

Stifterverband Abschlussbericht - “Semi-automatisch generierte Leseverständnistests”

Projektziel und Motivation

In vielen Lehrveranstaltungen wird zusätzlich zu den Vorlesungen ergänzende und vertiefende Literatur bereitgestellt. Dieses Angebot wird, auch nach eigener Erfahrung, von den Studierenden jedoch oft nicht wahrgenommen. Leseverständnistests gelten als ein Weg, Studierende zur Beschäftigung mit der zu einer Vorlesung bereitgestellten Literatur zu motivieren (Howard, 2004). Die Studierenden können so eigene Wissenslücken frühzeitig erkennen und setzen sich mit dem Vorlesungsstoff zusätzlich auseinander. Darüber hinaus stellen sie auch dem Lehrenden Feedback über mögliche Verständnisprobleme und Wiederholungsbedarf bei den Studierenden zur Verfügung. In der digitalen Hochschullehre bieten sich nun zwei Ansatzpunkte zur Verbesserung:

1) **Verbesserte Bereitstellung** Damit Leseverständnistests überhaupt wirksam werden können, müssen diese erst einmal vorhanden sein. Der Zusatzaufwand für Lehrende zur Generierung solcher Tests ist jedoch erheblich, weswegen diese oft nicht oder nicht in ausreichendem Umfang eingesetzt werden.

2) **Verbesserte Integration** Die Tests müssen möglichst gut in den Lehrablauf integriert werden. Freiwillige oder vertiefende Angebote werden kaum genutzt, wenn diese als nicht prüfungsrelevant erscheinen.

Unser Fragengenerierungsprojekt adressiert beide Punkte, indem zum einen ein System zur semi-automatischen Generierung von Leseverständnistests zur Verfügung gestellt wird, zum anderen die Integration in bestehende eLearning-Umgebungen sichergestellt und evaluiert wird.

Maßnahmen

Nun soll beschrieben werden, wie aus technischer Sicht zu einem bereitgestellten Text Fragen bestimmt werden. Hierzu wird anhand des folgenden Beispieltextes illustriert, welche Schritte bei der Erstellung von Fragen durchlaufen werden.

“KI gehört zu den neueren Gebieten in Wissenschaft und Technik. Ernsthafte Arbeiten begannen unmittelbar nach dem Zweiten Weltkrieg und ihr Name wurde 1956 geprägt. Zusammen mit der Molekularbiologie wird KI von Wissenschaftlern in anderen Disziplinen regelmäßig als das Gebiet bezeichnet “in dem ich am liebsten arbeiten würde”. Ein Student der Physik mag das berechtigte Gefühl haben, dass alle guten Ideen bereits von Galileo, Newton, Einstein und anderen vor ihm gedacht wurden. KI dagegen hält noch viele offene Türen für mehrere Vollzeit-Einsteins und -Edisons bereit.”

Dieser Text stammt aus der deutschen Übersetzung des für die Lehrveranstaltung „Grundlagen der künstlichen Intelligenz“ eingesetzten Lehrbuchs „Artificial Intelligence: A Modern Approach“ (Russell & Norvig, 2002). Nach der eher technischen Beschreibung des Systems wird auf die Sicht des Anwenders eingegangen, indem beschrieben wird, welche Möglichkeiten die grafische Oberfläche dem Benutzer für Erstellung, Verwaltung und Export von Fragen zur Verfügung stellt.

Extraktion von Fragen

Zu jedem eingelesenen Text werden als erster Schritt die wichtigsten darin enthaltenen Schlüsselwörter identifiziert, um Fragen stellen zu können, die zentral für den Text sind. Das kann sowohl bedeuten, dass gezielt nach einem solchen zentralen Konzept gefragt wird, dies also die Antwort auf eine Frage ist, oder dass eine Frage zu diesem Konzept gestellt wird, es also in der Frage selbst enthalten ist. Die Schlüsselwörter werden bestimmt, indem den Worten des Textes Werte zugeordnet werden, die ihre Besonderheit ausdrücken sollen. Zur Ermittlung dieser Werte wird die Häufigkeit von Wörtern im Text mit der in einem großen Hintergrundkorpus von Wörtern (Web1T5; Brants and Franz, 2006) verglichen. Bezogen auf den Beispieltext wären *KI* und *Wissenschaft* mögliche Kandidaten für Schlüsselwörter. Sätze, die Schlüsselwörter enthalten, werden dann intern als syntaktische Strukturen repräsentiert, um sie für das System verarbeitbar zu machen. Die Struktur eines Beispielsatzes aus dem Text ist in Abbildung 1 zu sehen. Aus dieser können jetzt unter Anwendung von Transformationsregeln Fragen generiert werden. Für das Englische wurden die hierfür genutzten Regeln von Heilman (2010) übernommen, während für die deutsche Sprache in Anlehnung an das System von Heilman neue Regeln formuliert wurden. Zunächst werden Teile der Satzstruktur markiert, die sich als Antwort auf eine Frage eignen. Im Anschluss werden zu diesen Antwortkandidaten durch weitere Regeln entsprechende Fragen generiert. Bezogen auf den in Abbildung 2 dargestellten Beispielsatz würde etwa *KI* als eine Antwort identifiziert, nach der mit *Was* gefragt werden kann, sodass sich die Frage *Was gehört zu den neueren Gebieten in Wissenschaft und Technik?* ergibt. Die Generation der Fragen folgt einem Übergenerationsansatz, sodass eine Kuration der erstellten Fragen durch den Benutzer erforderlich ist.

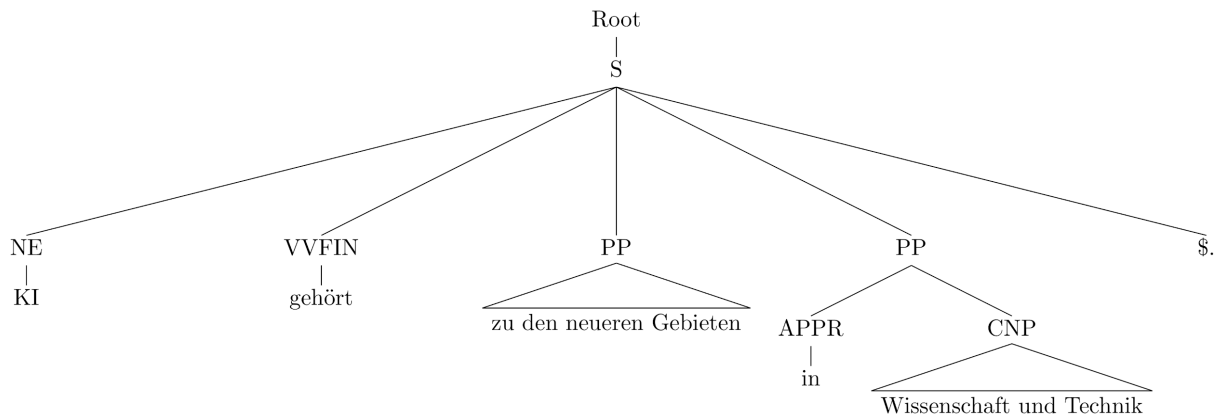


Abbildung 1: Beispiel für die syntaktische Struktur eines Satzes

Anwendung aus Nutzersicht

Die Anwendung lässt sich aus Sicht des Nutzers in zwei Schritte gliedern: das Einlesen von Texten und die anschließende Auswahl von Fragen. Um die Verständlichkeit der Oberfläche zu erhöhen gibt es für die einzelnen Bereiche kleine Fragezeichen-Icons, über die der Nutzer Informationen zur Benutzung des Systems erhalten kann.

Abbildung 4 zeigt die Oberfläche zum Hochladen von Texten. Der Benutzer hat hierüber die Möglichkeit, einen Ausgangstext in Form eines pdf- oder txt-Dokuments zur Verfügung zu stellen oder den Text direkt in ein Textfeld einzugeben. Es werden sowohl deutsche als auch englische Texte unterstützt und die hochgeladenen Texte werden zur Übersicht in einer Tabelle dargestellt. Das Klicken auf "Generate Questions" startet den Extraktionsprozess für alle diese Texte.

The screenshot shows the "Question Generation" interface. It has a "Text Selection" section with radio buttons for "English" and "German" (selected). Below this is an "Upload" button and a text area containing the sentence "KI gehört zu den neueren Gebieten in Wissenschaft und Technik. Ernsthaftige Arbeiten begannen unmittelbar nach dem Zweiten Weltkrieg und ihr Name wurde 1956 geprägt. Zusammen mit der Molekularbiologie wird KI von Wissenschaftlern in". There is another "Upload" button below the text area. To the right is a table with the following data:

Id	Language	Filename	Content
1	de	none	KI gehört zu den neueren Gebieten in Wissenschaft und Technik. Ernsthaftige Arbeiten begannen unmittelbar nach dem Zweiten Weltkrieg und ihr Name wurde 1956 geprägt. Zusammen mit der Molekularbiologie wird KI von Wissenschaftlern in

At the bottom right, there are "Generate Questions" and "Delete Text" buttons.

Abbildung 2: Menü zum Hochladen von Text

Nach erfolgter Extraktion der Fragen werden diese in der in Abbildung 3 gezeigten Oberfläche dargestellt. Die linke Spalte zeigt alle generierten Fragen, aus denen der Anwender unkompliziert diejenigen auswählen kann, die er als geeignet empfindet.

Die ausgewählten Fragen werden dann in einer separaten Liste gesammelt. Natürlich ist es auch möglich Fragen zu löschen oder anzupassen.

Standardmäßig wird die Übersicht generierter und ausgewählter Fragen mit einem einfachen Menü dargestellt, das es dem Nutzer erlaubt durch Seiten von generierten Fragen zu blättern. Dies gibt uns als Entwicklern die Möglichkeit die Fragen so zu ordnen, dass zunächst diejenigen angezeigt werden, von denen vermutet wird, dass der Benutzer sie am sinnvollsten finden wird. Alternativ zu dieser einfachen Oberfläche kann der Benutzer auch auf ein erweitertes Menü umschalten, das Möglichkeiten zur Filterung der generierten Fragen bietet. Dieses Menü ist rechts in Abbildung 3 zu sehen. Es erlaubt die Fragen per Freitextsuche zu filtern oder bestimmte Fragetypen und Schlüsselwörter auszuwählen, die die Fragen enthalten sollen. Dies kann zum Beispiel nützlich sein, wenn es ein spezielles Konzept gibt, zu dem eine Frage gestellt werden soll. Nach abgeschlossener Auswahl und Anpassung der Fragen können die ausgewählten Fragen auf unterschiedliche Art exportiert werden, zum Beispiel als einfaches Textdokument oder als Übungsaufgabe für Moodle oder Jack.

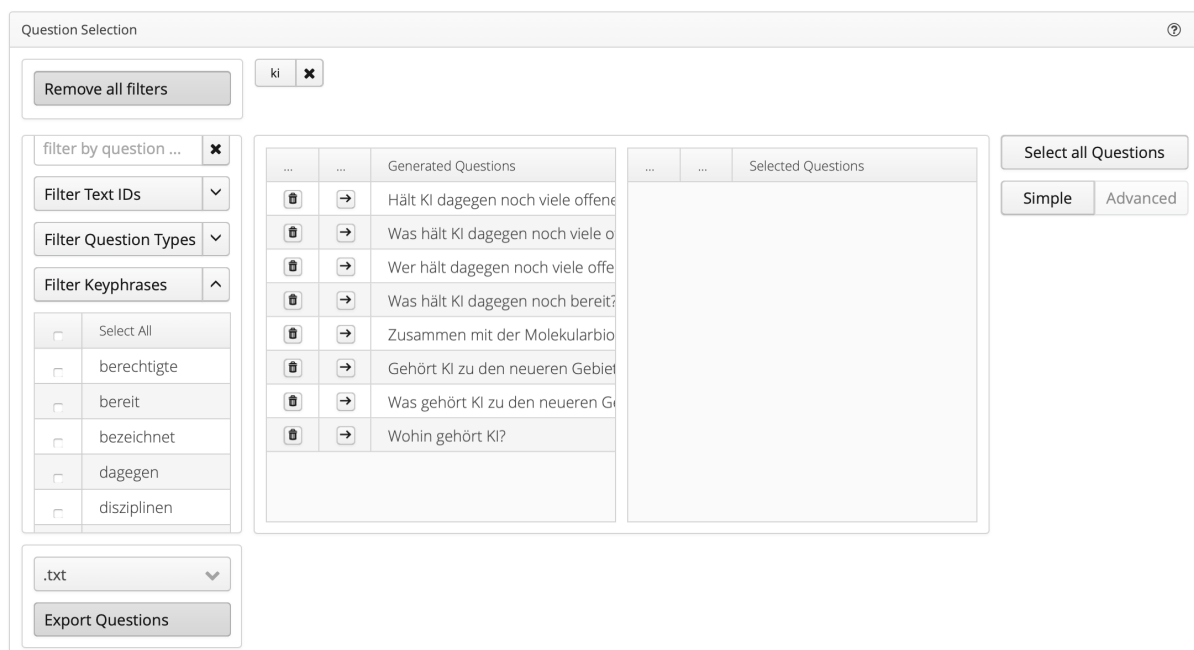


Abbildung 3: Erweitertes Menü zur Auswahl von Fragen.

Evaluation

Die Evaluation des Projekts wird im Wesentlichen zu Beginn des kommenden Sommersemesters stattfinden. Dazu sind zwei Maßnahmen angedacht:

- Einerseits werden wir die Usability des Systems in Userstudien evaluieren. Dazu werden interessierte Nutzer aus dem universitären Umfeld mit dem System interagieren und ihre Erfahrungen rückmelden, die dann wiederum in eine Verbesserung des Systems eingehen können. Evaluiert werden hier insbesondere die intuitive Benutzbarkeit des user-Interface, sowie mögliche technische Probleme.

- Des Weiteren sollen die im System generierten Fragen direkt im Vorlesungskontext erprobt werden. Dazu werden wir automatisch generierte Fragen zu verschiedenen vorlesungsrelevanten Texten in einem moodle-lesbaren Format exportieren und auf der Lernplattform den Studierenden bereitstellen.
Zum Vergleich setzen wir vom Menschen händisch generierte Fragen ein, die im vergangenen Sommersemester zu denselben Texten erstellt und eingesetzt wurden. Evaluiert werden soll dabei insbesondere, welche Fragen von den Studenten als hilfreicher und natürlicher beurteilt werden.

Fazit & Ausblick

Auf technischer Seite ist es uns gelungen ein bestehendes Fragengenerierungssystem (Heilman 2010) auf die deutsche Sprache anzupassen und in ein Userinterface zu integrieren. Diese Software ist als Webservice frei zugänglich und wird auch nach Projektende verfügbar sein.

Bei der ersten Erprobung wurde deutlich, dass die regelbasierte Erstellung von Fragen viele vergleichsweise triviale Fragen nach einzelnen Fakten generiert. Gleichzeitig sind aber oft die nicht-trivialen Fragen diejenigen, die mit einer intensiveren Beschäftigung mit dem Text verlangen.

Ein wünschenswerter nächster Forschungsschritt, der über die Laufzeit des Projekts hinausweist, ist daher die Ausweitung der Fragengenerierung auf nicht-triviale Fragentypen. Da bestehende Datensätze aus dem Bereich der Machine Comprehension, die oft als Trainingsmaterial für die Fragengenerierung genutzt werden, ebenfalls überwiegend einfache faktische Fragen enthalten, ist die Erstellung neuer Ressourcen für didaktisch sinnvolle Fragen eine spannende Frage für anschließende Forschungen.