on tools that favor scale, ubiquitous access and interoperability. Finally, we discuss three key applications of DOU, to a) help faculty members assess the relative utility of LMS services and legacy apps, b) aid instructional designers in measuring and improving the scope of interventions and LMS evangelism, and c) help LMS administrators identify the technology needs of actionable low-adoption cohorts. The framework forwards a *multistakeholder* view of LMS utilization, in that alongside learning analytics, it supports claim testing and cohort analysis for policy decision-making, an avenue with lesser treatment in educational research literature in the last decade [44].

4.2 Datasets and Methods

The primary dataset for Study I is Canvas LMS utilization metadata for nearly forty thousand courses offered in the fall and spring academic terms between, and including, 2017 and 2019. We used web scraping, entity resolution and topic modeling to parse and aggregate data from Virginia Tech course catalog, timetables, LMS site contents, and page request logs.

To answer our study research question (**RQI**), we begin by testing our hypotheses (**H1** - **H10**). DOU is ordinal and not normally distributed, so we use non-parametric Kruskal-Wallis H-test [40], in addition to an independent two-sample t-test, for hypotheses with discrete-valued meta-variables. We evaluate group differences in viewership and enrollment for each of low, medium and high DOUs using one-way ANOVA (F-test). To expand our analysis, we then claim-test each of the hypotheses against all constituent dimensions of

DOU. We combine these hypothesis tests with frequency and cohort analyses to examine the needs of the DOU use-case (adoption, support, learning outcomes).

4.3 Contributions and Implications for Practice

4.3.1 User Engagement Claims Framework

I developed an empirical framework called **Depth of Use** (DOU) that assigns ordinal rankings (low, medium, high) to a given course by aggregating the use of LMS-hosted apps. I hypothesis-tested DOU for its frequent course contexts on a dataset of nearly forty thousand courses commissioned on Canvas LMS between 2017 and 2019. I discovered that LMS use is near-consistently linked to better learning outcomes, and the needs for scale, interoperability and ubiquitous access drive faculty adoption of the learning management system.

4.3.2 Resource Allocation and Low-Adoption Cohort Discovery

The Depth-of-Use (DOU) framework has proven useful to instructional designers in evaluating the efficacy of faculty development programs and departmental resource allocation. It has also proved useful to LMS administrators in discovering actionable low-adoption course cohorts (for instance, **junk-drive**, **gradebook-only**, and **access-portal**).

4.3.3 Pandemic Response

In the spring of 2020, an institution-wide policy of emergency remote teaching was rapidly enacted by Virginia Tech IT leadership, in response to the COVID-19 pandemic. System administrators began with a DOU analysis conducted at the beginning of the term to determine key low-DOU course clusters (upper level STEM and general education coursework), frequent high DOU LMS features (typically the ones with lowest cognitive burden-of-discovery like files and gradebook), and frequent low DOU LMS features (quiz and assignment delivery). The administrators facilitated a rapid transition to remote teaching over a period of two weeks by focusing their support on low-DOU instructors. They designed training classes, in-person consultations, and in-depth documentation focusing on delivery and submission of assignments and quizzes via Canvas. The IT transition team was able to increase the total number of high DOU courses by over 49%. Three key takeaways emerge. First, the smallest gains were observed in files and gradebook modules, which hints at an abundance of junk-drive and gradebook-only courses prior to the transition. Instructors new to an LMS tend to first explore the tools they can utilize without significant cognitive effort. Second, the team found that post-transition, instructors' use of announcements (for bulk communication within the LMS) and discussion forums (as a replacement for in-class, face-to-face interaction) increased significantly (+44%, and +40%, respectively). Third, the increase in courses with high DOUs for assignment delivery (+15%), assignment submission (+24%), quiz delivery (+30%), and quiz submission (+36%) is often at the expense of low DOU courses in the same category. This suggests that in favoring online course assessments, Virginia Tech faculty responded to the transition team's focused development and support initiative. While an uptick in overall DOUs is expected during a transition to remote teaching, we contend that if the transition team had not been able to provide directed support based on DOU analysis, we would observe an abundance of low $\rightarrow$ medium DOU growth. The instances of low $\rightarrow$ high and medium $\rightarrow$ high DOU growth suggest that our analyses facilitated an actionable assessment of key faculty needs and a focused adjustment of the IT training and support regimen which maximized its impact.

4.3.4 Publications

The following manuscripts are accepted (**A**), under peer review (**PR**), or planned (**P**) within the scope of Study I.

(**ITiCSE'20**, **A**) **Taha Hassan**, Bob Edmison, Larry Cox, Matt Louvet, Daron Williams, and D. Scott McCrickard. "Depth of Use: An Empirical Framework to Help Faculty Gauge the Relative Impact of Learning Management System Tools." In Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education, pp. 47-53. 2020. [Acceptance Rate: 28%]

(**Studies in Higher Education, P**) **Taha Hassan**, Bob Edmison, Larry Cox, Matthew Louvet, Daron Williams, Brian Broniak and D. Scott McCrickard. "A measurement model for assessing the educational impact of learning management system tools at scale." In review for *Studies in Higher Education* [Impact Factor: 4.38].

Chapter 5

Identifying Modes of Editorial Trust in Recommendation

Recommendation algorithms frequently incorporate explicit and implicit signals of trust, for instance, the RS user's *interactional* awareness of their local neighborhood, or some consensus of the preferences of their self-reported, trustworthy friends in a social network at large. While algorithmic awareness of a user's neighborhood is important for producing accurate recommendations, real life recommendation tasks often involve user groups with differences in institutional or group-based roles, powers and prerogatives. Study II argues that recognizing stakeholder editorial power relationships, or *intentions* in the recommendation domain can highlight the broader context of trust in the recommendation algorithm. Study II examines Virginia Tech faculty and student preferences of editorial authority and trust in algorithmic agency, for a hypothetical 'Suggested Readings' recommender system aboard an LMS course site. Study III proposes a task-based metric of editorial authority (**E-Auth**), and evaluates the relationship between RS editorial task distribution and stakeholder trust in algorithmic agency. Using narrative analysis, the study also describes frequent editorial editorial roles, and their course contexts and stakeholder rationales, including subject expertise, scaling the learning environment, market needs and learner disengagement.

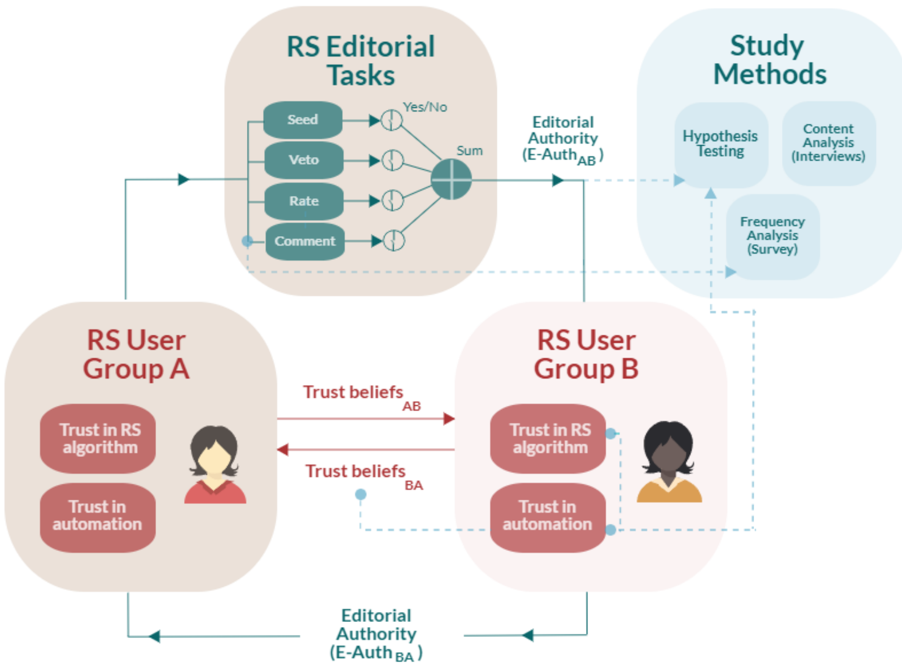

Figure 5.1: Study II overview: Recommendation stakeholders and their mutual trust beliefs (bottom), editorial trust intentions (top-left) and study methods (top-right).

5.1 Study Objectives

Outlined in the primary research question of Study II (section 1.2.2), the study objective is to understand the beliefs and intentions of trust between stakeholders in the domain of higher education, specifically in the context of use of an educational recommender system. We refer to this formulation as **multitarget** and **multistakeholder** trust. We aggregate key RS tasks (seed, edit, refresh, and delete) using the notion of *editorial authority* and measure its allocation among faculty, staff and students. We then use analysis of variance (ANOVA) hypothesis tests to understand the relationships between editorial authority, trust in algorithmic agency, delegation, and perceived user biases [30]. We contend that these differences in *editorial authority* - exemplified in an editor-consumer relationship between faculty and students, say - are related to the trust both assign to each other and the RS algorithm. For instance, faculty's willingness to incorporate student and TA feedback into the RS algorithm can point to a belief in editorial authority for multiple stakeholders, or regard for automation in the longer-term.

The study also seeks to identify RS editorial models frequently favored by the study participants, and their relationship with the trust exchanged between faculty, TAs, students, as well as the RS algorithm.

5.2 Datasets and Methods

Study II datasets include survey responses from 63 participants (27 faculty, 6 teaching assistants, 30 students), and transcripts and notes from follow-up semi-structured interviews with 18 (6 faculty, 3 teaching assistants, 9 students) of the aforementioned 63, all affiliated with Virginia Tech. These faculty members, students and teaching assistants represent 16, 12 and 5 departments, respectively. 46 of the survey respondents (17 faculty members, 6 TAs, 23 students) represent STEM disciplines of study. We recruited participants on a rolling basis between August 2020 and July 2021, using convenience sampling and voluntary response sampling on departmental mailing lists and Facebook groups.

For our analyses, Study II proposes a mixed-methods approach. We conducted hypothesis tests regarding the relationships between our study variables (relative trust in algorithms, automation, editorial authority) using one-way ANOVA (F-test). We also performed content analysis on survey responses and interview transcripts to identify frequent themes in the study participants' commentary regarding their preferred RS editorial tasks for each user role.

5.3 Contributions and Implications for Practice

5.3.1 Multistakeholder Trust

Study II intends to bridge individual and group notions of trust in recommender systems, by extending McKnight's typology of trust in information systems to all domain stakeholders, including the RS algorithm. Recommender systems for higher education also need to account for different stakeholder preferences in authoring and maintaining of recommendations. Study II proposes a simple metric of editorial authority (**E-Auth**) to reflect these intentions. Using online surveys, stakeholder focus groups, and semi-structured interviews, Study II a) isolates top N editorial preferences by group, and synthesizes the driving factors behind LMS and recommender system adoption, and b) proposes a consolidation with general UX models of recommender systems [38]. This should allow an assessment of the relationship between actual use of a content recommender prototype, user aspects (DOU, cognitive biases, general trust in technology, general trust in automation, delegation), and system aspects (output explanations, notification tiers).

5.3.2 Recommendation of Learning Resources

Recommender systems for education have fulfilled a variety of analytics tasks for the individual learner, interpretation and intervention, in-class and online [43]. However, there is a pressing need for the educational RS community to recognize platform-level changes in the domain. Recommender systems have to reckon with concerns of trust, efficacy, and interpretation *at scale*.

(a) Situating trust for the needs of an educational recommender system, demonstrating which dimensions matter and why,

(b) Discovering differences in group intentions, like students-as-viewers (SAV) and students-as-editors (SAE),

(c) Demonstrating the potential for actionable design guidelines, the "mapping", specifically to diversification, interpretation and disruption cues,

5.3.3 Web Mining for Bias Signatures

We employ datasets from social academic forums (*Koofers*, *RateMyProfessor*) to test claims about the relationship between student outcomes and instructor rankings at scale. A statistically significant relationship exists between rankings and course outcomes across hundreds of departments at Virginia Tech and its 25 peer institutions of higher learning. We are working to develop the connection between DOU, course context, and outcome bias, in order to test faculty claims about **student disengagement** and disenfranchisement often forwarded as a reason for lesser RS editorial authority allocated to students. Our interviews for Study II indicate that group disposition about trust includes suspicions of biases like outcome bias, and it appears to affect group intent.

5.3.4 Publications

The following manuscripts are accepted (**A**), under peer review (**R**) or planned (**P**) within the scope of Study II.

(**UMAP'21**, **A**) **Taha Hassan**, Bob Edmison, Timothy Stelter, and D. Scott McCrickard. "Learning to Trust: Understanding Editorial Authority and Trust in Recommender Systems for Education." In Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization, pp. 24-32. 2021. [Acceptance Rate: 23.3%]

(**WWW'19 Companion**, **A**) **Taha Hassan** and D. Scott McCrickard. "Trust and trustworthiness in social recommender systems." In Companion Proceedings of The 2019 World Wide Web Conference, pp. 529-532. 2019.

(**WebSci'19 Companion, A**) **Taha Hassan**. "On bias in social reviews of university courses." In Companion Publication of the 10th ACM Conference on Web Science, pp. 11-14. 2019.

(**ASONAM'19, A**) **Taha Hassan**, Bob Edmison, Larry Cox, Matthew Louvet, and Daron Williams. "Exploring the context of course rankings on online academic forums." In 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 553-556. IEEE, 2019.

Chapter 6

Facilitating Editorial Trust with Explanations

In our previous studies, we discover that faculty differ in their allocation of editorial tasks for educational recommendation. Some favor a *conservative* model, where students can view (and rate or comment on, in a subset of cases) suggested readings, but not create or remove them. Others lean towards an *egalitarian* model where students are actively involved in all or most authoring and feedback tasks. In Study III, we investigate if explanations, a transparency cue to describe stakeholder intent for recommendation users, can improve trust, increase delegation, and facilitate reflection.

Stakeholders cite obstacles to trust like perceived authoring burden, perceived risk of misinformation, and beliefs about relative competence and integrity. Study III proposes a live user study to explore the impact of three explanation types (guarantees, social proof, content-based) on stakeholder trust beliefs, trust intentions, tendency to reflect on domain practices, and task performance for a recommender system-aided literature review. We hypothesize that a proof of participation by students in editorial tasks can reduce perceptions of authoring burdens, misinformation risk, and mistrust of automation for both staff and students.

6.1 Study Objectives

Study III investigates whether the use of explanations in recommendation bridges the editorial trust divides in beliefs or intentions. Knijnenberg and Willemsen [38] describe a recommendation UX evaluation framework where the experiential (**EXP**) component of a recommender system informs the behavioral (interaction - **INT**). The experience of RS system aspects is further informed by a user's personal attributes (**PC**), for instance, trust or subject matter expertise. Our work in Study III intends to generalize this UX evaluation framework to a multistakeholder setting with potential power asymmetry and examine how the use of an explainable recommender system affects *initial* trust beliefs and intentions.

Three different types of explanations in the educational recommendation context are assessed in this study:

- Guarantee-based: 'faculty-recommended readings'
- Social proof-based: '30% of your classmates found these useful'
- Content-based: 'further reading on binary search from week 2'

The study objectives are as follows. The objectives proposed for the future are outside the present scope of the study, and not required for its completion. We hope to investigate them contingent on availability of time after the conclusion of Study III.

(I) Assess if explaining the rationale behind recommended readings can increase the levels of *initial* trust beliefs (competence, benevolence or integrity) overall, and for faculty and students whose editorial trust intentions (E-AUTH) are conservative vs. egalitarian,

(II) Assess if the editorial trust intentions (allocations of seed, veto, rate, and comment tasks) expressed by stakeholders change after interacting with the recommendations and their associated explanations,

(III) Assess if content based explanations are more effective than social proof, overall, or SAV/SAE, or policy preferences (based on **DM**) at the group level,

(IV) Assess if allocation of editorial tasks is linked to overall satisfaction, perceived efficacy of recommendations, and task performance,

(V) Assess if explanations are linked to higher system use (low or high DOU),

(VI) (Future) Assess if trust intentions mediate the relationship between beliefs and behaviors,

(VII) (Future) Assess if *dissonances* like a gap between beliefs and intentions, or intentions and behaviors is linked to low adoption or recommendation use,

(VIII) (Future) Assess if at the course level, there is a consensus between stakeholders on editorial trust intentions,

(IX) (Future) Assess if use/adoption and trust are orthogonal as design objectives, if guarantees help with adoption but not trust. Assess if trust leads to *passive* use of the recommender system.

6.2 Datasets and Methods

The primary dataset for Study III is behavioral data from online evaluation of a 'Suggested Readings' recommender system (LTI app) prototype. Study III proposes a mixed-methods approach, with hypothesis tests (one-way ANOVA, F-test) to compare the use, trust beliefs, and editorial trust intentions of all explanation types, between-groups and within-group before and after interacting with the educational recommender system. We also perform narrative analysis on pre-/post-study survey responses and interview transcripts to identify frequent themes in the study participants' commentary regarding their preferred explanation type.

6.3 Contributions and Implications for Practice

We assess the efficacy of explanations in guiding the evolution of *initial* editorial trust beliefs, intentions and behaviors. We then describe design guidelines to encourage ease of interpretation, diversification, and disruption in educational recommender systems, given DOU and faculty trust beliefs.

6.3.1 Dealing with scale

Reliance by students on recommended resources, and evidence that their trust is enhanced in egalitarian editorial arrangements could imply lesser reliance on teaching assistants. This may ease the burden of resource allocation at the department-level, increase student motivations to participate, and lessen faculty suspicions of outcome bias.

6.3.2 Value-sensitive design

Conservative and egalitarian course arrangements have different dispositional and attitudinal tendencies. Designing for these two cohorts requires respectful consideration of their preferred use-cases and initial trust beliefs.

6.3.3 Publications

The following manuscripts are accepted (**A**), under peer review (**R**) or planned (**P**) within the scope of Study III.

(**RecSys'22, P**) **Taha Hassan**, Bob Edmison, Larry Cox II, Matthew Louvet, Daron Williams, Brian Broniak, and D. Scott McCrickard. "The contours of prerogative: delegation and trust in multistakeholder recommender systems." In 2022 ACM Conference on Recommender Systems.

Chapter 7

Anticipated Conclusions

In the introduction to this dissertation proposal, we outline three primary research questions. The first research question (RQI) investigates the frequent use-contexts of a learning management system, the second (RQII) examines frequent modes of editorial trust for recommendation, and the third (RQIII) evaluates transparency cues to facilitate editorial trust. This section reviews the conclusions reached and anticipated from our investigations into each study research question.

7.1 Study I

Study I contends that a multitude of software services, and large volumes of user data make it difficult to assess and improve the usability of LMS-hosted educational apps, support pedagogy, and design meaningful interventions for at-risk students. To solve this problem, a critical first step is to concisely describe how these software services native to an LMS are utilized in aggregate by a university course. Study I thus proposes a novel method (**DOU**) to convert expert-sourced taxonomies of LMS use into a single ordinal (low, medium or high) ranking. The study then evaluates the relationship of LMS use and course attributes such as modality, participation, logistics, and outcomes for nearly 20,000 courses between 2017 and 2019 to discover frequent use-contexts (study methods detailed in our ITiCSE'20 paper).

The hypothesis tests in Study I demonstrate that the **DOU** framework enables evidence-based discovery of stakeholder needs in connection with the use of technology. For instance, we discover that needs for scale, ubiquitous access and interoperability drive faculty use of LMS tools. We also discover that DOU is frequently linked to better aggregate course outcomes (GPA, DFW), teaching support, and third-party app use. Finally, we conclude that instructional designers, LMS administrators and department leadership can use DOU-based analyses to inform LMS evangelism, design interventions, professional development, and pandemic response planning.

7.2 Study II

Study II contends that in higher education, interpreting the overall trust in technology requires accounting for all human and machine stakeholders and their institutional roles and prerogatives. Earlier in the project, Study I demonstrated how faculty's use of new educational technology can vary with their department, academic discipline, research duties, teaching workload, and institutional support. Study II hypothesizes that initial trust of new technology, specifically educational recommendation, can similarly vary with stakeholder and editorial tasks.

Our hypothesis tests and interviews in Study II reveal a spectrum of faculty's editorial trust intentions for students, ranging from conservative (students can view or rate recommended course readings) to egalitarian (students can perform recommendation authorship and editing tasks). Our preliminary study points to the link between editorial trust beliefs and intentions, and frequent stakeholder rationales of prerogative, perceived expertise, curriculum needs, risk of misinformation, and demands of the learning environment. The study conclusions are detailed in our UMAP'21 paper. For Study II, We propose a larger-scale investigation of stakeholder beliefs around trust, specifically perceptions of competence, benevolence and integrity [49].

7.3 Study III

Study III contends that explanations, a transparency cue to describe stakeholder intent for recommendation users, can improve editorial trust, increase delegation, and facilitate reflection. The study examines three broad types of explanations (guarantees, social proof, content tags), and seeks to investigate which stakeholders and editorial trust contexts respond to these explanations, and why. Evidence of the unique impact of explanations on different trust contexts can empower designers to create recommendation software reusable across departments and tailor it to specific pedagogies.

Study III hypothesizes that explanations would result in overall trust gains, and social proof and content-based explanations might result in larger gains relative to simple editor guarantees. It further hypothesizes that any potential trust gains from explanations will be significantly different for low and high DOU courses (Study I), and for conservative and egalitarian trust intentions (Study II).

Chapter 8

Timeline

In this section, we review the timeline of key milestones achieved and planned for all three studies in the proposed dissertation.

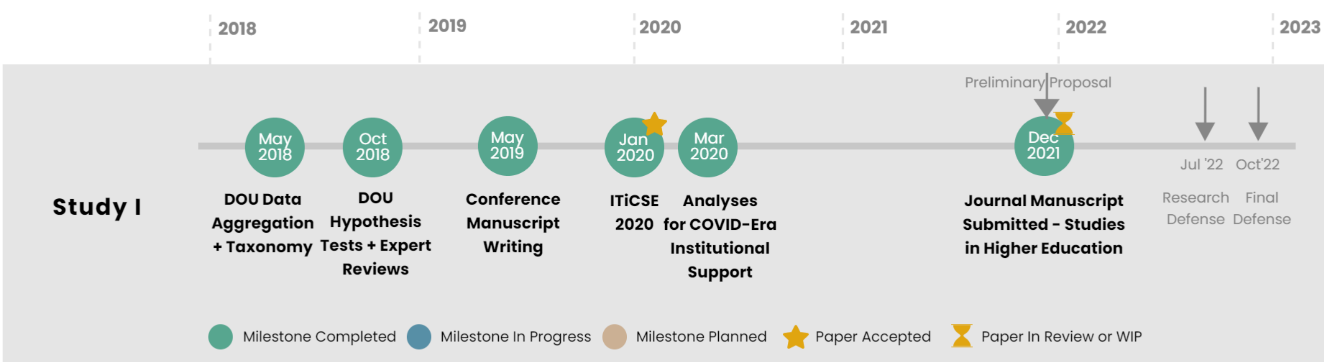

Figure 8.1: Study I Timeline

8.1 Study I

In Study I, we petitioned and aggregated all constituent signals of DOU at Virginia Tech between spring 2017 and summer 2018 using web scraping, entity resolution, and Canvas REST API. Using interviews and expert reviews with instructional designers, we compiled the DOU taxonomy and formulated the DOU estimation method in Fall 2018. We completed the first series of DOU analyses on 1327 Computer Science courses, and published our findings in a full paper at ITiCSE'21. In March of 2020, the COVID-19 pandemic spawned a new line of DOU analyses to support campus-wide remote teaching. A journal manuscript describing DOU analyses for nearly 20,000 Canvas courses, along with case studies in DOU-based administrative and design support was completed and submitted to *Studies in Higher Education* in fall of 2021.

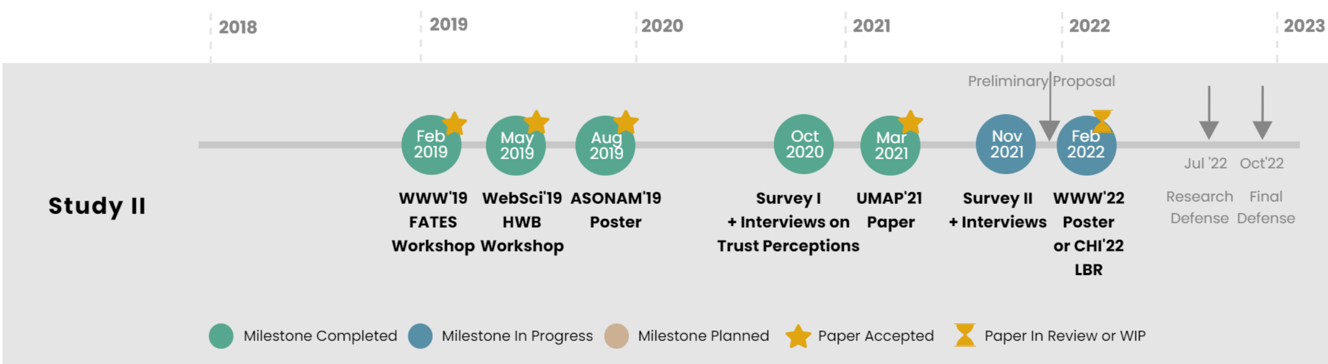

Figure 8.2: Study II Timeline

8.2 Study II

In Study II, we began by examining social media signatures of student motives, perceptions and expectations expressed in instructor ratings from two online forums (*Koofers* and *RateMyProfessors*). Findings from these inquiries appeared in two workshop papers at WWW 2019 and Web Science 2019 and a poster at ASONAM 2019. In Fall of 2020, we ran a survey on stakeholder trust perceptions and editorial task allocations, and conducted a preliminary round of interviews (N=63). Results from our analyses appeared in a full paper at UMAP 2021. A late-breaking work submission to CHI'22 (hedge: WWW'22 poster track) is planned for Jan 2022, based on an extended ongoing survey and a new round of stakeholder interviews.

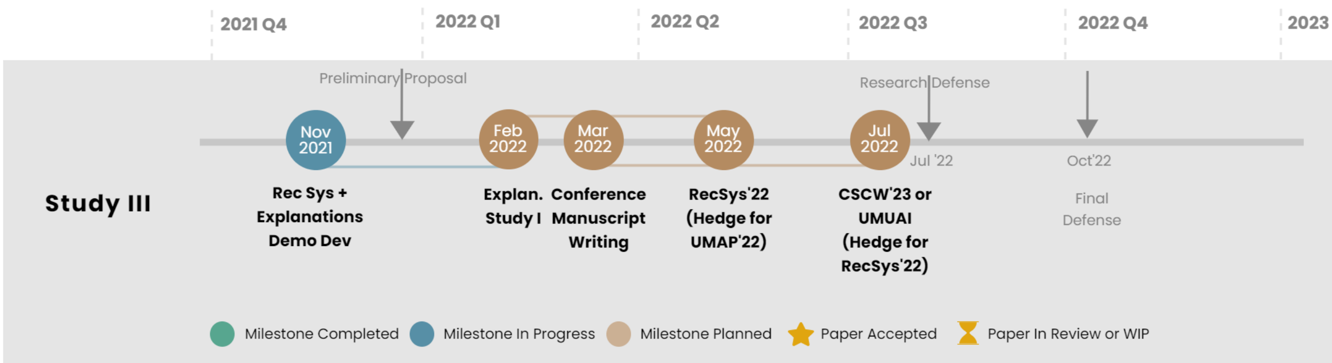

Figure 8.3: Study III Timeline

8.3 Study III

Study III seeks to investigate the effect of explanations on editorial trust beliefs and intentions. We propose to assess between-group differences for three explanation types in a live online study with an educational recommender system prototype. The study is due to begin start of Spring 2022 academic term (late Jan'22), pending ongoing development and testing of the prototype (Nov'21 - Jan'22). A conference manuscript is planned for RecSys'22 (deadline May'22) with writing to commence in Mar'22. Dissertation writing and a compilation of analyses from all three studies should proceed concurrently, with the research and final defense of this dissertation tentatively planned for July and October of 2022, respectively.

Table 8.1 summarizes the in-progress and planned milestones in the remainder of the dissertaion project studies.

Table 8.1: Dissertation Project Timeline (Dec 2021 - Oct 2022)

Month/Year	Project Task/Milestone	Study	Venue
Dec '21	Preliminary Proposal Defense		
Nov '21-Jan '22	IRB, Demo Development	III	
Feb '22	Conference LBR/Poster	II	CHI'22 (Hedge: WWW'22)
Jan-Apr '22	Live Recommender Use Study I	III	
May '22	Conference Manuscript	III	RecSys'22 (Hedge: CSCW'23)
Jul 22	Conference Manuscript	III	CSCW'23 (Hedge: UMUAI)
Jul '22	Research Defense		
Jul-Oct '22	Dissertation Writing		
Oct 2022	Final Defense		

Bibliography

[1] Himan Abdollahpouri, Gediminas Adomavicius, Robin Burke, Ido Guy, Dietmar Jannach, Toshihiro Kamishima, Jan Krasnodebski, and Luiz Pizzato. Multistakeholder recommendation: Survey and research directions. *User Modeling and User-Adapted Interaction*, 30(1):127–158, 2020.

[2] Tella Adeyinka. Reliability and factor analysis of a blackboard course management system success: A scale development and validation in an educational context. *Journal of Information Technology Education: Research*, 10:55–80, 2011.

[3] Tella Adeyinka and Stephen Mutula. A proposed model for evaluating the success of webct course content management system. *Computers in Human Behavior*, 26(6): 1795–1805, 2010.

[4] Kimberly E Arnold and Matthew D Pistilli. Course signals at purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd international conference on learning analytics and knowledge*, pages 267–270. ACM, 2012.

[5] Donovan Artz and Yolanda Gil. A survey of trust in computer science and the semantic web. *Journal of Web Semantics*, 5(2):58–71, 2007.

[6] Erik W Black, Kara Dawson, and Jason Priem. Data for free: Using lms activity logs to measure community in online courses. *The Internet and Higher Education*, 11(2): 65–70, 2008.

[7] Sviatoslav Braynov and Tuomas Sandholm. Contracting with uncertain level of trust. *Computational Intelligence*, 18(4):501–514, 2002.

[8] Iván Cantador and Pablo Castells. Group recommender systems: new perspectives in the social web. In *Recommender systems for the social web*, pages 139–157. Springer, 2012.

[9] María José Casany Guerrero, Marc Alier Forment, Nikolaos Galanis, Enric Mayol Sarroca, and Jordi Piguillem Poch. Analyzing moodle/lms logs to measure mobile access. In *UBICOMM 2012: The Sixth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, pages 35–40, 2012.

[10] Shruthi Chari, Oshani Seneviratne, Daniel M Gruen, Morgan A Foreman, Amar K Das, and Deborah L McGuinness. Explanation ontology: A model of explanations for user-centered ai. In *International Semantic Web Conference*, pages 228–243. Springer, 2020.

[11] Hamish Coates, Richard James, and Gabrielle Baldwin. A critical examination of the effects of learning management systems on university teaching and learning. *Tertiary education and management*, 11:19–36, 2005.

[12] Mihaela Cocea and Stephan Weibelzahl. Can log files analysis estimate learners' level of motivation? In *LWA*. University of Hildesheim, Institute of Computer Science, 2006.

[13] Mihaela Cocea and Stephan Weibelzahl. Cross-system validation of engagement prediction from log files. In *European Conference on Technology Enhanced Learning*, pages 14–25. Springer, 2007.

[14] Shane Dawson. 'seeing'the learning community: An exploration of the development of a resource for monitoring online student networking. *British Journal of Educational Technology*, 41(5):736–752, 2010.

[15] William H DeLone and Ephraim R McLean. Information systems success: The quest for the dependent variable. *Information systems research*, 3(1):60–95, 1992.

[16] William H Delone and Ephraim R McLean. The delone and mclean model of information systems success: a ten-year update. *Journal of management information systems*, 19 (4):9–30, 2003.

[17] edutechnica. 6th Annual LMS Data Update, 2018. URL https://edutechnica.com/2018/10/06/6th-annual-lms-data-update/.

[18] Asmaa Elbadrawy, R Scott Studham, and George Karypis. Collaborative multi-regression models for predicting students' performance in course activities. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, pages 103–107. ACM, 2015.

[19] Asmaa Elbadrawy, Huzefa Rangwala, Aditya Johri, and George Karypis. SDM 2017 Tutorials: Opportunities, challenges and methods for higher education data mining, 2017. URL https://www-users.cs.umn.edu/~elbad004/sdm2017_tutorial/higher_edm_SDM_2017.pdf.

[20] Nafsaniath Fathema, David Shannon, and Margaret Ross. Expanding the technology acceptance model (tam) to examine faculty use of learning management systems (lmss) in higher education institutions. *Journal of Online Learning & Teaching*, 11(2), 2015.

[21] Gianni Fenu, Mirko Marras, and Massimiliano Meles. A learning analytics tool for usability assessment in moodle environments. *Journal of e-Learning and Knowledge Society*, 13(3), 2017.

[22] Rebecca Ferguson. Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6):304–317, 2012.

[23] Martin Fishbein and Icek Ajzen. Belief, attitude, intention, and behavior: An introduction to theory and research. 1977.

[24] Gerhard Friedrich and Markus Zanker. A taxonomy for generating explanations in recommender systems. *AI Magazine*, 32(3):90–98, 2011.

[25] Mike Gartrell, Xinyu Xing, Qin Lv, Aaron Beach, Richard Han, Shivakant Mishra, and Karim Seada. Enhancing group recommendation by incorporating social relationship interactions. In *Proceedings of the 16th ACM international conference on Supporting group work*, pages 97–106, 2010.

[26] Fatih Gedikli, Dietmar Jannach, and Mouzhi Ge. How should i explain? a comparison of different explanation types for recommender systems. *International Journal of Human-Computer Studies*, 72(4):367–382, 2014.

[27] David Gefen. Reflections on the dimensions of trust and trustworthiness among online consumers. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 33(3):38–53, 2002.

[28] Junpeng Guo, Yanlin Zhu, Aiai Li, Qipeng Wang, and Weiguo Han. A social influence approach for group user modeling in group recommendation systems. *IEEE Intelligent systems*, 31(5):40–48, 2016.

[29] Taha Hasan, Naveed Arshad, Erik Dahlquist, and Scott McCrickard. Collaborative filtering for household load prediction given contextual information. In *Proceedings of the 2017 SIAM Workshop on Machine Learning for Recommender Systems (MLRec)*, 2017.

[30] Taha Hassan. On bias in social reviews of university courses. In *Companion Publication of the 10th ACM Conference on Web Science*, pages 11–14. ACM, 2019.

[31] Taha Hassan. Trust and trustworthiness in social recommender systems. In *Companion Proceedings of The 2019 World Wide Web Conference*, pages 529–532, 2019.

[32] Taha Hassan, Bob Edmison, Larry Cox II, Matt Louvet, Daron Williams, and D. Scott McCrickard. Depth of use: an empirical framework to help faculty gauge the relative impact of learning management system tools. In *Proceedings of the 25th ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '20)*. ACM, 2020.

[33] Taha Hassan, Bob Edmison, Timothy Stelter, and D Scott McCrickard. Learning to trust: Understanding editorial authority and trust in recommender systems for education. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, pages 24–32, 2021.

[34] Jiazhen He, James Bailey, Benjamin IP Rubinstein, and Rui Zhang. Identifying at-risk students in massive open online courses. In *AAAI*, pages 1749–1755, 2015.

[35] Curtis R Henrie, Robert Bodily, Kristine C Manwaring, and Charles R Graham. Exploring intensive longitudinal measures of student engagement in blended learning. *The International Review of Research in Open and Distributed Learning*, 16(3), 2015.

[36] Jiun-Yin Jian, Ann M Bisantz, and Colin G Drury. Foundations for an empirically determined scale of trust in automated systems. *International journal of cognitive ergonomics*, 4(1):53–71, 2000.

[37] Sepandar D Kamvar, Mario T Schlosser, and Hector Garcia-Molina. The eigentrust algorithm for reputation management in p2p networks. In *Proceedings of the 12th international conference on World Wide Web*, pages 640–651, 2003.

[38] Bart P Knijnenburg and Martijn C Willemsen. Evaluating recommender systems with user experiments. In *Recommender Systems Handbook*, pages 309–352. Springer, 2015.

[39] Bart P Knijnenburg, Niels JM Reijmer, and Martijn C Willemsen. Each to his own: how different users call for different interaction methods in recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems*, pages 141–148, 2011.

[40] William H Kruskal and W Allen Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260):583–621, 1952.

[41] Hao Ma, Irwin King, and Michael R Lyu. Learning to recommend with social trust ensemble. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 203–210, 2009.

[42] Leah P Macfadyen and Shane Dawson. Mining lms data to develop an "early warning system" for educators: A proof of concept. *Computers & education*, 54(2):588–599, 2010.

[43] Nikos Manouselis, Hendrik Drachsler, Riina Vuorikari, Hans Hummel, and Rob Koper. Recommender systems in technology enhanced learning. In *Recommender systems handbook*, pages 387–415. Springer, 2011.

[44] Florence Martin, Ting Sun, and Carl D Westine. A systematic review of research on online teaching and learning from 2009 to 2018. *Computers & education*, 159:104009, 2020.

[45] Paolo Massa and Paolo Avesani. Trust-aware collaborative filtering for recommender systems. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*, pages 492–508. Springer, 2004.

[46] Paolo Massa and Paolo Avesani. Trust-aware recommender systems. In *Proceedings of the 2007 ACM conference on Recommender systems*, pages 17–24, 2007.

[47] Riccardo Mazza and Christian Milani. Gismo: a graphical interactive student monitoring tool for course management systems. In *International Conference on Technology Enhanced Learning, Milan*, pages 1–8, 2004.

[48] Tanya J McGill and Jane E Klobas. A task–technology fit view of learning management system impact. *Computers & Education*, 52(2):496–508, 2009.

[49] D Harrison McKnight, Vivek Choudhury, and Charles Kacmar. Developing and validating trust measures for e-commerce: An integrative typology. *Information systems research*, 13(3):334–359, 2002.

[50] Lik Mui, Mojdeh Mohtashemi, and Ari Halberstadt. A computational model of trust and reputation. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, pages 2431–2439. IEEE, 2002.

[51] Eric WT Ngai, JKL Poon, and YHC Chan. Empirical examination of the adoption of webct using tam. *Computers & Education*, 48(2):250–267, 2007.

[52] Sevgi Ozkan and Refika Koseler. Multi-dimensional students' evaluation of e-learning systems in the higher education context: An empirical investigation. *Computers & Education*, 53(4):1285–1296, 2009.

[53] Cristóbal Romero and Sebastián Ventura. Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6):601–618, 2010.

[54] Cristóbal Romero, Sebastián Ventura, and Enrique García. Data mining in course

management systems: Moodle case study and tutorial. *Computers & Education*, 51(1): 368–384, 2008.

[55] Guido Rößling and Andreas Kothe. Extending moodle to better support computing education. *ACM SIGCSE Bulletin*, 41(3):146–150, 2009.

[56] Shunichi Seko, Takashi Yagi, Manabu Motegi, and Shinyo Muto. Group recommendation using feature space representing behavioral tendency and power balance among members. In *Proceedings of the fifth ACM conference on Recommender systems*, pages 101–108, 2011.

[57] Yue Shi, Martha Larson, and Alan Hanjalic. Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges. *ACM Computing Surveys (CSUR)*, 47(1):1–45, 2014.

[58] Nava Tintarev and Judith Masthoff. Designing and evaluating explanations for recommender systems. In *Recommender systems handbook*, pages 479–510. Springer, 2011.

[59] Patricia Victor, Martine De Cock, and Chris Cornelis. Trust and recommendations. In *Recommender systems handbook*, pages 645–675. Springer, 2011.

[60] Weiquan Wang and Izak Benbasat. Recommendation agents for electronic commerce: Effects of explanation facilities on trusting beliefs. *Journal of Management Information Systems*, 23(4):217–246, 2007.

[61] Richard E West, Greg Waddoups, and Charles R Graham. Understanding the experiences of instructors as they adopt a course management system. *Educational Technology Research and Development*, 55(1):1–26, 2007.

[62] Diane Wilcox, Jane Thall, and Oris Griffin. One canvas, two audiences: How faculty and students use a newly adopted learning management system. In *Society for Infor-*

mation Technology & Teacher Education International Conference, pages 1163–1168. Association for the Advancement of Computing in Education (AACE), 2016.

[63] Annika Wolff, Zdenek Zdrahal, Andriy Nikolov, and Michal Pantucek. Improving retention: predicting at-risk students by analysing clicking behaviour in a virtual learning environment. In *Proceedings of the third international conference on learning analytics and knowledge*, pages 145–149. ACM, 2013.