

# A System Architecture for Knowledge Exchange in the Industrial Domain

Fabian Quint<sup>1</sup>, Jörn Kreutel<sup>2</sup>, Frieder Loch<sup>1</sup>, Magnus Volkmer<sup>3</sup>, Peter Pollmanns<sup>4</sup>

Innovative Factory Systems, German Research Center for Artificial Intelligence (DFKI)<sup>1</sup>  
Department of Informatics and Media, Beuth University of Applied Sciences Berlin<sup>2</sup>  
Business Development & Innovation, Infoman AG<sup>3</sup>  
Software and Project Development, SemanticEdge GmbH<sup>4</sup>

## Abstract

Exchange and documentation of knowledge is vital in the industrial environment. The access to knowledge of colleagues can be essential for fulfilling several tasks. Just as well, workers can make experiences that are valuable for others. Within the project AmbiWise a collaboration system including adaptive and multimodal user interfaces for mobile interaction is developed to make knowledge easier accessible in companies. This paper presents the conception and implementation of the developed system architecture.

## 1 Introduction and Motivation

In daily work life several situations arise, where colleagues require knowledge of other colleagues to successfully fulfill their tasks, or made experiences by themselves informative and valuable for the colleagues. Especially, experts' knowledge of complex technical tasks is a valuable resource of today's companies and the documentation and exchange of this knowledge needs to be possible in an easy, seamless and non-disruptive manner. Within the project AmbiWise, new assistance and collaboration systems are developed which allow a culture of exchange and participation in companies. The technological basis for an improved accessibility to knowledge resources is formed by intuitive, adaptive and multimodal user interfaces (UI) which are developed for a mobile and context-sensitive interaction. This paper presents the concept of such a system, the requirements towards it and the resulting system architecture.

In the production domain the knowledge of standardized, planned processes is formally documented. Nevertheless the groundbreaking, situational knowledge is usually situated in the heads of the employees (Thom & Sollberger 2008). Especially, implicit knowledge, as described by (Polanyi 1985), which is gained by longtime employment is hard to verbalize

and is hence often missed capturing by traditional tools. For a general exchange of knowledge between employees, collaborative platforms based on Web 2.0 technologies have proved to be of value in several application scenarios (Stocker & Tochtermann 2011). Nevertheless, they have only shown limited success externalizing implicit knowledge, although this is from special interest for several scenarios (Probst et al. 2010). Allowing this kind of knowledge exchange in a production environment requires mobile, context-sensitive and bidirectional access to knowledge to provide on spot assistance to the worker. By using appropriate assistance modalities, production workers can have immediate access to companies' and colleagues' knowledge and transfer knowledge to the system by themselves.

## 2 Application Scenarios

A comprehensive requirement analysis is necessary as the basis to design an appropriate system. This analysis should involve all substantial stakeholders and their real working environment. The conducted analysis is described by (Weber et al. 2014). Two main scenarios were identified where the system should be applied to. These scenarios given by the industrial end users Daimler and Schaeffler, are described in section 3.1 and 3.2.

### 2.1 Assembly Processes

For Daimler the AmbiWise system is used to manage standardized procedures in assembly processes. The first task is to support audits of production stations based on these standards. These audits of production stations are conducted in a mobile way. At the beginning, the relevant documents are retrieved based on context information or manual selection. Then, the audit is conducted and deviations from the standard are identified, discussed and stored in the system. The second task is to provide mobile access to the latest production standards for qualification. This allows the personnel to examine standards and check if their knowledge is up to date or if they have to improve their knowledge. Learn instructors who manage individual qualification can use these documents to conduct training on spot. The third task is to adapt production standards. This happens, for example, when improvements were identified during an audit. These adaptations (e.g., editing text, updating pictures or adding videos) are created and transferred to the system on spot. Thereby, the editorial effort is minimized compared to handling these tasks on stationary computers.

### 2.2 Maintenance

For Schaeffler the AmbiWise system is used to document maintenance processes in a global production environment. The exchange and conservation of knowledge of maintenance standards is facilitated over different production sites. Primarily, experts' knowledge of complex maintenance tasks should be gathered to establish an extensive knowledge repository. Video recording is applied to document this knowledge. Maintenance personnel are thus able to document and explain procedures at the point of action with the least possible distraction from the actual work. To achieve this, the documented task is

automatically sent to the knowledge repository and processed there. Furthermore, the externalization of implicit knowledge is enabled, since it directly affects the adaption of the standard.

### 3 Concept

An overall concept was developed at the beginning of the project. The concept abstracts from the concrete scenarios and the results of the requirements analysis. The concept, as shown in figure 1, illustrates how an UI which is adapted to the available context (e.g. user, location, role, etc.) is used to retrieve or store knowledge from or to a knowledge repository. This repository also integrates existing sources, not yet enabled for collaboration and exchange in a production environment. The UI offers different modalities like speech, multi-touch, videos and pictures, to represent the knowledge in an appropriate form.

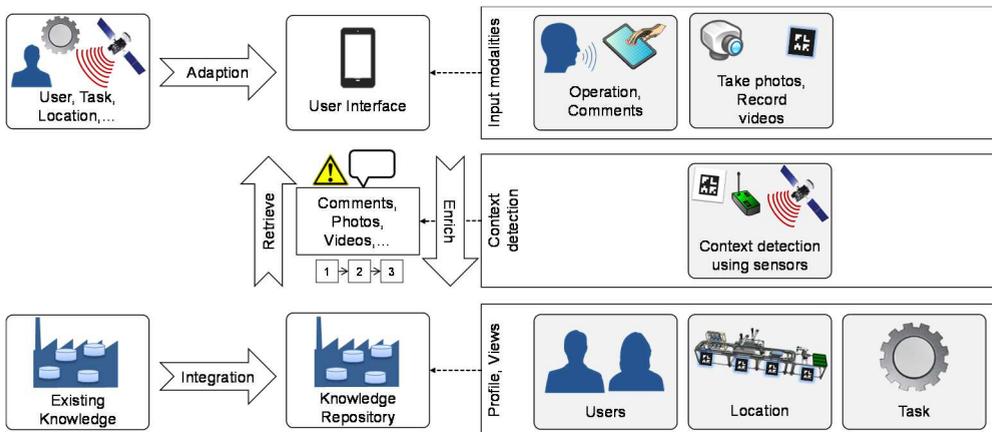


Figure 1: Concept

Information about the context is essential to access and store knowledge effectively. Several technologies are available which need to be chosen and adapted towards the respective scenarios. The variety of available sensors of mobile devices provides many possibilities to detect the context (e.g. camera for recording videos, scanning barcodes, microphone for speech recognition and recording). These technical capabilities allow both knowledge and context to be gathered immediately on spot. The presented concept supports different hardware platforms comprising for example tablets and glasses with different operating systems. Besides choosing the appropriate hardware, the software components which enable the different in and output modalities need to be designed according to the requirements as well. These software components are encapsulated, so they can be aggregated for different variants. Thereby, both scenarios can be covered with one system design although they have different requirements and task perspectives. The requirement and design of these encapsulated services is described in the following chapter.

## 4 System Design

Based on the overall concept a system design is developed hereafter. The considered requirements are discussed in section 4.1 to be met by the system architecture shown in section 4.2.

### 4.1 Requirements

The system design comprises a functional decomposition to address the requirements of the industry stakeholders and the projects' objective of creating a reusable and extensible system architecture. The system functionality is decomposed into particular, yet abstract functions that form the core of the system and address the requirements of the application scenarios. The architecture uses primary functions which are complemented by secondary functions that address cross-cutting concerns relevant for all primary functions. As shown in figure 2, primary functions include read and write access to the particular content, content search, and functions that allow communication between users (e.g. providing a workflow to all stakeholders that participate in the process of creating and releasing content). Secondary functions include for example context detection and speech recognition. A common foundation to all functions is provided by a shared model of the respective application scenario that is based on generic concepts of the production domain. It serves as the basis for content creation and content access, for context detection and for the recognition and interpretation of spoken input.

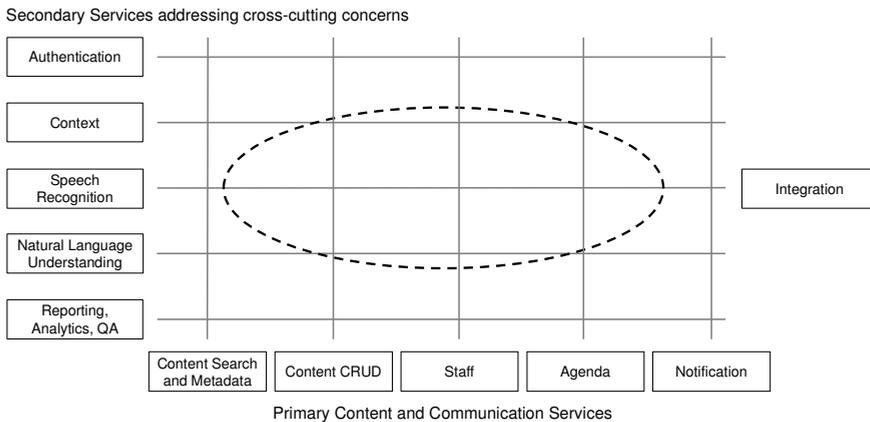


Figure 2: Decomposition of system functionality into primary services for content and communication (bottom) and services that address cross-cutting concerns (left)

Another requirement is to address a wide range of target platforms for running the mobile UI. Platforms range from legacy Windows 7 tablet computers with stylus input to mobile devices based on Windows 8. Hence, development of a single native mobile application for the most recent Windows platform version would neither meet the requirement to access the AmbiWise system from less recent devices, nor would it allow the seamless portability of the UI to alternative mobile platforms like Android or iOS.

## 4.2 Service Architecture

The software architecture adheres to the proposed decomposition of system functionality. Each functional aspect described above is modeled as a service for which a particular service interface can be identified. Service interfaces hide the actual implementation of a service from its users and allow to realize the overall system based on smaller functional units. These may be operated in a distributed way over various locations, rather than being assembled to a single monolithic runtime system. Interface design adheres to the principles of RESTful service interfaces (Fielding 2000). As for the overall service architecture, recently the paradigm of microservices (Fowler & Lewis 2014) has been positioned as an evolution of the last decade's paradigm of Services Oriented Architecture (SOA) (Arsanjani 2004) for realizing distributed systems. With regard to this distinction, the system employs a vertical decomposition of system functionality as proposed by the microservices approach, differentiating services that make up the primary system functionality. However, while aiming at minimizing dependencies between services, it assumes that an integration service acts as an orchestration layer that provides the functionality of the specified service interfaces to the layer of the UI. The integration service mediates between the UI and an underlying layer of more basic services that implement the actual backend functionality, e.g. using a particular enterprise content management system like Microsoft Sharepoint (Microsoft). This way, the UI components stay agnostic with respect to the particular backend systems which are employed in a particular deployment of the AmbiWise solution. Thus the project's idea of providing a reusable system architecture that can be integrated flexibly into various IT infrastructures is addressed.

All service interactions are mediated by a component that can be seen as a lightweight realization of what was foreseen as a "Service Broker" in SOA. It provides bindings of descriptions of the services to be called to particular service endpoints. Service description comprises, among others, a unique identifier of the abstract service interface and optional additional information regarding, e.g. the tenant for which the service is called. Service bindings contain, most importantly, the URL of the service endpoint and may include information regarding whether service calls require authentication or not. In the AmbiWise project the two industry partners Daimler and Schaeffler are modelled as tenants, where each tenant provides individual endpoints for the primary system functions to a centralized integration service. Once a service binding is available for a given service description, the actual service interaction will then be carried out directly between the service consumer and the given endpoint as a service provider, rather than using a centralized, yet usually heavy weight, enterprise service bus (ESB, see (Chappell 2004)) for communication.<sup>1</sup>

Given the idea of a distributed system architecture and the fact that content provided by the AmbiWise industry partners must be treated as confidential, one crucial issue for the system design is how service interaction can be secured. This should be realized in a generic way

---

<sup>1</sup> In a certain way, however, the mediation role of the integration service may be seen as functionally comparable to an ESB.

without each service having to implement its own mechanism for authenticating and authorizing users. Therefore, AmbiWise draws upon the OAuth 2.0 standard (Hardt 2012) that is widely used in consumer applications for “outsourcing” the authentication task to social network authentication services, like the ones of Facebook, Twitter or Google. In AmbiWise, authentication and authorization is realized by the cloud-based OAuth service of Microsoft Azure (Microsoft). There, access rights for both end users and services that may act in the name of an end user are managed. At runtime, each service call requires obtaining an access token from the authentication service. This is either done on the basis of the token that was received on end user authentication or using a token that was issued to a previous service call and that encodes the information about the actual end user. Access to an actual backend system involves a chain of such delegated authentication steps, starting at the UI, and allows the backend to finally verify whether the end user actually has the permission to access some content item or not.

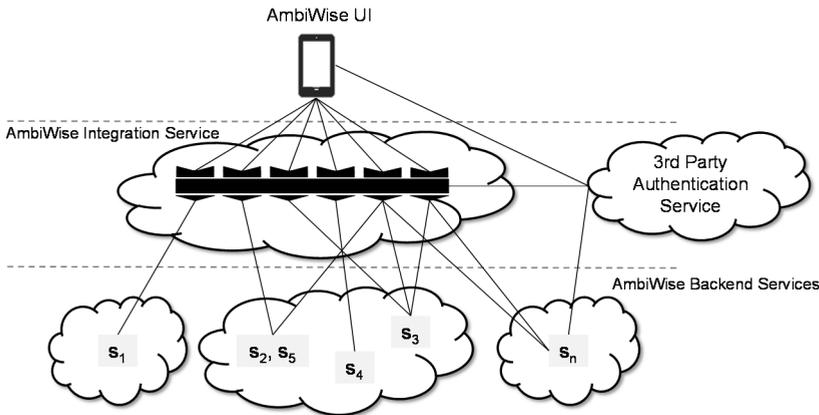


Figure 3: Overview of the AmbiWise system architecture using an integration service as mediator between UI and backend services and a 3<sup>rd</sup> party service for authentication and authorization

The decomposition of system functionality supports various deployment scenarios, ranging from on-premise solutions to mere cloud-based solutions and including hybrid variants in between. It does not only prove convenient for the organization of development activities, but also allows to address infrastructure circumstances and requirements of a wide range of potential application scenarios beyond the particular scope of the project.

### 4.3 User Interface Architecture

The realization of the UI has to take into account that various devices may be employed. For that reason the architecture foresees a two-layered approach where the core functionality of the UI is realized as a “fat client” web application that uses the system’s integration service for accessing backend functions. This application can be run on any recent browser including the versions of Microsoft Internet Explorer that are available for the legacy Windows 7 platform. It supports realizing all tasks of the different scenarios the system is designed for. For addressing the advanced UI functions, including context detection and speech

recognition and interpretation, this web application is wrapped into a native application for the Windows 8 platform, which supports the respective UI functionality. The wrapper interoperates with the web application using the provided interfaces of the native API for invoking JavaScript functions. For example, the generic mechanism for obtaining an access token for a service call is implemented by the web application and may be called by the wrapper, which will use the obtained token for calling the respective service. The implementation of the wrapper on the basis of a single native API allows to realize advanced UI features without being compromised by API restrictions imposed by cross platform frameworks like Cordova (Apache 2015) or Xamarin (Xamarin 2015). Thinking ahead, porting the UI to other platforms will allow to reuse the existing web application and will at the most require an implementation of the native wrapper for the target platform.

## 5 Summary & Outlook

This paper presented the architecture of a collaboration system targeting knowledge exchange and participation for the industrial domain. It gave an overview on the application scenarios which the system should facilitate. Based on a general concept, abstracting from the concrete scenarios, the considered requirements were summarized. The further derivation of the developed system architecture was shown. The motivation and benefits of the used service oriented architecture were discussed. The encapsulation of functionalities to services and their interaction was illustrated. Finally, the architecture of the UI designed to run on several operating systems and devices was introduced.

The current status of the system gave a proof of concept and will in near future be evaluated under real conditions for the two application scenarios. This evaluation ends the first phase of the AmbiWise project. In the second phase the focus is set on further UI modalities, which will enhance the system functionality and usability.

## 6 Bibliography

Apache (2015). *Cordova*. Available online at <https://cordova.apache.org>, checked on 6/17/2015.

Arsanjani, A. (2004). *Service-oriented modeling and architecture*. In *IBM developer works*, pp. 1–15.

Chappell, D. (2004). *Enterprise service bus*: O'Reilly Media, Inc.

Fielding, R. T. (2000). *Architectural styles and the design of network-based software architectures*. University of California, Irvine.

Fowler, M. & Lewis, J. (2014). *Microservices*. Available online at <http://martinfowler.com/articles/microservices.html>, checked on 6/17/2015.

Hardt, D. (2012). *The OAuth 2.0 authorization framework*. Available online at <http://tools.ietf.org/html/rfc6749.html>, checked on 6/17/2015.

Microsoft (2015). *Azure*. Available online at <https://azure.microsoft.com>, checked on 6/17/2015.

Microsoft (2015). *Sharepoint*. Available online at <https://products.office.com/en-us/sharepoint/collaboration>, checked on 6/17/2015.

Polanyi, M. (1985). *Implizites wissen*: Suhrkamp.

Probst, G., Raub, S. & Romhardt, K. (2010). *Wissen managen*. Wie Unternehmen ihre wertvollste Ressource optimal nutzen. 6., überarbeitete und erweiterte Auflage. Wiesbaden: Gabler Verlag / GWV Fachverlage GmbH Wiesbaden.

Stocker, A., Tochtermann, K. (2011). *Wissenstransfer mit Wikis und Weblogs: Fallstudien zum erfolgreichen Einsatz von Web 2.0 in Unternehmen*: Springer-Verlag.

Thom, N., Sollberger, B. A. (2008). *Das Zusammenspiel von Mensch, Kultur und Technologie für ein erfolgreiches Wissensmanagement*. In *Quo vadis Wirtschaftsinformatik?*, pp. 323–333.

Weber, H., Quint, F., Eierdanz, F., Ottersböck, N. & Gorecky, D. (2014). *Nutzerzentrierte Gestaltung von mobilen, kontext-sensitiven Assistenzsystemen für einen vereinfachten Wissensaustausch in komplexen Produktionszenarien*. 7. VDI/VDE Fachtagung USEWARE 2014. In *VDI-Berichte 2222*, pp. 133–143.

Xamarin (2015). *Xamarin*. Available online at <http://xamarin.com>, checked on 6/17/2015.

### **Acknowledgement**

The research described in this paper was conducted within the project AmbiWise funded by the German Ministry of Education and Research under grant number V4ISS035.

### **Contact Information**

Fabian Quint  
fabian.quint@dfki.de  
German Research Center for Artificial Intelligence  
Trippstadter Straße 122  
67663 Kaiserslautern

Joern Kreutel  
joern.kreutel@beuth-hochschule.de  
Beuth University of Applied Sciences  
Luxemburger Straße 10  
13353 Berlin

Frieder Loch  
frieder.loch@dfki.de  
German Research Center for Artificial Intelligence  
Trippstadter Straße 122  
67663 Kaiserslautern

Magnus Volkmer  
magnus.volkmer@infoman.de  
Infoman AG  
Meitnerstraße 10  
70563 Stuttgart

Peter Pollmanns  
peter.pollmanns@semanticedge.de  
SemanticEdge GmbH  
Kaiserin-Augusta-Allee 10-11  
10553 Berlin

# Multimodale Interaktion in der Produktionsstätte der Zukunft

Hermann Fürntratt<sup>1</sup>, Ferdinand Fuhrmann<sup>2</sup>

Intelligente Informationssysteme, DIGITAL, JOANNEUM RESEARCH<sup>1</sup>  
Intelligente Akustische Lösungen, DIGITAL, JOANNEUM RESEARCH<sup>2</sup>

## Zusammenfassung

In diesem Beitrag stellen wir ein neuartiges multimodales Interaktionskonzept für die Produktionsstätte der Zukunft vor. Ziel ist es den Beschäftigten in seiner Interaktion mit Menschen und Geräten sowohl physisch als auch kognitiv bestmöglich zu unterstützen. Ausgehend von Kontrollräumen, mit ihren besonderen Anforderungen an die Operatoren, erweitern wir unser Interaktionskonzept auf die gesamte Produktionshalle und zeigen anhand eines umgesetzten Prototyps die technische Machbarkeit. Probleme und Herausforderungen bei der Umsetzung, sowie Verbesserungsmöglichkeiten werden im Anschluss diskutiert.

## 1 Einleitung

Multimodale Interaktion gewinnt im Produktionsumfeld immer mehr an Bedeutung. Sie ermöglicht den wechselseitigen Informationsaustausch zwischen Mensch und Computer über natürliche Arten der menschlichen Kommunikation (Ratzka 2008). Die klassische Mensch-Maschine Schnittstelle – Maus und Tastatur – stößt hingegen immer mehr an ihre Grenzen. Wir sitzen zu lange tippend vor den Bildschirmen und obwohl es Versuche gibt, dem entgegen zu wirken (z.B. durch den Einsatz höhenverstellbarer Möbel), ist vorerst keine Besserung in Sicht. Mediziner warnen inzwischen eindringlich vor den gesundheitlichen Folgen dieser Entwicklung (Buckley et al. 2015).

Zusätzlich nimmt auch die kognitive Belastung der Arbeiter immer mehr zu (Spath et al. 2013). Am Beispiel eines industriellen Leitstandes etwa zeigt sich, dass die Datenmenge, die von den Operatoren überwacht wird und auf die im Störfall in kürzest möglicher Zeit reagiert werden muss, rapide gestiegen ist. Dazu kommt, dass die Benutzerschnittstellen für die Beobachtung und Bearbeitung dieser Daten in den seltensten Fällen aufeinander abgestimmt sind. Da der Datenstrom meist aus unterschiedlichsten Teilsystemen stammt, die aufgrund wirtschaftlicher und technischer Überlegungen meist stufenweise dem laufenden Gesamtsystem hinzugefügt wurden, verliert das Gesamtsystem seine ursprünglich

angedachte Intuitivität und wird über die Maßen komplex. Der Mensch steht nicht mehr im Mittelpunkt, sondern der Wettstreit der Systeme. Diese Tatsache lässt sich auf den Schreibtischen unschwer anhand der Vielzahl von Tastaturen und Mäusen – je ein Paar pro Teilsystem – nachvollziehen.

In der vorliegenden Arbeit versuchen wir daher Antworten auf folgende Fragen zu geben:

1. Wie kann ich den Arbeiter im Produktionsprozess wieder in den Mittelpunkt rücken, ihn körperlich und kognitiv entlasten?
2. Wie kann ich von einem bestehenden zu einem moderneren und intuitiveren Produktionssystem übergehen, ohne dass die Kosten für den Umbau den wirtschaftlichen Rahmen sprengen?

Im Folgenden wird das entwickelte Interaktionskonzept anhand des Beispiels Industrielitstand vorgestellt und mithilfe eines implementierten Prototyps die technischen Möglichkeiten sowie Usability Aspekte beleuchtet (Abschnitt 2). In Abschnitt 3 erweitern wir anschließend dieses Konzept auf die gesamte Produktion. In Abschnitt 4 werden schließlich die wichtigsten Erkenntnisse zusammengefasst und die nächsten Schritte diskutiert.

## 2 Interaktionskonzept

Ausgangspunkt für unsere Überlegungen ist der Leitstand, in dem die Kommunikation mit anderen Menschen, die Abfrage von Statusinformationen, das Steuern von Prozesszuständen, sowie das schnelle Reagieren in Alarmsituationen die wichtigsten Aufgaben sind.

### **Das Beispiel des Leitstandes**

Das entwickelte Gesamtinteraktionskonzept ist von der Vision geprägt, dass der Mensch mehr Freiraum in der Bedienung der Vielzahl an technischen Geräten im Leitstand erhalten soll. Durch die daraus resultierende Multimodalität und Intuitivität der Interaktion wird er sowohl physisch als auch kognitiv entlastet. In Zukunft agiert der Anwender (Operator) nicht mehr lokal über Maus, Tastatur und Telefon vor dem Bildschirm, sondern mobil und erweitert seine Interaktionsmöglichkeiten durch den Einsatz technischer Systeme, die ihm akustisch und visuell im Raum folgen.

Das hier entwickelte Konzept basiert auf den folgenden Interaktionsmechanismen:

- Gesteninteraktion – berührungsbasierte und berührungslose Sensoren ermöglichen eine intuitive und unmittelbare Steuerung des Systems und schränken den Menschen in seiner Mobilität nicht ein.
- Visuelle Interaktion – die dem Menschen übermittelte visuelle Information wird kontextsensitiv aufgearbeitet und auf das wesentliche reduziert.
- Akustische Interaktion – spezielle Technologien erlauben die Fokussierung des akustischen Signals auf bestimmte Zielpersonen sowie eine räumliche Kodierung der Information.



Abbildung 1: Gesamtübersicht eines möglichen zukünftigen Leitstandes.



Abbildung 2: Demonstrator-Setup mit gerichteten Lautsprechern und Operatorsessel.

Abbildung 1 zeigt die Gesamtübersicht des Leitstandes der Zukunft, in dem die Operatoren den Großteil ihrer Routinearbeit unterstützt von mehreren unterschiedlichen Sensoren erledigen. Alle wesentlichen Systemkomponenten sind in der Abbildung grafisch hervorgehoben. Die Gesteninteraktion folgt dem Grundprinzip: *Zeige mit einer Hand auf das, was du tun willst, und steure mit der anderen Hand, wie du es tun willst*. Dieses Prinzip wird sowohl für Armgesten (Makrogestik), welche über weitreichende Tiefensensoren eine Interaktion mit externen Geräten ermöglichen, als auch für Handgesten (Mikrogestik), die über im Sessel eingebaute Sensoren verarbeitet werden, angewandt. Ziel der Mikrogestik ist, dass der Operator in möglichst entspannter Haltung mittels kleinster Bewegungen der Finger das System steuert. Durch einen Drehsensor (IMU) im Operatorsessel lässt sich die aktuelle Zeigerichtung ermitteln und mit dem gespeicherten geometrischen Modell des Raumes und seiner darin enthaltenen Objekte abgleichen. Um Geräte steuern zu können, die keine visuelle oder akustische Rückmeldung geben können, sind Vibrationselemente als taktile Feedbackgeber in den Operatorsesseln eingebaut. Für das Manipulieren von Prozessvariablen steht dem Operator außerdem ein Multi-Touch Ein-/Ausgabegerät zur Verfügung, das auch als mobile Anzeige und Steuerung verwendbar ist.

Ein Lautsprechersystem, welches stark gerichteten Schall abstrahlen kann, versorgt den Operator mit akustischen Informationen. Um andere sich im Raum befindliche Personen nicht zu stören, werden die Lautsprecher den Ohren des Operators automatisch nachgeführt. Dafür wird das Head-tracking eines Tiefensensors zur Ansteuerung von zwei-motorigen Nachführungseinheiten verwendet. Weiters wird die abgestrahlte akustische Information über virtuelle Schallquellen räumlich positioniert. Dadurch ist es möglich, dem aus den Lautsprechern wiedergegebenen Schall eine räumliche Information zu geben. Daraus ergibt sich der für die Wahrnehmungspsychologie wichtiger Effekt, dass bei örtlicher Übereinstimmung von visueller und akustischer Information die menschliche Aufmerksamkeit ungleich höher ist als ohne solche Übereinstimmung (Schönhammer 2013). Analog dazu ist es dem Operator möglich, über ein Mikrofon-Array, das den Standort des Operators ermittelt und sein Sprachsignal automatisch aus anderen Geräuschquellen herausfiltert, mit anderen Gesprächspartnern zu kommunizieren.

Die Werkzeuge zum Manipulieren von Prozessvariablen werden kontextabhängig, je nach Anwendungsfall und Berechtigungsstufe des Operators, zur Verfügung gestellt. Dadurch

lässt die visuelle Information kompakt und intuitiv darstellen. Alle eingehenden Sensorinformationen durchlaufen außerdem einen intelligenten Vorverarbeitungsschritt. Dabei können kontextabhängig irrelevante Informationen gefiltert, unwichtige Meldungen verzögert oder zusammenhängende Hinweise zusammengefasst werden. Diese Reduktion an Daten führt zu einer weiteren kognitiven Entlastung des Operators.

### **Implementierung**

Für Demonstrations- und Evaluationszwecke wurde das entwickelte Interaktionskonzept für den Anwendungsfall Verkehrsleitstand in einem Prototypen implementiert. Dabei nimmt eine zentrale Serverkomponente alle Sensorströme in Echtzeit über LAN/WLAN entgegen. Die weiteren Steuermodule, etwa für die Maustreiber, die 3D-Raumkonfiguration, die grafische Anzeige von Daten, oder für eine Klanguasgabe sind verteilt installiert und bilden ein lose gekoppeltes Client-Server-Netzwerk. Abbildung 2 zeigt ein Bild des implementierten Arbeitsplatzes und eines sensorbestückten Operatorsessels.

### **Evaluation**

Die Interaktionsmechanismen sowie das gesamte Interaktionssystem wurden in zwei Benutzerstudien mit 24 bzw. 6 Teilnehmern evaluiert. Ziel war die Bewertung der einzelnen Ein-/Ausgabegeräte sowie des Interaktionskonzeptes für den Einsatz in einem Leitstand.

Im ersten Teil wurde über quantitative (Bearbeitungszeit) und qualitative (Interviews) Methoden die Bedienfreundlichkeit der einzelnen Benutzerschnittstellen evaluiert. Dabei erzielten vor allem die mobilen und touch-basierten Sensoren (Tablet und TrackPads) sehr gute Ergebnisse. Grund dafür war einerseits die direkte und intuitive Bedienung dieser Geräte, andererseits gaben fast alle Teilnehmer der Studie an, mit diesen Technologien bereits viel Erfahrung zu haben. Die berührungslosen Sensoren dagegen schnitten in den Bewertungen schlechter ab, obwohl die Teilnehmer durchwegs beeindruckt von der Unmittelbarkeit und gewonnenen Bewegungsfreiheit der Interaktion waren. Die Gründe dafür lagen einerseits in einer gewissen Berührungsangst vor der neuen Technologie, andererseits lies auch die Robustheit in der Gestenerkennung noch zu wünschen übrig.

Der zweite Teil dieser Studie umfasste die Evaluation des gesamten Interaktionssystems für den Einsatz in einem Leitstand. Hierfür wurden ausführliche Tests und Befragungen mit Operatoren einer Verkehrsleitzentrale durchgeführt. Besonders hervorgehoben wurde hier die mobile Variante des Leitstandes über den Tablet Computer, welche den Operatoren Bewegungsfreiheit während ihrer langen Schichtzeiten ermöglicht. Weiters wurde der Sensor-bestückte Operatorsessel sehr positiv bewertet. Mit ihm ist es den Operatoren möglich, verschiedene Komponenten ihres Leitstandes mit einer einzigen multimodalen Benutzerschnittstelle zu steuern. Besonders gut schnitt auch die Integration verschiedener Systeme in ein ganzheitliches Interaktionssystem ab. Dadurch wird mühsames Wechseln zwischen den einzelnen Systemen obsolet, was zur Optimierung von Arbeitsabläufen und zur Steigerung der Ergonomie führt. Schließlich wurde auch das akustische Interface mit seinen nachgeführten Lautsprechern und dem Mikrofonarray als sehr nützlich für den Leitstand eingestuft, da dadurch Mobilität und Ergonomie erhöht werden.

### 3 Die Produktion der Zukunft

War die konzeptionelle Umsetzung im Leitstand hauptsächlich auf das Steuern von Computeranwendungen und die freie Kommunikation ohne Headset ausgerichtet, lässt sich das vorgestellte Konzept der multimodalen und mobilen Interaktion auf die gesamte Produktion ausdehnen. In einer vernetzten und mit Sensoren verschiedenster Art bestückten Produktionsstätte kann sich der Mensch frei bewegen und mit den Geräten und Menschen in seiner Umgebung direkt interagieren.

Der Arbeiter kann über das Prinzip des Zeigens und Steuerns via Makro- und Mikrogestik eine Vielzahl verschiedenartigster Geräte im gesamten Produktionsumfeld bedienen. Er zeigt beispielsweise mit einer Hand auf die Beleuchtung und reguliert mit der anderen die Helligkeit in seiner Umgebung. In der Produktionshalle zeigt der Monteur auf den Industrieroboter und versetzt ihn mit einer Geste in den Wartungsmodus.

Auch die visuelle Interaktion kann hier angewandt werden. So wird auf dem mobilen Anzeigergerät (z.B. Tablet) immer die Information angezeigt die für den Menschen relevant ist – unnötige Informationen werden ausgefiltert. Befindet er sich direkt vor der Maschine, erhält er Manipulationswerkzeuge zum Steuern oder Warten. Demgegenüber lassen sich Statusinformationen von der gesamten Produktionsstrecke anzeigen, sollte er sich nicht in der Nähe einer Maschine befinden. Die Gesteninteraktion lässt sich mit der visuellen Interaktion in Form von *Augmented Reality* kombinieren: der Stahlarbeiter zeigt auf einen Kessel und bekommt durch eine Touchgeste die Statusinformation über diesen übersichtlich auf seinem Tablet präsentiert.

Schließlich lässt sich auch die akustische Schnittstelle in die Produktion der Zukunft integrieren. Über mehrere nachführbare Lautsprecher kann akustische Information zielgerichtet im gesamten Produktionsbereich an bestimmte Personen abgegeben werden. Weiters können über in der Produktionsstätte installierte Mikrofonarrays Sprachsignale von Menschen aufgefangen werden und so eine Kommunikation ohne tragbare Geräte ermöglicht werden. Da in vielen Produktionsstätten der Lärm von Maschinen diese Art der Kommunikation allerdings verhindert, ist ein Rückgriff auf kabellose Headsets problemlos möglich. So können auch in diesem Fall akustische Warnsignale räumlich kodiert werden sodass Richtungsinformation an den Menschen übermittelt wird. Eine Verbindung zwischen Headset und Gesteninteraktion ist ebenfalls sinnvoll: Der Techniker zeigt während der Störungsbehebung auf das rote Telefon an der Wand und baut mit einer Geste eine Telefonverbindung – ohne das bestehende Telefon zu verwenden – zum Kontrollzentrum auf.

### 4 Diskussion und Ausblick

Aus der Evaluation der Interaktionsmechanismen zeigt sich, dass Touch-basierte Geräte von den Testkandidaten anstandslos akzeptiert werden, während für die berührungslose Gesteninteraktion über Tiefensensoren für die Akzeptanz bei den Arbeitern zwei Voraussetzungen erfüllt sein müssen:

1. Ausreichende Robustheit der Gestenerkennung
2. Energieeffiziente intuitive Interaktion

Die Hardware-Entwicklung der Sensoren (z.B. höchstgenauer Tiefensensoren) schreitet zügig voran (Soli 2015). Die softwareseitige Algorithmetik zur robusten Erkennung von Körperbewegungen hinkt derzeit noch hinterher. Auch werden menschliche Eigenheiten im Zusammenhang mit HCI, wie etwa motorische Schwächen, Tremor, oder die Tatsache, dass wir eigentlich nicht dorthin zeigen, wo wir glauben hinzuzeigen (Mayer 2015), noch nicht ausreichend berücksichtigt.

Für den Einsatz in der Produktionsstätte jedoch, hat das vorgestellte Interaktionskonzept unserer Meinung nach großes Potential. Alle Funktionalitäten lassen sich mit geringem Aufwand in bestehende Infrastrukturen integrieren. Darüber hinaus ergibt sich durch den Aufbau eines Sensornetzwerkes, mit dem Arbeiter als menschliches Eingabemedium im Mittelpunkt die Möglichkeit, obsoleete Geräte sanft aus dem Betrieb herauszunehmen und generationsbedingt neue Arbeitsweisen zu etablieren.

### Literaturverzeichnis

Buckley, J. P., Hedge, A., Yates, T., Copeland, R. J., Loosemore, M., Hamer, M., Bradley, G. & Dunstan, D. W. (2015). *The sedentary office: a growing case for change towards better health and productivity. Expert statement commissioned by Public Health England and the Active Working Community Interest Company*. Br J Sports Med doi: 10.1136/bjsports-2015-094618.

Google ATAP: Project Soli (2015). Radar based interaction sensor for gesture recognition <https://plus.google.com/+GoogleATAP/videos> (zuletzt abgerufen am 22. 6. 2015)

Mayer, S., Wolf, K., Schneegass, S. & Henze, N. (2015). Modeling Distant Pointing for Compensating Systematic Displacements. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 4165-4168.

Ratzka, A. (2008). *Patterns for Robust and Flexible Multimodal Interaktion*. EuroPLoP, volume 610 of CEUR Workshop Proceedings, CEUR-WS.org.

Schönhammer, R. (Hrsg.) (2013). *Einführung in die Wahrnehmungspsychologie: Sinne, Körper, Bewegung. 2. Auflage*. Wien: facultas wuv, Universitätsverlag.

Spath, D. (Hrsg.), Ganschar, O., Gerlach, S., Hämmerle, M., Krause, T. & Schlund, S. (2013). *Produktionsarbeit der Zukunft – Industrie 4.0*. Stuttgart: Fraunhofer IAO.

### Kontaktinformationen

Hermann Fürntratt

Ferdinand Fuhrmann

JOANNEUM RESEARCH  
Intelligente Informationssysteme  
Leonhardstrasse 59  
8010 Graz

JOANNEUM RESEARCH  
Intelligente Akustische Lösungen  
Leonhardstrasse 59  
8010 Graz

hermann.fuerntratt@joanneum.at

ferdinand.fuhrmann@joanneum.at