

Vorhabenbeschreibung

## »Industrial Data Space«: Digitale Souveränität über Daten



### INDUSTRIAL DATA SPACE

ROBOTIK

RFID

TELEMATIK

SERVICES

DIENSTE

AUTONOMIK

CYBER PHYSICAL SYSTEMS

ASSISTENZSYSTEME

# Inhaltsverzeichnis

1	Ziele	4
1.1	Kontext des Vorhabens	4
1.2	Die Industrial Data Space Initiative	6
1.3	Forschungs- und Entwicklungsbedarf	8
1.3.1	Rolle der Daten	8
1.3.2	Wert der Daten	9
1.3.3	Industriebetrieblicher Handlungsbedarf	10
1.4	Ziele des Vorhabens	12
1.5	Bezug zu den förderpolitischen Zielen	14
2	Stand der Wissenschaft und Technik	15
2.1	Verwandte Initiativen	15
2.2	Technologien und eigene Vorarbeiten	22
2.3	Schutzrechte	27
3	Arbeitsplan	28
3.1	InDaSpace Architektur	28
3.1.1	Geschäftsarchitektur	30
3.1.2	Daten-Service-Architektur	31
3.1.3	Software-Architektur	31
3.1.4	Sicherheitsarchitektur	31
3.2	Implementierung der Kern-Softwarekomponenten	35
3.2.1	Überblick über Komponenten und ihre Aufgaben	35
3.2.2	Implementierung der InDaSpace Konnektoren (INIK, EXIK)	38
3.2.3	Aufbau des InDaSpace App-Stores	40
3.2.4	Internal InDaSpace Konnektor	42
3.2.5	External InDaSpace Konnektor	46
3.2.6	InDaSpace Broker	49
3.2.7	InDaSpace App Store	52
3.2.8	Funktionsweise und Zusammenspiel der Komponenten	53
3.3	Umsetzung ausgewählter Use Cases	55
3.4	Standardisierung	60
3.5	Zertifizierungskonzept	63
3.6	Methodik Business Innovation (für Anwender)	68
3.7	Handlungsempfehlungen Betrieb/Forschung	71
3.8	Institutionalisierung der Initiative	72
3.9	Projektmanagement	78
3.10	Balkenplan	79
3.11	Meilensteinplan	81

4	Verwertungsplan .....	83
4.1	Wirtschaftliche Erfolgsaussichten .....	83
4.2	Wissenschaftliche und/oder technische Erfolgsaussichten .....	87
4.3	Wissenschaftliche und wirtschaftliche Anschlussfähigkeit .....	88
4.4	Verwertungstabelle .....	89
5	Arbeitsteilung/Zusammenarbeit mit Dritten .....	91
5.1	Projektorganisation .....	91
5.2	Zusammenarbeit mit der Industrie .....	92
5.3	Arbeitsteilung innerhalb der Fraunhofer-Gesellschaft .....	93
5.4	Beteiligte Fraunhofer-Institute .....	95
5.4.1	Fraunhofer IAIS .....	95
5.4.2	Fraunhofer IML .....	95
5.4.3	Fraunhofer ISST .....	97
5.4.4	Fraunhofer AISEC .....	97
5.4.5	Fraunhofer IPA .....	98
5.4.6	Fraunhofer FIT .....	99
5.4.7	Fraunhofer FOKUS .....	100
5.4.8	Fraunhofer FKIE .....	101
5.4.9	Fraunhofer SIT .....	101
5.4.10	Fraunhofer IOSB .....	103
5.4.11	Fraunhofer IAO .....	104
5.4.12	Fraunhofer IESE .....	105
6	Notwendigkeit der Zuwendung .....	106
	Anhang .....	109
	Abkürzungsverzeichnis .....	109
	Glossar .....	109
	Literaturverzeichnis .....	110



# 1 Ziele

## 1.1 Kontext des Vorhabens

Die Digitalisierung ist die zentrale gesellschaftliche, betriebswirtschaftliche und technologische Entwicklung. Das 5G-Mobilfunknetzwerk, der 3D-Druck und smarte Services sind nicht nur technische Innovation, sondern verändern die Art und Weise, in der Menschen arbeiten und ihr Leben gestalten (vgl. [BHV14], [HKK14]).

Die Digitalisierung ist einerseits Konsequenz, andererseits »Befähiger« einer Reihe von Entwicklungen:

- *Mobilität*: Kunden erwarten Leistungserbringung, vornehmlich smarter Services, zu jeder Zeit an jedem Ort. Eine Ortsgebundenheit der Leistungserbringung wird kaum noch akzeptiert.
- *Globalisierung*: Viele Unternehmen agieren heute auf weltweiten Absatz- und Beschaffungsmärkten, um Größen-, Währungs- und Kostenvorteile nutzen zu können und gleichzeitig neue Wachstumspotentiale zu erschließen. Globalisierung ist nicht neu. Es gibt sie schon seit Jahrhunderten. War die Globalisierung jedoch in den 1980er und 1990er Jahren von Standardprodukten geprägt, die über klar definierte Kunden-Lieferanten-Beziehungen ausgetauscht wurden, so ist die Globalisierung heute geprägt von komplexen Produktions- und Service-Netzwerken (z. B. im Maschinen- und Anlagenbau sowie in der Automobilindustrie) und hoher Informationstransparenz.
- *»Sharing Economy«*: Zunehmend florieren Nutzungsmodelle, die auf dem Teilen von Ressourcen beruhen. Ein Grund dafür ist, besonders in städtischen Ballungsräumen westlicher Industrienationen, ein gesellschaftlicher Wertewandel, der dem Besitz an materiellen Gütern (z. B. von Autos, Häusern etc.) an sich einen weniger hohen Status zuweist, als in der Vergangenheit, auch weil ein hoher Anteil der Bedürfnisse der Bewohner bereits befriedigt ist. Auch Entwicklungen wie Crowd-Sourcing belegen den Wandel in gesellschaftlichen Werte- und individuellen Anreizsystemen.
- *Nachhaltigkeit*: Ausdruck eines Wertewandels in westlichen Industrienationen ist weiterhin die wachsende Bedeutung nachhaltiger Produkte und Dienstleistungen. Kunden sehen Nachhaltigkeit heute bereits als Qualitätsmerkmal. In Zukunft wird es ein Hygienefaktor sein, ohne den Produkte und Dienstleistungen zukünftig am Markt keine Chance mehr haben werden.
- *Privatheit*: Die Diskussion um Datenschutz und Privatheit im Internet ist von einer gewissen Widersprüchlichkeit geprägt. Einerseits wächst die Nutzung sozialer Netzwerke, von Suchmaschinen und Apps enorm, obwohl allen Nutzern bekannt ist, dass sie mit der Nutzung Rechte an ihren Daten abtreten. Andererseits wachsen die Forderungen nach mehr Datenschutz in gleichem Maße.

Diese Entwicklungen besitzen das Potential, angestammte Branchen in ihrer Mechanik grundsätzlich zu verändern und nicht nur für inkrementelle Weiterentwicklung zu sorgen. So ist bspw. zu hinterfragen, inwieweit die Automobilhersteller ihre dominierende Rolle in der Wertschöpfungskette als »Original Equipment Manufacturer« (OEM) halten können oder nicht vielmehr Gefahr laufen, zum Hardware-Lieferanten für Anbieter von Mobilitätsdiensten zu werden, die die obigen Entwicklungen möglicherweise besser adressieren, als traditionelle reine Automobilhersteller. Dieses disruptive Potential der erwähnten Entwicklungen wird

dadurch potenziert, dass die Schnittstelle zum Kunden heute zunehmend ebenfalls der Digitalisierung unterworfen ist – da Kunden zunehmend ohnehin dazu übergehen, Produkte im Internet zu kaufen, entfällt der Vorteil traditioneller Anbieter, eine physikalische Verkaufsfläche zur Verfügung zu haben und sich über diese von den Wettbewerbern abzuheben (Beispiel: amazon vs. traditioneller Buchhandel).

Aus diesen Entwicklungen ergeben sich unmittelbar Anforderungen an Produkte und Dienstleistungen (vgl. [HKK14], [PNH14]):

- *Individualisierung*: Produkte und Dienstleistungen sind individualisiert, d. h. sie richten sich an das Individuum mit seinen persönlichen Bedürfnissen. Nicht mehr die Organisation als Kunde steht im Vordergrund, sondern der einzelne Mensch in seinen Tätigkeiten, wie arbeiten, gesund werden oder bleiben, reisen, einkaufen usw.
- *Hybridität*: Die Grenzen verschwimmen nicht nur zwischen physischen Produkten und Dienstleistungen, sondern auch zwischen klassischen Angeboten und digitalen Dienstleistungen. Beispiele wie mytaxi.de, Air-BnB.com usw. zeigen, dass herkömmliche Leistungen (von A nach B kommen, übernachten etc.) digital angereichert werden.
- *Ende-zu-Ende-Unterstützung*: In der Vergangenheit bedienten Produkte Kundenbedarfe aus Sicht des Unternehmens. Zukünftig müssen Leistungsangebote den gesamten Kundenprozess adressieren, nicht nur die zufälligen Interaktionspunkte des Unternehmens mit dem Kunden. Gleichmaßen muss diese Durchgängigkeit auch in Richtung der Lieferanten und deren Prozesse gegeben sein.
- *Wirtschaftliches Ökosystem*: Für die umfassende Erfüllung der Kundenansprüche wird zunehmend das Zusammenspiel vieler Akteure zum Wohle der Ende-zu-Ende-Unterstützung ihrer Lebenslage notwendig oder zumindest von entscheidendem Vorteil. Bei der Orchestration der Leistungserbringung im wirtschaftlichen Ökosystem vertrauen Kunden denjenigen »Marken«, deren Wertesysteme die höchste Kongruenz zu ihren eigenen Wertevorstellungen aufweisen.

Wichtigster Erfolgsfaktor, um diese Anforderungen an Produkte und Dienstleistungen zu erfüllen, sind die dafür notwendigen Daten (seitens der Kunden sowie der Produkte und Dienstleistungen). Die Fähigkeit, Daten wie ein Wirtschaftsgut zu bewirtschaften und auf dieser Basis »smarte« Services anzubieten, ist für Unternehmen somit ein Wettbewerbsfaktor, der zunehmende existenzielle Voraussetzung darstellt.

Die Daten bilden auch das Fundament, die aus den neuen Produkten und Dienstleistungen resultierende Komplexität in der Produktion und in der Logistik beherrschbar zu machen. Industrie 4.0 ist in diesem Kontext ein Organisationsprinzip für den modernen Industriebetrieb, das auf Vernetzung von Menschen und Maschinen, der Autonomisierung von Prozessen und Systemen, der durchgängigen Informationstransparenz und Unterstützung von Entscheidungen beruht.

Somit sind die Daten das Bindeglied zwischen industrieller Produktion einerseits und neuen Leistungsangeboten über Smart Services andererseits. An dieser Scharnierfunktion ist der Industrial Data Space positioniert, wie Abbildung 1 zeigt.



sammenarbeit im Sinne eines gemeinsamen Wirkens sowie einen schnellen, automatisierten Datenaustausch ermöglichen, sind ein wesentliches Differenzierungsmerkmal der Unternehmen im internationalen Wettbewerb. Der Industrial Data Space soll hierzu horizontale Vernetzung und Virtualisierung ebenso wie den sicheren Zugriff von Apps auf private und globale Daten ermöglichen.

5. *Vertikale Integration als Basis attraktiver Dienste im Industrial Data Space:* Es besteht eine strategische Chance darin, vertikale, branchenspezifische Integration voranzutreiben, verbunden mit den jeweiligen Standortvorteilen in Produktion, Handel, Dienstleistung und Logistik. Der Software kommt auch in diesem Zusammenhang eine Schlüsselrolle sowohl beim Management als auch beim Design adäquater Systeme zu. Der Industrial Data Space bietet durch horizontale Vernetzung und Virtualisierung den Raum für hochwertige vertikale kontextuelle und kognitive Dienste.
6. *Transparenz und Marktorientierung:* Transparenz, Rückverfolgbarkeit und marktorientierte Mechanismen sind die Grundlage jeder vertrauenswürdigen Vernetzung von Daten und jedes wettbewerbsfähigen Services. Dies gilt insbesondere in einem kompetitiven Umfeld, wo verschiedene Unternehmen mit Daten und Diensten einerseits im Wettbewerb stehen, sich aber andererseits über den Industrial Data Space vernetzen können. Um dies sicherzustellen, beinhaltet der Industrial Data Space die notwendigen Mechanismen zur Zurechenbarkeit und Marktbildung (bspw. Online Ad Auctions, Real-Time Bidding und vieles mehr).
7. *Öffnung, Normung, Standards:* IT- Dienstleister legen oftmals eigene Technologien als Standards fest, um Marktanteile zu gewinnen und Kunden zu binden. Für langfristige, entwicklungs offene globale Interoperabilität und das von der Industrie eingeforderte Multi-Sourcing will der *Industrial Data Space* Öffnungskonzepte entwickeln, welche breiteren wirtschaftlichen Nutzen erzeugen. Eine Konsens-basierte internationale Normung soll ebenfalls vorangetrieben werden. Dies trägt zur Standortsicherung bei und vermeidet die Abhängigkeit von Monopolen.
8. *Der Industrial Data Space, Internet der Dinge und Industrie 4.0:* Die Initiative »Industrial Data Space« schafft eine Plattform und die zugehörigen Dienste, um einen sicheren Multi-Sourcing-Datenaustausch on Demand und auf Basis bestehender Netze zu ermöglichen. Der Industrial Data Space versteht sich als komplementärer Partner der echtzeitnahen und echtzeitfähigen Technologien im Sinne des Internet der Dinge oder der vierten industriellen Revolution (Industrie 4.0).

Der Industrial Data Space positioniert sich somit als alternativer und ergänzender Architektorentwurf im Vergleich zu bestehenden Konzepten, die sich entweder dadurch auszeichnen, dass Daten vollständig zentral und entsprechend monopolistisch verwaltet werden, oder, dass jede Verbindung individuell auszuhandeln ist und für genau einen Datenaustausch Gültigkeit hat. Die vom Internet gewohnte Flexibilität soll bei dieser Architektur erhalten bleiben. Hierfür bringt der Industrial Data Space umfangreiche Mechanismen mit sich, um Vertraulichkeit, Integrität und Verfügbarkeit der Daten und Dienste zu schützen. Die föderale Struktur des Industrial Data Space sichert den Dateneignern volle Souveränität über ihre Daten zu. Es ist also der Dateneigner, der über Verarbeitung seiner Daten entscheidet. Die Daten können bei Bedarf dezentral bei ihren Eignern verbleiben und werden nur im Rahmen der Nutzung unter Anwendung strenger Schutzmaßnahmen ausgetauscht. Die zentralisierte Datenhaltung, die in der Regel mit einem

Kontrollverlust über die eigenen Daten einhergeht, wird hierdurch optional. Abbildung 2 stellt die Positionierung des Industrial Data Space dar.

	Data Lake	INDUSTRIAL DATA SPACE	Pure Internet
			
	vollständig zentral, monopolistisch	föderal, sicher, „trusted“, Standard	vollständig dezentral, offen, unsicher
Datenhaltung	Zentrales Repository	Dezentral	Dezentral
Data Ownership	Zentral	Dezentral	Dezentral
Datenverknüpfung	Einzelanbieter	Föderal, on demand	Fehlt
Datensicherheit	Bilateral	Zertifiziertes Gesamtsystem	Bilateral
Marktstruktur	Zentraler Betreiber	Rollensystem	Unstrukturiert
Transportinfrastruktur	Internet	Internet	Internet

Abbildung 2: Positionierung des Industrial Data Space

Ziel des Integrated Data Space ist somit nicht die Definition neuer, sondern die Integration bestehender Prozesse, für die insbesondere an hohes Maß an Interoperabilität angestrebt werden muss. Der Industrial Data Space soll aus diese Weise als Ergänzung bzw. alternativer Architekturentwurf zu bereits bestehenden Initiativen (s. Abs. 2.1) die Vernetzung und die Zusammenarbeit der Unternehmen fördern, indem Daten und Dienste zu neuen Leistungsangeboten bzw. gar Geschäftsmodellen verknüpft werden. Novum im Vergleich zu den bestehenden Ansätzen ist, dass Anbieter und Nutzer von Daten oder Diensten auch ihre bestehenden IT-Infrastrukturen nutzen können, dass sie Vertrauensschutz hinsichtlich der Identität der teilnehmenden Akteure und des Datenaustausches und der Datennutzung genießen und dass die Gesamtarchitektur des Industrial Data Space keine Partikularlösung darstellt, sondern offen für alle Branchen und Unternehmen ist.

### 1.3 Forschungs- und Entwicklungsbedarf

#### 1.3.1 Rolle der Daten

Die Bedeutung der Daten für Geschäftserfolg steigt kontinuierlich seit der Einführung der elektronischen Datenverarbeitung und der Automatisierung von Produktionsprozessen in Unternehmen. Zwar diskutieren Forschung und Wirtschaft bereits seit den 1980er Jahren die Rollen von Daten als Anlagegut. Aber im Zuge der Digitalisierung wird erstmals von »Data-Driven Business« gesprochen [0014]. Die Digitalisierung ist ohne Daten nicht möglich (vgl. [LEO14]).

Abbildung 3 zeigt die Rolle der Daten im Zeitverlauf.

- *Daten als Prozessergebnis:* Zwar werden Informationssysteme bereits seit den 1960er und 1970er Jahren zur Unterstützung von Unternehmensfunktionen eingesetzt, aber die Daten hatten vornehmlich eine unterstützende Rolle. Warenwirtschaftssysteme dienten bspw. der Unterstützung der Lagerprozesse an einem Standort. Wollte der Lagermitarbeiter wissen, ob von einem bestimmten Teil noch Lagerbestand ver-

fürbar war, musste er nicht mehr im Regalgang nachschauen, sondern konnte das System befragen. Der Wertbeitrag für das Unternehmen entstand jedoch ausschließlich durch das physische Produkt.

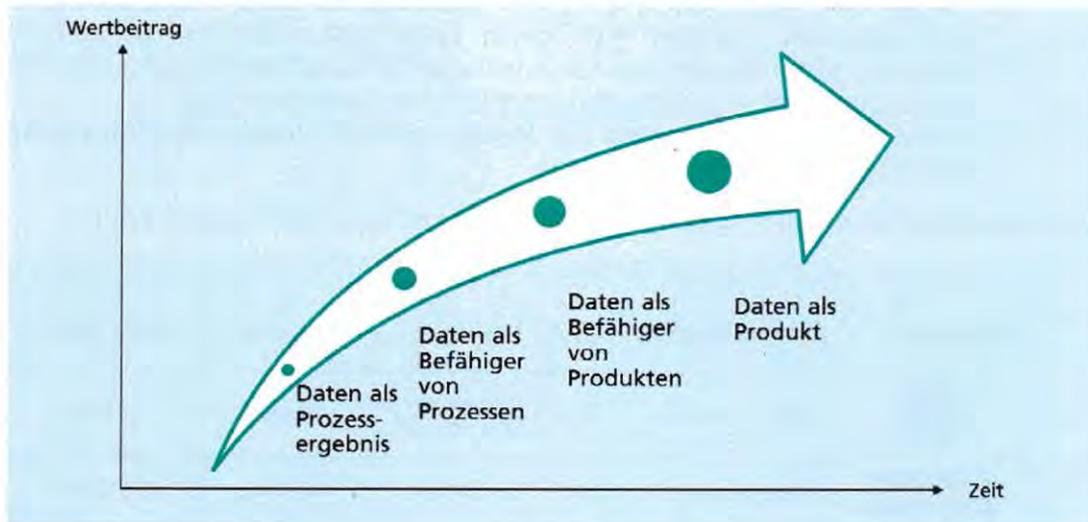


Abbildung 3: Rolle der Daten in der betrieblichen Leistungserstellung

- *Daten als Befähiger der Prozesse:* Mit der Verbreitung von Manufacturing Resource Planning (MRP)- und Enterprise Resource Planning (ERP)-Systemen in den 1980er und 1990er Jahren entwickelten sich Daten zum Befähiger des unternehmensweiten Geschäftsprozessmanagements. Ohne konsistente, in nahezu Echtzeit verfügbare Daten wäre die Einführung weltweit (oder mindestens regional) standardisierter Prozesse wie Order-to-Cash, Procure-to-Pay etc. nicht möglich gewesen. Daten wurden zur strategischen Ressource für Operational Excellence in der Produktion, Logistik und im Service.
- *Daten als Befähiger der Produkte:* Seit dem Jahrtausendwechsel bieten Unternehmen zunehmend Produkte an, die ohne Daten hoher Qualität nicht möglich sind. Beispiele sind miCoach von adidas, die Leasing- und Fleet Management-Modelle von Hilti und sämtliche Smart Services für unsere persönlichen Lebenslagen.
- *Daten als Produkt:* In den letzten Jahren entstehen Marktplätze für Daten. Abrufe von Daten-APIs werden nach Volumen und Zeiteinheit abgerechnet. Die Daten sind nicht mehr Befähiger der Produkte, sondern das Produkt selbst.

Die einzelnen Entwicklungsstufen in der Rolle der Daten sind nicht disjunkt, sondern sind parallel in Unternehmen zu finden. Deswegen kommt es auch zu dem »Datenparadoxon«, dass Daten einerseits Konsequenz der o.a. Entwicklungen wie Digitalisierung und Industrie 4.0 sind (denn Maschinen, Smart Services etc. produzieren eine steigende Menge an »Big Data«). Andererseits sind Daten Ressource der Leistungserstellung oder sogar das Produkt selbst.

### 1.3.2 Wert der Daten

Mit der sich wandelnden Rolle der Daten ändert sich auch der Wert der Daten für die Unternehmen. Die Bewertung von Daten richtet sich dabei nach den Prinzipien der Bewirtschaftung immaterieller Güter. Grundsätzlich haben Unternehmen drei Möglichkeiten:

- *Herstellkosten*: Die Kosten für die Beschaffung und/oder die Erzeugung der Daten bilden den Datenwert.
- *Nutzwert*: Der Beitrag der Daten zur Prozessleistung stellt den Wert der Daten dar. Hierunter fallen bspw. Beiträge zur Steigerung der Kundenzufriedenheit, zur Vermeidung von Lagerbeständen und damit zur Reduktion gebundenen Kapitals sowie der effizientere Einsatz von Vertriebspersonal in Geschäftsmodellen mit Direktvertrieb.
- *Marktwert*: Daten werden am Markt verkauft, sodass der Verkaufspreis den Wert der Daten bestimmt.

Alle drei Modelle werden in der Praxis eingesetzt, wie Abbildung 4 zeigt.

Unternehmen	Leistungsangebot	Land	Datenklasse	Wertansatz	Wert pro Datensatz
	Supermarktkette	US	Kundendaten inkl. Einkaufsprofile	Marktwert	1,6 EUR <sup>1</sup>
	Social Network	US	Benutzerdaten	Marktwert	225 USD <sup>2,3</sup>
	Automatisierungstechnik	DE	Teilestammdaten	Herstellkosten	500 bis 5.000 EUR <sup>4</sup>
	Agrochemie	CH	Materialstammdaten	Nutzwert	184 CHF <sup>5</sup>

1) <http://www.wsj.de/nachrichten/SB11446175161338053998704580212211843086060>  
 2) <http://en.wikipedia.org/wiki/Facebook>; 890 million daily active users.  
 3) <http://www.ft.com/cms/s/0/ecc0f050-37a3-11e4-bd0a-00144feabdc0.html#axzz3RH6OPOTH>; Markt kapitalisierung von 200 Mrd. USD.  
 4) Vgl. Otto, Boris: Managing the business benefits of product data management: the case of Festo. In: Journal of Enterprise Information Management 25 (2012), Nr. 3, S. 272-297, DOI: 10.1108/17410391211224426; 5.000 EUR pro Neuanlage, 500 EUR jährliche Pflegekosten.  
 5) <http://www.marketwatch.com/investing/stock/syl/financials> (Abruf am 9.2.15); Umsatz: 13,85 Mrd. CHF; Zahl Materialstammdaten: ca. 1,5 Mio; Kostensenkungspotential durch hohe Datenqualität gemäß Experteninterview: 2 Umsatzprozent.

Abbildung 4: Wert von Daten und seine Ermittlung

Unternehmen sind mit einer Vielzahl von offenen Fragen konfrontiert. Die Beantwortung dieser Fragen ist von strategischer Bedeutung, weil die o.a. Entwicklungen grundlegende Änderungen für Branchen und Geschäftsmodelle mit sich bringen und im Zweifel die Unternehmensexistenz bedrohen.

### 1.3.3 Industriebetrieblicher Handlungsbedarf

Die oben genannten Entwicklungen haben Implikationen auf die gesamte Unternehmensarchitektur, also auf die Geschäftsstrategie, die Geschäftsprozesse, auf das Datenmanagement und die Informations- und Kommunikationssysteme. Tabelle 1 skizziert den Handlungsbedarf anhand der Fragestellungen, mit denen sich Projekt- und Linienverantwortliche der Unternehmen aktuell beschäftigen.

Unternehmensarchitekturebene	Offene Fragen
Geschäftsstrategie	<ul style="list-style-type: none"> <li>- Was ist der Ende-zu-Ende-Prozess unserer Kunden?</li> <li>- Welche hybriden Produkte und Dienstleistungen bieten wir an?</li> <li>- Wie entwickeln wir diese Produkte und Dienstleistungen?</li> <li>- Was ist unsere Stellung im »Ökosystem«? Mit wem sollten wir kooperieren?</li> <li>- Welche (digitalen) Fähigkeiten benötigt un-</li> </ul>

	ser Unternehmen? Wie bauen wir diese auf?
Geschäftsprozesse	<ul style="list-style-type: none"> <li>- Wie ändert sich das Geschäftsprozessmanagement mit Industrie 4.0 und Konsumentenzentrierung? Sind Prozesse noch planbar oder nur noch stochastisch?</li> <li>- Welche Prozesse sind in der Zukunft benötigt? Decken herkömmliche Prozessmodelle diese Anforderungen ab?<sup>1</sup></li> </ul>
Datenmanagement	<ul style="list-style-type: none"> <li>- Wer entwickelt Datenprodukte?</li> <li>- Wer ist Eigner der »Data Supply Chain«?</li> <li>- Wie sieht eine Datenarchitektur aus, die sowohl interne als auch externe Daten beinhaltet?</li> <li>- Wem gehören die Daten? Was darf unser Unternehmen mit Daten aus dem »Ökosystem« tun?</li> <li>- Wie ist Datenqualität zu sichern für unsere Daten, aber auch für Daten von extern?</li> <li>- Welche Plattformen unterstützen Data Services?</li> </ul>
IuK-Systeme	<ul style="list-style-type: none"> <li>- Welche Technologien müssen wir zukünftig verwenden?</li> <li>- Müssen wir die Software für »smarte« Services selbst entwickeln?</li> <li>- Wie spielen »Legacy«-Systeme wie ERP, CRM etc. mit neuen Technologien zusammen? Wie schaffen wir den Parallelbetrieb?</li> </ul>

Tabelle 1: Typische Fragen in Unternehmen

Grundsätzlich unterstützt eine datenzentrierte Unternehmensarchitektur bei der Strukturierung und Planung des Handlungsbedarfs.

Abbildung 5 zeigt die verschiedenen Ebenen einer datenzentrierten Unternehmensarchitektur sowie die Positionierung des Industrial Data Space. Der Industrial Data Space fokussiert auf die Ebene der Datendienste und ihrer Beziehungen zu den anderen Architekturebenen.

<sup>1</sup> Herkömmliche Prozessmodelle bilden heutige Anforderungen nicht mehr ab. Im Handel dominieren bspw. Prozessmodelle wie das Handels-H oder SAP Business Maps, in denen aber die Rolle von »Returns« oder »Churn Prevention« beim Online-Handel nicht enthalten sind.

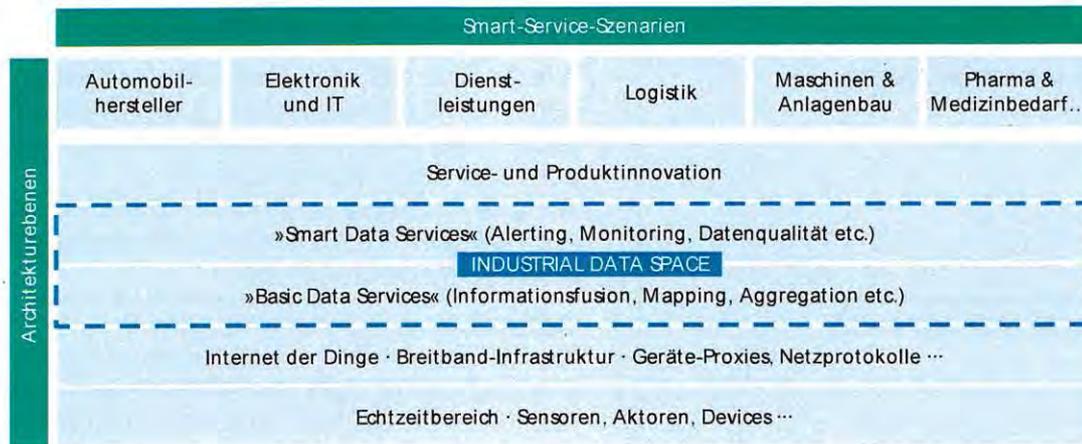


Abbildung 5: Datenzentrierte Unternehmensarchitektur

#### 1.4 Ziele des Vorhabens

Aufgrund der Bedeutung der Digitalisierung und der technologischen Entwicklungen der letzten Jahre in Bereichen wie dem mobilen Internet, Cloud Computing, Embedded Systems oder Big Data formen sich zurzeit weltweit verschiedene Konsortien und Initiativen (siehe Kapitel 2.1).

Der Industrial Data Space soll als Ergänzung bzw. alternativer Architekturentwurf zu diesen Initiativen die Vernetzung und die Zusammenarbeit der Unternehmen fördern, indem Daten und Dienste zu neuen Leistungsangeboten bzw. gar Geschäftsmodellen verknüpft werden. Novum im Vergleich zu den oben genannten Ansätzen ist, dass Anbieter und Nutzer von Daten oder Diensten auch ihre bestehenden IT-Infrastrukturen nutzen können, dass sie Vertrauensschutz hinsichtlich der Identität der teilnehmenden Akteure und des Datenaustausches und der Datennutzung genießen und dass die Gesamtarchitektur des Industrial Data Space keine Partikularlösung darstellt, sondern offen für alle Branchen und Unternehmen ist.

Das hiermit beantragte Forschungs- und Entwicklungsvorhaben zielt auf die vorwettbewerbliche Vorbereitung des Aufbaus des Industrial Data Space ab. Im Einzelnen sollen hier die folgenden wissenschaftlich-technischen Ziele erreicht werden:

- Entwurf, Beschreibung und prototypische Umsetzung einer Referenzarchitektur für den Industrial Data Space. Die Referenzarchitektur beschreibt als konzeptionelles Modell nicht nur die software-technische Basis des Industrial Data Space, sondern auch die notwendigen Schutz-, Governance-, Kooperations- und Kontrollmechanismen für die sichere Verarbeitung und den sicheren Austausch der Daten.
- Prototypische Implementierung der Referenzarchitektur in ausgewählten Use Cases.
- Entwurf und Fortschrieb einer »Standardisierungslandkarte«.
- Entwurf des Geschäftsmodells der Betreibergesellschaft des Industrial Data Space.
- Entwurf des Zertifizierungskonzepts und eines Geschäftsmodells für die Zertifizierungsstelle des Industrial Data Space.
- Entwicklung einer Methodik für Anwender des Industrial Data Space, um Geschäftsstrategien entsprechend der neuen Anforderungen der Digitalisierung zu transformieren.

- Ableitung von Handlungsempfehlungen für den Betrieb des Industrial Data Space sowie Identifikation neuer Forschungsfelder zur nachhaltigen Entwicklung des Industrial Data Space.

Innovationen gegenüber dem bekannten Stand der Technik und dafür zu schaffen- des Wissen: Das Vorhaben Industrial Data Space adressiert das Fehlen einer ganzheitlichen Architektur, welche neben den zuvor aufgeführten Punkte im Besonderen den Datenaustausch und die Vernetzung zwischen Unternehmen regelt und dabei Datenschutz und Datensicherheit sicherstellt sowie die weitere Nutzung bestehender IT-Infrastrukturen der Unternehmen gewährleistet. Die Architektur hat das Ziel, Daten und Dienste verschiedener Unternehmen zu verknüpfen und dadurch neue Geschäftsmodelle entstehen zu lassen. Die Architektur besteht aus den Teilarchitekturen:

1. Die Governance-Architektur adressiert Fragen der »Spielregeln« und der Zusammenarbeit der Teilnehmer am InDaSpace.
2. In der Daten- und Servicearchitektur werden die grundlegenden Datenmodelle, Austauschformate und Protokolle im Rahmen von festgelegten Architekturmustern und das Data-Governance beschrieben.
3. Die Softwarearchitektur leistet einer Zerlegung der Gesamtarchitektur in Architekturkomponenten, welche auf der Basis von Architekturmustern interagieren.
4. Die Sicherheitsarchitektur beschreibt die Mechanismen zur Absicherung des Umgangs mit sensiblen Daten, die im Kontext von abrechenbaren Dienstleistungen zur Verfügung gestellt werden.

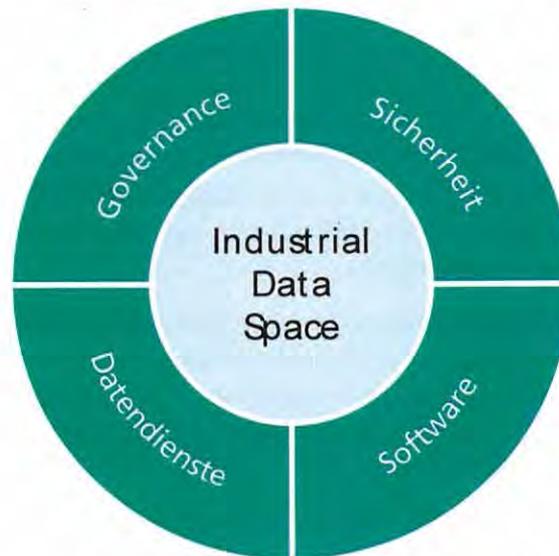


Abbildung 6: Teilarchitekturen des Industrial Data Space

Abbildung 6 stellt die vier verschiedenen Teilarchitekturen grafisch dar.

Das Vorhaben adressiert die folgenden offenen Forschungsfragen, welche sich auf die ganzheitliche Architektur und deren vier Teilarchitekturen beziehen:

1. Welche Aufgaben und Komponenten umfasst eine ganzheitliche Architektur mit dem Ziel einen Datenaustausch und eine Vernetzung zwischen Unternehmen zu ermöglichen und dabei Datenschutz und Datensicherheit sicher zu stellen?
2. Wie interagieren die Architekturkomponenten der Gesamtarchitektur?

3. Welche Gouvernance-Aufgaben, Teilnahme-Bedingungen und Regeln müssen für die Nutzung und die Teilnehmer am InDaSpace definiert und erfüllt werden?
4. Welche Datenmodelle, Standards und Protokolle müssen von der Daten- und Servicearchitektur für den InDaSpace berücksichtigt werden?

Welche Datenschutz- und Datensicherheits-Mechanismen und Standards müssen für die Absicherung der Daten und des Datenaustausch definiert und umgesetzt werden?

Gewählter Lösungsansatz und Begründung:

Das Vorhaben greift die Vorarbeiten aus den verschiedenen Konsortien und Initiativen (siehe Kapitel 2.1) auf. Diese Initiativen erarbeiten allerdings nur Lösungen und Lösungsansätze für einzelne Bereiche (siehe Kapitel 1.5) und Teilarchitekturen des Industrial Data Space. Der Industrial Data Space wird daher insbesondere die ganzheitliche Betrachtung in der Vordergrund stellen und einen ergänzenden ganzheitliche bzw. alternativen Architekturentwurf zu diesen Initiativen entwickeln, welche insbesondere die Vernetzung und die Zusammenarbeit der Unternehmen fördern, indem Daten und Dienste zu neuen Leistungsangeboten bzw. gar Geschäftsmodellen verknüpft werden. Die ganzheitliche Betrachtung ermöglicht im Vergleich zu den oben genannten Ansätzen ist, dass Anbieter und Nutzer von Daten oder Diensten auch ihre bestehenden IT-Infrastrukturen nutzen können. Im Weiteren steht der Vertrauensschutz hinsichtlich der Identität der teilnehmenden Akteure, des Datenaustausches und der Datennutzung besonders im Vordergrund der Entwicklung des Industrial Data Space. Außerdem ist durch die ganzheitliche Betrachtung sichergestellt, dass die Gesamtarchitektur des Industrial Data Space keine Partikularlösung darstellt, sondern offen für alle Branchen und Unternehmen ist.

#### 1.5 Bezug zu den förderpolitischen Zielen

Das Vorhaben Industrial Data Space unterstützt unmittelbar die im Rahmen des »Zukunftsprojekts Industrie 4.0« definierten Handlungsfelder »Standards und IT-Architekturen« und »IT-Sicherheit«.

Im InDaSpace werden Referenzarchitektur, Betriebsmodelle und die entwicklungstechnischen Grundlagen für Softwarekomponenten einer Plattform zum Datenaustausch und Kooperation von Industrieunternehmen entstehen, die auf der Basis des Einsatzes von Standards erhebliche Optimierung der Prozesse in der Produktion und Logistik erlauben.

Sowohl durch technische Methoden als auch durch die Zertifizierung von Teilnehmern und Services nimmt die Sicherheit des Gesamtsystems einen zentralen Stellenwert der F&E-Arbeiten ein.

## 2 Stand der Wissenschaft und Technik

### 2.1 Verwandte Initiativen

Aufgrund der Bedeutung der Digitalisierung und der technologischen Entwicklungen der letzten Jahre in Bereichen wie dem mobilen Internet, Cloud Computing, Embedded Systems oder Big Data formen sich zurzeit weltweit Konsortien im Kontext des Vorhabens. Es wurde von den Antragstellern bereits eine umfassende, detaillierte Umfeldanalyse durchgeführt, in der die relevante Konsortien und Allianzen in Deutschland, Europa und global betrachtet und thematisch zugeordnet wurden (vgl. Abbildung 7), um eine Abgrenzung bzw. Positionierung des Industrial Data Space zu ermöglichen. Die Umfeldanalyse beinhaltet ein Vergleichsschema mit 25 Kriterien zu Formalien und Technologie der als relevant identifizierten 92 Initiativen.

Zu den formellen Aspekten zählten der Umfang der Initiative (Thema, Wirkungsbereich, Anwendungsfelder und Produkte/Ergebnisse), ihre Mitglieder (in Deutschland und Key Player weltweit), die Organisation (Rechtsform, Sitz, URL, Gründungsjahr, Governance Model und Business Model), sowie textuelle Beschreibungen. Im technologischen Teil wurden die Architektur (Designprinzip, Komponenten und Kommunikation), Daten (Art der Daten und Datenhaltung), Services (Schnittstellen und Basisdienste), Technik (Software und Erweiterbarkeit) sowie Security-Aspekte (Authentisierung, Autorisierung und Zugriffskontrolle und Encryption) betrachtet, sofern es dazu öffentliche Informationen gab.

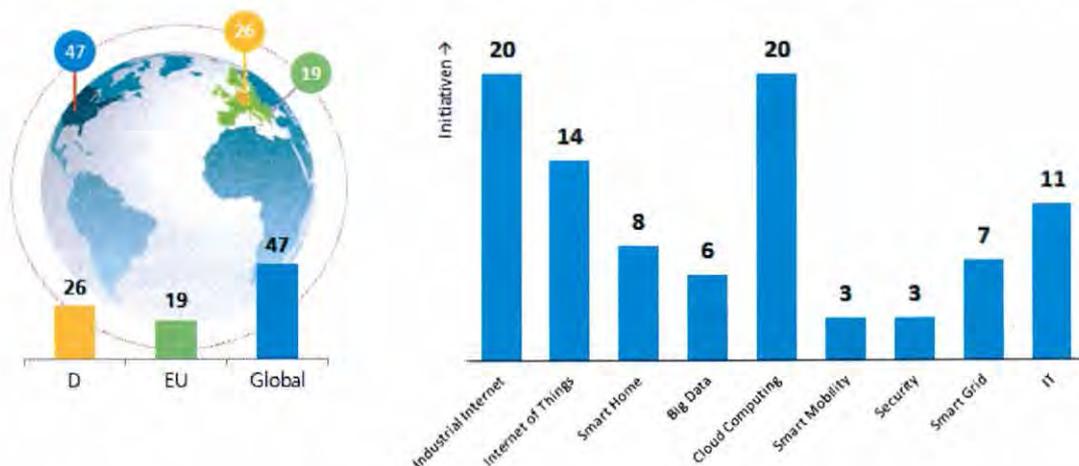


Abbildung 7: Überblick aktueller Digitalisierungsinitiativen

Die Bandbreite der betrachteten Initiativen reicht von schwergewichtigen internationalen Allianzen über nationale Vereinigungen, Verbundprojekte und Konsortien bis hin zu Initiativen einzelner Unternehmen. Die Relevanz für den Industrial Data Space leitet sich jeweils aus thematischen oder organisatorischen Elementen der jeweiligen Initiative ab. Die wichtigsten Initiativen wurden anschließend näher betrachtet und gegenüber dem Industrial Data Space in den Dimensionen Business und Technologie positioniert. Wie Abbildung 8 zeigt, sind diese Initiativen entweder sehr technologisch getrieben, z.B. um konkrete Systeme zu vernetzen, oder sie zielen auf neue Geschäftsmodelle ab (ohne jedoch ausreichend die technologischen Rahmenbedingungen zu berücksichtigen). Bisher schafft es keine der existierenden Initiativen, diese beiden Aspekte (Geschäftsmodelle und technologische Rahmenbedingungen) auf eine Art und Weise gleichermaßen miteinander

zu verbinden, die das Erreichen der Ziele des hier geplanten Vorhabens ermöglichen würde. Dies stellt einen Grund für die Notwendigkeit einer neuen Initiative dar.

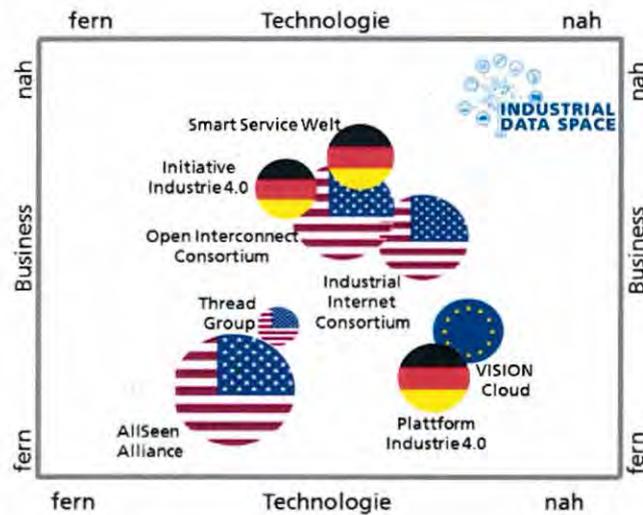


Abbildung 8: Positionierung des Industrial Data Space zu wichtigen Initiativen weltweit

In einem zweiten Teil der Umfeldanalyse wurden über 50 relevante Datenstandards und -Kandidaten (Business Message Standards, Ontologien, Klassifikations-schemata und Identifikationsschemata) für die verschiedenen intendierten An-wendungsfelder untersucht. Diese sollen bei der Realisierung des Industrial Data Space berücksichtigt werden.

Nachfolgend werden die wichtigsten Initiativen in Deutschland und darüber hin-aus kurz skizziert, die in engerem Bezug zum Industrial Data Space gesehen wer-den können.

Die deutsche Initiative Industrie 4.0 ermöglicht gemäß der Umsetzungsempfeh-lungen für das Zukunftsprojekt Industrie 4.0 von acatech »eine firmenübergrei-fende Vernetzung und Integration über Wertschöpfungsnetzwerke« [ACA13]. Sie strebt eine Standardisierung der Mechanismen der Zusammenarbeit an und möchte die auszutauschenden Informationen festlegen. Eine Referenzarchitektur dient als allgemeines Muster für die Produkte und Dienstleistungen aller zusam-menarbeitenden Unternehmen und bildet den Rahmen für die Strukturierung, Entwicklung, Integration und den Betrieb der im Rahmen von Industrie 4.0 rele-vanten technischen Systeme. Unterstützt wird dies durch Software-Applikationen und Software-Services (vgl. exemplarisch Abbildung 9).

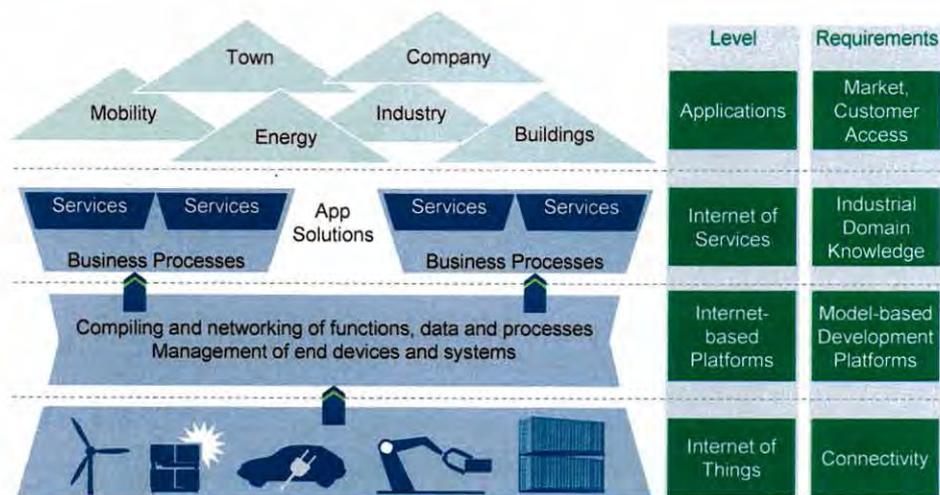


Abbildung 9: Unternehmensreferenzarchitektur für Industrie 4.0 [ACA13]

Die Plattform Industrie 4.0 ist ursprünglich eine Kooperation der drei Industrieverbände BITKOM, VDMA und ZVEI. Sie verfügt über ein starkes Konsortium mit ca. 40 Mitgliedern mit Global Playern aus der IT und knüpft an das Zukunftsprojekt Industrie 4.0 an, welches im Rahmen der »Hightech-Strategie 2020« im November 2011 von der Bundesregierung verabschiedet wurde, um mittels geeigneter Maßnahmen, Deutschland in die Lage zu versetzen, bis 2020 Leitanbieter für cyber-physische Systeme (CPS) zu werden. Arbeitsgruppen zu Referenzarchitektur, Standardisierung, Forschung, Sicherheit, Recht und Aus-/Weiterbildung arbeiten in den Themenfeldern Produktionstechnik, Maschinenbau, Verfahrenstechnik, Automatisierungstechnik. Ein erster Entwurf der Referenzarchitektur RAMI 4.0 wurde bereits vorgestellt (vgl. Abbildung 10).

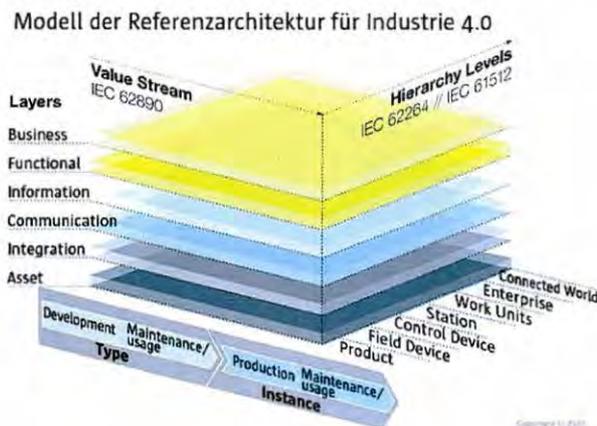


Abbildung 10: Referenzarchitektur RAMI der Plattform Industrie 4.0

Die ebenfalls deutsche Initiative Smart Service Welt der acatech und der gleichnamige Technologiewettbewerb des BMWi adressieren Internetbasierte Dienste für die Wirtschaft. In der Smart Service Welt werden intelligente Dienstleistungen und Produkte, die im Betrieb mit dem Internet verbunden sind, zur Grundlage für zahlreiche neue daten- und dienstbasierte Geschäftsmodelle. Abbildung 11 zeigt das Schichtenmodell digitaler, aufeinander aufbauender Plattfortmtypen der acatech, in der die verschiedenen Layer von technischer Infrastruktur über physis-

sche, softwaredefinierte und Service-Plattformen zu einem digitalen Ökosystem führen sollen.



Abbildung 11: Schichtenmodell der Smart Service Welt [SSW15]

Im Rahmen des Spitzencluster Wettbewerbs des BMBF wird in der zweiten Wettbewerbsrunde der Software-Cluster in der Region um Darmstadt, Walldorf, Kaiserslautern, Karlsruhe und Saarbrücken gefördert. Der Software-Cluster forscht u.a. in vier Verbundprojekten an dem Thema emergente Software insbesondere im Fokus der Dienstleistungserbringung im zukünftigen Internet. Dabei passt sich emergente Software an dynamisch an neue Anforderungen an und erfüllt damit eine wichtige Anforderung an moderne digitale Unternehmen. Das Projekt InDiNet entwickelt eine Plattform und Mehrwertdienste für innovative Dienstleistungen im zukünftigen Internet (siehe Abbildung 12). Dabei steht der Fokus auf Interoperabilität von Services und der Kombination von einzelnen Services zu komplexen Geschäftsprozessen und damit zu neuen Produkten oder Dienstleistungen auf Basis der Unified Service Description Language (USDL).

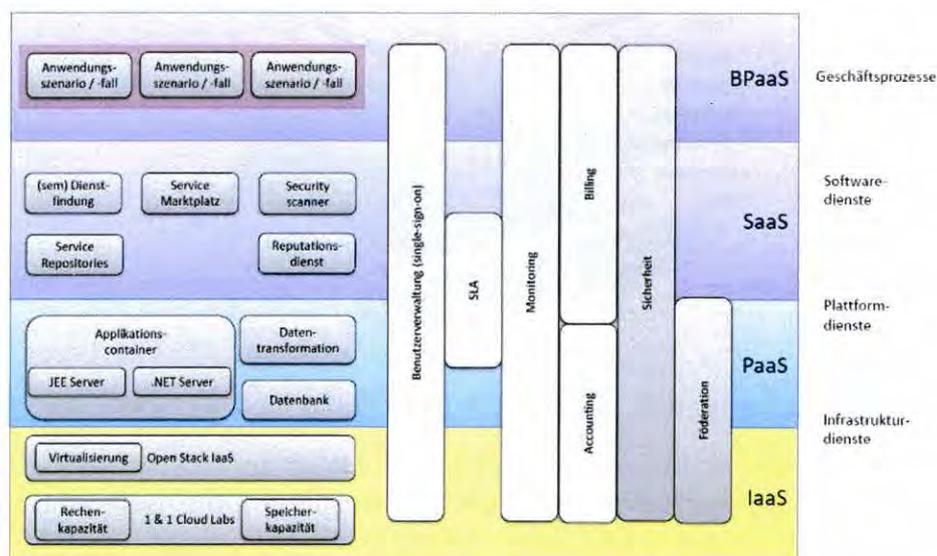


Abbildung 12: InDiNet Architektur (Quelle: www.conweaver.com)

Das Projekt SINNODIUM des Software-Clusters führt die Ergebnisse der übrigen Projekte zusammen und bildet unterschiedliche Anwendungsszenarien (Smart Retail, Smart Production und Smart Services) ab. Dazu werden Plattformen für

Infrastruktur und Entwicklung bereitgestellt und mit emergenten Software-Komponenten und Services bestückt, die die Basis für hochindividualisierte Geschäftsprozesse bilden (siehe Abbildung 13). Damit fokussiert sich der Software-Cluster auf softwaretechnische Themen zur funktionalen Kombination von Software-Komponenten. Die Rolle der Daten steht dabei im Fokus der Komponenten und damit nicht im thematischen Fokus des Software-Clusters.

Ferner gibt es noch die Verbundprojekte EMERGENT (Grundlagen emergenter Software) und SWINNG (Prozessinnovationen in der Software-Industrie), die aber zunächst weniger relevant für den InDaSpace sind.

Der Software-Cluster hat während seiner Laufzeit seit 2011 eine große Menge an Unternehmen aus dem ITK-Bereich als Partner gewonnen. Mit dem Business Portal wurde weiterhin eine zentrale Anlaufstelle für interessierte Unternehmen geschaffen.

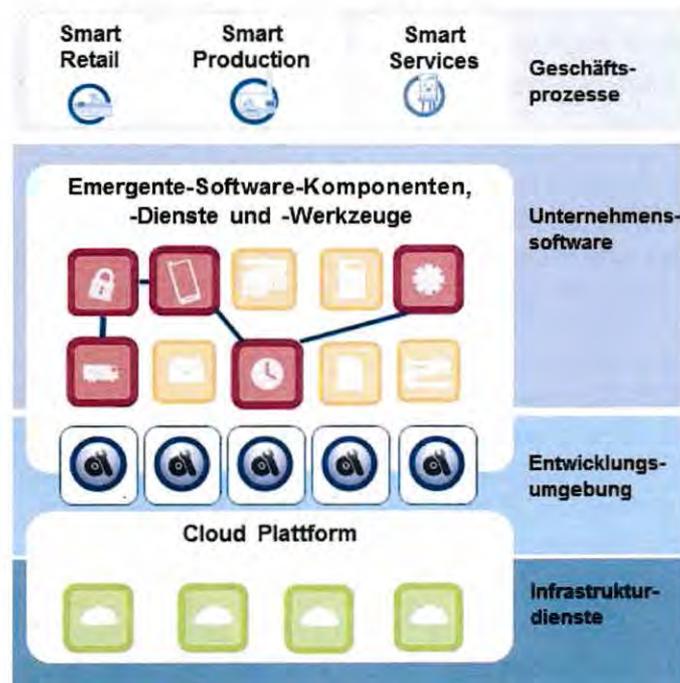


Abbildung 13: SINNODIUM Architektur (Quelle: [www.software-cluster.org](http://www.software-cluster.org))

Während sich der Software-Cluster auf die dynamische und flexible Anpassung von Geschäftsprozessen durch emergente Software-Komponenten fokussiert, basiert der Ansatz des InDaSpace auf der effizienten Nutzung von Daten. Wichtige Ziele des Industrial Data Space, die vom Software-Cluster nicht adressiert werden, beinhalten die folgenden:

- Unterstützung einer sicheren standardisierten unternehmensübergreifenden (und branchenübergreifenden) Kommunikation unter Erhalt der Datensouveränität
- Datenzentrierte Gesamtarchitektur unter Berücksichtigung betriebswirtschaftlicher Fragen
- Allgemein neue effiziente Kooperationsmöglichkeiten

Somit werden sich potentielle Nutzergruppen und deren Geschäftsmodelle unterscheiden, aber gut ergänzen, da emergente Software und Prozesse von dem sicheren Datenraum InDaSpace profitieren werden:

- Der Softwarecluster adressiert innovative Geschäftsmodelle für die wirtschaftliche Verwertung von Software (für die Softwareindustrie), sowie Software passend zum Geschäftsmodell und Business DANN.
- Demgegenüber adressiert InDaSpace speziell datenzentrierte Geschäftsmodelle, unternehmens- und branchenübergreifende und auch weltweite, kollaborative Geschäftsmodelle, Geschäftsmodelle der Zertifizierung und die Weiterentwicklung bestehender Geschäftsmodelle, sowie das Ziel, Daten und Dienste verschiedener Unternehmen zu verknüpfen und dadurch neue Geschäftsmodelle entstehen zu lassen.
- InDaSpace-Nutzergruppen wollen speziell ihre Datensouveränität erhalten, ihnen ist Datenstandardisierung und unternehmens/branchenübergreifende Kommunikation wichtig.
- Ferner benötigen potentielle Betreiberorganisationen Kompetenzen der Datenanalyse.

Nichtsdestotrotz sind natürlich die erarbeiteten Ergebnisse des Software-Clusters von sehr großer Relevanz für den Industrial Data Space, und es ist geplant, diese Ergebnisse soweit wie möglich zu verwenden und auf ihnen aufzubauen. Dies wird zum Einen dadurch unterstützt, dass sich Teilnehmer des Software-Clusters auch am Industrial Data Space beteiligen. Darüberhinaus ist geplant, vierteljährliche Workshops zur Abstimmung mit Repräsentanten des Software-Clusters durchzuführen, um auf diese Weise sicherzustellen, dass die beiden Vorhaben wechselseitig von den erzielten Ergebnissen profitieren können, und um eine gemeinsame Zielrichtung abzustimmen. Zur Unterstützung dieser Aktivitäten kann dabei, falls notwendig, auch auf die im Industrial Data Space vorgehaltenen „NN-Mittel“ zurückgegriffen werden.

Das Projekt SemProm verfolgte von 2008 bis 2011 die Erschließung des Nutzenpotenzials der semantischen Interoperabilität von Softwarekomponenten entlang der gesamten Wertschöpfungskette bis zum Endkunden. Dies wurde durch die Weiterentwicklung von mobilen, eingebetteten und funkbasierten Elementen ermöglicht. Im Sinne eines „Internet-der-Dinge“ sind diese in der Lage, alle Bestands- und Betriebsdaten eines Produktes zu erfassen und mit der Umgebung, den Nutzern und anderen Produkten auszutauschen. Der Industrial Data Space fokussiert nicht auf die Daten einer bestimmten Klasse (u.a. Internet-der-Dinge Daten) sondern hat die allgemeine und effiziente Nutzung sowie den Austausch von Daten aus unterschiedlichen Quellen und Systemen im Fokus mit dem Ziel, kooperative Geschäftsmodelle zu ermöglichen. Eine semantische Integration, Vokabulare und ein automatisches Mapping stehen nicht im direkten Fokus. Werden entsprechende Ansätze benötigt, wird auf entsprechende Vorarbeiten (u.a. SemProm) zurückgegriffen.

Im internationalen Umfeld stehen die nachfolgenden Initiativen in einem Bezug zum Industrial Data Space.

Das Industrial Internet Consortium hat zum Ziel Anwendungsfelder und Demonstrationsszenarien für die Verbindung physischer und digitaler Produkte und Dienstleistungen für digitale Unternehmen zu entwickeln. Dabei steht auch die Entwick-

lung einer Referenzarchitektur und von Frameworks zur Interoperabilität im Fokus. Das Industrial Internet Consortium differenziert dabei die vier unterschiedlichen Perspektiven Business, Usage, Functional und Implementation (siehe Abbildung 14) Damit soll die Adaption von innovativen Technologien und Methoden für Unternehmen erleichtert werden. Das IIC sieht sich als Plattform zum Erfahrungsaustausch zwischen weltweit führenden Industrieunternehmen, das auch gemeinsame Testbed-Implementierungen anregt, aber keine eigene Intellectual Property, Systeme oder Standards selbst entwickeln will. Mehrere deutsche Unternehmen, die auch am Industrial Data Space Interesse haben, und auch der Fraunhofer-IuK-Verbund sind Mitglieder des IIC.

Das Open Interconnect Consortium entwickelt Spezifikationen für zuverlässige Interoperabilität im Internet of Things, Dazu haben sich internationale Unternehmen, wie Samsung oder Intel zusammengeschlossen. Die Thread Group konzentriert sich auf den Bereich Smart Home, arbeitet aber ebenfalls an der Interoperabilität im Internet of Things. An dieser Initiative ist unter anderem Google oder der deutsche TÜV Rheinland beteiligt.

Als nonprofit Organisation arbeitet die AllSeenAlliance an einem Framework für das Internet of Everything, also dem Zusammenschluss von Geräten, Diensten und Systemen im Kontext der

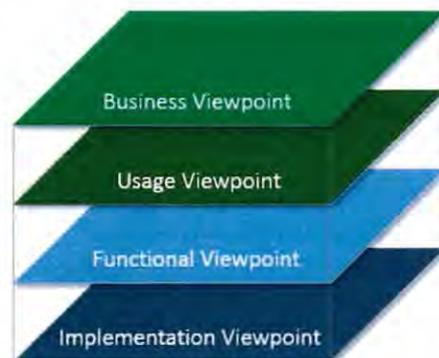


Abbildung 14: Perspektiven in der Referenz Architektur des Industrial Internet Consortium (Quelle: <http://www.iiconsortium.org/IIRA-1-7-ajs.pdf>)

Industrie und des Smart Home. In verschiedenen Working Groups werden Komponenten Standards entwickelt, wobei sich die bestehenden Working Groups auf Basistechnologien und Smart Home konzentrieren. Auch die Showcases beziehen sich hauptsächlich auf Unterhaltungselektronik.

Der ENX Verband ist eine Non-Profit-Organisation europäischer Automobilhersteller, -Zulieferer und Verbände und fungiert als rechtliches und organisatorisches Dach für den Netzwerkstandard ENX (European Network Exchange). Er bietet beteiligten Unternehmen eine Plattform zum sicheren Informationsaustausch auf Basis verschiedener IP-gesegnete Protokolle und Applikationen und reduziert Kosten und Komplexität im unternehmensübergreifenden Datenaustausch in der Automobil-Branche zur Ermöglichung kooperativer Produktionsmodelle. Im Gegensatz zum InDaSpace liegt der Fokus auf nur einer Branche und auf gesicherte, konsolidierte, standardisierte Kommunikation (Branchennetzwerk), weniger auf den kommunizierten Daten. Wegen seiner starken Verbreitung sollte ENX jedoch für den InDaSpace berücksichtigt werden: Im März 2013 nutzen über 1.500 Unternehmen der Automobilindustrie (und anderer Branchen) in über 36 Ländern

das weltweit verfügbare Netzwerk. Auch Unternehmen in Japan und Nordamerika sind mittlerweile angeschlossen.

Weiterhin zu berücksichtigen sind relevante Vorarbeiten in den BMBF-Kompetenzzentren Big Data (Berlin, Leipzig) und Sicherheit (Darmstadt, Karlsruhe, Saarbrücken). Dabei ist das geplante Vorhaben mit zweien dieser Zentren über SIT und IOSB direkt verbunden ist, und bildet zu den Big Data-Zentren insofern ein Komplement, als sich diese auf High-Performance-Massendatenverwaltung und -Analytik konzentrieren, während im geplanten Vorhaben der P2P-Austausch zwischen autonomen Playern im Vordergrund steht.

Bestehende Initiativen leisten einen wichtigen Beitrag für die Entwicklung der Gesamtarchitektur des Industrial Data Space, weil sie einerseits ein konzeptionelles Fundament schaffen, auf dessen Basis der Industrial Data Space errichtet werden kann, und andererseits Analysen der technischen Machbarkeit und Reife (z. B. von Standards) liefern, die zu einer Beschleunigung der Arbeiten im hier skizzierten Vorhaben führen.

Zudem adressiert der Industrial Data Space einen Bedarf der Praxis und eine Lücke in der Forschung, die von keiner der verwandten Initiativen adressiert wird. So verbleiben viele Initiativen auf konzeptioneller Ebene und verzichten auf Pilotimplementierungen. Andere Vorhaben adressieren zudem lediglich einzelne Branchen und Anwendungsdomänen. Und kaum eine Initiative verfolgt das Ziel einer Gesamtarchitektur, die explizit nicht nur die technische Grundlage adressiert, sondern auch betriebswirtschaftliche Fragen (z. B. zur Governance, zum Kooperationsmodell und zum Wert der Daten) aufgreift. Und schließlich nimmt der Industrial Data Space konsequent die Perspektive der Daten ein.

Der Industrial Data Space adressiert das Fehlen einer ganzheitlichen Architektur, welche neben den zuvor aufgeführten Punkte im Besonderen den Datenaustausch und die Vernetzung zwischen Unternehmen regelt und dabei Datenschutz und Datensicherheit sicherstellt sowie die weitere Nutzung bestehender IT-Infrastrukturen der Unternehmen gewährleistet. Im Mittelpunkt des Vorhabens steht die Beantwortung der Frage, wie mit Daten in kollaborativen Szenarien umgegangen werden muss. Die Frage wird mit der Entwicklung einer ganzheitlichen Architektur adressiert, welche aus den folgenden Teilarchitekturen besteht:

1. Die Geschäftsarchitektur adressiert Fragen der »Spielregeln« und der Zusammenarbeit der Teilnehmer am Industrial Data Space.
2. In der Daten- und Servicearchitektur werden grundlegenden Datenmodelle, Austauschformate und Protokolle im Rahmen von festgelegten Architekturmustern und das Data-Governance beschrieben.
3. Die Softwarearchitektur definiert die einzelnen Architekturkomponenten der Gesamtarchitektur und deren Interaktion.
4. Die Sicherheitsarchitektur beschreibt die Mechanismen zur Absicherung des Umgangs mit sensiblen Daten, die im Kontext von abrechenbaren Dienstleistungen zur Verfügung gestellt werden.

## 2.2 Technologien und eigene Vorarbeiten

Nachfolgend sind einige der grundlegenden Ansätze und Technologien benannt, die bei der Realisierung des Industrial Data Space Anwendung finden sollen. Einen Teil dieser Technologien hat Fraunhofer aktiv mitentwickelt.

Software-technische Herausforderungen: Wie in jedem großen Entwicklungsvorhaben sind hohe Ansprüche an die resultierende Software zu erfüllen. Dazu zählen unter anderem Zuverlässigkeit, Leistung, Effizienz, Wartbarkeit und Skalierbarkeit. Besonders hervorzuheben ist die Architektureigenschaft "Flexibilität" d.h. das neue oder geänderte Anforderungen berücksichtigt werden müssen. Dies wird durch moderne, mehrschichtige und lose gekoppelte Architekturen erreicht, die relativ einfach erweitert oder angepasst werden können.

Als agiler Softwareentwicklungsprozess soll in diesem Projekt Scrum verwendet werden [Pic08]. Dieses auf wenigen Regeln basierende Vorgehensmodell erfreut sich wachsender Beliebtheit in der Softwaretechnik und löst ältere Modelle wie V- oder Wasserfall-Modell ab. Die aus dem agilen Manifest<sup>2</sup> stammenden vier Kernanforderungen werden im Scrum-Modell entsprechend adressiert. Scrum unterscheidet drei zentrale Rollen: einen Product Owner, die Entwickler und den Scrum Master. Der Product Owner entscheidet und verantwortet die Eigenschaften und die zeitliche Fertigstellung des (Software-) Produktes. Die Anforderungen werden in einem Katalog gesammelt, der als Backlog bezeichnet wird. Die Entwickler benennen den Arbeitsaufwand für einzelne Produkt-Eigenschaften und realisieren diese. Zeitrahmen für die Realisierung ist ein festes Intervall von wenigen Wochen das Sprint genannt wird. Ziel jedes Sprints ist es eine neue Variante des Produkts zu liefern, die für eine Nutzung geeignet ist. Der Scrum Master moderiert und betreut das Vorgehen innerhalb des Scrum-Teams. Alle Rollen in Scrum sind unabhängig und sollten keine hierarchischen Beziehungen untereinander enthalten.

Die zügige Umsetzung von Anforderungsänderungen durch die sprint-orientiert Priorisierung des Backlog ist durch Scrum gewährleistet. Für die plangerechte Fertigstellung von Meilensteinen (insbesondere solcher, die mehr als einen Sprint für die Fertigstellung benötigen) muss der Scrum Entwicklungsprozess um ein ganzheitliches Controlling erweitert werden. Ein solches Controlling steuert die langfristige Erreichung von Projektziele im Zusammenspiel mit den kurzfristigen Prioritäten eines Sprints. Diese Herausforderung zusammen mit einem auf mehrere Standorte verteiltem heterogenem Entwicklungsteam muss in einer auf dieses Projekt zugeschnittenen Scrum-Variante berücksichtigt werden.

Strategisches Datenmanagement überträgt Management-Konzepte für die Bewirtschaftung physischer Güter auf immaterielle Datengüter. Grundlegendes Prinzip ist die Bewirtschaftung der Daten hinsichtlich Kosten, Zeit und Qualität. Zu bekannten Ansätzen gehören das Total Data Quality Management, das in den 1990er Jahren am Massachusetts Institute of Technology (MIT) entwickelt wurde [WLPS98], Strategic Data Planning [GKQW92], Data Warehouse Quality (DWQ) [JJQV99, JLVV03], sowie das EFQM Excellence Model for Corporate Data Quality Management [EFQM13]. Im EU-Grundlagenprojekt wurde unter Leitung des Fraunhofer FIT ein Qualitätsmodell für die integrierte Informationssysteme (Data-Warehouse-Systeme) entworfen. Das Vorgehensmodell zum Datenqualitätsmanagement wurde vom FIT auch auf datenstrom-basierte Big-Data-Anwendungen für Verkehrsinformationssysteme übertragen [GWQ11]. Das EFQM Excellence Model for Corporate Data Quality Management ist Ergebnis des Kompetenzzentrums Corporate Data Quality (CC CDQ), einem Konsortialforschungsprojekt zur Entwicklung von Methoden und Lösungen zum qualitätsorientierten Datenmanagement.

---

<sup>2</sup> Vgl. <http://www.agilemanifesto.org>

an dem auch das Fraunhofer-Institut für Materialfluss und Logistik mitwirkt. Aktuelle Forschungsthemen umfassen die ökonomische Bewertung von Daten und Datenarchitekturen für die Digitalisierung. Beispielsweise gelten die traditionellen Kategorien für Wirtschaftsgüter auch für Daten, sodass zwischen öffentlichen Gütern (z. B. Wetterdaten), Club- bzw. Gemeinschaftsgütern (z. B. gemeinsame Stammdaten einer Lieferkette) und privaten Gütern (z. B. Zeichnungsdaten eines Maschinenbauers) unterschieden werden kann.

Linked Data ist ein pragmatischer Ansatz zur Realisierung eines semantischen Daten-Web. Linked Data basiert auf dem Grundprinzip Entitäten und Datensätze mit URIs zu identifizieren, diese URIs über das HTTP-Protokoll de-referenzierbar zu machen, Beschreibungen der Entitäten [REDACTED] auszuliefern und diese Beschreibungen mit Verweisen (Links) zu anderen Datensätzen und Entitäten zu versehen (vgl. [LW09]). Mittlerweile wurden 50 Milliarden Fakten frei im Web entsprechend den Linked Data Prinzipien publiziert (vgl. [BEM+13]). Im Rahmen des vom Fraunhofer IAIS koordinierten LOD2 Projektes wurde eine umfassende Unterstützung des Linked Data Lebenszyklus [ALN11] entwickelt, die Werkzeuge für der Phasen Extraktion, Speicherung, Verlinkung, Authoring, Anreicherung, Qualitätskontrolle, bis hin zur Visualisierung umfasst (vgl. [ABD+12]). Unternehmen (z.B. Daimler, Google, BBC) haben begonnen Linked Data zu nutzen um interne und externe Daten zu verknüpfen und zu integrieren.

Datenaustauschformate: Der Datenaustausch im IDS erfolgt durch die Verwendung von gemeinsamen Vokabularen für die unterschiedlichen Domänen. Diese Vokabulare werden als grundlegendes Datenformat [REDACTED] verwenden.



- Vokabulare: Im IDS werden zwei unterschiedliche Typen von Vokabularen definiert. Einerseits werden für domänenunabhängige Beziehungen im IDS

<sup>3</sup> <https://developers.google.com/protocol-buffers/>

<sup>4</sup> <https://thrift.apache.org/>

zentrale Vokabulare

Diese sind teilweise weit verbreitet und haben sich als generische Vokabulare für verschiedene Einsatzumgebungen etabliert. Andererseits werden für die Beschreibung von domänenspezifischen Daten Vokabulare definiert, die auf existierenden Industrie 4.0 Standards basieren. In diesem Zusammenhang werden derzeit die folgenden Standards untersucht: AutomationML<sup>9</sup> ist ein XML-basiertes Datenformat zur Beschreibung von Daten für die Planung von Fertigungs- und Automatisierungsanlagen mit dem Ziel der Realisierung der digitalen Fabriken der Zukunft. eClass<sup>10</sup> ist eine industrieübergreifende Standardisierung zur Klassifikation und detaillierten Beschreibung von Produkten. SCOR<sup>11</sup> ist das Supply Chain Operations Reference Modell, welches die Beschreibung von Zuliefererbeziehungen, Geschäftsprozessen und Leistungsmaßen ermöglicht. OPC UA<sup>12</sup> ist ein Kommunikationsstack zum Austausch von Daten zwischen industriellen Maschinen. Im Gegensatz zu den vorher beschriebenen Standards beschreibt OPC UA auch die Serialisierung und den Transport der Daten.

Hardware-basierte Sicherheit. Sicherheit basierend auf Hardware-Modulen, z.B. HSMs oder TPMs oder Smartcard-Chips ermöglicht sichere Schlüsselspeicherung, einen sicheren Bootvorgang und eine Überprüfung der Integrität von dritter Seite (Remote Attestation). Die im Rahmen dieses Projektes erarbeitete Lösung kann in erster Ausbaustufe mit bereits standardisierten TPMs nach Version 1.2 (vgl. [TCG12]) realisiert werden und basiert damit auf Industriestandards. In weiteren Schritten wird die Lösung auf das angekündigte TPM 2.0 Framework angepasst, wobei auch gegebenenfalls notwendige Erweiterungen des aktuellen Standards zu betrachten sind.

Vertrauenswürdige Client- und Dienstplattform. Zur Isolation der Software-Komponenten auf den Konnektoren wird auf eine Betriebssystem-Level Virtualisierung, basierend auf dem Linux-Kernel, zurückgegriffen. Das Fraunhofer AISEC verfügt über eine eigens entwickelte Lösung zur Absicherung von mobilen Endgeräten (vgl. [WSH+13]) und hat eine Lösung zusammen mit Industriepartnern zur Marktreife gebracht (vgl. [FAT15]). Weiterhin gibt es bereits etablierte Lösungen, die ähnlichen Ansätzen folgen und als Open-Source Lösung verfügbar sind (vgl. [DOC15], [RKT15]). Auch Integritätsnachweise auf Basis dieser Technologien sind erprobt (vgl. [WWH+14]), müssen aber in neuem Kontext umgesetzt werden. Neben der Betriebssystem-level Virtualisierung kann im Server-Umfeld, wo große Datenmenge verarbeitet werden sollen und keine aufwändige grafische Darstellung erfolgen soll, in einer späteren Ausbaustufe auch eine Systemvirtualisierung mit sehr kleiner TCB basierend auf dem L4 Microkernel eingesetzt werden. Hier

<sup>9</sup> <https://www.automationml.org>

<sup>10</sup> <http://www.eclass.de/>

<sup>11</sup> <http://www.apics.org/sites/apics-supply-chain-council/frameworks/scor>

<sup>12</sup> <https://opcfoundation.org/about/opc-technologies/opc-ua/>

könnten vertrauenswürdige Anwendungen zur Datenverarbeitung und -auswertung direkt als minimale L4-Tasks ausgeführt werden und somit stark in Ihren Kommunikationsmöglichkeiten beschnitten werden um auch hier einen ungewollten Abfluss von Informationen zu verhindern.

Policies und Usage Control. Definitionen von Policies für die Zugriffskontrolle sind als Spezifikationen verfügbar<sup>13</sup> und Policies zur Nutzungskontrolle von Daten sind seit etlichen Jahren Gegenstand der Forschung<sup>14</sup> [HPB+07]. Ihre Umsetzung erfordert jedoch vertrauenswürdige Plattformen und Mechanismen zum Austausch von Daten und Policies ([CL08], [GBA14]). Fraunhofer AISEC hat eine solche vertrauenswürdige Plattform bereits für mobile Plattformen entwickelt [WSH+13]. Die Ergebnisse bisheriger Forschung werden mit diesen Entwicklungen im InDaSpace zusammengeführt, so dass eine sinnvolle Usage Control möglich wird.

Standards für elektronischen Datenaustausch. Seit vielen Jahren existieren Standards für den elektronischen Datenaustausch wie EDIFACT, RossetaNet oder ebXML. *UN/EDIFACT* ist ein branchenübergreifender internationaler Standard für das Format elektronischer Daten im Geschäftsverkehr. *ebXML* ist eine 1999 gestartete, gemeinsame Initiative von UN/CEFACT und OASIS, deren Ziel die Entwicklung eines technischen Rahmens zur Nutzung von XML für elektronische Geschäftsprozesse ist. Leider hat ebXML das ursprüngliche Ziel der Senkung der Eintrittsbarrieren für klein- und mittelständische Unternehmen (KMU) und Entwicklungsländer nur begrenzt erreicht, wie z.B. die kritische Analyse von Rawlins EC Consulting zeigt<sup>15</sup>. Insbesondere für dynamische Wertschöpfungsketten und informellere Daten- und Informationsflüsse zwischen Unternehmen hat sich gezeigt, dass XML und transaktions-basierte Datenaustauschprozesse immer noch zu aufwendig und unflexibel sind. InDaSpace stellt daher eine optimale Ergänzung zu Standards wie EDI, ebXML usw. dar, da die Einrichtung eines InDaSpace-Endpunktes relativ einfach ist und neue Informationsstrukturen zum Datenaustausch nicht zentral standardisiert werden müssen

Automatisches Mapping und Integrations-Workflows. Die Vielzahl und Heterogenität von Datenquellen verlangt im Industrial Data Space eine modellgetriebene Automatisierung der für Datenfusion und Datenanalyse notwendigen Transformationen nicht nur auf Modellebene, sondern auch auf Datenebene; dabei sind nicht nur strukturierte Daten, sondern zunehmend auch Text- und Multimedia-Daten einzubeziehen. Bisherige Lösungen bearbeiten dieses Problem überwiegend aus Sicht eines zentralen Servers und für relationale oder XML-Daten. Am Fraunhofer FIT wurden in Verallgemeinerung von Vorarbeiten der IBM auf den Fall heterogener Datenmodelle Methoden und Werkzeuge entwickelt, die Modell-Mappings bis hin zur Generierung von effizienz-optimiertem Datentransformationscode weitgehend automatisieren, aber darüber hinaus auch Ziele wie etwa die Minimierung des integrierten Ergebnis-Schemas berücksichtigen [KCL+09, JJQ14]. Sowohl aus Gründen des Dateneigentums als auch aus Effizienzgründen und zur Vorverarbeitung multimedialer Daten ist zudem im Industrial Data Space eine Repräsentation und situationsoptimierte Auswahl mehrschrittiger Verarbeitungsstrategien erforder-

<sup>13</sup> OASIS. eXtensible Access Control Markup Language (XACML) Version 3.0. <http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.pdf>, 2015

<sup>14</sup> Alexander Pretschner, Manuel Hilty, Davin Basin. Distributed usage control. In Communications of the ACM - Privacy and security in highly dynamic systems, 2006

<sup>15</sup> Rawlins EC Consulting 2007, ebXML - A Critical Analysis, <http://www.rawlinsecconsulting.com/ebXML/index.html>

derlich; am Beispiel des Use Case verteilte Hochdurchsatz-Medizinforschung wurde am FIT das ZETA-Workflowsystem entwickelt und industriell demonstriert, welches einen Ausgangspunkt für die Lösung dieser Fragestellung bieten kann.

Compliance und Transparenz. Bei unternehmensübergreifenden Geschäftsprozessen und dem damit verbunden Austausch von Daten sind verschiedene Compliance-Auflagen zu berücksichtigen. So muss z.B. geprüft werden, welche Daten das Unternehmen überhaupt verlassen dürfen, oder ob z.B. eine Anonymisierung notwendig ist. Diese Problematik ist im Industrial Data Space ebenfalls zu adressieren. Um eine möglichst automatisierte Prüfbarkeit zu erreichen, kann hier auf bestehende Vorarbeiten (vgl. [WWH+12], [HWP+2014], [Jü14], [ACJ15]) des Fraunhofer ISST zurückgegriffen werden, die z.T. im Rahmen des Attract-geförderten Projektes „Architectures for Auditable Business Process Execution - Monitoring, Control, and Compliance (APEX)“ entstanden sind.

Automatische Synthese. Die automatische Synthese von komplexen, typkorrekten Programmen aus Komponentensammlungen erlaubt die Generierung von smarten Datenadaptern (Adaptercode) inklusive der algebraischen Datentransformationen sowie deren Protokollen auf der Basis von semantischen Spezifikationen. Staged Composition Synthesis (SCS) [DMR15,BDD+14] bietet ein Modell für die Synthese, welches bei der Codesynthese die Implementierungssprache (Java, C#, XML) und die Metasprache für die Synthesekontrolle (funktional) durch Stufen trennt jedoch eine logische sowie semantische Verbindung präsentiert.

### 2.3 Schutzrechte

Es sind keine dem Gegenstand des Vorhabens entgegenstehende Schutzrechte bekannt.

### 3 Arbeitsplan

Nachfolgend wird der Arbeitsplan des Vorhabens beschrieben. Abbildung 15 gibt einen Überblick über die verschiedenen Arbeitspakete.

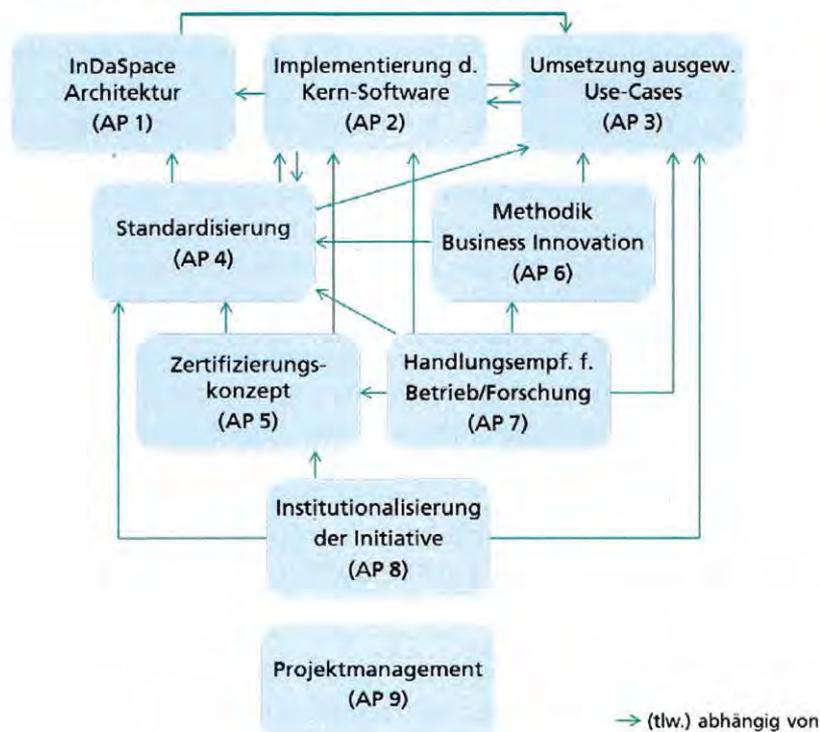


Abbildung 15 Übersicht Arbeitspakete

#### 3.1 InDaSpace Architektur

Die InDaSpace Architektur gliedert sich insgesamt in Teilarchitekturen. Zu den Teilarchitekturen gehören die Geschäftsarchitektur, die Daten- bzw. Daten-Service-Architektur, die Software-Architektur und die Sicherheitsarchitektur. Durch das Vorgehensmodell Scrum und regelmäßige Meilensteine wird die Architektur in Abhängigkeit des Backlog und des Projektfortschritts iterativ fertiggestellt. Im Rhythmus der Sprints werden neue oder geänderte Anforderungen aufgenommen und in der Architektur berücksichtigt. Die Architekturdokumente werden regelmäßig aktualisiert und für die Entwicklungsteams veröffentlicht. In den Meilensteinen wird dieser Zyklus zusätzlich formal dokumentiert.

#### Arbeitspaketbeschreibung

AP-Nr.: 1	InDaSpace Architektur
Startdatum:	M1
Enddatum:	M18
<p>1. Ziel des Arbeitspaketes</p> <ul style="list-style-type: none"> <li>- Gesamtlösung unter Berücksichtigung von geschäftlichen, fachlichen und technischen Aspekten wird konzipiert</li> <li>- Organisatorische und technische Architekturmuster werden evaluiert und ausgewählt</li> <li>- Referenzarchitektur bestehend aus den Teilarchitekturen Geschäftsarchitektur, Daten-Service-Architektur, Softwarearchitektur und Sicherheitsarchitektur wird bereitgestellt.</li> </ul>	

<p>2. Voraussetzung (Input)</p> <ul style="list-style-type: none"> <li>• Anforderungen aus den Use-Cases</li> </ul>	
<p>3. Lösungsweg</p> <ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 1: Geschäftsarchitektur</i> <ul style="list-style-type: none"> <li>- Ableitung von Architekturmerkmalen aus den Anforderungen der Industriepartner und InDaSpace Grundprinzipien</li> <li>- Modellierung von Interaktionen der InDaSpace Akteure und Festlegung von Interaktionsregeln</li> <li>- Ableitung von Anforderungen an Datenqualität, -sicherheit und Verfügbarkeit</li> <li>- Konsolidierung der Teilergebnisse in der Geschäftsarchitektur</li> </ul> </li> </ul>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 2: Daten- und Service Architektur</i> <ul style="list-style-type: none"> <li>- Ableitung von Architekturmerkmalen aus den Anforderungen der Industriepartner, den Use-Cases und InDaSpace Grundprinzipien</li> <li>- Definition des Datenmodells für die Datenintegration und den Datenaustausch.</li> <li>- Festlegung einer Daten-„Hochsprache“ für die semantische Beschreibung von Datenquellen.</li> <li>- Definition eines Metadaten-Modells zur Beschreibung von Mappings zwischen Datenmodellen</li> <li>- Definition einer Strategiesprache (Integrations- und Verteilungsworkflow) für Datentransformations- und Austauschstrategien unter Berücksichtigung der Kommunikationsinfrastruktur</li> <li>- Festlegung von Datenformaten, Protokollen und Schnittstellen für die verschiedenen Datenabrufszszenarien</li> <li>- Evaluierung und ggf. Auswahl von service-orientierten Architekturmustern.</li> <li>- Konsolidierung der Teilergebnisse in der Daten- und Servicearchitektur.</li> </ul> </li> </ul>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 3: Softwarearchitektur</i> <ul style="list-style-type: none"> <li>- Ableitung von Architekturmerkmalen aus den Anforderungen der Industriepartner, den Use-Cases und InDaSpace Grundprinzipien</li> <li>- Modellierung der Softwareeigenschaften auf der Basis der funktionalen und nicht-funktionalen Anforderungen</li> <li>- Ergänzung und vollständige Beschreibung der zur entwickelnden Softwarekomponenten</li> <li>- Evaluierung und ggf. Auswahl von Architekturmustern.</li> <li>- Festlegung und Beschreibung der Datentransformations- und Austauschprozesse einschließlich der Verteilungsoptionen</li> <li>- Festlegung und Beschreibung der durch die Software unterstützten Prozesse</li> <li>- Entwicklung interner Test- und Monitoring-Funktionen und deren Integration in die InDaSpace-Architektur. Hier insbesondere die Definition geeigneter Managementschnittstellen und deren Datenformate. Dies sind interne Schnittstellen die nicht publiziert werden aber ebenso gesichert werden müssen wie die anderen InDaSpace Schnittstellen (Grundschutz der Infrastruktur).</li> <li>- Spezifikation der Daten, die über das Monitoring abgerufen</li> </ul> </li> </ul>	

<p>werden können (z.B. Zugriffsüberwachung / Logging der Zugriffe auf InDaSpace Schnittstellen).</p> <ul style="list-style-type: none"> <li>- Konsolidierung der Ergebnisse in der Softwarearchitektur</li> </ul>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 4: Sicherheitsarchitektur</i> <ul style="list-style-type: none"> <li>- Ableitung von Architekturmerkmalen aus den Anforderungen der Industriepartner, den Use-Cases und InDaSpace Grundprinzipien</li> <li>- Grundschatz in der Architektur: Spezifikation der Architektureigenschaften und Mechanismen, die den Grundschatz in der Architektur sichern</li> <li>- Entwicklung eines mehrstufigen Sicherheitskonzeptes anhand von verschiedenen Schutzklassen für Daten bzw. Use-Cases</li> <li>- Modellierung der verschiedenen Sicherheitsstufen und Ableitung eines Sicherheitskonzeptes auf technischer Ebene (weitere Ebenen werden in der Geschäftsarchitektur berücksichtigt).</li> <li>- Entwicklung eines geeigneten Rollenmodells für den InDaSpace; geeigneter Policies für den Zugriff von Rollen auf Datenobjekte und einer Policy-basierten Architektur für die Durchsetzung des Rollenmodells (Policy Enforcement); der Vertrauensmodelle, insbesondere zu den Modellen zu Übertragung von Vertrauen (z.B. Übertragung von Vertrauen in die herausgebende Instanz in Vertrauen in die Korrektheit der Daten);</li> <li>- Abbildung der Sicherheitsarchitektur auf das Unterliegende Netzmodell zur Definition geeigneter unterstützender Netzdienste (z.B. durch geeignete Orchestrierung im Kontext Softwaredefinierter Netze um Isolation der Domänen untereinander und Verschleierung der Kommunikation bzw. deren Metadaten zu erreichen); als Betrag zur Erstellung von Best Practice Dokumenten zur Unterstützung der industriellen Anwender in der Implementierung dieser Ansätze.</li> <li>- Konsolidierung der Ergebnisse in der Sicherheitsarchitektur.</li> </ul> </li> </ul>	
<p>Personalaufwand gesamt (in PM)</p>	
<p>4. Ergebnis (Output)</p> <ul style="list-style-type: none"> <li>- Finales Dokument für die Geschäftsarchitektur liegt vor.</li> <li>- Technologien, Software und Protokolle sind evaluiert und ausgewählt.</li> <li>- Finales Dokument für die Daten- und Servicearchitektur liegt vor.</li> <li>- Finales Dokument für die Softwarearchitektur liegt vor.</li> <li>- Finales Dokument für die Sicherheitsarchitektur liegt vor.</li> <li>- Finales Dokument für die Gesamtarchitektur liegt vor.</li> </ul>	
<p>5. Sonstige Ressourcen</p> <ul style="list-style-type: none"> <li>▪ Keine</li> </ul>	

### 3.1.1 Geschäftsarchitektur

Die Geschäftsarchitektur adressiert Fragen der »Spielregeln« und der Zusammenarbeit der Teilnehmer am InDaSpace. Sie besteht ihrerseits aus drei Teilarchitekturen:

Erstens legt die Data-Governance-Architektur Regeln für die Bewirtschaftung und Nutzen der Daten im Industrial Data Space fest. Sie regelt die Anforderungen an Datenquellen, die für den InDaSpace zugelassen sind, sowie für die Provider, also

die Eigner der Datenquellen. Ferner trifft sie Aussagen zur Datenqualität, regelt bspw. ob es Datenqualitätssicherungsmechanismen im InDaSpace gibt und wer diese festlegt. Außerdem setzt sie fest, ob es im InDaSpace Data Stewards gibt und, wenn ja, für welche Datenklassen mit welchen Aufgaben.

Zweitens trifft die Kooperationsarchitektur Aussagen darüber, wie die einzelnen Teilnehmer im InDaSpace zusammenarbeiten. Das betrifft Fragen, ob Provider bzw. Datenquellen sichtbar sind im Registry und - falls ja - für wen. Sie regelt auch den Umfang mit öffentlich verfügbaren Daten und deren Nutzung. Zudem adressiert sie Haftungsfragen im Falle von Datendefekten bzw. macht Vorgaben für »Roll-Back-Mechanismen« etc.

Drittens definiert die Geschäftsmodellarchitektur für den InDaSpace im fachlichen Sinne die Prozesse fest, wie Zugang und Mitwirkung zum InDaSpace gewährt wird, ob es Zugangskontrollen bzw. »Quality Gates« gibt und wie Zahlungsflüsse für die Nutzung der Daten abgewickelt werden.

Die Daten- bzw. Daten-Service-Architektur unterscheidet Basisdaten- und Mehrwertdatendienste.

Basisdatendienste machen u.a. die Daten verfügbar, transformieren sie auf Zielformate und gewährleisten den sicheren Datenaustausch.

Mehrwertdatendienste hingegen ermöglichen Datenanalysen bzw. Alerting-Funktionen bei Abweichung des Werts der Daten von einem Schwellwert (z. B. bei der Überwachung der Kühlkette in der Lebensmittelindustrie).

### 3.1.2 Daten-Service-Architektur

In der Daten und Servicearchitektur werden grundlegenden Datenmodelle, Austauschformate und Protokolle im Rahmen von festgelegten Architekturmustern beschrieben. Im Kontext des Industrial Data Space werden dazu Schnittstellen für die Anbindung von Unternehmenssystemen und Datenbanken festgelegt, Datenformate und Vokabulare für die semantische Beschreibung von Daten zur Unterstützung der Datenintegration festgelegt sowie die Einbindung von Datendiensten detailliert beschrieben.

### 3.1.3 Software-Architektur

Die Softwarearchitektur leistet einer Zerlegung der Gesamtarchitektur in Architekturkomponenten, welche auf der Basis von Architekturmustern interagieren. Den Architekturkomponenten werden im zweiten Schritt Softwarekomponenten zugeordnet, deren Beziehungen modelliert werden. Von der Software unterstützte Abläufe werden aufgezeigt und funktionale und nicht-funktionale Eigenschaften abgeleitet. Aufgrund der Bedeutung der Sicherheitsaspekte für den InDaSpace, werden diese im Rahmen einer separaten Sicherheitsarchitektur gesondert behandelt.

### 3.1.4 Sicherheitsarchitektur

Die Sicherheitsarchitektur beschreibt die Mechanismen zur Absicherung des Umgangs mit sensiblen Daten, die im Kontext von abrechenbaren Dienstleistungen zur Verfügung gestellt werden. Überdies soll die Sicherheitsarchitektur des InDaSpace jedoch auch als Alleinstellungsmerkmal dienen, das den InDaSpace von anderen, bestehenden Plattformen in positiver Weise abhebt. Im Gegensatz zu anderen Ansätzen soll der InDaSpace gewährleisten, dass Daten-Eigentümer,

wenn sie dies wünschen, die volle Kontrolle über ihre Daten behalten können und sie nicht physisch aus ihrer Vertrauensdomäne herausbewegen müssen. Gleichzeitig soll der InDaSpace flexibel genug sein, um ein Hosten von Daten-Providern in der Cloud zu ermöglichen. Die Sicherheitsarchitektur adressiert dabei folgende Punkte:

- Erarbeitung eines Konzepts und Bereitstellung einer technologischen Lösung in Form von Referenzarchitekturen zum Herstellen eines angemessenen Schutzniveaus für den InDaSpace und zur Sicherstellung des Vertrauens der Nutzer in den InDaSpace
- Konzeption von technologischen Lösungen, die schrittweise erweiterbar sind und unterschiedliche Vertrauens- und Offenheitsstufen unterstützen: von einfachen Lösungen, die schnell unter Nutzung vorhandener Web-Standards umzusetzen sind, aber nur für eingeschränkte Szenarien mit Vertrauensmodellen, die vertraglich abgesichert sein müssen, abzielen, bis hin zu offenen Interaktionsstrukturen, deren Absicherung durch dedizierte, hardware-basierte Konzepte erfolgen.
- Entwicklung von Alleinstellungsmerkmalen, durch die neue Standards in der Nutzung verteilter Daten im Umfeld Industrie 4.0 gesetzt werden.

Unter anderem sollen folgende Sicherheitsaspekte untersucht werden und ggf. Berücksichtigung finden:

#### 3.1.4.1 Verbindungssicherheit

Verbindungen zwischen Endpunkten (d.h. Konnektoren) müssen vertraulich sein und gegen Manipulationen durch beispielsweise Man-in-the-Middle-Angriffe geschützt sein. Die Endpunkte müssen authentisch sein, sodass Verbindungen nicht irrtümlicherweise mit anderen Kommunikationspartnern als den beabsichtigten zustande kommen (Spoofing). Diese Anforderung gilt sowohl für Datenströme als auch für Nachrichten- basierte Kommunikation.

#### 3.1.4.2 Nachrichtensicherheit

Verbindungssicherheit schützt lediglich die Kommunikation zwischen zwei Endpunkten, ist jedoch agnostisch gegenüber Zuständen und Nachrichten. Daher ist es zusätzlich zur Verbindungssicherheit erforderlich, Nachrichten gegen Abhören und Manipulation zu sichern. Neben der Authentizität, also der Echtheit der Kommunikationspartner ist es überdies erforderlich, verlässliche Identitätsinformationen in Form von Attributen zu erhalten, an die Zugriffskontrolle und ggf. weitere Parameter der Nachrichtensicherheit (Lebensdauer, Nutzungsbeschränkung) geknüpft werden können.

#### 3.1.4.3 Identitätsnachweise für InDaSpace-Teilnehmer

InDaSpace-Teilnehmer müssen zum Zwecke der Abrechnung, der Nachrichtensicherheit und Zugriffskontrolle eindeutig identifizierbar sein. Grundsätzlich kann schon der Besitz eines kryptographischen Schlüssels wie z.B. bei der Verwendung von Zertifikaten für den Aufbau von TLS-Verbindungen als Identität eines Endpunktes verstanden werden. Der Nachteil hierbei ist jedoch, dass ein Datum (der Schlüssel) mehreren Zwecken dient (Verschlüsselung, Identifizierung, Autorisierung), wodurch sich verschiedene Probleme ergeben. Stattdessen müssen Identitätsinformationen in Form von Attributen zur Verfügung gestellt werden, die den jeweiligen InDaSpace-Teilnehmer beschreiben. Attribute können z.B. wiederum den Public Key beinhalten, aber auch Informationen über den Speicherort von Daten, das Sicherheitslevel des jeweiligen Konnektors, zertifizierte Eigenschaften des InDaSpace-Teilnehmers, etc. Dies können in einer klassischen PKI in der Re-

gel nur statische Attribute sein. In einem System mit dynamischer Beglaubigung von Attributen wären auch flüchtige Tags und Attribute möglich, beispielsweise der aktuelle Trustlevel.

#### 3.1.4.4 Sicheres Schlüsselmanagement

Die InDaSpace-Teilnehmer werden mit kryptographischem Schlüsselmaterial ausgestattet sein, mit dem sie sich gegenüber einem Identity-Provider oder dem InDaSpace-Broker authentifizieren und mit dem sie sichere Sitzungen mit anderen InDaSpace-Komponenten aufbauen. Das Management dieser Schlüssel umfasst die Erstellung, die Revokation bzw. Erneuerung der Schlüssel, die Validierung ihrer Authentizität, sowie die sichere Speicherung und Anwendung des Schlüsselmaterials. Die Prozesse für die Erstellung und Verteilung des Schlüsselmaterials sollten eng an die der Registrierung der InDaSpace-Teilnehmer geknüpft werden. Eine Variante ist ein zentraler Registrierungsprozess, den alle InDaSpace-Teilnehmer durchlaufen müssen, und in dessen Rahmen sie Schlüsselmaterial erhalten, das von einer zentralen vertrauenswürdigen Certification Authority (CA) erstellt wurde. Dieses Modell entspricht einer klassischen Unternehmens-PKI. Es hat den Vorteil, dass durch den Registrierungsprozess ein einheitliches Sicherheitsniveau sämtlicher InDaSpace-Teilnehmer garantiert werden kann und Schlüssel leicht revoziert werden können. Allerdings hat die Vergangenheit gezeigt, dass zentrale CAs ein äußerst anfälliger Single-Point-of-Failure sein können. Zudem stellt eine solche Architektur eine hohe Eintrittshürde für die Teilnahme am InDaSpace dar. Ein offeneres Konzept sieht vor, dass sich jeder InDaSpace-Teilnehmer sein eigenes Schlüsselmaterial erstellt und damit direkt zur Kommunikation mit InDaSpace-Komponenten befähigt wird. Ausgehend hiervon können sich InDaSpace-Teilnehmer Attribute durch einen vertrauenswürdigen Identity Provider (IdP) ausstellen und an ihr Schlüsselmaterial binden lassen. Durch die Anzahl und Qualität der Attribute können abgestufte Vertrauenslevel realisiert werden, so dass einerseits eine vollständig anonyme Teilnahme am InDaSpace möglich ist (durch die ggf. nur kostenlose Daten verwendet werden können), aber auch eine voll qualifizierte und zertifizierte Identität erstellt werden kann (mit der ggf. höherwertigere Daten genutzt werden können). Grundfunktionen für das Speichern und die Verwendung der Schlüssel kann der InDaSpace-Konnektor bieten. Für höhere Sicherheitsstufen, die für den Zugriff auf sensitive Daten erforderlich sein könnten, werden jedoch zusätzlich auch Hardware-basierte Sicherheits-Lösungen benötigt, um ein angemessenes Schutzniveau zu erreichen.

#### 3.1.4.5 Zugriffskontrolle / Usage Control

Der Zugriff auf die Daten eines InDaSpace-Providers ist an verschiedene Voraussetzungen gebunden. So müssen die Daten ggf. zunächst käuflich erworben werden, bzw. der Consumer muss gewisse Mindeststandards bzgl. Sicherheit und Datenverarbeitung einhalten. Es soll dem jeweiligen InDaSpace-Provider überlassen sein, die Richtlinien zur Kontrolle des Zugriffs auf Basis der Identitätsattribute und ggf. weiterer Bedingungen selbst zu definieren. Die Prüfung und Durchsetzung kann ihm ggf. durch eine Komponente innerhalb des InDaSpace-Konnektors abgenommen werden, so dass es zum Erstellen eines InDaSpace-Providers lediglich erforderlich ist, die entsprechenden Richtlinien zu definieren. Neben der reinen Zugriffskontrolle, die binäre Entscheidungen im Sinne von erlaubt/nicht erlaubt trifft, ist überdies eine Usage Control erforderlich, die die Nutzung von Daten unter bestimmten Voraussetzungen erlaubt. Diese Voraussetzungen müssen an den Consumer kommuniziert werden und deren Einhaltung ggf. durch einen vertrauenswürdigen InDaSpace-Konnektor auf dessen Sicherheitsanforderungen und -

konzepte sichergestellt werden. Dies gilt auch und besonders für InDaSpace-Konnektoren, auf denen Applikationen und Dienste laufen. Dies ist zum Beispiel der Fall, wenn ein Konnektor eingesetzt wird, um Mehrwertdienste zu betreiben. Diesem können beispielsweise Daten vom Dateneigner übertragen werden. Um die Weitergabe zu verhindern, werden Attribute an diese Daten angehängt. Durch diese Attribute wird der Konnektor so konfiguriert, dass eine Nutzung der Daten dem angegebenen Profil entspricht.

#### 3.1.4.6 Sichere Ausführungsumgebung für Konnektoren

Grundsätzlich ist es technisch möglich, InDaSpace-Endpunkte (d.h. Konnektoren) auf unsicheren Plattformen auszuführen. Dies hätte jedoch einige problematische Implikationen, die für die InDaSpace-Teilnehmer ggf. nicht wünschenswert sind:

- Die vollständige und korrekte Aufzeichnung der Datennutzung und damit das Accounting können im Falle eines angreifbaren Konnektors nicht garantiert werden.
- Falls das verwendete Schlüsselmaterial kompromittiert wird, ist es möglich, dass die Identität des angegriffenen InDaSpace-Konnektors gefälscht und Dienste im Namen und auf Kosten des Teilnehmers genutzt werden.
- Eine ggf. im Konnektor stattfindende Vorverarbeitung der Daten wäre nicht vertrauenswürdig, ihre Ergebnisse damit nicht verlässlich
- Ein durch den Konnektor angeblich umgesetztes Sicherheitsniveau wäre nicht verlässlich

Durch die Etablierung von sicheren Ausführungsumgebungen für Konnektoren lassen sich die oben genannten Probleme verhindern. Eine solche Ausführungsumgebung kann in verschiedenen Sicherheitsausbaustufen existieren und ist in der Lage, z.B. Java-Code oder Anwendungscontainer, die gegen definierte Schnittstellen implementiert wurden, auszuführen. D.h. die ausgeführten Komponenten sind voneinander isoliert und haben keinen Zugriff auf das darunterliegende System, da ihnen nur die definierten und kontrollierten Schnittstellen zur Verfügung stehen.

#### Remote Attestation

Durch Remote Attestation kann ein InDaSpace-Konnektor einem Kommunikationspartner nachweisen, dass er sich in einem vertrauenswürdigen Zustand befindet. Dadurch kann sich bspw. ein Daten-Provider vor der Preisgabe von Daten an den Consumer davon überzeugen, dass der Consumer tatsächlich in Form eines zertifizierten und nicht-kompromittierten InDaSpace-Konnektors realisiert ist und die Sicherheitsanforderungen des Providers einzuhalten vermag. Usage Control Sind InDaSpace-Konnektoren in einer vertrauenswürdigen Ausführungsumgebung realisiert, deren Integrität durch Remote Attestation nachgewiesen werden kann, so können zusätzlich zu den bisherigen Maßnahmen zur Zugriffskontrolle auch Consumer-seitige Obligations durchgesetzt werden. Dies bedeutet, dass Daten-Provider die Nutzung ihrer Daten an Bedingungen knüpfen können, die durch die Ausführungsplattform auf Seiten des Consumers durchgesetzt werden. Ein einfaches Beispiel ist die zeitlich limitierte Nutzung der Daten. D.h., die Daten werden durch den Provider bereitgestellt und an die Bedingung geknüpft, nach Ablauf eines Zeitraumes gelöscht werden zu müssen. Da die Datenlöschung auf Seiten des Consumers erfolgen muss, muss sichergestellt werden, dass sich auf Seiten des Consumers eine vertrauenswürdige Komponente befindet, die zum einen verhindert, dass die Daten im Nutzungszeitraum die geschützte Umgebung ver-

lassen und außerdem mit Ablauf der Frist gelöscht werden. Darüber hinaus muss sie die Löschung attestieren und dem Provider nachweisen. Erst durch die Existenz einer sicheren Ausführungsumgebung und der Durchsetzung von Obligations auf Consumer-Seite können daher Szenarien realisiert werden, in denen Daten verwendet werden, die Löschfristen oder anderweitigen Einschränkungen bzgl. ihrer Weitergabe unterliegen, wie es z.B. für personenbezogene Daten der Fall ist. App-Store für InDaSpace-Konnektoren. Ein weiterer Use-Case, der durch die Ausführung von InDaSpace-Konnektoren in einer sicheren Umgebung ermöglicht wird, ist die Bereitstellung von Apps, also Anwendungen zur Datenverarbeitung direkt im InDaSpace-Konnektor. Da der InDaSpace Rohdaten in großer Menge und in unterschiedlichen Formaten zur Verfügung stellen wird, ist es sinnvoll, die Vorverarbeitung der Daten nahe an die Datenquelle zu legen. Dies ist insbesondere für echtzeitkritische Anwendungen erforderlich, bringt aber auch den Sicherheitsgewinn, dass keine Rohdaten preisgegeben werden müssen, sondern lediglich deren Aggregate. Ein Beispiel für eine solche App ist Complex Event Processing (CEP). Hierbei registrieren Daten-Consumer Abfragen auf Datenströmen und werden per Callback benachrichtigt, sobald die Abfrage erfüllt ist.

### 3.2 Implementierung der Kern-Softwarekomponenten

Das folgende Kapitel beschreibt die Arbeitsschritte zur Konzeption und Implementierung der Softwaremodule des InDaSpace.

#### 3.2.1 Überblick über Komponenten und ihre Aufgaben

Abbildung 16 zeigt die vorgesehenen Komponenten, welche im Kontext des Industrial Data Space realisiert werden sollen. Diese befähigen die Akteure ohne großen Aufwand und unter Einhaltung der Werteverprechen am InDaSpace teilzunehmen. Dabei setzt der InDaSpace Standards in den Bereichen Unternehmensübergreifende Datenintegration und Sicherheit. Die Liste der Komponenten kann sich auf Basis der Ergebnisse aus dem Arbeitspaket 1 InDaSpace Architektur verändern.

Die Auswahl der Komponenten erfüllt die u.a. folgenden Bedingungen:

- Der InDaSpace stellt Softwarekomponenten für jede notwendige Rolle im Daten-Ökosystem zur Verfügung
- Die InDaSpace-Software beseitigt Hürden bei der Datenintegration durch die Bereitstellung von Standardvokabularen und Werkzeugen
- Die InDaSpace Software ist modular aufgebaut und ermöglicht den Anschluss vielfältiger Backend-Systeme (wie ERP-Systeme, ESBs, Cloud-Lösungen etc.)

Softwaretechnische Herausforderungen: Als agiler Softwareentwicklungsprozess soll in diesem Projekt Scrum verwendet werden (für weitere Informationen dazu vgl. Abs. 2.2).

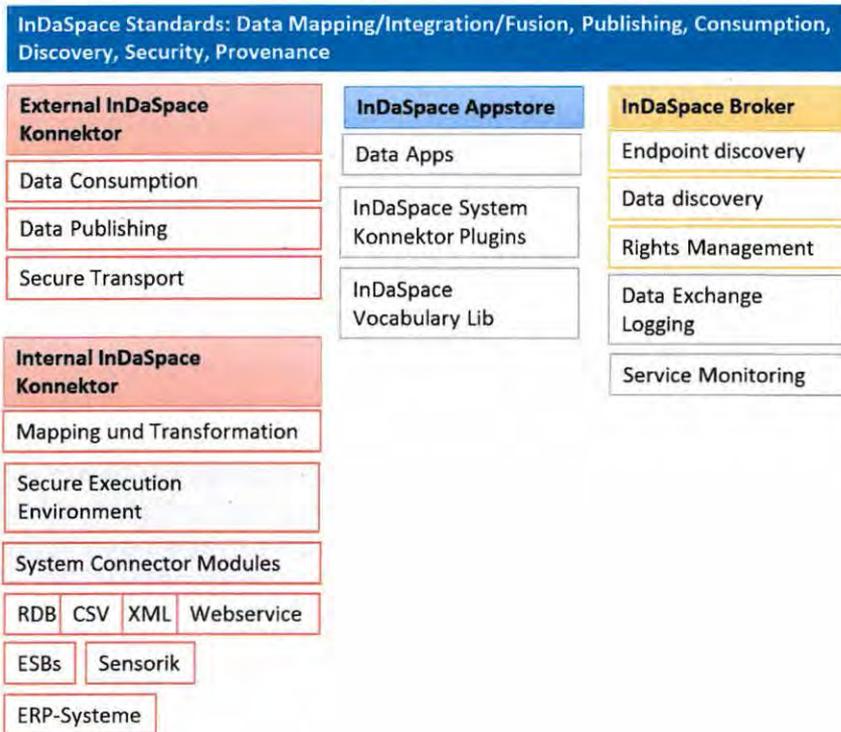


Abbildung 16 Übersicht - Komponenten und funktionale Blöcke

### Arbeitspaketbeschreibung

AP-Nr.: 2	Implementierung der Kern-Softwarekomponenten
Startdatum:	M1
Enddatum:	M36
<p>1. Ziel des Arbeitspaketes</p> <ul style="list-style-type: none"> <li>- Kern-Softwarekomponenten werden konzipiert</li> <li>- Kern-Softwarekomponenten sind für ausgewählte Use Cases im Pilotbetrieb verfügbar</li> <li>- Kern-Softwarekomponenten sind als Referenzimplementierung des InDaSpace verfügbar</li> </ul>	
<p>2. Voraussetzung (Input)</p> <ul style="list-style-type: none"> <li>- Auswahl bzw. Priorisierung der Use Cases</li> <li>- Übergeordnete Systemarchitektur</li> <li>- (Zwischen)Ergebnisse aus AP Architektur (Auswahl Technologien, Software, Protokolle)</li> <li>- (Zwischen)Ergebnisse aus AP Standardisierung</li> </ul>	
<p>3. Lösungsweg</p> <ul style="list-style-type: none"> <li>▪ Arbeitsschritt 1: <i>Durchführung von Voruntersuchungen</i> <ul style="list-style-type: none"> <li>- Basis-Anforderungsanalyse gemäß ausgewählter Use Cases</li> <li>- Analyse der Daten aus ausgewählten Use Cases</li> <li>- Analyse der Systeme aus ausgewählten Use Cases</li> </ul> </li> <li>▪ Arbeitsschritt 2: <i>Implementierung der InDaSpace Konnektoren (INIK, EXIK) (Details siehe AP 2.2)</i> <ul style="list-style-type: none"> <li>- Bereitstellung einer sicheren Ausführungsumgebung für Application Container</li> </ul> </li> </ul>	

- Implementierung des Internal InDaSpace Konnektors (INIK) (Details siehe AP 2.2.1)
  - Datenanalyse und Vorverarbeitung
  - Erarbeitung des InDaSpace internen Datenschemas
  - Basisdienste/-applikationen zum Mapping und zur Transformation von Daten
  - Adapter zu unternehmensinternen Systemen (Nutzer-spezifische Systemkonnektoren)
  - Werkzeug/Dienst zur Batchverarbeitung von Transformationen
  - Erstellung von Dokumentationen des INIK
- o Implementierung des External InDaSpace Konnektors (EXIK) (Details siehe AP 2.2.2)
  - *Definition und Implementierung des Datenmodells*
  - Implementierung der Ingest Mechanismen
  - Module zur Zugriffskontrolle und -steuerung (Data Consumption) (Select, Harvest)
  - Realisierung der Broker Anbindung
  - Realisierung von Sicherheitsmechanismen für Authentifizierung, Autorisierung sowie Transport und Nachrichtensicherheit
  - Erstellung von Dokumentationen des EXIK
  - Detailspezifikation und Implementierung einer Sprache für Anfragen und Mappings zwischen EXIKs, sowie für verteilbare Ausführungspläne
  - Realisierung einer Data-Transformation-Engine zum Mapping zwischen heterogenen Datenmodellen
  - Erstellung von Dokumentationen des EXIK
- Entwicklung eines InDaSpace Interfaces zum Datenaustausch
- Methodik: Anforderungsanalyse, agile Softwareentwicklung, Testing (fortlaufend)

▪ *Arbeitsschritt 3:*

*Implementierung des InDaSpace Broker (Details siehe AP 2.3)*

- Registrierungs- und Verwaltungsmodul implementieren (Datenanbieter, Datenangebote)
- Ingest Modul und Aufbau des Persistence Layers (Klassifizierung und Indizierung von Daten)
- Realisierung eines Metadaten-Repositories zur Ablage von Datenmodellen und Mappings
- Website zum Publizieren aufbauen (inkl. Access, Search/Data Discovery)
- Unterstützung des Matchmakings/Vertragsabschlusses als Modul integrieren
- API entwickeln und publizieren (Inhalte, Formate, Datenanbieter, Abrufkonditionen)
- Komponenten zum Protokollieren der Nutzungsinformationen (Datenabrufe, Abrechnung) implementieren (Data Exchange Logging)
- Komponente zum Service Monitoring implementieren
- Erstellung von Dokumentationen zum Broker
- Methodik: Anforderungsanalyse, agile Softwareentwicklung, Testing (fortlaufend)

<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 4:</i> <i>Aufbau des InDaSpace App Store (Details siehe AP 2.4)</i> <ul style="list-style-type: none"> <li>- Technische Testbed Infrastruktur für InDaSpace App Store aufbauen</li> <li>- Technische Bedingungen an Apps definieren (Einhaltung von Schnittstellenspezifikationen, Sicherheitsanforderungen)</li> <li>- Definition einer Sprache bzw. Modells zur Beschreibung der Daten-Operationen, die durch eine App durchgeführt werden</li> <li>- Prozesse zum Hochladen von Apps festlegen (Sicherstellen der Hersteller-Identität, Code-Integrität, automatische Prüfung der Apps)</li> <li>- Erstellung von Dokumentationen zum App Store</li> </ul> </li> </ul>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 5:</i> <i>Qualitätssicherung (Integrierte Systemtests)</i> <ul style="list-style-type: none"> <li>- Identifizierung und Planung der übergeordneten Systemtests</li> <li>- Implementierung und Durchführung der übergeordneten Systemtests</li> <li>- Beseitigung von Fehlern, auch nach Rückmeldung aus den Use Cases</li> <li>- Implementierung von Testfunktionen. Realisierung der funktionalen Teile als a) interne (persistente) Testkomponente und b) als "Bibliothek" für Anwender Apps die Nutzer zur Selbstprüfung in ihrer InDaSpace Instanz anwenden können (Zielsetzung: Möglichkeit der Entwicklung von Test-Apps als Klient der Test und Monitoring Interfaces der Komponenten).</li> <li>- Validierung der Testfunktionen anhand von Testcases (z.B. unter Nutzung der Testsuites die als in Arbeitspaket 5 bereitgestellt werden als Referenz).</li> <li>- Implementierung von Schnittstellen zum Infrastruktur Monitoring, insbesondere zur Anbindung von Monitoring Probes außerhalb der Software-Kernkomponenten (z.B. Einbindung von Nagios o.ä. etablierten Systemen)</li> </ul> </li> </ul>	
Personalaufwand gesamt (in PM)	
4. Ergebnis (Output) <ul style="list-style-type: none"> <li>- Softwarekomponente INIK liegt vor.</li> <li>- System Connector Plugins zu 3 ausgewählten Systemen liegen vor.</li> <li>- Softwarekomponente EXIK liegt vor.</li> <li>- Softwarekomponente Broker liegt vor.</li> <li>- Softwarekomponente App Store liegt vor</li> </ul>	
5. Sonstige Ressourcen <ul style="list-style-type: none"> <li>- Kosten Entwicklungsumgebung</li> <li>- Kosten Pilotumgebung</li> <li>- Reisekosten für Projekttreffen</li> <li>- Softwarelizenzkosten?</li> </ul>	

### 3.2.2 Implementierung der InDaSpace Konnektoren (INIK, EXIK)

Der InDaSpace unterstützt seine Akteure mit Softwarekomponenten, welche die Teilnahme am InDaSpace ermöglichen und die Wertversprechen des InDaSpace durchsetzen.

Erfahrungsgemäß sollte die Einstiegshürde für die Teilnahme am InDaSpace (Aufbau von InDaSpace Infrastruktur) niedrig gehalten und auf die Realität in den Un-

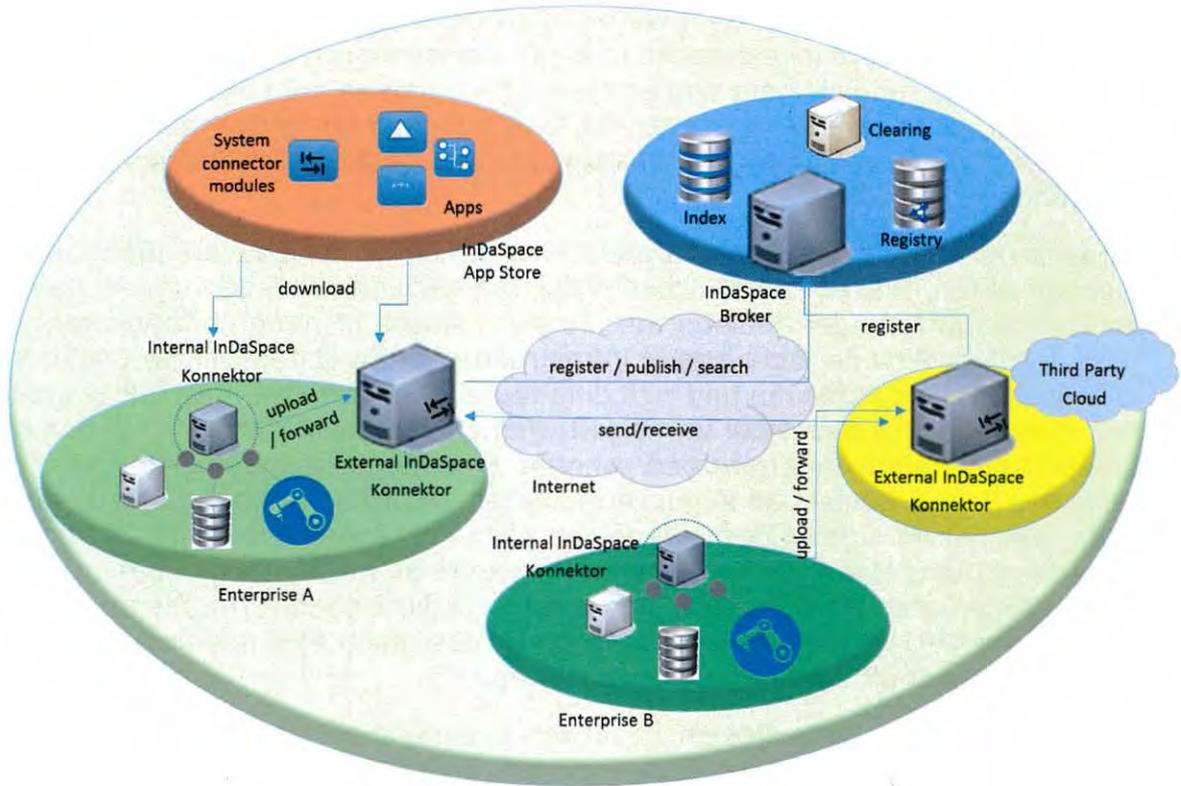
ternehmen abgestimmt werden. Während große Unternehmen mit IT-Abteilungen ausgestattet sind und Investitionen in die IT-Infrastruktur einplanen, stellt die gleiche Anforderung kleinere und mittlere Unternehmen mit Fokus auf industrielle Produktion vor schwierige Aufgaben. Es sind jedoch unter Umständen gerade diese Unternehmen, welche für die Wertschöpfung essentielle Datensätze beisteuern können.

Eines der Grundprinzipien des InDaSpace ist der Erhalt der Souveränität über die eigenen Daten. Dieses Prinzip schließt die Übertragung von Eigentumsrechten an zentrale Instanzen oder Anbieter aus. Es muss jedoch nicht immer bedeuten, dass der Data Provider eine eigene Infrastruktur zur Zwischenspeicherung von InDaSpace Daten aufbauen und sich eine teure IT-Sicherheitsinfrastruktur aneignen muss und im Zweifelsfall Vertragsstrafen erleidet, weil die SLA`s für den Datenabruf nicht eingehalten werden konnten. Eine mögliche Lösung ist, dass die Rolle des Data Providers an einen zertifizierten Dienstleister übergeben werden kann, welcher im Auftrag des Dateneigners am InDaSpace teilnimmt. Da verschiedene Dienstleister konkurrierende Angebote an KMUs machen um für diese als Auftrags-Daten-Provider tätig zu werden und die Eigentumsrechte an den Daten bei den KMU verbleiben ist sichergestellt, dass auch KMU in einem solchen Szenario die Hoheit über Ihre Daten behalten.

In der Architektur wurde diesem Gedanken Rechnung getragen, indem der softwaretechnische Endpunkt des Data Providers (InDaSpace Konnektor) in zwei Komponenten aufgeteilt wurde. Eine Komponente dient der Selektion und Transformation von Unternehmensdaten und ist in die IT des Dateneigners eingebettet (Internal InDaSpace Konnektor). Die zweite Komponente (External InDaSpace Konnektor) veröffentlicht die Daten im InDaSpace und läuft entweder beim Dateneigner oder bei einem zertifizierten und autorisierten Anbieter. Dieser Anbieter übernimmt dann vertraglich festgelegt die juristischen und operativen Verpflichtungen des Dateneigners in der Rolle des Data Provider im InDaSpace, unterwirft sich also ebenfalls den gemeinsamen Regeln im InDaSpace. Diese Verpflichtungen werden beispielsweise Anforderungen an die Qualität der Angebotenen Daten und des Quality of Service beinhalten

Abbildung 17 illustriert das Zusammenspiel von Unternehmen und zentralen Komponenten des InDaSpace. Enterprise A veranschaulicht die Situation in einem Unternehmen, welches die InDaSpace Infrastruktur selbst betreibt. Enterprise B veranschaulicht die Situation in einem Unternehmen, welches die InDaSpace-Teilnahme an einen Dienstleister vergeben hat.

Datenintensive Auswertungs- und Analyse-Operationen sollten aus Performance-Gründen und zur Reduktion der Kommunikationsaufwände möglichst nah bei den eigentlichen Datenquellen stattfinden. Aufgrund von Ressourcenbeschränkungen oder Sicherheitsanforderungen kann es dennoch notwendig sein, Teile der Daten-Operationen auf anderen EXIK-Instanzen auszuführen (z.B. in der Third-Party-Cloud). Die Daten-Transformations- und -Verarbeitungs-Logik des InDaSpace muss also eine flexible Verteilung der Datenoperationen erlauben. Dazu ist es notwendig, die deklarativ-spezifizierten Operationen (z.B. Anfragen, Transformationen, Analysen) in verteilbare Ausführungspläne zu übersetzen.



Industrial Data Space

Abbildung 17 InDaSpace Szenario

### 3.2.3 Aufbau des InDaSpace App-Stores

Der InDaSpace-App-Store (siehe auch Abbildung 18) bietet die Möglichkeit, Software-Module in die Konnektoren herunterzuladen und in der vertrauenswürdigen Ausführungsumgebung auszuführen.

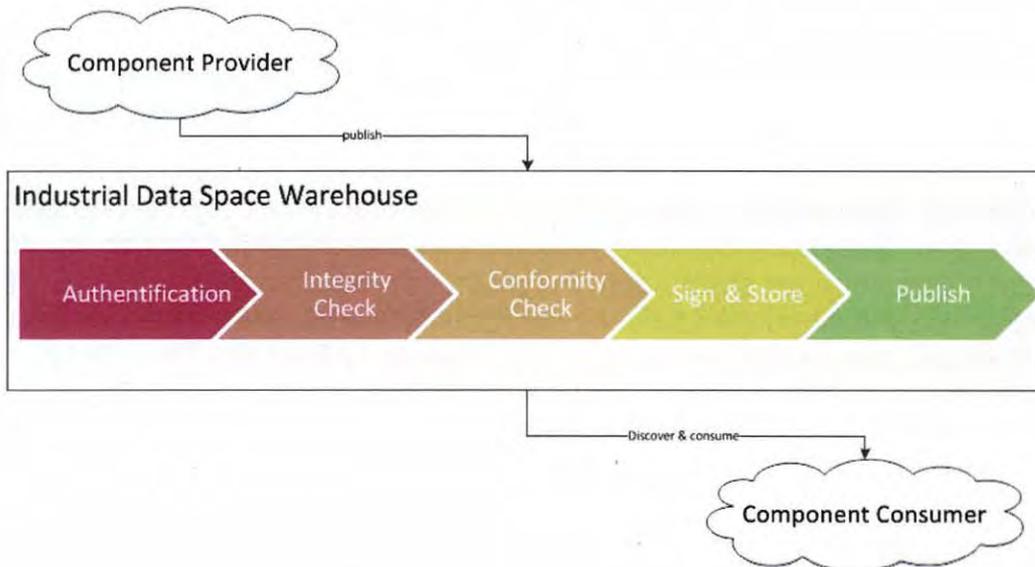


Abbildung 18 InDaSpace App Store Prozess

Diese Apps können von beliebigen InDaSpace-Teilnehmern zur Verfügung gestellt und bei Abrechnung der Nutzung einen Teil der Wertschöpfung darstellen. Damit dies reibungslos erfolgen kann, müssen Anforderungen an derartige Module gestellt werden. Dies können zum Beispiel definierte Schnittstellen sein, die zur Verfügung stehen. Aber auch Anforderungen an die Implementierung können definiert werden. Hierfür müssen Prozesse definiert werden, die jede App durchlaufen muss:

- Überprüfung der Identität des Herstellers
- Validierung der Code-Integrität
- Automatische Prüfung der Apps

Diese Prüfprozesse müssen nach der Definition auch technisch implementiert werden.

#### Arbeitspaketbeschreibung

AP-Nr.: 2.2	Implementierung der InDaSpace Konnektoren (INIK, EXIK)
Startdatum:	M4
Enddatum:	M36
<p>1. Ziel des Arbeitspaketes</p> <ul style="list-style-type: none"> <li>- Gemeinsame Grundarchitektur für eine sichere Ausführungsumgebung für INIK und EXIK schaffen</li> <li>- Softwarekomponenten entwickeln, die sowohl seitens des Internal InDaSpace Konnektors als auch seitens des External InDaSpace Konnektors nachgenutzt werden können.</li> <li>- InDaSpace Konnektoren sind als Softwarekomponente(n) verfügbar</li> </ul>	
<p>2. Voraussetzung (Input)</p> <ul style="list-style-type: none"> <li>- Ergebnisse aus Voruntersuchung liegen vor (Basis-Anforderungen)</li> </ul>	
<p>3. Lösungsweg</p> <ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 1:</i> <i>Bereitstellung einer sicheren Ausführungsumgebung für Application Container</i> <ul style="list-style-type: none"> <li>- Konzeption und Aufbau der Ausführungsumgebung</li> <li>- Anpassungen zur sicheren Ausführung von Datenapplikationen (Data Apps) durchführen</li> <li>- Anpassungen zur sicheren Ausführung von Complex Event Processing (CEP) durchführen</li> <li>- Methodik: Anforderungsanalyse, agile Softwareentwicklung, Testing (fortlaufend)</li> </ul> </li> <li>▪ <i>Arbeitsschritt 2:</i> <i>Implementierung des Internal InDaSpace Konnektors (INIK) (Details siehe AP 2.2.1)</i> <ul style="list-style-type: none"> <li>- Methodik: Anforderungsanalyse, agile Softwareentwicklung, Testing (fortlaufend)</li> </ul> </li> <li>▪ <i>Arbeitsschritt 3:</i> <i>Implementierung des External InDaSpace Konnektors (EXIK) (Details siehe AP 2.2.2)</i> <ul style="list-style-type: none"> <li>- Methodik: Anforderungsanalyse, agile Softwareentwicklung, Testing (fortlaufend)</li> </ul> </li> <li>▪ <i>Arbeitsschritt 4:</i> <i>Entwicklung eines InDaSpace Interfaces zum Datenaustausch</i> <ul style="list-style-type: none"> <li>- Konzeption und Entwicklung einer API zum Datenaustausch zwischen internem und externem InDaSpace Konnektor</li> </ul> </li> </ul>	

<ul style="list-style-type: none"> <li>- Definition eines Austauschformats zur Übertragung von Daten zwischen INIK (send) und EXIK (receive)</li> <li>- Definition eines Austauschformats zur Übertragung und zum Empfangen von Daten zwischen INIK und diversen Enterprise Komponenten (send/receive)</li> <li>- Definition eines Austauschformats zur Übertragung und zum Empfangen von Daten, Daten Apps zwischen INIK und InDaSpace App Store (send/receive)</li> <li>- Definition eines Austauschformats zum Datenaustausch zwischen EXIK und InDaSpace Broker (send/receive)</li> <li>- Definition eines Austauschformats zur Übertragung von Daten zwischen verschiedenen EXIKs (send/receive)</li> <li>- Auswahl und Anpassung bzw. Entwicklung von sicheren Transport Mechanismen</li> <li>- Erstellung einer API Dokumentation</li> <li>- Methodik: Anforderungsanalyse, agile Softwareentwicklung, Testing (fortlaufend)</li> <li>- Bereitstellung und prototypische Integration von Softwarekomponenten die den Zugriff auf interne Test- und Monitoring-Funktionen implementieren (Interface-Implementierung).</li> <li>- Implementierung von Testfunktionen für diese Komponenten (Implementierung durch die Komponente, Nutzung durch den Klienten der Interfaces) zur Unterstützung von nicht-intrusivem Performance Monitoring und zur Überprüfung der Implementierung von Sicherheitsrichtlinien beim Zugriff auf Interfaces und Daten</li> <li>- Definition und Implementierung von Datenmappings für die InDaSpace Konnektoren: Auswahl der Mapping-Technologie, Identifizierung und Implementierung relevanter fremder Datenmodelle, Festlegung und Implementierung von semantischen Datenmappings zwischen fremden Datenmodellen und dem InDaSpace-Datenmodell.</li> </ul>	
<p>Personalaufwand gesamt (in PM)</p>	
<p>4. Ergebnis (Output)</p> <ul style="list-style-type: none"> <li>- Softwarekomponente INIK liegt vor.</li> <li>- System Connector Plugins zu 3 ausgewählten Systemen liegen vor.</li> <li>- Softwarekomponente EXIK liegt vor.</li> </ul>	
<p>5. Sonstige Ressourcen</p> <ul style="list-style-type: none"> <li>▪ keine</li> </ul>	

### 3.2.4 Internal InDaSpace Konnektor

Der Internal InDaSpace Konnektor (INIK) ist eine von zwei notwendigen Komponenten (Verortung durch gelbe Markierung in Abbildung 19), um Data Provider und Data Consumer an den InDaSpace anzuschließen.

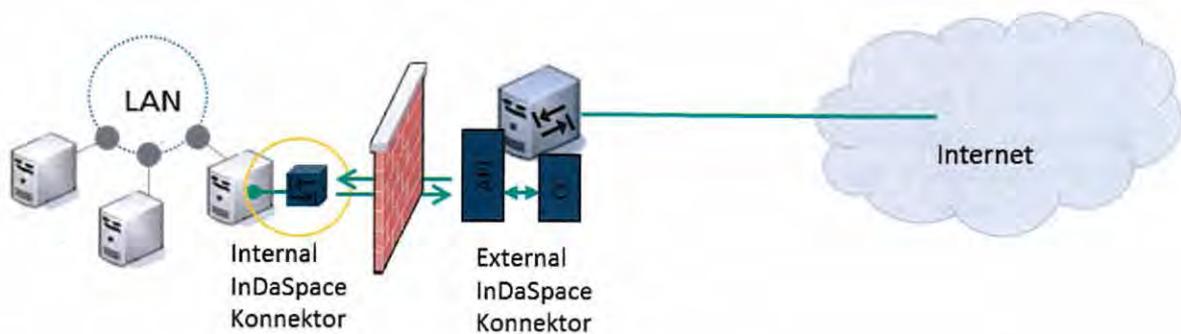


Abbildung 19 Verortung des Internal InDaSpace Konnektors

Der INIK realisiert in diesem Kontext folgende Funktionen:

Komponentensteckbrief

Name	Internal InDaSpace Konnektor
Funktionalitäten	<p>Bereitstellung einer sicheren Ausführungsumgebung für Application Container. Die Container beinhalten Module wie Datenverarbeitungs-, Datenveredlungs- und Datenanreicherungs-Applikationen (Data Apps) bzw. Complex Event Processing (CEP). Zugriff auf die „Data Apps“ (Module) erfolgt über den InDaSpace App Store.</p> <p>Anbindung vorgefertigter oder nutzerspezifischer Adapter zu unternehmens-internen Systemen (z. B. SAP, ESBs, Sensorik, etc.) als Module der Application Container.</p> <p>Bereitstellung von Basisdiensten/-applikationen zum Mapping und zur Transformation von Daten auf ein InDaSpace-empfohlenes Format. Lauffähig in Application Container.</p> <p>Zugriff auf empfohlenen Vokabulare und Schemata aus dem InDaSpace App Store.</p> <p>Übertragung der Daten in Richtung External InDaSpace Konnektor</p>
Verortung	Betrieb durch Data Provider, Data Consumer innerhalb des Enterprise LAN.
Art der Realisierung	Eigenständige Server-Komponente
Schnittstellen	<p>External InDaSpace Connector (send)</p> <p>InDaSpace App Store (send/receive)</p> <p>Diverse Enterprise Komponenten (send/receive)</p>
Besonderheiten	Notwendige Softwarekomponente zur Teilnahme an InDaSpace für Data Provider / Data Consumer

Die Abbildung 20 verdeutlicht den Aufbau des INIK als Komponente mit zwei wesentlichen Schichten. Der Communication Layer bietet ein Interface für System-Konnektoren und eine Schnittstelle zum External InDaSpace Konnektor, über welche Daten hochgeladen oder sichere Verbindungen (Tunnel) für Streams und Bus-Protokoll aufgebaut werden können.

Der Application Container Layer bietet zum einen eine Ausführungsumgebung für Basis-Applikationen wie Daten-Mapping und -Transformation, welche durch den Nutzer konfiguriert werden, um die Daten auf eine InDaSpace empfohlene Format abzubilden. Die Konfiguration des Mapping und der Transformation wird durch empfohlene Daten-Schemata und Vokabularen aus dem InDaSpace App Store unterstützt. Darüber hinaus ermöglicht der Application Container Layer die

Verankerung von zentral verfügbaren (z.B. InDaSpace App Store) oder nutzerspezifischen Systemkonnektoren. Diese Konnektoren realisieren die Anbindung von Daten aus IT-Systemen, Datenbanken (Office Floor) oder Sensorik und Bus-Systemen (Shop Floor). Zudem können komplexere Data Apps (auch von Drittanbietern) aus dem App Store geladen und in der Ausführungsumgebung (Container) verwendet werden.

In dieser Komponente soll keine dauerhafte, maximal eine temporäre Speicherung der Daten stattfinden.

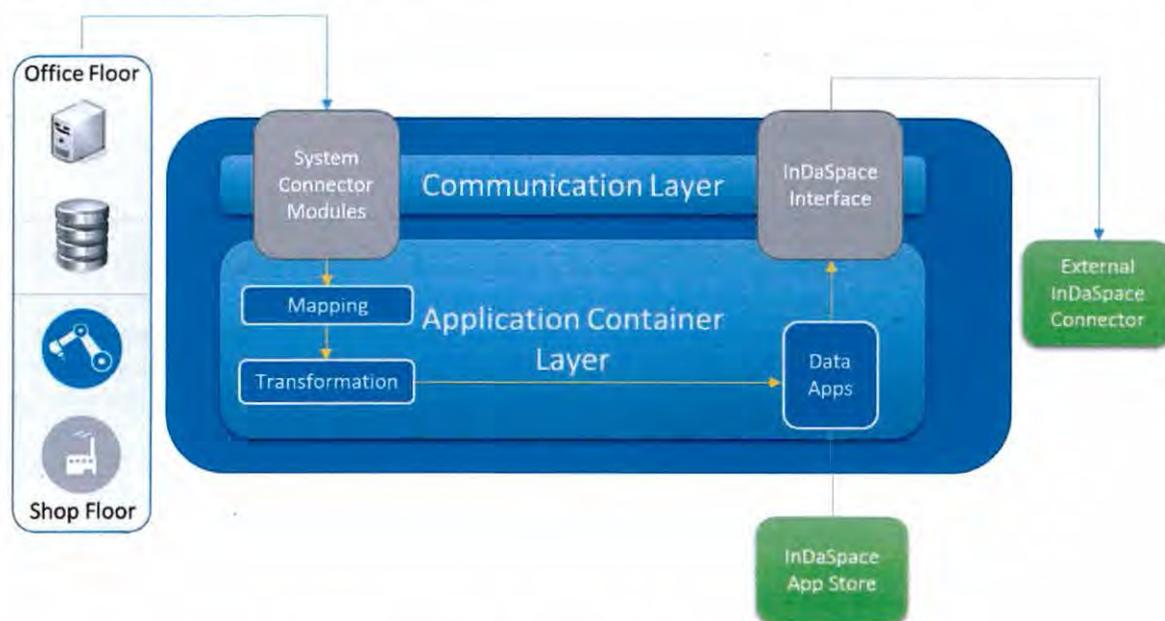


Abbildung 20: Internal InDaSpace Konnektor

Die Komponente INIK wird innerhalb des Unternehmensnetzes installiert und betrieben. Dies hat mehrere Gründe. Zu einen ist realisiert die Komponente über eine die Einbindung von System-Konnektoren eine direkte Verbindung zur internen Datenquellen (Datenbanken, Live-Systemen). Zum anderen unterstützt die Komponente wichtige Basis-Vorverarbeitungsschritte wie Mapping und Transformation sowie optionale Verarbeitungsschritte wie beispielsweise Anonymisierung oder Aggregation, welche durchgeführt werden sollten, bevor die Daten das Unternehmen verlassen. Schließlich hat eine Verarbeitung der Daten nahe an der Datenquelle wohlbekannte Vorteile.

#### Arbeitspaketbeschreibung

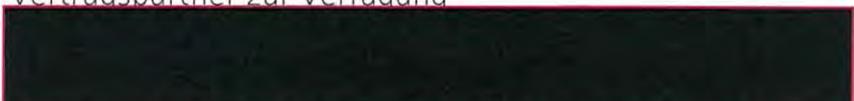
AP-Nr.: 2.2.1	Implementierung des Internal InDaSpace Konnektors (INIK)	
Startdatum:	M4	
Enddatum:	M36	
<ol style="list-style-type: none"> <li>1. Ziel des Arbeitspaketes <ul style="list-style-type: none"> <li>- Alle Basiskomponenten des INIK sollen konzipiert und entwickelt werden.</li> <li>- Das interne Datenschema soll festgelegt werden.</li> <li>- Beispielhafte Unternehmensdaten sollen für die ausgewählten Use Cases über den INIK ausgetauscht werden können.</li> </ul> </li> </ol>		
<ol style="list-style-type: none"> <li>2. Voraussetzung (Input) <ul style="list-style-type: none"> <li>- Auswahl von unternehmens-internen Systemen gemäß Use Cases</li> </ul> </li> </ol>		

<ul style="list-style-type: none"> <li>- Konzeptionelle Mappings werden seitens der Unternehmen erstellt bzw. unterstützt.</li> <li>- Testdaten müssen vorliegen</li> <li>- (Zwischen)Ergebnisse aus AP Standardisierung [REDACTED]</li> </ul>	
<p>3. Lösungsweg</p>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 1:</i> <i>Datenanalyse und Vorverarbeitung</i> <ul style="list-style-type: none"> <li>- Festlegung der Austauschformate (XML, CSV, Webservice etc.)</li> <li>- Festlegung der verschiedenen Metadatenformate/-strukturen der unternehmens-internen Daten</li> </ul> </li> </ul>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 2:</i> <i>Erarbeitung des InDaSpace internen Datenschemas</i> <ul style="list-style-type: none"> <li>- Festlegung der erforderlichen Metadatenstruktur des internen Repräsentationsformats</li> <li>- Evaluierung vorhandener Datenschemata</li> <li>- Definition des InDaSpace internen Datenschemas</li> </ul> </li> </ul>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 3:</i> <i>Basisdienste/-applikationen zum Mapping und zur Transformation von Daten</i> <ul style="list-style-type: none"> <li>- Beratung in der Erstellung von konzeptionellen Mappings für Unternehmensdaten</li> <li>- Implementierung von technischen Mappings zur Transformation von Daten (Aufbau einer Transformation Library)</li> <li>- Qualitätssicherung für Beispieltransformationen</li> <li>- Massenverarbeitung mit speziellem Werkzeug (s. Arbeitsschritt 5)</li> <li>- Methodik: Anforderungsanalyse, agile Softwareentwicklung, Testing (fortlaufend)</li> </ul> </li> </ul>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 4:</i> <i>Adapter zu unternehmens-internen Systemen (Nutzerspezifische Systemkonnektoren)</i> <ul style="list-style-type: none"> <li>- Entwicklung von System Connector Plugins zu System 1 (Relationale Datenbank)</li> <li>- Entwicklung von System Connector Plugins zu System 2 (ERP-System)</li> <li>- Entwicklung von System Connector Plugins zu System 3 (Sensorik)</li> <li>- Festlegung der Austauschformate (XML, OPC UA, EDI, ebXML etc.)</li> <li>- Methodik: Anforderungsanalyse, agile Softwareentwicklung, Testing (fortlaufend)</li> </ul> </li> </ul>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 5:</i> <i>Werkzeug/Dienst zur Batchverarbeitung von Transformationen</i> <ul style="list-style-type: none"> <li>- Entwicklung eines Softwareprogramms zur Massenverarbeitung von Transformationen</li> </ul> </li> </ul>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 6:</i> <i>Erstellung von Dokumentationen des INIK</i> <ul style="list-style-type: none"> <li>- Beschreibung des InDaSpace internen Datenschemas</li> <li>- Anleitung für konzeptionelle Mappings</li> <li>- Anleitung zur Nutzung der System Connector Plugins</li> <li>- Anleitung zur Nutzung des Werkzeugs zur Massenverarbeitung von Transformationen</li> <li>- Anleitung zur Integration weiterer Daten Apps</li> </ul> </li> </ul>	

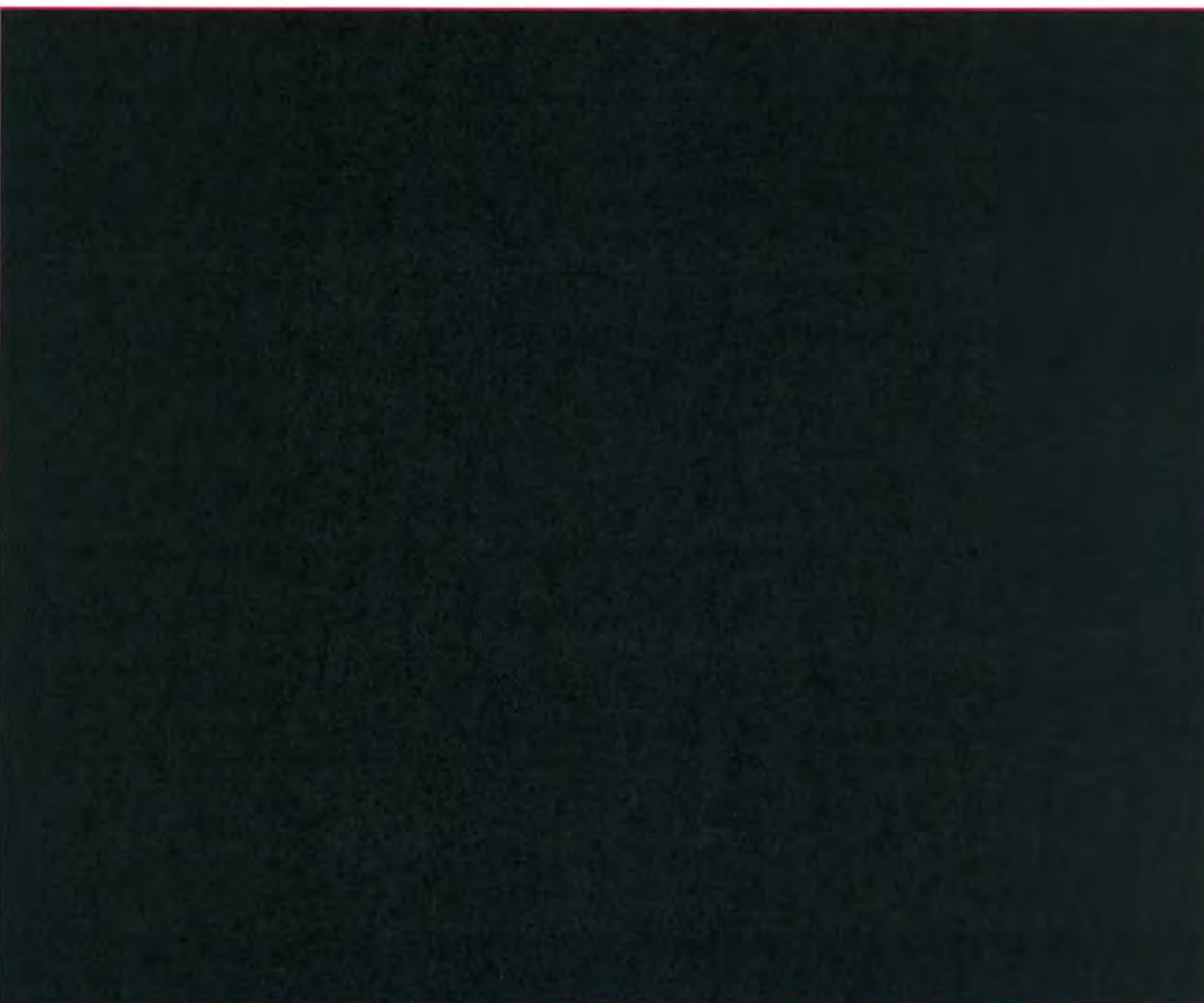
4. Ergebnis (Output)	
<ul style="list-style-type: none"> <li>- Softwarekomponente INIK inklusive Dokumentationen liegt vor: <ul style="list-style-type: none"> <li>- Internes Datenschema liegt vor.</li> <li>- Vorgehen für konzeptionelle Mappings liegt vor:</li> <li>- Transformation Library für x Transformationen liegt vor.</li> <li>- System Connector Plugins zu 3 ausgewählten Systemen liegen vor.</li> <li>- Werkzeug zur Batchverarbeitung von Transformationen liegt vor.</li> <li>- Möglichkeit der Erweiterung um weitere Data Apps</li> </ul> </li> </ul>	
5. Sonstige Ressourcen	
<ul style="list-style-type: none"> <li>- Software Lizenzen (XML Spy)</li> </ul>	

### 3.2.5 External InDaSpace Konnektor



Name	External InDaSpace Konnektor
Funktionalitäten	<p>Publiziert Datenangebot (Datenbeschreibung, Konditionen)  Registriert das Datenangebot bei InDaSpace Broker  Stellt (evtl. gespeicherte) Daten zum sicheren Abruf durch Vertragspartner zur Verfügung</p>  <p>Bereitstellung einer sicheren Ausführungsumgebung für Application Container. Die Container beinhalten Module wie Datenverarbeitungs-, Datenveredlungs- und Datenanreicherungs-Applikationen (Data Apps) bzw. Complex Event Processing (CEP). Zugriff auf die „Data Apps“ (Module) erfolgt über den InDaSpace App Store.</p>
Verortung	Betrieb durch Data Provider, Data Consumer (z.B. in der DMZ) oder als Service durch Third Parties (in sicherer Cloud-Umgebung).
Art der Realisierung	Eigenständige Server-Komponente
Schnittstellen	Internal InDaSpace Connector (receive) InDaSpace Broker (send/receive)

	External InDaSpace Connector (send/receive)
Besonderheiten	Notwendige Softwarekomponente zur Teilnahme an InDaSpace für Data Provider / Data Consumer



Im Application Container Layer stehen Container für das sichere Ausführen von Datenapplikationen zur Verfügung. Standard-Module regeln den Datenzugriff. Weiterhin sind Modulen zur Zugriffskontrolle und -steuerung dort angesiedelt.

Zudem ist es analog zum INIK möglich, komplexere Data Apps (auch von Drittanbietern) aus dem App Store zu geladen und in der Ausführungsumgebung (Container) zu verwenden.

Vom Data Provider veröffentlichte Daten, können über einen Ingest-Prozess im Persistence Layer abgelegt werden. Dies gilt sowohl für die Metadaten (Datenangebote) als auch wenig dynamische Echtdateien.

Die Entscheidung für eine mögliche Speicherung der Daten wurde getroffen, um Herausforderungen bei der Anbindung von nicht weiter spezifizierten produktiven Systemen zu umgehen. Als Herausforderung sind z.B. Performanzaspekte und Compliance-Aspekte für Unternehmensprozesse (oftmals ist ein von außen gesteuerter Abruf untersagt) und InDaSpace-spezifische Prozesse zu nennen. Gespeichert werden durch Vertragspartner abrufbare Datensätze in einem In-

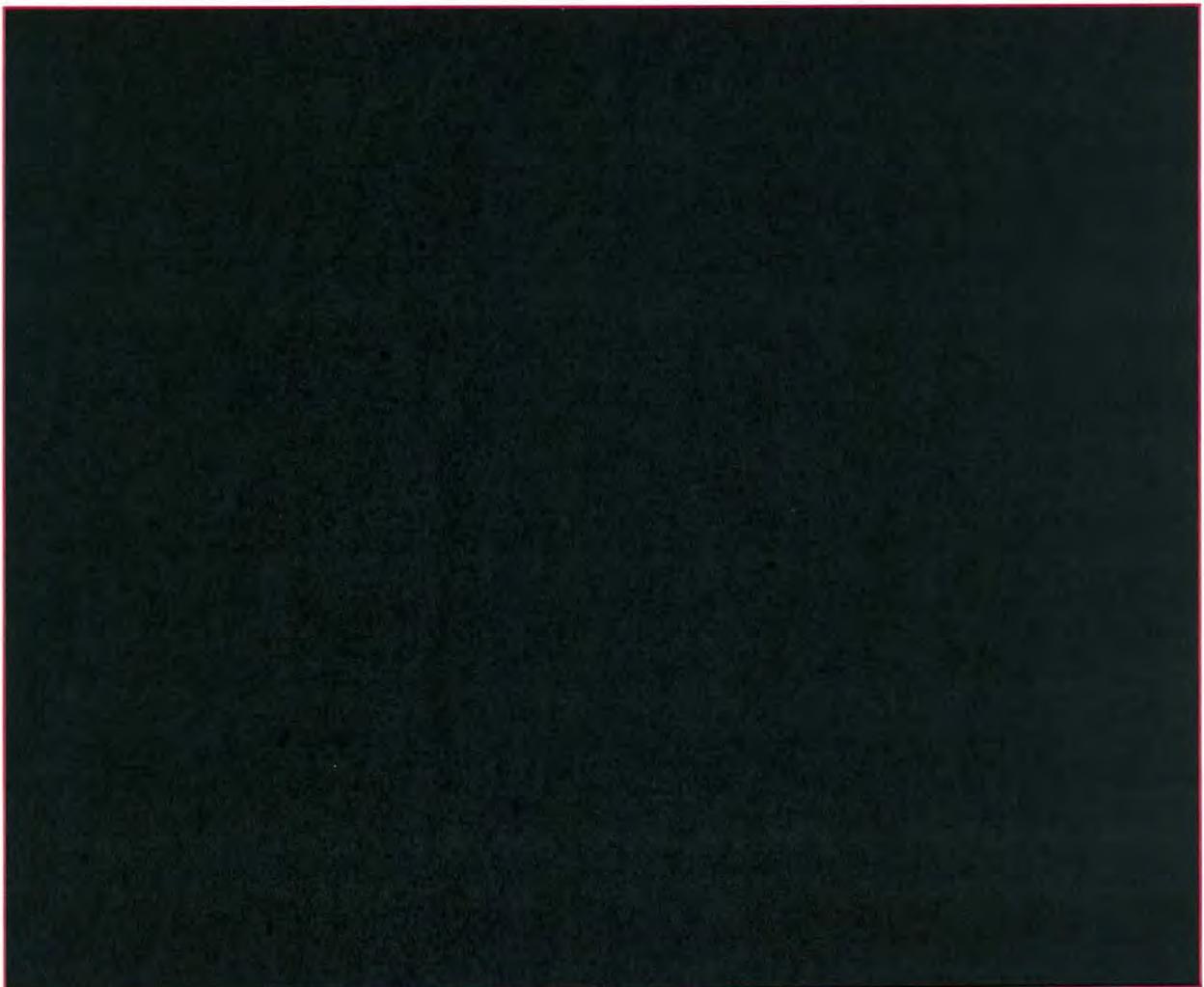
DaSpace empfohlenen Format und Metadaten zu den angebotenen Datenquellen. Die Zwischenspeicherung ist nicht zu empfehlen für hoch-dynamische Datensätze und Use-Cases mit hohen Anforderungen an Aktualität der Daten. In diesen Fällen sollten die Daten direkt an den Data Consumer weitergeleitet werden (Forward).

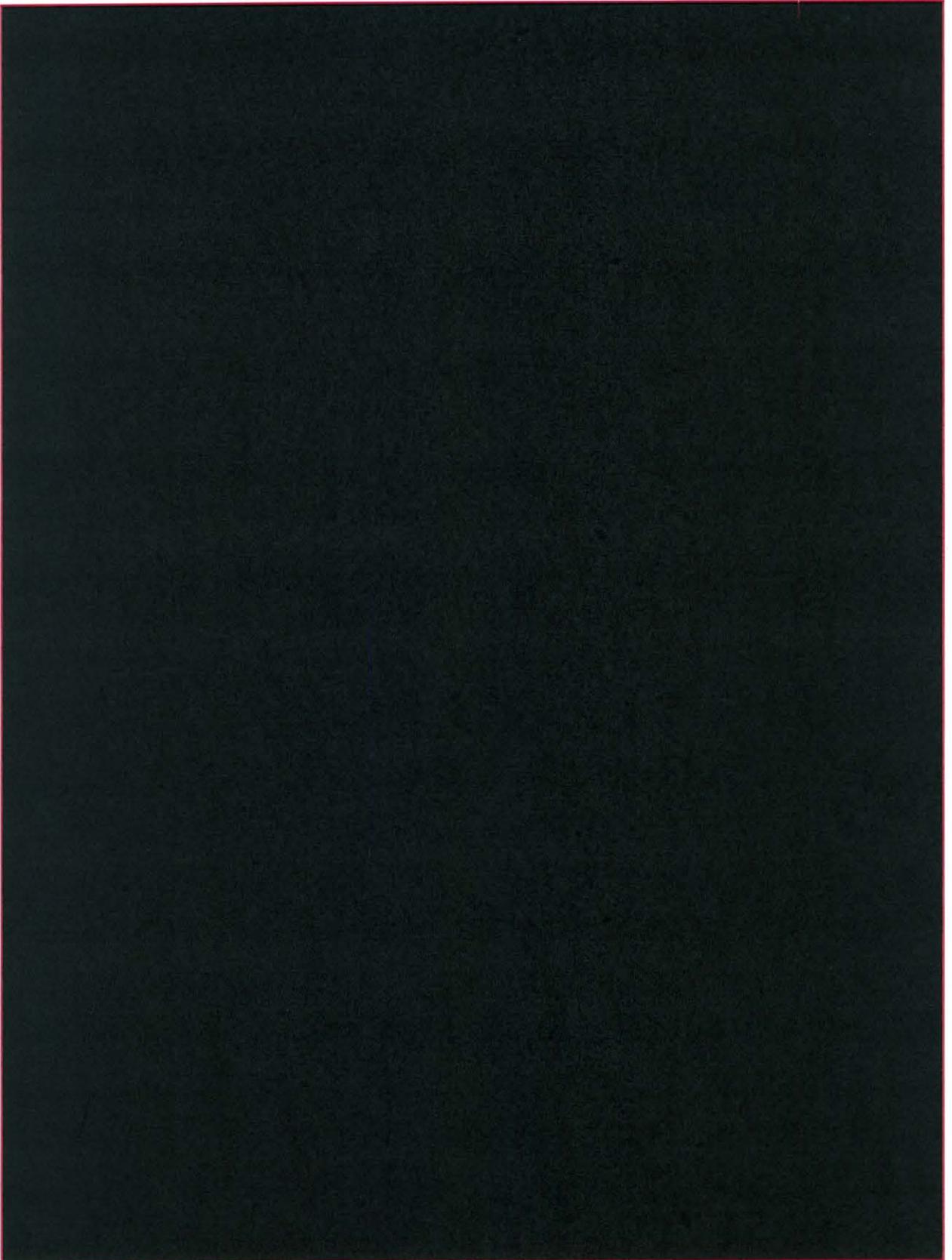
Der EXIK kann publizierte Daten des Data Providers speichern. Im Sinne der Daten Souveränität liegt es nahe, dass die Komponente vom Data Provider gehostet und betrieben wird. Auf der anderen Seite kann der Data Provider die Speicherung und Verwaltung der Daten auch von einem Dienstleister durchführen lassen ohne das Prinzip der Datensouveränität aufzugeben. Dies ist insbesondere wahrscheinlich, wenn der Data Provider keine eigene Infrastruktur aufbauen will, um am InDaSpace teilzunehmen (z.B. attraktiv für kleine und mittlere Unternehmen). Der External InDaSpace Konnektor ist somit geeignet um durch Third Party Cloud Anbieter instanziiert zu werden und als Dienstleistung am Markt angeboten zu werden.

#### Arbeitspaketbeschreibung

AP-Nr.: 2.2.2	Implementierung des External InDaSpace Konnektors (EXIK)	
Startdatum:	M4	
Enddatum:	M36	
<ol style="list-style-type: none"> <li>1. Ziel des Arbeitspaketes <ul style="list-style-type: none"> <li>- Alle Basiskomponenten des EXIK sollen konzipiert und entwickelt werden.</li> <li>- Das interne Datenschema soll festgelegt werden.</li> <li>- Beispielhafte Unternehmensdaten sollen für die ausgewählten Use-Cases über den EXIK ausgetauscht werden können.</li> </ul> </li> </ol>		
<ol style="list-style-type: none"> <li>2. Voraussetzung (Input) <ul style="list-style-type: none"> <li>- Auswahl von unternehmens-internen Systemen gemäß Use-Cases</li> <li>- Testdaten müssen vorliegen</li> </ul> </li> </ol>		
<ol style="list-style-type: none"> <li>3. Lösungsweg <ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 1:</i> <i>Definition und Implementierung des Datenmodells</i> <ul style="list-style-type: none"> <li>- Festlegung des Datenmodells und des Mengengerüsts für den Zugriff</li> <li>- Auswahl von Storage-Technologie</li> <li>- Implementierung des Datenmodells</li> </ul> </li> <li>▪ <i>Arbeitsschritt 2:</i> <i>Implementierung der Ingest Mechanismen</i> <ul style="list-style-type: none"> <li>- Konzeption und Implementierung eines Modules für den Ingest und das Forwarding von Daten</li> </ul> </li> <li>▪ <i>Arbeitsschritt 3:</i> <i>Module zur Zugriffskontrolle und -steuerung (Data Consumption) (Select, Harvest)</i> <ul style="list-style-type: none"> <li>- Aufbau eines Basismoduls zur Übertragung von Daten als Datei oder Dokument im nicht-Echtzeitkritischen Bereich</li> <li>- Aufbau eines Moduls zum Tunneln von Streams bzw. Machine-to-Machine Kommunikation und entsprechenden Protokollen.</li> </ul> </li> <li>▪ <i>Arbeitsschritt 4:</i> <i>Realisierung der Broker Anbindung</i></li> </ul> </li> </ol>		

<p>Registered and Loan Clearing Dienst</p> <ul style="list-style-type: none"> <li>- Realisierung der Workflows für die Prozesse Veröffentlichung und Datenabruf als Modul im Application Container Layer</li> </ul>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 5:</i> <i>Realisierung von Sicherheitsmechanismen für Authentifizierung, Autorisierung sowie Transport und Nachrichtensicherheit</i> <ul style="list-style-type: none"> <li>- Konzeption und Implementierung der Sicherheitsmodule</li> </ul> </li> </ul>	
<ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 6:</i> <i>Erstellung von Dokumentationen des EXIK</i> <ul style="list-style-type: none"> <li>- Beschreibung Datenmodells</li> <li>- Bereitstellung der Schnittstellenbeschreibung</li> <li>- Bereitstellung der Prozessbeschreibung</li> </ul> </li> </ul>	
<p>4. Ergebnis (Output)</p> <ul style="list-style-type: none"> <li>- Softwarekomponente EXIK inklusive Dokumentationen liegt vor: <ul style="list-style-type: none"> <li>- Internes Datenschema liegt vor.</li> <li>- Möglichkeit der Erweiterung um weitere Data Apps</li> <li>- Module zum Datentransfer sind vorbereitet</li> </ul> </li> </ul>	
<p>5. Sonstige Ressourcen</p> <ul style="list-style-type: none"> <li>- keine</li> </ul>	





Arbeitspaketbeschreibung

AP-Nr.: 2.3	Implementierung des InDaSpace Broker
Startdatum:	M4
Enddatum:	M36
<p>1. Ziel des Arbeitspaketes</p> <ul style="list-style-type: none"> <li>- Alle Basiskomponenten des Broker sollen konzipiert und entwickelt werden.</li> <li>- Das interne Datenschema soll festgelegt werden.</li> <li>- Beispielhafte Unternehmensdaten sollen für die ausgewählten Use-Cases über den Broker verfügbar gemacht werden können.</li> </ul>	
<p>2. Voraussetzung (Input)</p> <ul style="list-style-type: none"> <li>- Schnittstellenbeschreibung EXIK</li> <li>- Architekturkonzept für Broker und EXIK</li> </ul>	
<p>3. Lösungsweg</p> <ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 1:</i></li> </ul> <div style="background-color: black; width: 100%; height: 100%; margin: 5px 0;"></div> <ul style="list-style-type: none"> <li>▪ <i>Arbeitsschritt 2:</i> <i>Registrierungs- und Verwaltungsmodul implementieren (Datenanbieter, Datenangebote)</i> <ul style="list-style-type: none"> <li>- Entwurf und Implementierung einer Schnittstelle zum Registrieren von Daten Provider</li> <li>- Entwurf und Implementierung einer Schnittstelle inkl. Funktionalität zur Verwaltung von Datenangeboten durch Daten Provider</li> </ul> </li> <li>▪ <i>Arbeitsschritt 3:</i> Ingest Modul und Aufbau des Persistence Layers (Klassifizierung und Indizierung von Daten) <ul style="list-style-type: none"> <li>- Umsetzung des in AP 1 definierten Metadaten-Modells in der Persistenzschicht</li> <li>- Implementierung eines Moduls zum Ingestieren und Verwalten von Daten in der Persistenzschicht</li> <li>- Implementierung der Suchlogik auf den veröffentlichten und angebotenen Daten</li> </ul> </li> <li>▪ <i>Arbeitsschritt 4:</i> <i>Website zum Publizieren aufbauen (inkl. Access, Search/Data Discovery)</i> <ul style="list-style-type: none"> <li>- Anfertigung eines Web-Designs</li> <li>- Konzeption der Webseitennavigation</li> <li>- Implementierung von Features für die Suchfunktionalität</li> </ul> </li> <li>▪ <i>Arbeitsschritt 5:</i> <i>Unterstützung des Match-Making/Vertragsabschlusses als Modul integrieren</i> <ul style="list-style-type: none"> <li>- Konzeption von Schnittstellen zur Unterstützung des Vertragsabschlusses zwischen Provider und Consumer</li> </ul> </li> <li>▪ <i>Arbeitsschritt 6:</i> <i>API entwickeln und publizieren (Inhalte, Formate, Datenanbieter, Abrufkonditionen)</i></li> </ul>	

- Implementierung eines API zum maschinenbasierten Match-Making	
▪ <i>Arbeitsschritt 7:</i> <i>Komponenten zum Protokollieren der Nutzungsinformationen (Datenabrufe, Abrechnung) implementieren (Data Exchange Logging)</i>	
▪ <i>Arbeitsschritt 8:</i> <i>Komponente zum Service Monitoring implementieren</i>	
▪ <i>Arbeitsschritt 9:</i> <i>Erstellung von Dokumentationen zum Broker</i>	
Personalaufwand gesamt (in PM)	
4. Ergebnis (Output) Softwarekomponente Broker inklusive Dokumentationen liegt vor: - Internes Datenschema liegt vor. - Webseite und API ermöglichen das Auffinden von Datenquellen	
5. Sonstige Ressourcen ▪ Keine	

### 3.2.7 InDaSpace App Store

Der InDaSpace App Store ist eine Komponente, welche das Registrierung und Publikationen Applikationen zum Vorteil der InDaSpace Akteure ermöglicht. Initial werden System-Konnektoren zur Anbindung von Standard Systemen an den InDaSpace, Data Apps zur Verarbeitung, Analyse oder Anreicherung von Daten und empfohlene Schemata/Vokabulare zur Transformation von unternehmensspezifisch ausgezeichneten Daten angeboten. Die beschriebenen Angebote können sowohl von Data Providern als auch von Data Consumern genutzt werden. Die Module werden zur Nutzung heruntergeladen und kommen im INIK zum Einsatz. Applikationen sind hier eine lediglich symbolhafte Bezeichnung für ganze Dienste, die bereit gestellt werden können, oder hinzuladbare Komponenten, wie beispielsweise eine Berechnungsmethode.

Komponentensteckbrief:

Name	InDaSpace APP Store
Funktionalitäten	Publiziert Angebot an Data Apps Publiziert Angebot an empfohlenen Schemata und Vokabularen Publiziert Angebot an System Konnektoren Unterstützt Prozess zur Registrierung von Data Apps und System Konnektoren durch Drittanbieter Unterstützt Prozess zur Registrierung von Schemata und Vokabularen
Verortung	Betrieb durch zertifizierte InDaSpace App Store Anbieter.
Art der Realisierung	Eigenständige Server-Komponente
Schnittstellen	Internal InDaSpace Connector (send/receive)
Besonderheiten	-