

# Kontextsensitive Informationsvisualisierung mit kompositen Rich Internet Applications für Endnutzer

## Kurzfassung der Dissertation

zur Erlangung des akademischen Grades  
Doktoringenieur (Dr.-Ing.)

vorgelegt an der  
Technischen Universität Dresden  
Fakultät Informatik

eingereicht von  
**Dipl.-Medieninf. Martin Voigt**  
geboren am 06. August 1982 in Großröhrsdorf

### **Gutachter:**

Prof. Dr.-Ing. Klaus Meißner (Technische Universität Dresden)  
Prof. Dr.-Ing. Jürgen Ziegler (Universität Duisburg-Essen)



Dresden im November 2014

# 1 Motivation und Herausforderungen

Das stetig wachsende Datenaufkommen ist eine wesentliche Herausforderung für den Menschen in allen Bereichen des beruflichen und privaten Alltags [Agr12]. 90% der weltweiten Daten wurden in den letzten beiden Jahren generiert [SIN13]. Problematisch ist, dass die menschliche Aufnahmefähigkeit limitiert bleibt [Key12]. Um trotzdem relevante Informationen zu identifizieren und auch zu verstehen, nehmen Techniken und Anwendungen zur **Informationsvisualisierung (InfoVis)** einen immer größeren Stellenwert ein. Sie sollen helfen, die Interpretation der Daten zu vereinfachen, um letztlich neue Erkenntnisse zu gewinnen. Verbreitete Werkzeuge, die den Endanwender adressieren, sind Microsoft Excel aber auch webbasierte Anwendungen wie IBM Many Eyes oder Tableau Public.

Dennoch hat sich die Vision der „InfoVis for and by the masses“ aufgrund dreier Hürden noch nicht durchgesetzt [Gra10]. Zunächst hat der Endanwender das Problem, die für ihn möglicherweise interessanten Daten zu identifizieren, zu verstehen und zu selektieren. Danach ist das Erstellen einer korrekten und effektiven Abbildung der Daten auf einzelne visuelle Attribute ein Hindernis. Letztlich existiert die Schwierigkeit, dass der Nutzer das erzeugte Bild auf Basis seines Vorwissens interpretieren muss. Verschärft wird die Situation dadurch, dass sich Fehlentscheidungen in den einzelnen Schritten akkumulieren und letztlich eine korrekte Interpretation bzw. ein Erkenntnisgewinn nahezu unmöglich wird.

Hieraus kann man schließen, dass eine zielführende und effektive Visualisierung entweder weiterhin nur für Experten möglich ist, oder dass Werkzeuge geschaffen werden müssen, die keine oder nur kleine Eintrittsbarrieren für Endnutzer aufbauen [Gra12, S. 146]. Ein gutes Beispiel für Letzteres ist die Plattform [visual.ly](http://create.visual.ly/)<sup>1</sup>. Der Nutzer wird in dieser stark geführt und muss zumeist nur die Datenquelle und ein Template für die Visualisierung auswählen, welches genau auf eine Aufgabe, wie etwa das Visualisieren des Lebenslaufs, zugeschnitten ist. Die gute Unterstützung des Endnutzers geht hier jedoch vor allem zu Lasten der Generalität der Werkzeuge. So ist bspw. die Auswahl an Datenformaten und Quellen eingeschränkt, nur wenige, sehr konkrete Aufgaben werden unterstützt, die Visualisierungstechniken sind meist fest vorgegeben oder die Anwendung unterscheidet nicht zwischen den Gegebenheiten unterschiedlicher Endgeräte.

Gerade der Nutzung mobiler Endgeräte sowie der allgegenwärtigen Verfügbarkeit der Werkzeuge wird bei der visuellen Analyse eine wachsende Bedeutung zugeschrieben [Fin12]. Die Erweiterung und Anpassung der Werkzeuge, um etwa eine neue Visualisierungstechnik für Smartphones bereitzustellen, ist meistens sehr aufwändig. Entwickler müssen nicht nur die Visualisierung implementieren, sondern auch die Abbildungsregeln der Daten auf die visuellen Strukturen und Attribute definieren – und dies für jegliche Kombinationen zwischen Daten und Visualisierungen. Ein möglicher Lösungsansatz für die skizzierten Probleme ist die Realisierung von „Ad-hoc Systems“ [Kei10, S. 88], die sich u. a. auf Softwarekomponenten zur visuellen Analyse stützen. Diese Vision geht mit den Zielen der **Service-oriented Architecture (SOA)** und der kompositen Webanwendungen einher, die insbesondere den *Long Tail* [And06] von

---

<sup>1</sup> <http://create.visual.ly/>

Softwarelösungen adressieren. Sie zielen darauf ab, Daten, Geschäftslogik und auch **User Interface (UI)**-Bestandteile zu kapseln und mittels fest definierter Schnittstellen bereitzustellen. Auf deren Basis können Daten und **UI**-Widgets je nach Anwendungsfall teils automatisch kombiniert werden. Dies erhöht die Wiederverwendbarkeit und spart Zeit und Entwicklungskosten. Unter Zuhilfenahme von Modellen, die die Anwendungslogik sowie den -kontext beschreiben, ist es zudem möglich, eine komposite **Rich Internet Application (RIA)** angepasst an die vorliegende Situation und Laufzeitumgebung zu nutzen [Pie12]. Diese Kontextsensitivität ist in heutigen **InfoVis**-Werkzeugen kaum gegeben [Ebe04].

Auch wenn die genannten Konzepte passend für den Anwendungsfall der **InfoVis** erscheinen, so adressieren sie aufgrund des Vorgehensmodells oftmals nur den Softwareentwickler [Cas11]. Um dem Endanwender Zugang zu den adaptiven, kompositen **RIA** zu verschaffen, mangelt es an einem Integrationsprozess, der den speziellen Anforderungen des Visualisierungsworkflows gerecht wird, und dabei dessen Komplexität durch intelligente Techniken reduziert, so dass die bereits genannten Visualisierungshürden gemindert bzw. gar eliminiert werden.

Aktuelle Ansätze im Bereich der **SOA** und auch der kompositen **RIA**, wie **CRUISe** [Pie12], verwenden semantische Modelle, um die Komplexität einer Domäne von der eigentlichen Softwareentwicklung zu trennen [Rui06]. Neben der Nutzung als Informationsquelle oder als Teil der Geschäftslogik werden sie auch als Grundlage für intelligente Benutzerschnittstellen verwendet. Für den Einsatz von semantischen Modellen zur Unterstützung der **InfoVis** ergeben sich jedoch zwei grundlegende Probleme. Einerseits mangelt es an Ontologien, die die teilweise interdisziplinären Konzepte und Zusammenhänge der Visualisierungsdomäne beschreiben. Eine große Herausforderung ist dabei, die immer konkurrierenden Ansichten und Definitionen abzuwägen und hierbei die Community in den Formalisierungsprozess mit einzubeziehen. Andererseits fehlt es an einem Konzept, wie das formalisierte Wissen innerhalb des Visualisierungsprozesses eingesetzt werden kann [Wan09], so dass der Endanwender nicht mit den theoretischen Grundlagen von Ontologien in Berührung kommt.

Aus der Motivation wird deutlich, dass komposite **RIA** unter der Verwendung von semantischen Modellen ein vielversprechendes Softwareparadigma darstellt, um Endanwendern in verschiedenen Kontexten eine geeignete **InfoVis** ihrer Daten zu ermöglichen. Doch wirft deren Verwendung in dieser Domäne eine Reihe komplexer Forschungsfragen auf, die durch bestehende Ansätze bislang nicht ausreichend beantwortet werden. Die vorliegende Arbeit widmet sich somit den drei folgenden grundlegenden wissenschaftlichen Herausforderungen:

- Die Spezifikation eines Semantik-gestützten **Visualisierungsprozesses**, der Endanwendern die kontextsensitive **InfoVis** ermöglicht.
- Die Konzeption einer modularen, erweiterbaren **Visualisierungsontologie (VISO)** zur Repräsentation des Visualisierungswissens als Grundlage für den definierten Prozess.
- Die Schaffung einer plattformunabhängigen **Architektur für die komponentenbasierte InfoVis**, die den Anforderungen an ein webbasiertes Visualisierungssystem genügt und den definierten Semantik-gestützten Prozess realisiert.

Die folgenden drei Abschnitte geben einen Überblick zu den neuen wissenschaftlichen Konzepten der drei Problemfelder. Anschließend wird kurz auf die Umsetzung und Validierung eingegangen, ehe abschließend die Kernbeiträge prägnant zusammengefasst werden.

## 2 Semantik-gestützter Visualisierungsprozess für komposite Rich Internet Anwendungen

Das maßgebliche Konzept der Dissertation stellt der mehrstufigen, Semantik-gestützten **InfoVis**-Prozess dar. Aufgrund geeigneter **UI**- und Interaktionsparadigmen sowie von Automatismen, die in verschiedenster Art und Weise eine gemeinsame, formale Wissensbasis nutzen, wird es für den Endanwender möglich, die drei essentiellen Barrieren des **InfoVis**-Prozesses – Datenauswahl, Erstellung von Abbildungsregeln und Interpretation der Visualisierung – zu überwinden. Abbildung 2.1 stellt den Prozess im Überblick dar. Er besteht auf der einen Seite aus fünf Schritten, die die Anwender durchführen müssen, um aus den Daten Erkenntnisse zu gewinnen, die auf der anderen Seite durch Hilfsfunktionen unterstützt werden. Die drei genannten Hürden sind in Abbildung 2.1 rot hervorgehoben. Im Folgenden werden die bestimmenden Teilkonzepte und wissenschaftlichen Beiträge der einzelnen Prozessschritte kurz erläutert.

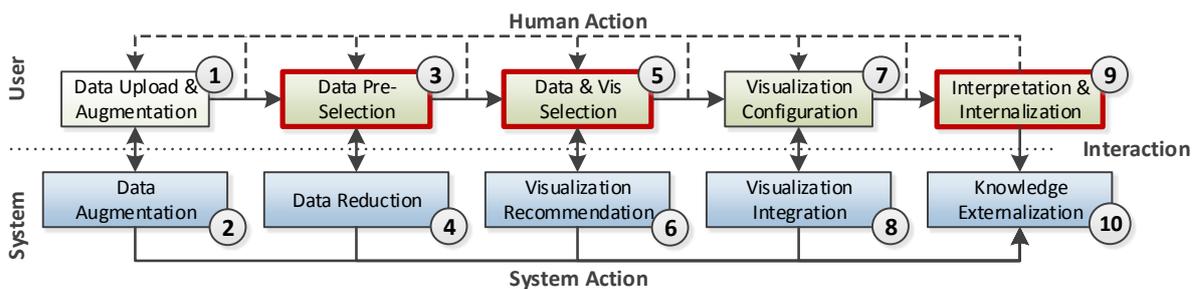


Abbildung 2.1: Überblick zum Semantik-gestützten Visualisierungsprozess

**Transformation und Auszeichnung der Daten** Nachdem die zu visualisierenden Daten dem Prozess zugeführt wurden ①, durchlaufen diese eine mehrstufige, semiautomatische Pipeline ②. Hierbei werden die Daten aus unterschiedlichsten Quellen und Formaten zunächst automatisch in ein einheitliches, semantisches Datenformat transformiert, was die spätere Annotation der Rohdaten um semantische Konzepte deutlich vereinfacht bzw. gar erst ermöglicht. Das Hauptaugenmerk liegt jedoch auf der Datenanalyse, um in folgenden Prozessschritten die notwendigen Metainformationen und damit ein Verständnis über die Daten zur Verfügung stellen zu können. Beispielsweise werden fehlende Datentypen, **InfoVis**-relevante Metriken oder Schlüsselkonzepte des Datensatzes erkannt. Ein Eingreifen des Datenanbieters ① ist nur notwendig, um die generierten Ergebnisse zu beurteilen und ggf. an Spezifika des Datensatzes anzupassen, wenn die Annotationen nicht den Vorstellungen entsprechen.

**Endnutzer-gerechte Filterung der Daten** Aufbauend auf den identifizierten Metainformationen sowie auch auf Basis von gesammelten Informationen zum Nutzerverhalten beinhalten die Schritte ③ und ④ Konzepte, die den Endanwender bei der Datenfilterung und -auswahl unterstützen. Hierzu werden facetthierarchische Filter definiert sowie Zusatzinformationen und Empfehlungen

berechnet, die teils allgemeingültig für den Datensatz aber auch teilweise auf den Nutzer zugeschnitten sind. Sie alle sind wichtige Grundlagen einer kompositen, auf Endanwender zugeschnittenen Benutzerschnittstelle, die die Exploration im Datenraum, die Auswahl von interessanten Konzepten und damit insgesamt die Filterung großer Datensätze ermöglicht.

**Kontextsensitive Auswahl von InfoVis-Komponenten** Prozessschritt ⑤ stellt die zweite und höchste Hürde des Visualisierungsprozesses dar: Die Abbildung der gewählten Daten auf visuelle Strukturen und Attribute. Hierzu wird mit dem Konzept des **Weighted Faceted Browsing (WFB)** eine intelligente, für Endnutzer einfach zu erschließende Benutzerschnittstelle vorgeschlagen. Diese nutzt die errechneten Vorschläge eines mehrstufigen Empfehlungsalgorithmus ⑥, der die Semantik der Komponenten, der Daten sowie Kontextinformationen zum Anwender und seines Endgerätes ausnutzt. Durch die Kombination beider Konzepte kann der Endnutzer iterativ seine teils unsicheren Suchkriterien definieren, wobei er sofort Feedback über mögliche Visualisierungstechniken erhält. Weiterhin wird durch die Verwendung der facettierten Filterung sowie durch die automatische Berechnung von Abbildungsregeln von Daten auf Komponenten seine kognitive Last und das notwendige visualisierungsspezifische Vorwissen minimiert.

**Integration und Konfiguration der kompositen InfoVis** Für die automatische Integration ⑧ der vom Endnutzer selektierten InfoVis- und Filterkomponenten kann mit wenigen Ausnahmen auf das existierende Vorgehen von CRUISe [Pie12] zurückgegriffen werden. So ist es für Endanwender möglich, webbasierte Komponenten zu nutzen, ohne Programmierkenntnisse einsetzen zu müssen. Einen wissenschaftlichen Beitrag stellt jedoch die Klassifikation und Vorstellung von Adaptionstechniken für eine komposite InfoVis-RIA dar ⑦. Neben der Wiederverwendung bereits existierender Techniken werden auch neue Konzepte, wie z. B. für die Etablierung von Koordinationsmechanismen zwischen Komponenten, definiert. Insgesamt machen die vorgestellten Techniken die iterative, semiautomatische Anpassung der Anwendung für den Endnutzer zur Laufzeit möglich.

**Unterstützung bei der Nutzung der kompositen InfoVis** Weiterhin werden fünf Konzepte für die User Assistance vorgestellt, um den Anwender bei der Bedienung und Interpretation seiner kompositen InfoVis zu helfen ⑨. Hiermit werden Antworten auf die drei identifizierten Hilfesituationen, mangelndes Verständnis 1) zur Komponente, 2) zur Komposition oder 3) zu den Daten, gegeben. Durch das Gesamtkonzept bietet sich dem Endnutzer die Chance, die Visualisierung effektiver zu nutzen und zu verstehen. Gleichzeitig wird aber auch der Komponentenentwickler entlastet, da viele der Aufgaben zur Erstellung einer Hilfe von der Plattform übernommen werden. Neben den einheitlichen UI- und Interaktionsmetaphern für alle Hilfsfunktionen ist gerade die Nutzung des VISO-Vokabulars ein großer Gewinn für Nutzer und Entwickler. Letzterer muss sich nicht über die richtige Bezeichnung und Erläuterung von Konzepten Gedanken machen, und für den Endanwender sind alle Komponenten konsistent beschrieben.

**Wiederverwendung von Visualisierungswissen** Letztlich wird ein Konzept zur Externalisierung und Wiederverwendung von im InfoVis-Prozess geschaffenem Wissen aufgestellt ⑩. Unter Letzterem ist vor allem die Effizienz von Daten-Komponenten-Kombinationen zu verstehen. Zur

Wissensgewinnung werden explizite Bewertungen der Endnutzer mittels einer binären Skala sowie implizite, aus Interaktionen abgeleitete Ratings eingesetzt. Die Wiederverwendung des externalisierten Wissens erfolgt insbesondere im Empfehlungsalgorithmus ⑥. Hat der Nutzer für eine Daten-Komponenten-Kombination noch keine explizite Bewertung abgegeben, so wird versucht, diese durch **Collaborative Filtering (CF)** oder aus den impliziten Ratings zu errechnen. Für ein laufendes **InfoVis**-System wird es somit ohne Beteiligung von Experten möglich, Fakten dazuzulernen und so bessere Komponentenvorschläge zu generieren.

### 3 Formalisierung von Visualisierungswissen als Grundlage für InfoVis-Systeme

Ausgehend von den Anforderungen des Semantik-gestützten **InfoVis**-Prozesses aus Kapitel 2 wird eine **VISO** konzipiert, um **InfoVis**-Wissen interpretierbar durch Mensch und Maschine zu formalisieren. Grundlage ist hierzu einerseits ein Vorgehensmodell, um dem iterativen Entwicklungsprozess die notwendige Struktur zu geben, und andererseits eine Analyse von existierenden Klassifikationen der **InfoVis**-Domäne. Aus Letzterer ist ersichtlich, dass grundsätzlich sieben Wissensgebiete in der **InfoVis** existieren, die bisher jedoch nie in einem formalen Wissensmodell gemeinsam betrachtet wurden. Auch wird das eingangs der Arbeit dargestellte Problem bestätigt, dass die große Anzahl bestehender Klassifikationen wenig formal sind.

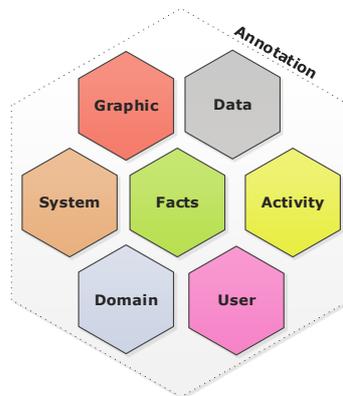


Abbildung 3.1: Konzeptioneller Überblick der **VISO**-Module

Abbildung 3.1 zeigt den Aufbau des modularen **VISO**-Konzepts, das aus sieben, sich mehr oder weniger referenzierenden Teilmodellen besteht, auf deren Inhalt im Folgenden kurz eingegangen wird. Das *Facts*-Modul ist symbolisch in der Mitte angeordnet, da es zur Formalisierung von Faktenwissen auf die stabilen Konzepte aller Teilmodelle zurückgreifen kann.

*Daten* In diesem Modul existieren Konzepte, Relationen und Eigenschaften, um Datenstrukturen, wie etwa Graphen, und Datenvariablen zu definieren und näher zu beschreiben. Hierzu gehört bspw. die Skala oder die Kardinalität einer Variable.

*Graphisches Vokabular* Dieses Teilmodell dient insbesondere zur Definition von graphischen

Repräsentationen, wie Karten oder bestimmten Diagrammen, samt deren zumeist kompositen Struktur. Weiterhin werden auch visuelle Variablen und deren Abhängigkeiten formalisiert.

*Menschliche Aktivität* Teile dieses Wissensgebietes werden vor allem unter Zuhilfenahme der Aktivitätstheorie [Bar97] hierarchisch hinsichtlich Aufgaben, Aktionen und Operationen strukturiert.

*Nutzer-, System- und Domänenwissen* Als Ontologie formalisiertes Wissen zum Nutzer- und Systemkontext kann im Rahmen der VISO gut wiederverwendet werden. Es wird auch kein Wissen einer Fachdomäne modelliert, sondern Konzepte und Relationen geschaffen, um VISO-Inhalte mit bestehenden Ontologien zu verknüpfen.

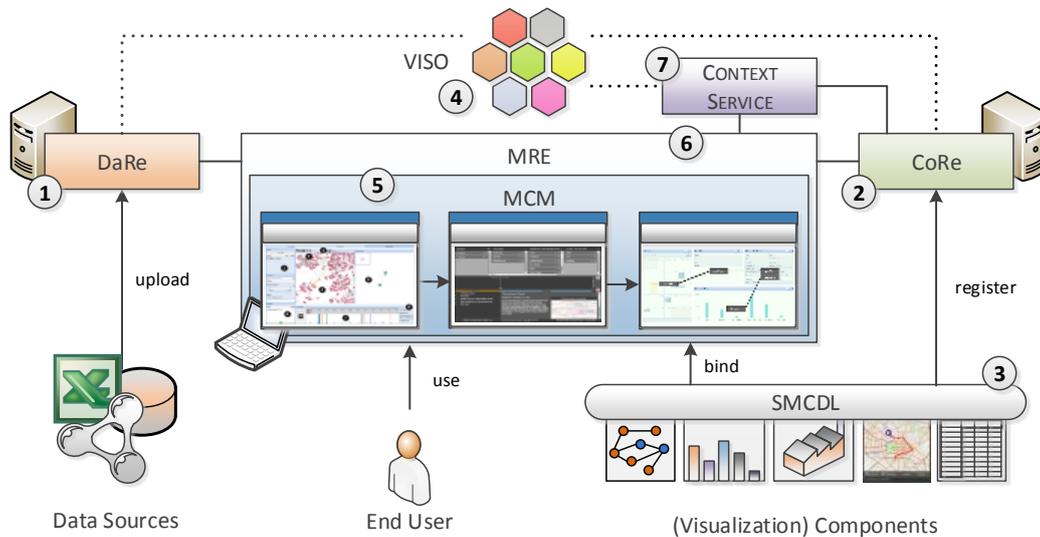
*Faktenwissen* Dieses Modul enthält einerseits ein Vokabular zur Beschreibung von Faktenwissen, dass viele Konzepte der anderen Module referenziert, und andererseits auch Instanzdaten aus bspw. empirischen Studien.

Die bisherige Realisierung der frei zugänglichen VISO erfolgt mit Hilfe der W3C-Standards [RDF Schema \(RDFS\)](#) und [Web Ontology Language \(OWL\)](#). Ein Grundpfeiler der VISO und ein Abgrenzungskriterium zu anderen Ontologien mit ähnlichem Fokus ist die Wiederverwendung und Zitierung der zugrundeliegenden Literatur und damit der Definition der Begriffe innerhalb des Modells, was besonders das Verständnis der VISO-Interessierten aber auch der Endnutzer fördert. Hierzu wurde auch eine webbasierte Dokumentation samt eines Forums realisiert, so dass Domänenexperten ohne Semantic-Web-Kenntnisse das Wissensmodell zielgerichtet nutzen, diskutieren und so zur Weiterentwicklung beitragen können.

## 4 Softwarearchitektur für Ontologie-gestützte, komponentenbasierte InfoVis-Systeme

Zur Realisierung des konzipierten Semantik-gestützten InfoVis-Prozesses (vgl. Kapitel 2) bedarf es neben einer formalen Wissensbasis (vgl. Kapitel 3) auch einer flexiblen, intelligenten Softwarearchitektur für die webbasiert InfoVis. Für diese stellen die Konzepte komponentenbasierter RIA aus CRUISe [Pie12] eine wesentliche Grundlage dar, da sie viele Anforderungen, wie die semantische Beschreibung und Verwaltung von Komponenten oder deren kontextsensitive Integration zur Laufzeit, bereits adressieren.

Abbildung 4.1 gibt einen Überblick zu den wesentlichen Konzepten der Softwarearchitektur. Das [Data Repository \(DaRe\)](#) ist ein Web Service ①, der alle Funktionen rund um die Haltung, Analyse und Bereitstellung der zu visualisierenden Daten beinhaltet. Es offeriert eine gemeinsame Datenschnittstelle samt einheitlichem Datenmodell aller Nutzdaten für die anderen Architekturbestandteile. Dem gegenüber steht das [Component Repository \(CoRe\)](#) ②, welches u. a. die Verwaltung von und Suche nach Komponenten basierend auf deren Beschreibung gewährleistet. Diese erfolgt mittels der erweiterten [Semantic Mashup Component Description Language \(SMCDL\)](#) ③. Auch wenn durch sie UI-Komponenten sowohl als auch



**Abbildung 4.1:** Überblick zur komponentenbasierten Softwarearchitektur

solche ohne **UI** beschrieben werden können, werden im Rahmen der Arbeit insbesondere erstere adressiert, da die Daten zentral über das **DaRe** bereitgestellt werden. Des Weiteren sollte dem Endnutzer stets eine adäquate visuelle Repräsentation geboten werden, um eine Komponente effektiv bedienen zu können. Dies betrifft insbesondere die sogenannten Logikkomponenten zur Manipulation und Filterung der Daten, um sie hinsichtlich des **Coordinated Multiple View (CMV)**-Konzepts in der kompositen **InfoVis** nutzen zu können. Die Daten und auch die Komponenten bzw. deren Beschreibungen sind mit Konzepten der **VISO** (4) ausgezeichnet, so dass gezielt Abbildungsregeln identifiziert und bewertet werden können. Die komposite Anwendung, die letztlich den mehrstufigen **InfoVis**-Prozess aus Kapitel 2 realisiert, wird plattformunabhängig mittels eines **Mashup Composition Model (MCM)** (5) beschrieben. Dieses Modell wird durch das **Mashup Runtime Environment (MRE)** (6) interpretiert, welches die für den aktuellen Kontext (7) passenden Komponenten integriert. Zur Sammlung, Konsolidierung und Verwaltung von Kontextinformationen kann das bereits in **CRUISe** integrierte Framework **CroCo** [Pie12, S. 163ff.] ohne Einschränkungen verwendet werden.

## 5 Umsetzung und Validierung der Konzepte

Um die Umsetzbarkeit und Validität der definierten Konzepte zu prüfen, wurden zunächst die einzelnen Architekturkonzepte aus Kapitel 4 prototypisch implementiert. Hierbei stand in erster Linie die Implementierung des **DaRe** als Java-basiertem Web Service im Mittelpunkt, da dieser die zentralen, zeitkritischen Aufgaben der Verwaltung, Analyse und Augmentierung der zu visualisierenden Daten beinhaltet. Weiterhin wurden die funktionalen Erweiterungen der **CRUISe**-Plattform realisiert. Hierzu zählen u. a. die Anpassungen an der **SMCDL** und am **MCM**. Neben der damit einhergehenden Anpassung an der Modellverwaltung stellt die wesentliche, in

das **CoRe** integrierte Funktion der definierte kontextsensitive Empfehlungsalgorithmus dar. Für diesen wurde in Benchmarks festgestellt, dass der Abruf von annotierten Daten aus dem **DaRe** zeitkritisch ist, so dass zur Reduzierung des Problems erfolgreich Caching-Strategien realisiert wurden. Funktionen, wie die Komponentenintegration, die Etablierung der Koordination zwischen den Widgets, Hilfsfunktionen oder die Bewertung der genutzten Komponenten wurden in der **Thin Server Runtime (TSR)** umgesetzt, welche eine umfangreiche, rein clientseitige Implementierung der **MRE** darstellt. Insgesamt wurde gezeigt, dass **InfoVis**-Komponenten eine generische Schnittstellenbeschreibung der Daten aufweisen und kontextsensitiv passende Abbildungsregeln von Daten auf die Komponenten generiert werden können. Dies kann zu einer breiten Wiederverwendbarkeit der gekapselten Visualisierungstechniken führen.

Auch die Realisierbarkeit des Gesamtprozesses aus Kapitel 2 wurde erfolgreich durch die technische Umsetzung zweier **CRUISe**-Webanwendungen nachgewiesen. Die erste dient dabei zur Bereitstellung der Daten und zur Bearbeitung der Annotationen durch den Data Provider (vgl. Abbildung 2.1 Prozessschritt ①). Die zweite, wesentlich komplexere Anwendung (**VizBoard**) adressiert direkt den Endanwender und dient der Validierung der Schritte ③ bis ⑨. Durch die Umsetzung der in der Konzeption vorgeschlagenen **UI** zur Datenvorfilterung kann der Endanwender in einem semantischen Datensatz browsen, ihn anhand von vordefinierten Facetten filtern und so relevante Konzepte identifizieren. Das Konzept des **WFB** als Such-Frontend für den Empfehlungsalgorithmus wurde ebenso erfolgreich umgesetzt. Die hierzu durchgeführte formative Nutzerstudie zeigt, dass Endanwender gut ihre teils unsicheren Suchkriterien definieren können. Nach der Auswahl einer oder mehrerer Komponenten werden diese ohne Zutun des Endanwenders integriert, wobei ein automatisches Mapping der Daten auf die Schnittstelle der Komponenten erfolgt. Um Nutzern den Einstieg in die komposite, webbasierte **InfoVis** zu erleichtern, wurde der Großteil der vorgeschlagenen Assistenzfunktionen erfolgreich realisiert. Leider wurde bei der Validierung mit Endnutzern festgestellt, dass sie auf die Hilfe nur in ca. 1/3 der Anwendungsfälle zurückgreifen. An dieser Stelle ist die Integration von proaktiven Hilfemechanismen vorstellbar, doch deren Konzeption und Realisierung ist ein Forschungsthema für sich. Insgesamt konnte aber gezeigt werden, dass durch die neuen Konzepte nun auch Endanwender in die Lage versetzt werden, effizient ohne Programmier- und Visualisierungswissen mit Techniken des **End User Development (EUD)** komposite Webanwendung zur **InfoVis** zu erstellen.

Letztlich stand die Frage im Raum, ob die vorgeschlagenen und letztlich integrierten Widgets effektiv und aufgabenangemessen sind. Aus diesem Grund wurde eine abschließende Evaluation durchgeführt. Der Fokus der Untersuchung lag hierbei auf der Identifikation essentieller Konzepte in der Vorfilterung und der Wahl effektiver **InfoVis**-Komponenten. Es wurde jedoch nicht untersucht, ob die Endanwender die präsentierten Daten korrekt interpretieren, da diese Analyse zu vielschichtig und nicht Kern der Dissertation ist. Die Grundidee der Evaluation war, Visualisierungsaufgaben durch Endnutzer mit Unterstützung von **VizBoard** sowie von Experten ohne Assistenzsysteme zu lösen. Da für diese Analyse die Anzahl der Probanden und die Wahl adäquater Experten wichtige Einflussfaktoren sind, deren notwendige Zahl bzw. Güte indes schwer zu bestimmen ist, sind die Ergebnisse eher als Tendenz denn als ein allgemeingültiges Ergebnis zu verstehen. Die Effektivität der Probanden ist als gut zu bewerten. In der Vorfilterung haben 13 von 15 Probanden (ca. 86,7%) ein wichtiges Konzept identifiziert. Und auch im **WFB** wählten im Schnitt 12 von 15 Probanden (80%) eine auch von Experten empfohlene **InfoVis**-Komponente zur Lösung der jeweiligen Aufgabe. Da sich die Effizienz der vorgeschlagenen Konzepte nur in einer vergleichenden Studie adäquat untersuchen lässt, es

aber an passenden Konkurrenzlösungen mangelt, wurden die Probanden lediglich nach ihrer subjektiven Meinung gefragt. Im Vergleich zu Microsoft Excel, mit dem alle Teilnehmer mindestens Grundkenntnisse aufweisen können, sagten 80% aus, dass sie mit VizBoard bei mindestens gleich guten Ergebnissen schneller ans Ziel gelangen. So lässt sich zusammenfassend sagen, dass die in VizBoard realisierten Konzepte den Endanwender tendenziell schneller zu einer zielführenden **InfoVis** verhelfen.

## 6 Zusammenfassung und Ausblick

In der vorliegenden Arbeit wurde ein umfassendes, innovatives Konzept für die kontextsensitive **InfoVis** mit **RIA** für Endanwender vorgestellt. Ausgangspunkt ist der durchgängige, an die Visualisierungspipeline angelehnte Semantik-gestützte **InfoVis**-Prozess. In diesem wird es aufgrund von **UI**- und Interaktionsparadigmen sowie Automatismen, die eine gemeinsame, formale Wissensbasis nutzen, für den Endanwender möglich, die drei essentiellen Barrieren des **InfoVis**-Prozesses – Datenauswahl, Erstellung von Abbildungsregeln und Interpretation der Visualisierung – zu überwinden. Weiterhin wurde eine modulare **VISO** konzipiert und prototypisch realisiert, um für den neuartigen Prozess eine adäquate, formale Wissensbasis zu schaffen. Durch ihr Fundament in einer tiefgründigen Recherche sowie durch die auf Semantic-Web-Standards basierende Veröffentlichung kann ihr Einsatz über diese Dissertation hinaus erfolgen. Schließlich wurden umfassende Konzepte einer Softwarearchitektur für Ontologie-gestützte, komponentenbasierte **InfoVis**-Systeme vorgestellt. Mit diesen steht nun erstmals für eine webbasierte **InfoVis**-Lösung ein ganzheitliches Architekturkonzept zur Verfügung, das die Ausführbarkeit der Anwendungen in der heute existierenden, heterogenen Landschaft der (mobilen) Endgeräte gewährleisten kann. Insgesamt wurden durch diese Dissertation die folgenden wissenschaftlichen Kernbeiträge geschaffen:

- Konzeption eines ganzheitlichen, Semantik-gestützten **InfoVis**-Prozesses zur Unterstützung des Endanwenders. Hier ordnen sich folgende Beiträge unter:
  - Spezifikation von Mechanismen zur Identifikation und Filterung interessanter bzw. relevanter Konzepte in großen, semantischen Datensätzen
  - Konzeption eines Facetten-basierten, iterativen Suchkonzepts bei einem teilweise nur vagen Verständnis der Suchkriterien
  - Spezifikation eines Algorithmus zur kontextsensitiven, gewichteten Suche und Empfehlung von **InfoVis**-Widgets
  - Konzeption von Unterstützungsmechanismen für die Nutzungsphase einer kompositen **InfoVis**
  - Definition von Methoden zur Gewinnung und Wiederverwendung von Nutzer-spezifischem **InfoVis**-Wissen
- Identifikation, Formalisierung, Dokumentation und Publikation von **InfoVis**-Wissen in Form einer modularen Ontologie

- Erweiterung und Adaption einer modell- und komponentenbasierten **RIA**-Plattform zur Nutzung in der **InfoVis**-Domäne

Durch die Implementierung entscheidender Konzepte sowie einer beispielhaften **InfoVis**-Anwendung wurde die Tragfähigkeit der geschaffenen Konzepte nachgewiesen. Anhand einer Vielzahl von formativen sowie einer summativen Nutzerstudien konnte validiert werden, dass sich aus den geschaffenen Lösungen tendenziell Vorteile für den Endanwender bei der Erstellung einer webbasierten **InfoVis** ergeben. Trotz kleiner Mängel hinsichtlich der Gebrauchstauglichkeit, die auf den prototypischen Charakter der Implementierung zurückzuführen sind, erweisen sich die Konzepte als tragfähig und bilden damit eine solide Grundlage für Erweiterungen durch angeschlossene Forschungsprojekte. Hierfür sind verschiedene Stoßrichtungen denkbar. Ein Beispiel sind proaktive, intelligente Mechanismen zur Unterstützung der Endanwender, die u. a. kontextsensitiv Hinweise zur Bedienung geben. Eine andere, naheliegende Forschungsrichtung ist die Übertragung der geschaffenen Konzepte auf kollaborative **InfoVis**-Szenarien.

## Literaturverzeichnis

- [Agr12] AGRAWAL, Divyakant; BERNSTEIN, Philip; BERTINO, Elisa; DAVIDSON, Susan; DAYAL, Umeshwar; FRANKLIN, Michael; GEHRKE, Johannes; HAAS, Laura; HALEVY, Alon; HAN, Jiawei; JAGADISH, H. V.; LABRINIDIS, Alexandros; MADDEN, Sam; PAPA-KONSTANTINOY, Yannis; PATEL, Jignesh M.; RAMAKRISHNAN, Raghu; ROSS, Kenneth; SHAHABI, Cyrus; SUCIU, Dan; VAITHYANATHAN, Shiv und WIDOM, Jennifer: Challenges and Opportunities with Big Data, Techn. Ber., Computing Community Consortium (2012), URL <http://cra.org/ccc/docs/init/bigdatawhitepaper.pdf>  
(Zitiert auf Seite 1)
- [And06] ANDERSON, Chris: *The Long Tail: Why the Future of Business Is Selling Less of More*, 1401302378, Hyperion (2006)  
(Zitiert auf Seite 1)
- [Bar97] BARDRAM, Jakob E.: I love the system – I just don't use it!, in: *GROUP '97: Proceedings of the international ACM SIGGROUP conference on Supporting group work*, ACM, New York, NY, USA, S. 251–260  
(Zitiert auf Seite 6)
- [Cas11] CASATI, Fabio: How End-User Development Will Save Composition Technologies from Their Continuing Failures, in: MariaFrancesca Costabile; Yvonne Dittrich; Gerhard Fischer und Antonio Piccinno (Herausgeber) *End-User Development*, Bd. 6654 von *Lecture Notes in Computer Science*, Springer Berlin Heidelberg (2011), S. 4–6, URL [http://dx.doi.org/10.1007/978-3-642-21530-8\\_2](http://dx.doi.org/10.1007/978-3-642-21530-8_2)  
(Zitiert auf Seite 2)
- [Ebe04] EBERT, Achim: *Context-Sensitive Visualization*, Dissertation, Universität Kaiserslautern, Fachbereich Informatik (2004)  
(Zitiert auf Seite 2)

- [Fin12] FINK, Elissa: Top 10 Trends for Business Intelligence: Our Outlook on 2013, [http://cdnlarge.tableausoftware.com/sites/default/files/blog/top10\\_12.10.12.pdf](http://cdnlarge.tableausoftware.com/sites/default/files/blog/top10_12.10.12.pdf) (2012), besucht am 13.11.2014  
(Zitiert auf Seite 1)
- [Gra10] GRAMMEL, Lars; TORY, Melanie und STOREY, Margaret-Anne: How Information Visualization Novices Construct Visualizations, in: *Proc. Info Vis 2010*  
(Zitiert auf Seite 1)
- [Gra12] GRAMMEL, Lars: *User Interfaces Supporting Information Visualization Novices in Visualization Construction*, Dissertation, University of Victoria (2012), URL [http://larsgrammel.de/publications/grammel\\_2012\\_phd\\_thesis.pdf](http://larsgrammel.de/publications/grammel_2012_phd_thesis.pdf)  
(Zitiert auf Seite 1)
- [Kei10] KEIM, Daniel A.; KOHLHAMMER, Jörn; ; MANSMANN, Florian; MAY, Thorsten und WANNER, Franz: Visual Analytics, in: Daniel Keim; Jörn Kohlhammer; Geoffrey Ellis und Florian Mansmann (Herausgeber) *Mastering the Information Age - Solving Problems With Visual Analytics*, ISBN 978-3-905673-77-7, Eurographics Association (2010), S. 7–18, URL <http://www.vismaster.eu/news/mastering-the-information-age/>  
(Zitiert auf Seite 1)
- [Key12] KEY, Alicia; HOWE, Bill; PERRY, Daniel und ARAGON, Cecilia: VizDeck: self-organizing dashboards for visual analytics, in: *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, SIGMOD '12, ACM, New York, NY, USA, S. 681–684, URL <http://doi.acm.org/10.1145/2213836.2213931>  
(Zitiert auf Seite 1)
- [Pie12] PIETSCHMANN, Stefan: *Modellgetriebene Entwicklung adaptiver, komponentenbasierter Mashup-Anwendungen*, Dissertation, Technische Universität Dresden (2012), URL <http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa-101559>  
(Zitiert auf Seiten 2, 4, 6 und 7)
- [Rui06] RUIZ, Francisco und HILERA, José R.: Using Ontologies in Software Engineering and Technology, in: Coral Calero; Francisco Ruiz und Mario Piattini (Herausgeber) *Ontologies for Software Engineering and Software Technology*, Springer (2006), S. 49–102  
(Zitiert auf Seite 2)
- [SIN13] SINTEF: Big Data, for better or worse: 90% of world's data generated over last two years., <http://www.sciencedaily.com/releases/2013/05/130522085217.htm> (2013), besucht am 13.11.2014  
(Zitiert auf Seite 1)
- [Wan09] WANG, Xiaoyu; JEONG, Dong Hyun; DOU, Wenwen; LEE, Seok-Won; RIBARSKY, William und CHANG, Remco: Defining and applying knowledge conversion processes to a visual analytics system. *Computers & Graphics* (2009), Bd. 33(5): S. 616 – 623, URL <http://www.sciencedirect.com/science/article/B6TYG-4WS2HX2-2/2/cfee587700dd22e2c170a87ebce42ddd>  
(Zitiert auf Seite 2)