

I22 Extraktion

TÄTIGKEITEN UND EDV-ERFAHRUNG					
Tätigkeiten					
<i>Gegenstände</i>	<i>Methoden/Verfahren</i>	<i>Ziele/Fragen</i>	<i>Institutioneller Rahmen</i>	<i>Quelle</i>	
digital Edition, digitale Archive, Lehre Masterstudium Wahlfachmodul Informationsmodellierung	angewandte digitale Editionstechnik		Universität, Forschungszentrum, Leiter/in	\I22-1\	
Wissensrepräsentation, ontologiebasierte Modellierung von Wissensdomänen, Annotationssprachen	Datenmodellierung, Semantisierung	nachhaltige Datenqualität und -nutzbarkeit	Universität, Forschungszentrum	\I22-2\	
Erfahrung					
	<i>Beschreibung</i>	<i>Bewertung</i>		<i>Quelle</i>	
	digitale Editionstechnik, digitale Archive	technische Erfahrung zentral, Leiter des Zentrums		\I22-1\	
	Beschäftigung mit Annotationssprachen wie XML und TEI	ganz zentrale Beschäftigung --> große Erfahrung		\I22-2\	
	Eclipse gewohnt	Erfahrung in dem Bereich hoch		\I22-9\	
BEDARFSKATEGORIEN					
K 1 Bedienbarkeit					
<i>Benennung</i>	<i>Beschreibung</i>	<i>Bewertung</i>	<i>Bezug</i>	<i>Quelle</i>	<i>Reduktion</i>
vom Interviewer thematisiert, vom Befragten konkretisiert	Grundvoraussetzung: Fachwissenschaftler müssen kein XML können, arbeiten mit Word und OpenOffice, man kann nicht jedem Fachwissenschaftler auf's	grundlegender Aspekt, wichtiger Punkt	Produktion/Manipulation, Modularität/Anpassbarkeit	\I22-6\	K 1.4b Arbeiten wie in der „normalen“ Textverarbeitung ermöglichen, Vermeidung von spitzen Klammern, XML-Generierung im

I22 Extraktion

	Auge drücken, TEI bzw. XML zu kodieren				Hintergrund Bedienbarkeit der Auszeichnungsebene
selbst thematisiert	Am Anfang steht ein didaktisch moderierter iterativer Prozess, in dem Fachwissenschaftler darauf kommen, was sie wollen können von so einem System, XML als wichtiges Instrument, das uns in der Informationsverarbeitung den Text zurückgegeben hat, Funktionalitäten am semantisierten Text erkennbar, die auch die Bedeutung für das wissenschaftliche Denken zeigen	Wichtiger Aspekt	Produktion/Manipulation, Kollaboration/Kooperation	\I22-7\	K 1.1g Schulungen für technisch weniger versierte Nutzer, Schritt-für-Schritt-Anleitung, Heranführen/Einbinden der Community, didaktischer Prozess, Mehrwerte vermitteln, Testprojekte Benutzerführung / Erlernbarkeit
vom Interviewer thematisiert, vom Befragten konkretisiert	Usability, englischsprachige Benutzeroberfläche, kulturräumliche Spezifika können verloren gehen,	zufriedenstellend, gut, aber aufgrund von Eclipse-Erfahrung "vorbelastet", englischsprachige Benutzeroberfläche kein Problem, aber zusätzliche Ebene, die verunsichern kann	Produktion/Manipulation	\I22-9\	K 1.1d mit mehr EDV-Erfahrung (z.B. mit Eclipse) schnell erlernbar und gut bedienbar Benutzerführung / Erlernbarkeit und K 1.2d Deutschsprachige

I22 Extraktion

					Oberfläche nützlich/notwendig Orientierung / Übersichtlichkeit / Oberflächen- gestaltung
K 2 Rezeption/Zugriff					
<i>Benennung</i>	<i>Beschreibung</i>	<i>Bewertung</i>	<i>Bezug</i>	<i>Quelle</i>	<i>Reduktion</i>
selbst thematisiert	für "prädikate", logische Abfragen Daten erschließen	zentrale Beschäftigung	Produktion/Manipulation,	\I22-2\	K 2.2a Spezifizierbarkeit des Zugriffs, fachspezifische Zugriffsmöglichkeiten, verschiedene Zugriffsperspektiven durch Auszeichnungen und Präsentationsformen (z.B. parallele Ansicht von Faksimile und Transkription im TBLE) Zugriffsformen, Rezeptions- und Recherche- möglichkeiten
vom Interviewer thematisiert, vom Befragten konkretisiert	in der Transliteration des Textes mit Curser bewegen und dabei kontextbezogen Editionskommentare und Lupenansicht der Handschrift sichtbar, wird		Produktion/Manipulation	\I22-12\	K 2.2a Spezifizierbarkeit des Zugriffs, fachspezifische Zugriffsmöglichkeiten, verschiedene Zugriffsperspektiven durch Auszeichnungen

I22 Extraktion

	möglich durch Modellierung				und Präsentationsformen (z.B. parallele Ansicht von Faksimile und Transkription im TBLE) Zugriffsformen, Rezeptions- und Recherche-möglichkeiten
vom Interviewer thematisiert, vom Befragten konkretisiert	Datenorganisation zur Orientierung, Bilden von inhaltlichen Konfigurationen, Modellierung durch den Nutzer	ganz wichtige Form	Kollaboration/Kooperation	\I22-13\	K 2.2b Leichte Zugänglichkeit, Übersicht durch Browser, Suchmasken, Ontologien, Klassifizierungen, Indizierungen, Strukturierungen Zugriffsformen, Rezeptions- und Recherche-möglichkeiten
vom Interviewer thematisiert, vom Befragten konkretisiert	Suche nach grammatischen Strukturen durch entsprechende Auszeichnung möglich, Möglichkeit für neue Erkenntnisse, durch Modellierung bis auf Glyphen-Ebene sind Untersuchungen nach Verzierungsformen, Diakritika-Formen, Superskripten und	potenzial für neue Erkenntnisse	Produktion/Manipulation	\I22-17\	K 2.2a Spezifizierbarkeit des Zugriffs, fachspezifische Zugriffsmöglichkeiten, verschiedene Zugriffsperspektiven durch Auszeichnungen und Präsentationsformen (z.B. parallele Ansicht von Faksimile und Transkription im TBLE)

	Subskripten möglich, Synopsendarstellung von Fassungsvarianten, auf Basis von TEI lassen sich auch neue Repräsentationsformate in Form von Analyseansichten generieren				Zugriffsformen, Rezeptions- und Recherche-möglichkeiten
K 3 Produktion/Manipulation					
<i>Benennung</i>	<i>Beschreibung</i>	<i>Bewertung</i>	<i>Bezug</i>	<i>Quelle</i>	<i>Reduktion</i>
vom Interviewer thematisiert, vom Befragten konkretisiert	Modellierung von Daten und Semantisierung von Texten i.S.v. Wissensrepräsentation und Repräsentation von Wissensprodukten als künftig wichtige Aufgabe von Geisteswissenschaftlern, wichtige Funktionalität von Systemen	Nachhaltigkeit durch Datenqualität, Technologien verändern sich hingegen, künftig wichtig, müssen sich Geisteswissenschaftler stärker damit auseinandersetzen		\I22-2\	K 3.1c Auszeichnungsmöglichkeiten / Informationsanreicherung zur Bearbeitung neuer Fragestellungen, tiefere Auszeichnung, spezifischere Auszeichnung je nach Forschungsinteresse (z.B. Musikwissenschaft), semantische Auszeichnung, Informationsmodellierung XML-Produktion
vom Interviewer thematisiert, vom Befragten konkretisiert	für Word und OpenOffice werden Dokumentvorlagen definiert, mit denen Fachwissenschaftler arbeiten, wir generieren	diesen Weg möchte ich gehen, in vielen Projekten ganz wichtig	Bedienbarkeit, Modularität/Anpassbarkeit, Kollaboration/Kooperation	\I22-6\	K 3.1b XML-Auszeichnungs- bzw. Produktionshilfe, Tools zur digitalen Herstellung und Verfügbarmachung von

I22 Extraktion

	automatisiert TEI-Dokumente daraus über OpenDocument-Format,				Texten, XML-Schemata vorgefertigt für verschiedene individuelle Bedürfnisse, Konvertierung in XML/TEI, Baseline Encoding, Streaming Editor, Unterstützung bei zunächst geringerer Auszeichnungstiefe XML-Produktion
selbst thematisiert	XML als Informationsverarbeitung am Text	Potenzial, wichtige Erkenntnis	Rezeption/Zugriff	\I22-7\	K 3.1c Auszeichnungsmöglichkeiten / Informationsanreicherung zur Bearbeitung neuer Fragestellungen, tiefere Auszeichnung, spezifischere Auszeichnung je nach Forschungsinteresse (z.B. Musikwissenschaft), semantische Auszeichnung, Informationsmodellierung XML-Produktion
vom Interviewer thematisiert, vom Befragten konkretisiert	Auszeichnung als Basis neuer Zugriffsformen, z.B. auch neue Repräsentationsformate in Form von Analyse-	potenzial für neue Erkenntnisse, neues wissenschaftliches Arbeiten	Rezeption/Zugriff	\I22-17\	K 3.1c Auszeichnungsmöglichkeiten / Informationsanreicherung zur Bearbeitung neuer

I22 Extraktion

	Ansichten				Fragestellungen, tiefere Auszeichnung, spezifischere Auszeichnung je nach Forschungsinteresse (z.B. Musikwissenschaft), semantische Auszeichnung, Informationsmodellierung XML-Produktion
K 4					
Modularität/Anpassbarkeit					
<i>Benennung</i>	<i>Beschreibung</i>	<i>Bewertung</i>	<i>Bezug</i>	<i>Quelle</i>	<i>Reduktion</i>
selbst thematisiert	Verarbeitbarkeit der modellierten Daten mit verschiedenen Frameworks und Technologien		Kollaboration/Kooperation, Produktion/Manipulation	\I22-2\	K 4.3a Anschließbarkeit/ Nachnutzung von Ressourcen, Formate mappen Inhalte
selbst thematisiert	standardisiertes Tool als modulhafte, erweiterbare Plattform, wo unterschiedlichste Werkzeuge einfließen können	zutiefst sinnvoll	Kollaboration/Kooperation	\I22-3\	K 4.2a Anschließbarkeit und Kombinierbarkeit von Tools und Services, auch von selbstentwickelten – modulares Konzept Tools und Services
selbst thematisiert	Modularität/Anpassbarkeit von Tools an spezifische Annotationsformen, z.B.	diesen Weg gehen wir	Produktion/Manipulation, Bedienbarkeit	\I22-6\	K 4.2b Anpassbarkeit von Tools an das jeweilige Projekt /

I22 Extraktion

	das Leidener Klammersystem, Tool zur Überführung in ein XML-Dokument für die Weiterverarbeitung				die jeweiligen Gegenstände Tools und Services
vom Interviewer thematisiert, vom Befragten konkretisiert	proprietäre Formate in ein nachhaltiges Format überführen, manchmal "Software-Archäologie", aber in die gewohnte Arbeitsumgebung hineinimplementieren	Gutteil unserer Arbeit	Organisation/Verwaltung	\I22-14\	K 4.3a Anschließbarkeit/ Nachnutzung von Ressourcen, Formate mappen Inhalte
K 5 Kollaboration/Kooperation					
<i>Benennung</i>	<i>Beschreibung</i>	<i>Bewertung</i>	<i>Bezug</i>	<i>Quelle</i>	<i>Reduktion</i>
selbst thematisiert	Geisteswissenschaftler müssen Daten modellieren, Weiterverarbeitung mit Frameworks und Technologien durch Informationstechniker		Produktion/Manipulation, Modularität/Anpassbarkeit	\I22-2\	K 5.1a Austausch von Ressourcen, Nachnutzung Austausch von Inhalten und Werkzeugen
selbst thematisiert	Wir verwenden TEI		Produktion/Manipulation	\I22-2\	K 5.1d Standards, TEI, Unabhängigkeit von Datenformaten Austausch von Inhalten und Werkzeugen
selbst thematisiert	standardisiertes Tool, "nicht jeder erfindet ständig die Welt neu", grafischer Editor wirft	zutiefst sinnvoll, grafischer Editor zu wenig standardbasiert	Modularität/Anpassbarkeit, Bedienbarkeit	\I22-3\	K 5.1c Austausch von Arbeitsmitteln (Tools, Formaten, Skripten usw.)

I22 Extraktion

	allerdings natives Format aus				und K 5.1d Standards, TEI, Unabhängigkeit von Datenformaten Austausch von Inhalten und Werkzeugen
selbst thematisiert	Fachwissenschaftler und Techniker im iterativen Prozess, Techniker erkennen oft nicht die methodologischen Probleme von Geisteswissenschaftlern, unterschätzen die Bedeutung von XML, Geisteswissenschaftler denken weniger kategorial, können am Text arbeiten und Bedeutung für wissenschaftliches Denken erkennen	bedeutender Prozess	Produktion/Manipulation	\I22-7\	K 5.2b Zusammenarbeit zwischen Institutionen, Projekten, auch intern zwischen verschiedenen Stellen, mit Technikern, Workflow Zusammenarbeit
selbst thematisiert	Ressentiments, was das Einstellen von Inhalten angeht, noch keine realen Projekte in TG beim Workshop	etwas "überrascht"	Organisation/Verwaltung	\I22-10\	K 5.1a Austausch von Ressourcen, Nachnutzung und K 5.1b Austausch von Arbeits- bzw. Forschungsergebnissen, Bereitschaft dazu

					Austausch von Inhalten und Werkzeugen und K 5.3a Partizipation/Nutzung als Voraussetzung und K 5.3b Community-Building erforderlich, Umdenken erforderlich, durch Überzeugungsarbeit, Veranstaltungen und Erfahrungsaustausch anregen Community-Building-Aspekte
vom Interviewer thematisiert, vom Befragten konkretisiert	Zusammenarbeit zwischen Bibliothek und Fachwissenschaft, Fachwissenschaftler müssten fachbezogene Metadaten beisteuern, Bibliothek die deskriptiven, die Ressource erschließenden Metadaten, Tool sollte Formen des Zusammenarbeitens unterstützen	schwierig, aber wichtig	Produktion/Manipulation	\I22-11\	K 5.2b Zusammenarbeit zwischen Institutionen, Projekten, auch intern zwischen verschiedenen Stellen, mit Technikern, Workflow Zusammenarbeit

I22 Extraktion

selbst thematisiert	bei vielen Editionsprojekten arbeiten mehrere Personen an einem Dokument	muss unterstützt werden	Organisation/Verwaltung	\I22-11\	K 5.2a Zusammenarbeit in Teams von verschiedenen Orten aus, auch an einem Dokument, verteilte Bearbeitung verschiedener ‚Schichten‘ eines Dokuments usw. Zusammenarbeit
vom Interviewer thematisiert, vom Befragten konkretisiert	Fachwissenschaftler sollten bei der Inhaltsmodellierung immer beraten werden, aber es sollte eine Ebene geben, auf der sie selbst modellieren können	Wichtig	Organisation/Verwaltung, Produktion/Manipulation	\I22-13\	K 5.2b Zusammenarbeit zwischen Institutionen, Projekten, auch intern zwischen verschiedenen Stellen, mit Technikern, Workflow Zusammenarbeit
vom Interviewer thematisiert, vom Befragten konkretisiert	adäquate, flexible Strukturen bzw. Institutionalisierungen zwecks Standardisierung schaffen, Community-Aufbau durch Integrationsfiguren, Erfordernisse an Annotationssprachen zusammen definieren	wichtig, muss sein, es ist Unsinn, wenn jeder alles neu erfindet	Produktion/Manipulation	\I22-15\	K 5.1d Standards, TEI, Unabhängigkeit von Datenformaten Austausch von Inhalten und Werkzeugen und Community-Building erforderlich, Umdenken erforderlich, durch Überzeugungsarbeit,

I22 Extraktion

					Veranstaltungen und Erfahrungsaustausch anregen Community-Building-Aspekte
vom Interviewer thematisiert, vom Befragten konkretisiert	Einbindung in die Lehre, Verknüpfung mit Projekten als Teil des Gesamtkonzepts, Nachwuchsgewinnung, Community-Aufbau, Integration	funktioniert gut, mehr Studenten, langsam fließt etwas zurück		\I22-16\	Community-Building erforderlich, Umdenken erforderlich, durch Überzeugungsarbeit, Veranstaltungen und Erfahrungsaustausch anregen Community-Building-Aspekte
vom Interviewer thematisiert, vom Befragten konkretisiert	interdisziplinäre Synergien durch Austausch auf Auszeichnungsebene	unbedingt, schon immer gesagt, großes Potenzial	Produktion/Manipulation	\I22-18\	K 5.2c Interdisziplinäre Zusammenarbeit Zusammenarbeit
selbst thematisiert	Standardisierung von Workflows, Community-Building auch interdisziplinär, Austausch von Werkzeugen - in einem flexiblen Rahmen, in dem gleichzeitig nicht jeder die Welt neu erschaffen muss	größter Fortschritt, den TG ermöglichen könnte	Produktion/Manipulation	\I22-19\	K 5.1c Austausch von Arbeitsmitteln (Tools, Formaten, Skripten usw.) und K 5.1d Standards, TEI, Unabhängigkeit von Datenformaten Austausch von Inhalten

					und Werkzeugen
K 6 Organisation/Verwaltung					
<i>Benennung</i>	<i>Beschreibung</i>	<i>Bewertung</i>	<i>Bezug</i>	<i>Quelle</i>	<i>Reduktion</i>
selbst thematisiert	filebasierte Strukturierung von Informationen im System, Gegenmodell mit FEDORA objektorientiert, unterschiedliche Datenströme zu einem Objekt zusammenfassen, darauf basierend Methoden definieren und so Inhalte erschließen, logische Einheiten von Transliterationen als TEI, Bildern als Zip und über XSLT-Transformation zu einer HTML- oder PDF-Seite, außerdem Zugriffsrechte auf die unterschiedlichen Entitäten eines Objekts definieren, Objekt-Typen mit Grundstrukturen definieren, auch für Publikation relevant	zu stark filebasiert, müsste in Richtung Objektorientierung gehen	Kollaboration/Kooperation, Rezeption/Zugriff	\I22-4\	K 6.1b Datenspeicherung und –verwaltung, Langzeitarchivierung, verbunden mit Bearbeitungsmöglichkeiten und Nachnutzung Organisation und Verwaltung von Inhalten und K 6.2a Workflow-Organisation, Projekt-Verwaltung Organisation und Verwaltung von Arbeitsprozessen
selbst thematisiert	Archiv als Datenbasis aus TEI-modelliertem Inhalt und Referenzen auf Stylesheets zur Transformation, aus der dynamisch Websites und PDFs generiert werden,	Vorteil, positiv		\I22-5\	K 6.1b Datenspeicherung und –verwaltung, Langzeitarchivierung, verbunden mit Bearbeitungsmöglichkeiten und

I22 Extraktion

	Trennung von Inhalt und Repräsentation, Vorteil: Inhalt schon erstellen, bevor Layout geklärt ist, gleichzeitige Veränderung von tausenden Dokumenten möglich				Nachnutzung und K 6.1d Eigene Sichtbarkeit, Publikationsmöglichkeit Organisation und Verwaltung von Inhalten
selbst thematisiert	Nachhaltigkeit selbstreferenzielle Dokumente, die alle Informationen enthalten, die zum Verständnis notwendig sind, Ressentiments von Fachwissenschaftlern: gedrucktes Buch liegt vor, digitale Strukturen können verschwinden, Hybridedition in einer Übergangsphase, in der Vertrauen wachsen kann, nachhaltige Zitierbarkeit digitaler Ressourcen durch z.B. Persistent Identifier	wichtige Aspekte in einem digitalen Archiv, da wird es hingegen, aber es ist ein langer Weg	Rezeption/Zugriff	\I22-8\	K 6.1b Datenspeicherung und –verwaltung, Langzeitarchivierung, verbunden mit Bearbeitungsmöglichkeiten und Nachnutzung Organisation und Verwaltung von Inhalten
selbst thematisiert	Versionierung von Dokumenten	ganz wichtiger Aspekt	Kollaboration/Kooperation	\I22-11\	K 6.1a Versionenverwaltung Organisation und Verwaltung von Inhalten

I22 Extraktion

selbst thematisiert	Akzeptanz im Hinblick auf Datensicherheit, Export muss ermöglicht werden, nachhaltige Zitierbarkeit auch von verschiedenen Entitäten, z.B. einzelne Münze usw.	größte Problematik		\\I22-20\	<p>K 6.1c Adressierung, Zitierbarkeit, Urheberrecht, Leistungsanerkennung und K 6.1e Export und Löschen von Inhalten</p> <p>Organisation und Verwaltung von Inhalten</p>
---------------------	--	--------------------	--	-----------	---