

A7134: Textbook Open Knowledge Network

Michael Genesereth
genesereth@stanford.edu

Richard Baraniuk
richb@rice.edu

Mark Musen
musen@stanford.edu

Craig Heller
hcheller@stanford.edu

Vinay K. Chaudhri
vinayc@stanford.edu



Overview

College students today face the challenge of mastering concepts in the new subject areas, relating those concepts across multiple disciplines, and one size fits all nature of textbooks. Intelligent Textbooks (ITB) using Artificial Intelligence (AI) and knowledge graphs (KG) solve these problems by allowing students to dynamically interact with the textbook content, increasing their ability to understand concepts, increasing engagement, and thereby, improving academic performance. ITBs offer students easy access to definitions and descriptions of concepts, make connections across different sections of the syllabus, and allow students to pose their own questions.

Description

Initial trials of ITBs that utilize KGs have been found to improve student grade outcomes by a full letter grade over the control group that was using a conventional textbook. ITBs have been found especially helpful for underperforming students, thus, broadening participation.

The potential of ITBs to facilitate better learning has been extremely difficult to realize without major investments of time, money, and expertise. The reason is that KGs are currently constructed using human subject-matter experts in a process that is extremely expensive and time consuming. Due to the large investment required, publishers and ed tech providers keep their KGs proprietary, eliminating their utility outside of the scope of the project for which they were created.

The primary outcome of our project is an open source Textbook Open Knowledge Network (TOKN) that can be freely used for creating ITBs and a variety of education technology applications.

A secondary outcome is a novel process and tools for creating the KGs that combine automatic construction of a KG with validation by humans to ensure high accuracy. We envision a community of educators who would co-create TOKN, and eventually take the ownership for its future development and evolution.

We will use the new approach for KG construction developed in our project for Biology and Psychology undergraduate textbooks that are published by the open source publisher OpenStax. We will eventually integrate the ITB technology into the full OpenStax library of 41 textbooks, which has the potential to influence millions of students. OpenStax has further developed a robust ecosystem of 54 commercial partners (including every major publisher save one) who use OpenStax materials within their educational technologies and a further 30 major college and university systems who are using OpenStax to institutionalize open educational resources (OER) at their campuses.

Our KG construction tools will be integrated with Stanford's Protégé environment. Protégé is one of the most widely used knowledge authoring environments in the world today. In 2019 alone, Protégé was downloaded 145,000 times, and its web version hosts over 70,000 projects with over 50,000 user accounts.

Differentiators

An ITB relies on an explicit representation of knowledge in a textbook that matches human understanding and enables precise reasoning with it. The current narrative in AI is dominated by machine learning and natural language processing that achieve scale by sacrificing either accuracy or expressiveness. Such a compromise is acceptable in applications such as search, recommendation

systems, machine translation, etc. In education, our domain models must be accurate and expressive. The textbooks need to be nearly 100% accurate. Any computer-based tools that will eventually be as good as human tutors have to use an explicit model of the knowledge of the domain. These characteristics are also shared by many other applications, for example, income tax calculations and automated enforcements of a contract.

Our team is a world leader in research, application and education of formal knowledge representation methods. We have created strong partnerships with non-profit and for-profit organizations in the field of textbook publishing, educational technology, and academic search, thus, substantially enhancing the probability of success of our Phase II effort.

Road Map

Activity	Milestone
KG Construction Tool	9 months
KG Tool Evaluation for Biology	12 months
KG Tool Evaluation for Psychology	15 Months
ITB incorporating the KG	18 months
ITB Evaluation	24 months

There will be two primary products of our project: a KG construction tool, and TOKEN that contains KGs for a few chapters from Biology and Psychology textbooks from OpenStax. We will evaluate the KG construction tool for OpenStax Biology and Psychology textbooks, and measure the time and effort required per chapter. We will evaluate the resulting KGs by incorporating them into an ITB and measure its impact on student learning and engagement.

Partnerships

The foundation of the team is a partnership between Stanford University and OpenStax at Rice University. We bring together strong prior research on ITBs and KGs from Stanford with the vision to reimagine and reinvent textbooks from OpenStax.

To ensure that our innovations impact the commercial textbook industry, we have formed a collaborative relationship with Macmillan Learning. They will evaluate ITBs for two of their textbooks, gather user feedback, and help guide our future development.

To ensure that TOKEN can be useful outside the context of an ITB, and for educational technology products, we have teamed with Educational Testing Service (ETS). ETS will specify the requirements of KGs for item generation, and once the KG is developed, evaluate it for item generation and question scoring outside the context of an ITB.

To leverage and contribute to publicly available Big Data, we have formed a collaboration with Microsoft Academic Graph (MAG). In our collaboration with MAG team, we will investigate if their technology can help bootstrap the taxonomy for TOKEN, and if TOKEN could be incorporated into MAG ensuring its wide usage and contribution to the task of academic search.

Intellectual Property

TOKEN, the KG construction platform, and new ITB applications will be released under an open source license. The existing ITB platform, called Inquire, is owned by Vulcan Inc, and Stanford has rights to sublicense it. Automatic Item Generation algorithms are owned by ETS and will be licensed to the Stanford team.