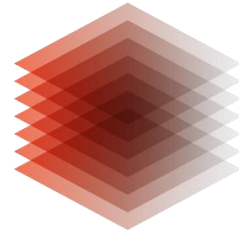

LEIBNIZ-INFORMATIONSZENTRUM
TECHNIK UND NATURWISSENSCHAFTEN
UNIVERSITÄTSBIBLIOTHEK



TIB

Projektvorhaben TEASER

Entwicklung einer Plattform für die kollaborative Thesauruspflege und -verknüpfung am Beispiel „Industrie 4.0“

Dr. Anna Kasprzik
Göttingen, den 11. Mai 2017
Workshop SI & IT der VZG (Projekt coli-conc)

TEASER: Thesaurus Engineering and Alignment Service Environment and Repository



- Projektpartner: TIB Hannover, TH Wildau, Fraunhofer IOSB
- Geplante Laufzeit: 3 Jahre
- Drittmittelbewilligung steht momentan noch aus

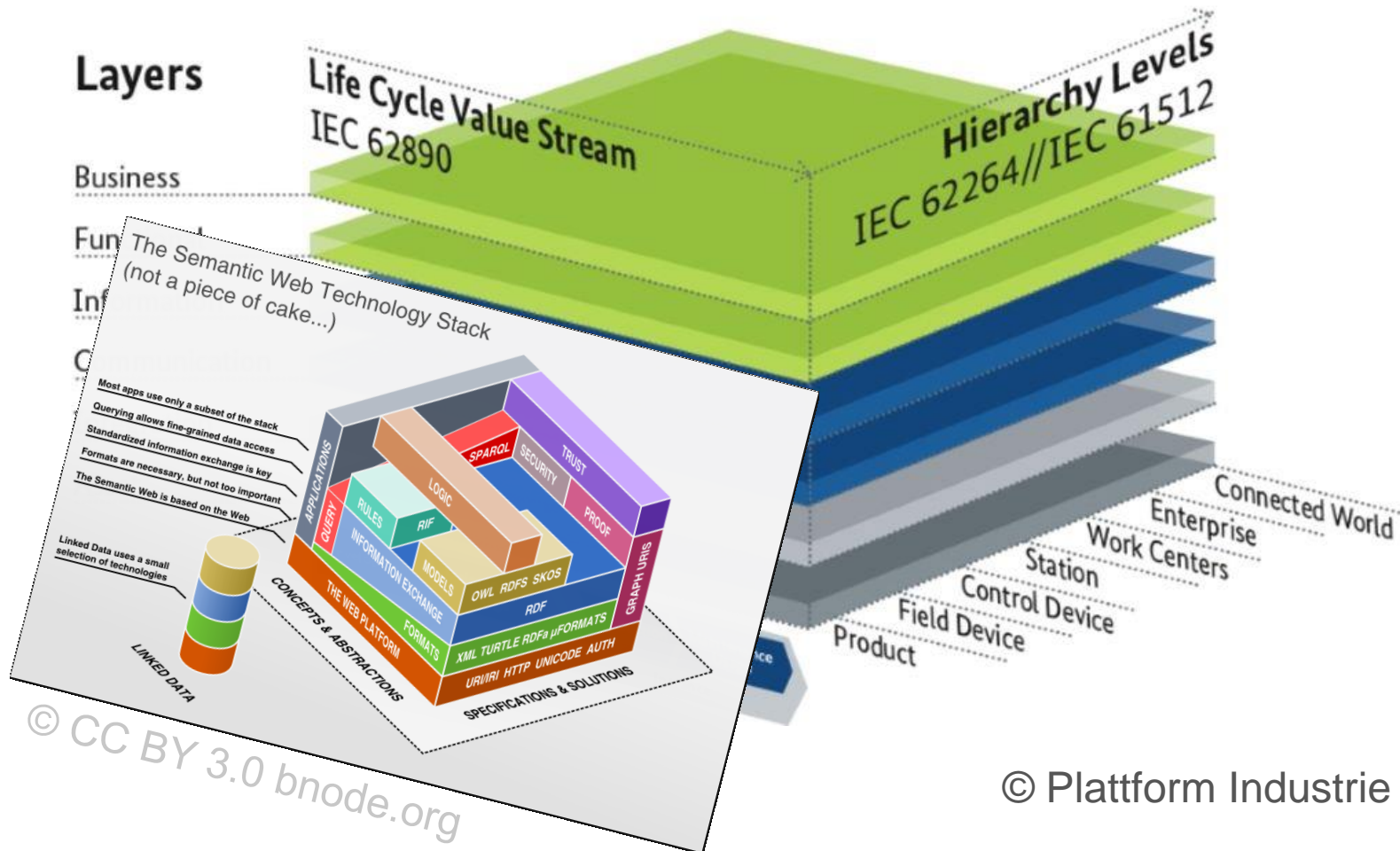


Ausgangslage

- Es gibt weiterhin relativ wenig ausgereifte Fachvokabulare, auch im technisch-naturwissenschaftlichen Bereich
 - Fachvokabulare liegen oft nicht in Standardformaten oder generell **nicht nachnutzbar** bei Institutionen als interne Arbeitsinstrumente vor
 - Eine Vernetzung mit anderen Vokabularen bleibt für kleinere und weniger technikaffine Informationseinrichtungen eine Herausforderung
 - Im Zeitalter von Linked Open Data und Semantic Web haben wir jedoch prinzipiell die Möglichkeiten zur Vernetzung
-  **ALSO** besteht Bedarf an niederschwelligen Angeboten, um eine solche Vernetzung voranzutreiben, auch über Fachgrenzen hinweg, um interdisziplinäre Forschungskontexte bedienen zu können
- Ein Beispiel für ein interdisziplinäres Forschungs- und Anwendungsfeld in Wissenschaft und Wirtschaft ist **Industrie 4.0**

Industrie 4.0: RAMI 4.0, Glossar „Industrie 4.0“

Referenzarchitekturmodell Industrie 4.0 (RAMI 4.0)



Definition Thesaurus (Semantic Web)

- Ein **Thesaurus** im Sinne des **Semantic Web** ist eine Sammlung von **Konzepten**, die mit **Termen (Labeln)** versehen sind und durch Ober- bzw. Untermengenrelationen, Synonymie und weitere semantische Relationen **zueinander in Beziehung gesetzt** sind.
- Ein gängiger Standard für die Darstellung solcher Thesauri ist **SKOS (Simple Knowledge Organization System)** mit seinen Weiterentwicklungen, namentlich SKOS-XL und iso-thes

(siehe dazu auch den neuen ISO-Standard 25964 zu Thesauri und der Interoperabilität mit anderen Vokabularen)

Eigene Vorarbeiten: Die TIB

ist eine Infrastruktureinrichtung der wissenschaftlichen Informationsversorgung mit nationalen Aufgaben. Im Bereich “Wissensorganisationssysteme” und “Linked (Open) Data” engagiert sie sich durch diverse Projekte und Dienstleistungen:

- Bereitstellung von Linked Open Data aus dem Literaturrechercheportal und dem AV-Portal für Videomaterial zur Nachnutzung durch Bibliotheksnutzer
- Weiterentwicklung von VIVO, einem Linked-Open-Data-basierten Forschungsinformationssystem mit integrierter Ontologie zur Datenablage
- **Projektreihe „Fachontologie Technik“**
Weiterentwicklung des Thesaurus „Technik und Management“ (TEMA)
 - Extraktion von Termen aus Volltexten zur Anreicherung von TEMA und GND
 - Entwicklung in Richtung einer Fachontologie im Rahmen eines Kooperationsprojekts mit der WTI-Frankfurt eG am Sandkastenbeispiel „Elektromobilität“

Der TEMA-Thesaurus

- Seit 2013 Weiterentwicklung durch die TIB in Kooperation mit der WTI
- Abgedeckte **Fachgebiete**: Maschinen- und Anlagenbau, Elektrotechnik, Elektronik, Umwelttechnik, Informations- und Kommunikationstechnik, Betriebsführung und -organisation => damit auch **Querschnittsthemen** wie Industrie 4.0, Elektromobilität und IT-Sicherheit
- **bilingual**: deutsch und englisch
- ca. 53.465 Konzepte mit ca. 172.614 Termen
- wird (in Verbindung mit WTI-Fachinformationsdatenbanken) an ca. 130 Universitäten und Fachhochschulen in Deutschland sowie in neun Hochschulen in Österreich und der Schweiz eingesetzt

Eigene Vorarbeiten: Die TH Wildau

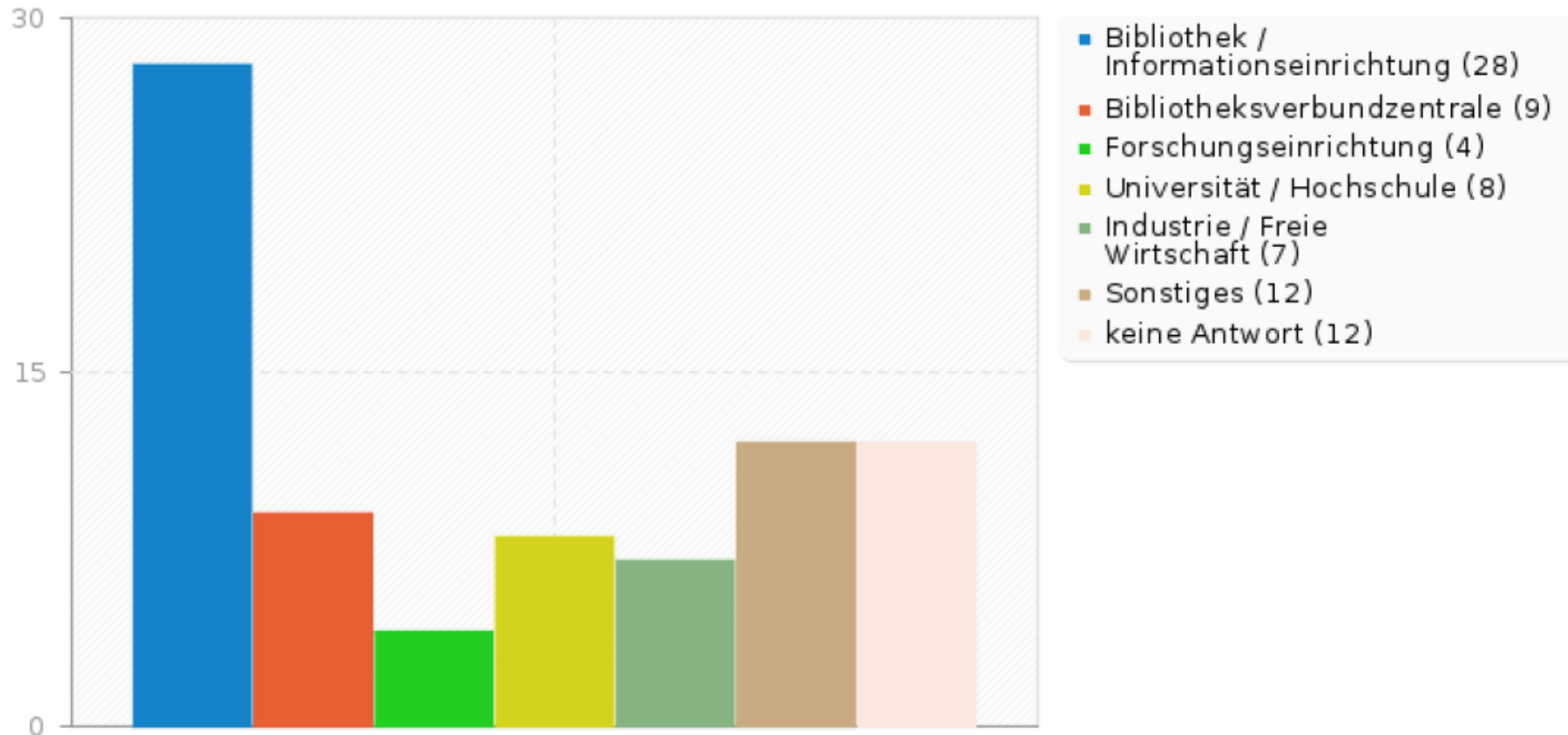
- Interdisziplinäre Forschungsaktivitäten zum Thema „Produktion im Wandel – Industrie 4.0“: Kompetenzzentrum „CPPS/iC3@Smart Production“
- Seit 2013, an der Hochschulbibliothek der TH Wildau:
 - Erstellung eines **standardisierten, onlinefähigen Thesaurus** für die **Fachgebiete** Informatik, Wirtschaft, Technik & Ingenieurwissenschaften und Verkehr
 - **bilingual**: deutsch und englisch
 - 667 Konzepte, 837 deutsche Nicht-Deskriptoren, 835 englische Nicht-Deskriptoren sowie 310 deutschsprachige Definitionen
 - modelliert und bereitgestellt über das Vokabular-Managementsystem „iQvoc“ (Open Source) in **SKOS**
 - kommt im Discovery-System der Bibliothek zum Einsatz

- entwickelt Leitsysteme für industrielle Produktionsanlagen und Systemlösungen für ein durchgängiges **Datenmanagement** auf allen Ebenen der **industriellen Automatisierung** und im “Internet of Things“
- verfügt über **Projekterfahrung** in der Unterstützung bei der Erstellung, Weiterentwicklung, dem Alignment von und Reasoning mit **Ontologien** (BMW-Förderprogramm THESEUS, WP3 Cluster „Ontology Management“)
- **Wissensmanagement-Framework WebGenesis®** mit Schnittstellen und Funktionalitäten zum Import, Editieren und Verknüpfen von Ontologien und SKOS-Thesauri (eigene und über SPARQL-Endpoints zugängliche)
- **Fachexpertise zum Thema “Industrie 4.0”:**
 - Mitwirkung im VDI/VDE-GMA-Fachausschuss 7.21 „Industrie 4.0“
Leitung der UAG „Begriffe“ (Glossar „Industrie 4.0“)
 - Mitwirkung im VDI/VDE-GMA-Fachausschuss 7.20
„Semantik und Interaktionen von Industrie 4.0 Komponenten“
 - DIN SPEC 16593 „Referenzmodell für Industrie 4.0 Service-Architekturen“

Institutionen brauchen Informationen und Zugang zu ...

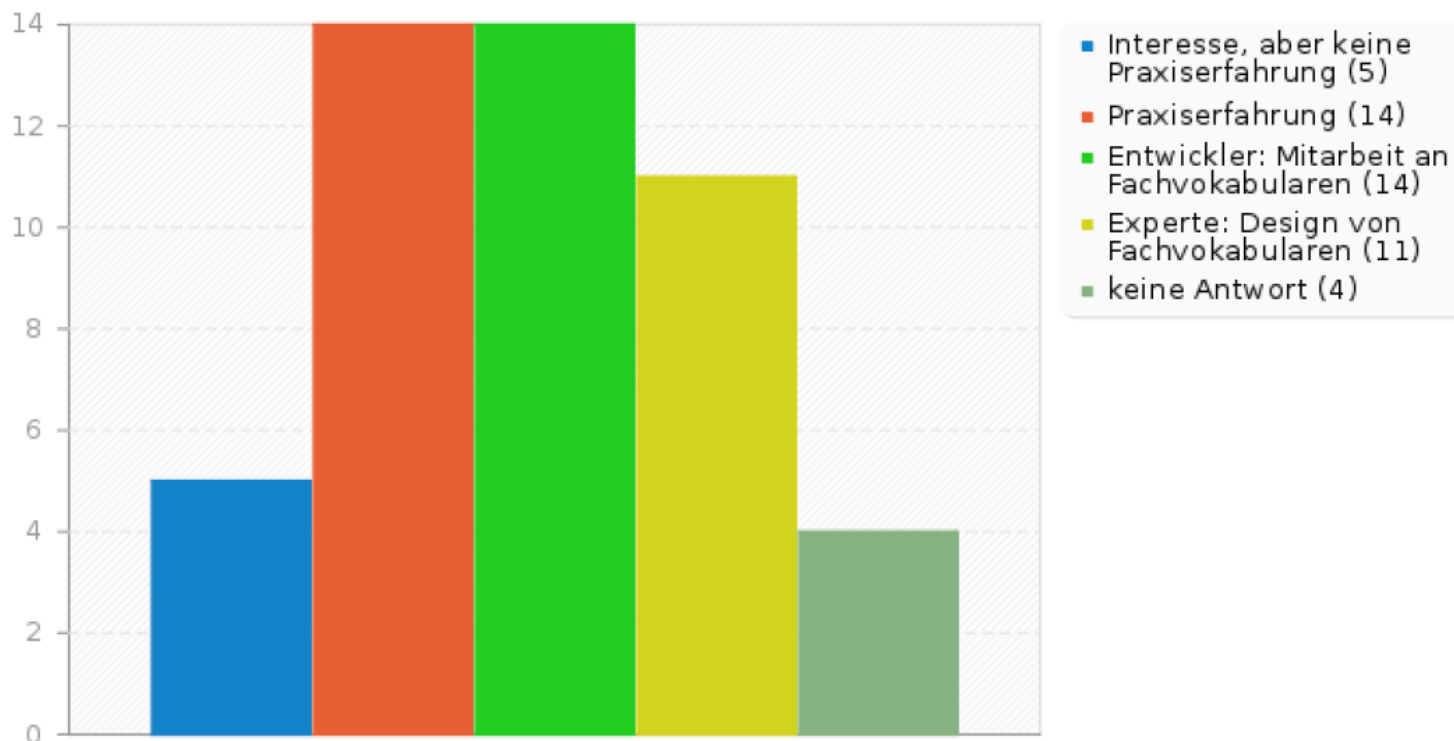
- bewährten Methoden („Best Practices“)
- gängigen und geeigneten Formaten und weiteren Standards im Bereich Fachvokabularerstellung, -pflege und -vernetzung
- Werkzeugen zur Umwandlung in Standardformate und zur Weiterverarbeitung und Integration ins eigene System
- anderen Thesauri und zu technischen Möglichkeiten, Thesauri kollaborativ weiterzuentwickeln, miteinander zu verflechten und die Ergebnisse dieser Aktivitäten zu veröffentlichen

Bedarfsanalyse/Umfrage: Befragte

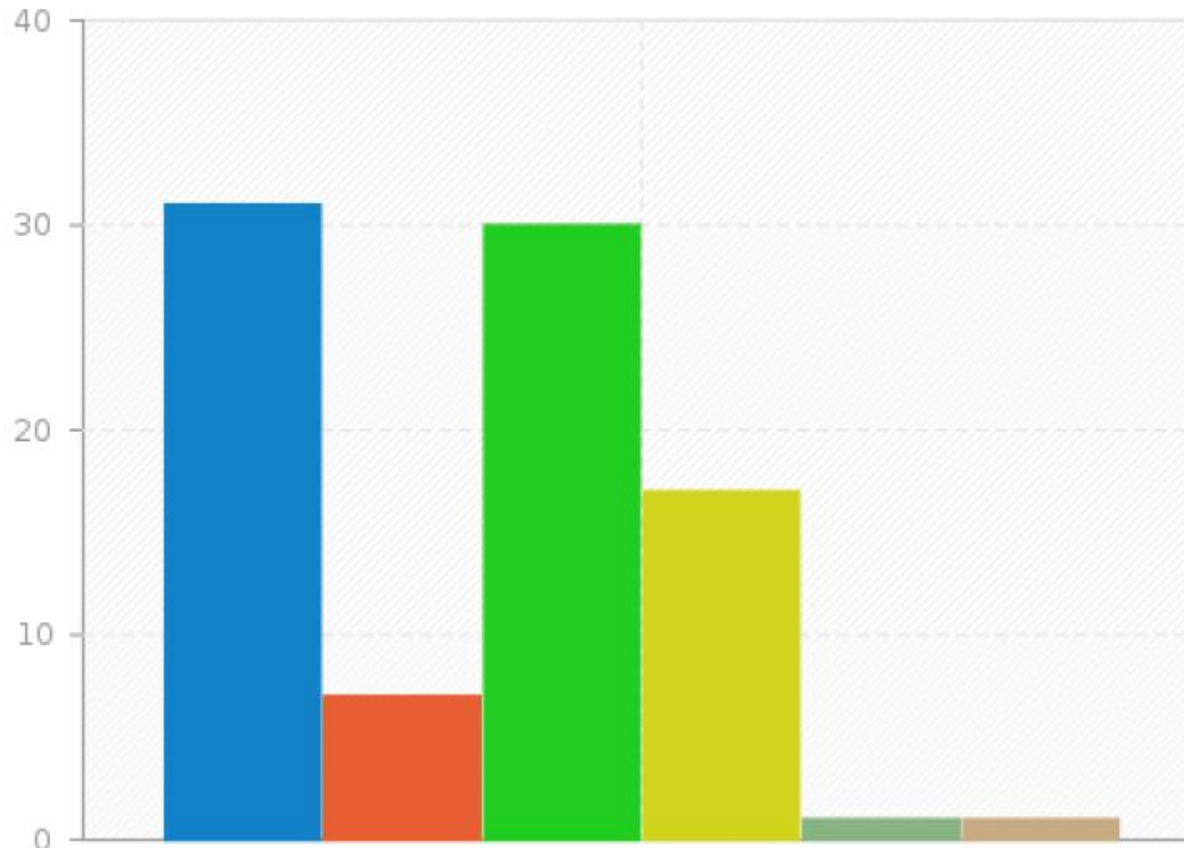


Bedarfsanalyse/Umfrage: Kompetenzen

Wie schätzen Sie Ihre eigene Kompetenz im Umgang mit Fachvokabularen ein?



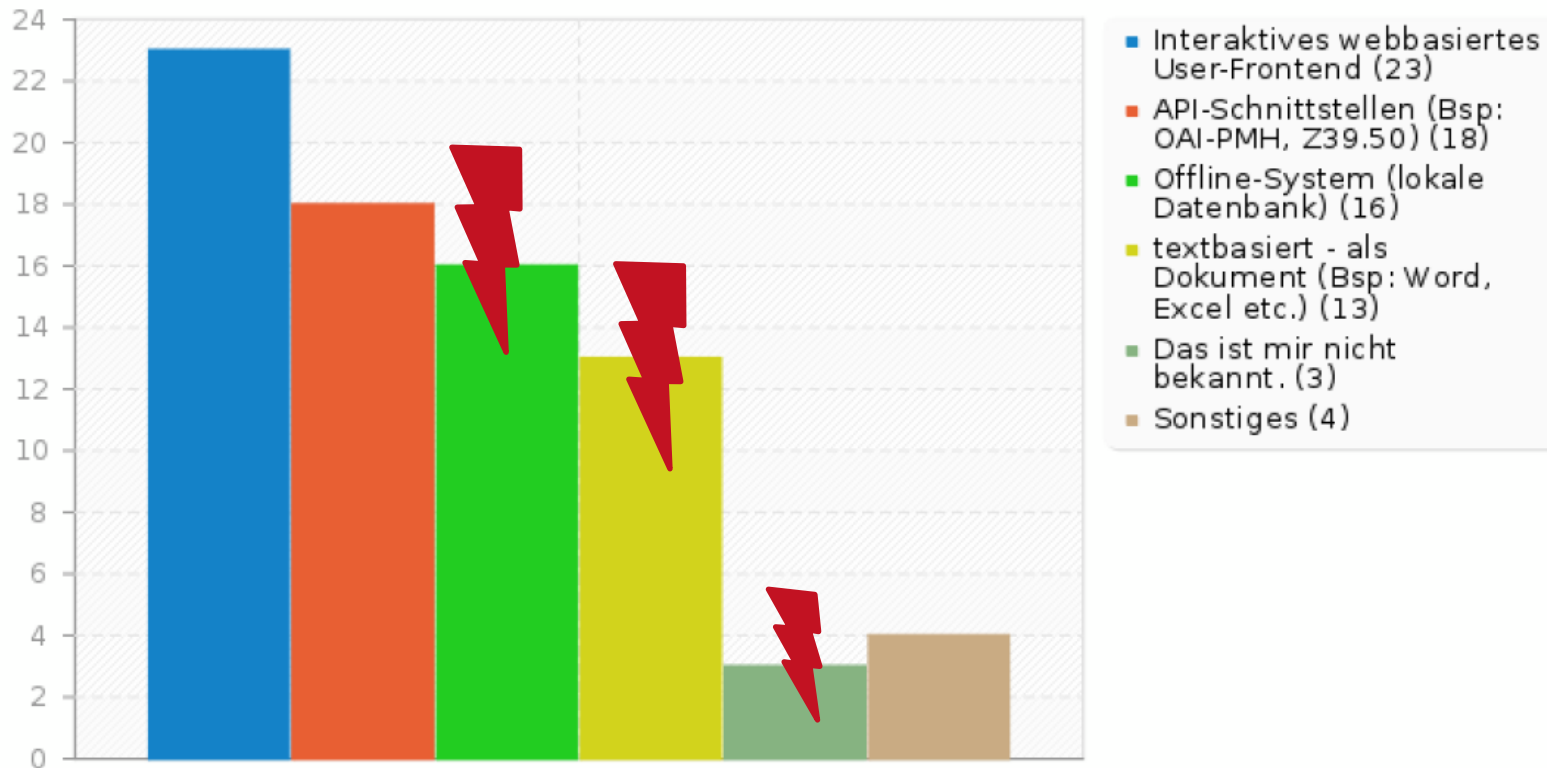
Bedarfsanalyse/Umfrage: Anwendung 1



- Meine Einrichtung hat ein oder mehrere Fachvokabulare entwickelt. (31)
- Meine Einrichtung hat ein oder mehrere Fachvokabulare extern erworben. (7)
- Meine Einrichtung verwendet eine oder mehrere frei verfügbare Fachvokabulare. (30)
- Wir bieten ein oder mehrere Fachvokabulare an. (17)
- Die Herkunft des verwendeten Fachvokabulars ist mir nicht bekannt. (1)
- Sonstiges (1)

Bedarfsanalyse/Umfrage: Anwendung 2

Mittels welcher technischen Lösung nutzen Sie Ihr Fachvokabular?



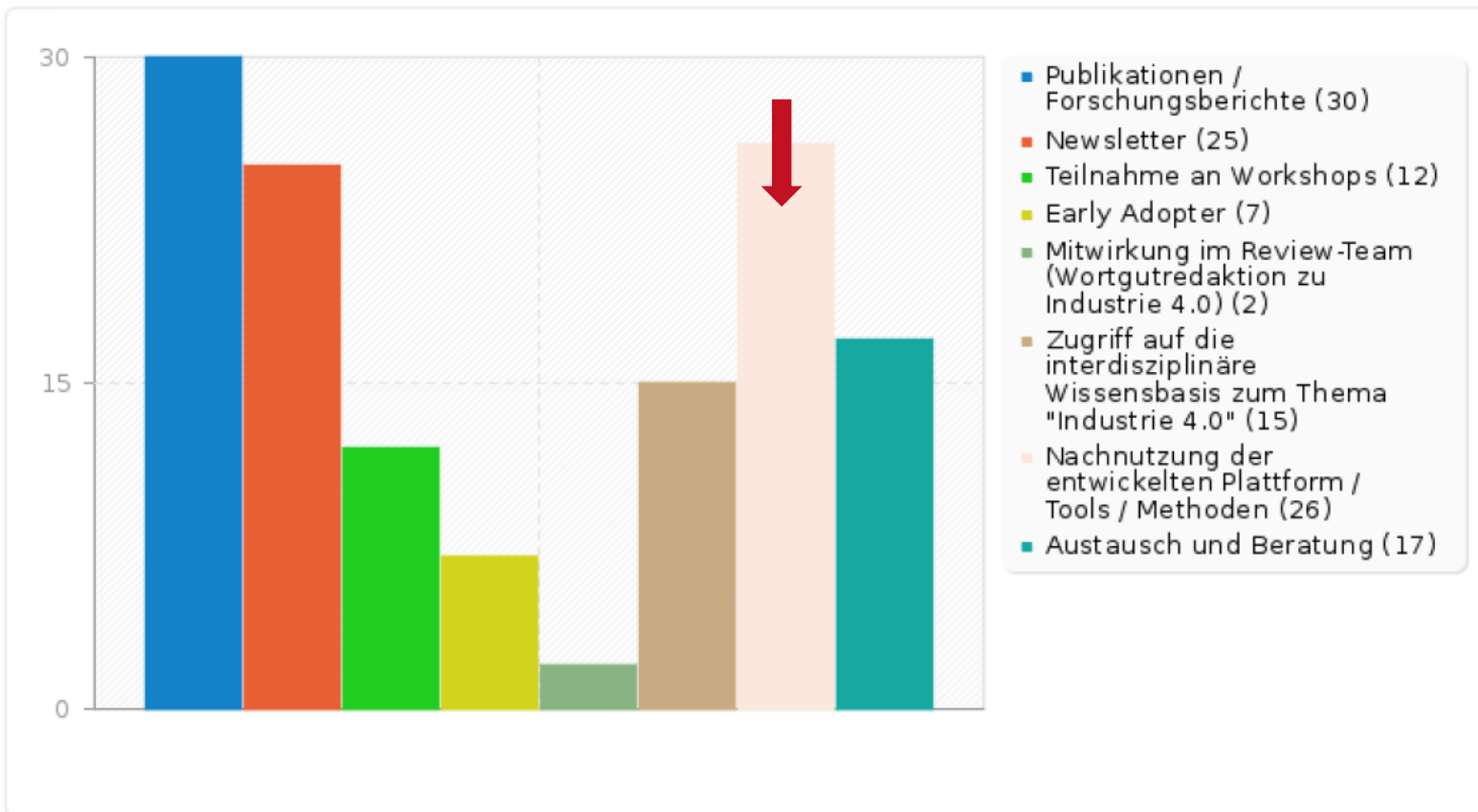
Bedarfsanalyse/Umfrage: Relevanz

Die folgenden Angebotsarten wurden insbesondere von Thesaurusentwicklern durchgängig als „sehr relevant / interessant“ eingestuft:

- ein gemeinsamer Zugriff auf mehrere Thesauri
- ein zentrales Portal für frei verfügbare Vokabulare
- eine unterstützende Suchfunktion für Querschnittsthemen
- eine frei zugängliche technische Plattform für die Integration von Thesauri
- überprüfte Methoden für die Verknüpfung von Fachthesauri

Bedarfsanalyse/Umfrage: Fachgemeinschaft

In welcher Form wünschen Sie Kontakt mit dem Projekt?



Geplantes Vorhaben

- **Inhaltlich:** Für den Bereich „Industrie 4.0“ relevante Ausschnitte aus dem TEMA-Thesaurus („Technik und Management“; TIB mit WTI Frankfurt eG) und aus dem Thesaurus für Informatik, Wirtschaft, Technik/Ingenieurwissenschaften und Verkehr der TH Wildau werden exemplarisch in SKOS überführt und mittels semantischer Relationen verknüpft
- **Technisch:** Hierfür wird der Prototyp einer modular aufgebauten, erweiterbaren Plattform zur Vorhaltung der Fachvokabulare implementiert; auf dieser werden auch Werkzeuge zur Umwandlung von Wortgutsammlungen in SKOS-Thesauri, deren Verknüpfung, Aktualisierung und Weiterverarbeitung angeboten
- **User experience:** Mehrwert und Nutzerfreundlichkeit werden im Dialog mit den fachlichen Zielgruppen kontinuierlich evaluiert und dokumentiert
- **Verstetigung:** Erarbeitung eines Konzepts zur zukünftigen Finanzierung und Organisationsform der Plattform und der angebotenen Dienstleistungen

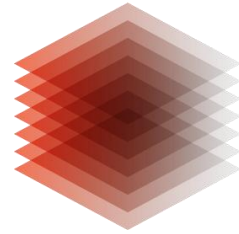
Nutzungsszenarien: Unterstützung für ... bei ...

- Infrastruktureinrichtungen:
 - Ausbau von **intelligenten Navigationsmöglichkeiten** über Fachthesauri
 - **Einspielen** von externen Vokabularen in das eigene **Bibliothekssystem** bzw. Verknüpfung mit bereits verwendeten Vokabularen
 - Förderung der **Interoperabilität** durch die **Erweiterung eigener Vokabulare** unter Berücksichtigung aktueller **Semantic-Web-Standards**
- Wissenschaftliche Fachgemeinschaft:
 - Intelligent suchen mit Hilfe von **interaktiven Visualisierungen** mit Definitionen und Auswahlmöglichkeiten für Schlagwörter samt Anzeige assoziierter Bestände
 - **Suchterweiterung** insbesondere in interdisziplinären Kontexten
- Thesaurus- und Ontologieersteller in der angewandten Wissenschaft:
 - **Offener Zugang** zu **nachnutzbaren Thesauri, Dienstleistungen und Werkzeugen** zur Erstellung, kollaborativen Weiterentwicklung, Veröffentlichung und Verknüpfung von Fachvokabularen

Geplante Interaktion mit der Fachgemeinschaft, Verstetigung

- Aufbau und Pflege einer Webpräsenz zu den Projektaktivitäten
- Austausch mit der wissenschaftlichen Community im Bereich „Industrie 4.0“ und mit laufenden nationalen und internationalen Forschungsprojekten sowie Anwendergruppen von Softwareprojekten
- Durchführung von Schulungen und Workshops bezüglich der Plattform-Komponenten für verschiedene Nutzergruppen (Wissenschaft / Infrastruktureinrichtungen / Thesaurusentwickler)
- Entwicklung eines nachhaltigen Organisations- und Finanzierungsmodells (Prüfung von Vereins-/Genossenschaftsmodellen ggf. mit kostenpflichtigen Diensten im Bereich Beratung, Datenpflege oder Infrastructure-as-a-Service) mit dem Ziel der Überführung in den stetigen Betrieb und der Sicherung der Nachnutzbarkeit für Wissenschaft und Gesellschaft

LEIBNIZ-INFORMATIONSZENTRUM
TECHNIK UND NATURWISSENSCHAFTEN
UNIVERSITÄTSBIBLIOTHEK



TIB

DANKE

Kontaktdaten

Dr. Anna Kasprzik

T 0511 762-14219, anna.kasprzik@tib.eu