Visual Analytics Supporting Knowledge Management:

A Case Study of Germany's Federal Employment Agency

Eldar Sultanow
Capgemini
Bahnhofstraße 11C, 90402 Nuremberg
Germany
eldar.sultanow@capgemini.com

Marinho Tobolla
IT Dept. of the Federal Employment Agency
Südwestpark 26, 90449 Nuremberg
Germany
marinho.tobolla@arbeitsagentur.de

André Ullrich
University of Potsdam
August-Bebel-Straße 89, 14482 Potsdam
Germany
aullrich@lswi.de

Gergana Vladova
University of Potsdam
August-Bebel-Straße 89, 14482 Potsdam
Germany
gvladova@lswi.de

ABSTRACT

The Federal Employment Agency (FEA) spent seven years developing (and continues to develop) a mission-critical software system responsible for one hundred thousand transactions every day while ensuring the safe processing of more than EUR 25 billion a year. This system comprises more than 718,000 lines of code and the development team consists of approximately 90 developers. As is normal for very large private or public projects, the knowledge of external experts/consultants is involved. As a result, isolated, highly specialized knowledge is developed and not sufficiently shared. That is why an early warning KMS that incorporates visual analytics has been developed from the bottom up as an answer to specific challenges within the very knowledge intensive project. Furthermore, the well-established, structured knowledge management framework of the FEA helps support and establish appropriate activities to meet these challenges. This paper describes the motivation, challenges, specifics, and implementation of this KMS pilot system.

CCS CONCEPTS

• **Computing methodologies → Artificial intelligence**; *Knowledge representation and reasoning*; **Human-centered computing → Visualization;** *Visualization application domains*; Information visualization;

KEYWORDS

Knowledge Management, Public Sector, Mission-critical knowledge, Big Data Visualization, D3.JS, Angular 4, Cassandra

1 INTRODUCTION

Germany's Federal Employment Agency employs more than 96,000 people and operates one of the largest IT-infrastructures in Germany, with more than 160,000 networked PC-workplaces, 9,000 servers, two main data centers, and several regional ones. The huge amount of processed data includes email, bank transfers, mail, and print products.

Over the last seven years, the FEA has developed (and continues to develop) the mission-critical software system ALLEGRO. ALLEGRO stands for "Arbeitslosengeld II Leistungsverfahren Grundsicherung Online" (Unemployment Benefits II Performance Benefits Basic Provision Online). The system is responsible for one hundred thousand transactions a day and ensures the safe processing of more than EUR 25 billion a year. It comprises more than 718,000 lines of code, written by a development team numbering approximately 90. The software is responsible for retrieving, administrating, and processing data; calculating unemployment benefits and periods; payment orders for benefits under the Second Book of the Social Code (SGB II); reporting and payment to social insurance funds; preparing decisions; and central and decentralized printing.

As is normal for very large private or public projects, the knowledge of external experts/consultants is involved. Due to alternating ramp-up and ramp-down phases and personnel changes, knowledge bearers fluctuate and knowledge flows are highly volatile. The situation is characterized by knowledge flow peaks and troughs, and knowledge objects (artifacts, documents, source code, etc.) may be left untouched for periods. Thus, familiarization with these knowledge objects is complicated and time-consuming, lacking the necessary relational and causal knowledge that is or was available as tacit knowledge.

Based on the idea of knowledge as an object (e.g. [1-3]) to be generated, identified, stored, and manipulated, unveiling knowledge drain while identifying and focusing on critical knowledge bearers may help large, long-term projects to reduce development times and costs. To this end, knowledge must be identified in a way that makes it effortlessly identifiable, internalizable, and interpretable for relevant stakeholders such as knowledge and project managers and software developers. On the other hand, there is little value in identifying, generating, and storing massive amounts of information and knowledge on the chance it might be relevant to a project [4]. The right information and knowledge must be identified at the right time, since only useful information will be used to find solutions for present or future challenges [5]. Thus, people and the culture of public sector employees have been identified as key factors for future research on public sector knowledge management [6].

Public sector projects involve many external consultants. Ramp-ups, ramp-downs, and departing consultants cause project relevant knowledge to fluctuate and drain off. This creates a need for an early warning system, which can identify knowledge monopoles, flows, and critical knowledge bearers while providing a way forward. Knowledge management research in the public sector remains limited [7] and should be conducted as special research rather than as a part of private sector management research. Although the aforementioned knowledge types are identical for both the private and the public sectors, public sector knowledge management cannot simply adapt private sector thinking [8]. On the contrary, the practices of sharing and transferring knowledge should be adapted to the specifics of the public sector [9, 10]. In particular, customized knowledge management systems that suit unique bureaucratic hierarchies and cultural features should be developed [11]. Compared to the private sector, the pressure to compete and efforts to cut costs are less important, and knowledge sharing is less evident [10]. Furthermore, public sector organizations differ from private sector organizations in their goals, environment, and political influences [9].

Therefore, this paper aims to present a knowledge management system (KMS) using a case study, highlighting public authorities' knowledge positions and their critical knowledge resources by visualizing the development, atrophy/degeneration, and endangerment of knowledge in such highly business critical projects.

This paper is organized as follows: Section 2 provides theoretical background on knowledge management, knowledge management systems, and challenges during their implementation. Section 3 introduces the methodical approach and Section 4 describes the KMS pilot system. Section 5 illustrates its functionalities within various applications. Section 6 provides an evaluation while Section 7 outlines limitations. The contribution closes with a discussion and outlook (Section 8).

2 Background

Knowledge and information play pivotal roles for both private companies and the public sector. Particularly for the latter, high-volume information transfer, and knowledge and information allocation among diverse administrative units and external partners, present major knowledge management challenges [12]. Knowledge management is the integration of tools and methods that "harnesses the value of knowledge and engages it in processes with people, processes, and organizational infrastructure" [13, 14]. Thus, a promising and sustainable knowledge management framework addresses a range of taxonomical aspects of knowledge to foster its distribution within an organization:

- Tacit knowledge, which is “personal, context specific, and very difficult to communicate” [15];
- Explicit knowledge, which can be distributed in a formal and systematic language;
- Individual knowledge "possessed" by a single entity;
- Collective knowledge, as well goal-oriented transfer and interplay [16];
- Procedural knowledge, which is important in large projects, since it may advise how to handle diverse kinds of challenges;
- Causal knowledge, the "knowing-why"; and
- Relational knowledge, which provides answers about interactions and interdependencies [5].

The individually distributed and "hidden" tacit knowledge of people is especially critical for the successful execution of large projects [17, 18], such as in software development for public authorities.

Knowledge may be viewed from diverse perspectives: as a state of mind, as an object, as a process, as a condition of possessing access to information, or as a capability [5].

The first view understands knowledge as a state of knowing and focuses on enabling individuals to expand and apply personal knowledge [19]. Considering knowledge as an object grasps knowledge as something that can be generated, identified, stored, and manipulated [3]. Here, the role of information technology (IT) involves gathering, storing, and transferring knowledge. Alternately, knowledge can be described as the simultaneous process of knowing and acting [20]. According to this view, knowledge enables acting by applying expertise. The view of knowledge as a condition of access to information focusses on the organization of knowledge and the accessibility of knowledge objects [2]. IT needs to provide effective search, visualization, and retrieval mechanisms. Alternatively, knowledge can be seen as a capability with the potential to influence future actions, emphasizing the capacity to use information [21]. The processing and transfer of explicit and tacit knowledge requires the development of different methods and approaches based on their specific features. Explicit knowledge (or information) can be transferred by communication, by numbers, by pictures, or by language. It can be processed, altered, and learned together [22], [23]. Tacit knowledge, on the other hand, is based on, but not equated with, information and is person-bound and very difficult to articulate.

Procedural knowledge, contextual knowledge (about relevant legal and political aspects and decisions), and content knowledge (about facts and rules) have been highlighted as essential [24, 25] to the public sector in particular. According to [24], knowledge management is shaped by certain features of the public sector, e.g. a focus on savings, high employee turnover, and the public sector’s role as a service provider for citizens and enterprises, whereby the quality of the services depends on the quality of process relevant data and information.

Ihringer [12] identifies the following knowledge management instruments as typical for the public sector: a list of experts, web-based portals, document management, business intelligence, decision support, controlling systems. However, she sees solutions for bringing experts together and supporting collaboration and communication processes as of primary importance. According to Barachini [26], focusing on people is also a key factor and a big challenge for public sector knowledge management future research. One explanation is that individuals generally do not offer knowledge freely. Furthermore, there may be differences in the employee characteristics of private and public sector organizations [27]. Another important point is the resistance encountered in public sector organizations, when attempting to adapt the cultural characteristics of the private sector [28, 6].

Ihringer [12] lists four challenges to the development of appropriate knowledge management solutions for public sector organizations.

1) The consideration of the core knowledge management aspects: technology, people, and organization. Focusing on one aspect alone is not sufficient for proper knowledge management and to gain or sustain competitive advantages. Instead, it is the interaction between technology, people, and techniques that enables effective knowledge management [29]. In addition, the strategy of the organization and public sector-specific goals should be taken into consideration.

2) The establishment and support of networking and semi-structured knowledge transfer activities;

3) The establishment of learning processes; and

4) The development of customized technical solutions: This last point is particularly relevant to this contribution by describing the development of a cloud-based knowledge management system for one public sector authority.

KMS, Cloud Computing and Visual Analytics

Given the complexity and variety of knowledge management aspects in organizations and the huge amounts of data and information involved, information systems are often implemented to support organizational knowledge management with platforms for knowledge exchange, retrieval, storage, usage, and visualization. Such platforms are usually referred to as knowledge management systems (KMS). A KMS is an information and communication system that combines and integrates functions for the structured and contextualized handling of explicit and tacit organizational knowledge [30]. Thus, KMSs are a class of information systems applied to manage organizational knowledge. They are developed to support and enhance the organizational processes of knowledge creation, storage/retrieval, transfer, and application [5]. Their basic functions include content management, information retrieval, visualization and aggregation of knowledge, and collaboration. Because collaboration is often characterized by distributed work with no time limits and cloud computing facilitates scalability, cost-efficiency, availability, and location independence, there is a tendency towards cloud-based KMSs.

Cloud-based KMSs enable information search and retrieval at any time and from any location, as well as knowledge sharing and reuse in distributed environments that are not feasible in many conventional knowledge management approaches [31]. A cloud-based KMS also enables the handling of big data and the application of analyses [32]. Cloud-based KMSs are more effective and user-oriented solutions for organizations. They are generally provided to users as Software as a Service (Saas), whereby the provider uses Infrastructure as a Service (IaaS) to host the cloud-based KMS. The third basic principle of cloud computing (Platform as a Service) allows user-individual customization of the KMS. In addition to these conventional layers, Tsui et al. [31] propose a new service layer named Knowledge-as-a-Service (KaaS), which facilitates the management of personal knowledge, i.e. information retrial, evaluation, and knowledge organization. Cloud computing enables new KMS models, integrating additional systems, collaborating with other organizations, and facilitating knowledge exchange [33]. The cloud offers infrastructure services (e.g. storage and communication), knowledge services (e.g. knowledge creation, sharing, and reuse), and platform services (e.g. databases) [34] and allows knowledge workers to integrate external content via Web 2.0 tools and build up their own knowledge facilitating environment. They can easily access various Cloud service platforms and resources through the Internet to obtain their KM demand [35]. Resources are accessed via user interfaces, which offer intelligent guidance via the knowledge creation process.

Visual analytics is an "iterative process that involves information gathering, data, preprocessing, knowledge representation, interaction, and decision making" [36]. To solve a given problem, "visual analytics combines the strengths of machines with those of humans". In the presented case the problem is the volatility of external knowledge that is critical to ALLEGRO. Our developed pilot will demonstrate the power to resolve this problem by combining both human and machine abilities. Data mining is a key pillar of visual analytics and "automatically extracts valuable information from raw data by means of automatic analysis algorithms" [37, 38]. Cloud computing has created new possibilities for data mining extraction and analytical processes. The processing algorithms are computationally very intensive and have high hardware requirements, which easily can be covered by cloud computing. This explains why visual analytics is increasingly offered as cloud service.

The scope of visual analytics in this project is illustrated by Figure 1. Code Statistics, commits, and reviews provide the basis for statistical analytics. Interaction, and cognitive and perceptual science play a role in using the graphical interface of our KMS, which visualizes knowledge areas and hotspots on the basis of large amounts of data.

Information analytics is used to identify which artifacts (code fragments, documents, confluence articles, etc.) belong to which knowledge area. In the context of knowledge discovery, our pilot identifies bottlenecks, a gap between knowledge holders and seekers, and the need for knowledge management measures. Based on this, our pilot system issues early warnings. Presentation, production, dissemination, and data management also lie within our visual analytics scope, since a major topic addressed by this paper is the presentation and the data layer of our KMS pilot system.

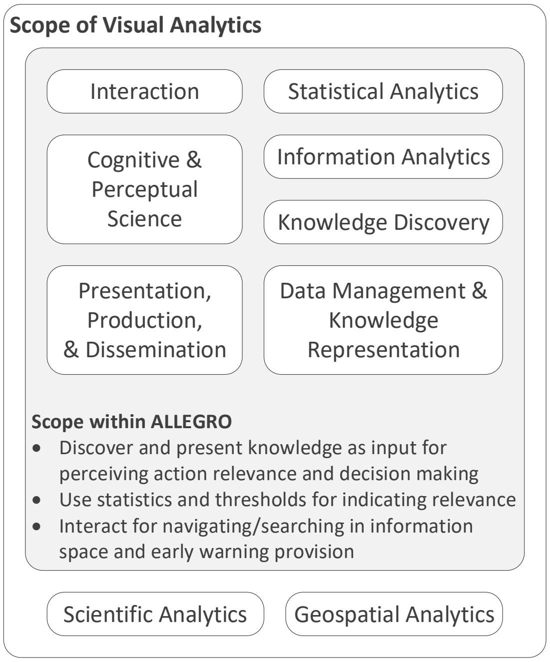


Figure 1: Scope of Visual Analytics [36] applied to ALLEGRO

3 METHODICAL PROCEDURE

As described above, the FEA is responsible for a huge amount of highly relevant critical data and information, while acting as both an employer and a service provider. Against this background, an appropriate knowledge management framework has been established. This provides a well-structured approach for the identification of knowledge management challenges and the development of best possible strategies and concepts.

This case study represents an interesting knowledge management phenomenon. In order to develop a way of dealing with huge amounts of specific information and data (as one knowledge management topic), the FEA started developing the ALLEGRO core system. However, given the project specifics (e.g. the importance of tacit knowledge and collaboration, relevance, and fluctuation of external experts), important new knowledge management challenges have been identified – management of ALLEGRO-specific knowledge and information, and ensuring the success of this very knowledge intensive software development process.

As depicted in Figure 2, the methodological procedure of the ALLEGRO project has three main steps: interviews and participating observations (1), identification of critical knowledge and risks (2), development of a visualization and early warning system (3). The methodological procedure follows the usual steps of process-oriented knowledge management in public organizations (cf. [39], [40], [41]). The main phases of the structured, process-oriented knowledge management approach are presented at the top of Figure 2. These phases have a universal and general character, and can and should be periodically passed though in order to identify new and relevant knowledge management challenges. The lower boxes within Figure 2 describe concrete knowledge management steps and activities in the context of ALLEGRO. Evaluating the current state of the software development project has involved interviews and participating observations. As part of the second phase, the importance of external knowledge and the risks of knowledge loss have been identified. Characteristic risk addresses the peaks and troughs of knowledge flows and knowledge objects (artifacts, documents, source code, etc.), project-relevant knowledge loss caused by changes to external consultants, and shortcomings in the identification of relevant knowledge and knowledge bearers and, therefore, rising knowledge gaps. The empty boxes in these two phases represent further phase-relevant instruments and aspects, which are part of the general knowledge management framework of the FEA but not relevant to this knowledge management project. Developing an early warning and visualization KMS is a strategic knowledge management goal for ALLEGRO in the third phase. Because the ALLEGRO team is not involved in these strategic steps, the steps are marked in red and with broken lines. In contrast, the operative phases of the pilot project are marked in red. They will be described in further detail in the next section.

4 THE PILOT SYSTEM

The pilot system enables the investigation and visualization of several knowledge management issues within the software development process of the ALLEGRO system. This KMS pilot is a web application running on a Jetty server that uses Angular 4 in combination with D3.JS. D3 (Data-Driven Documents) is a JavaScript library for visualizing large data using web standards. It provides graph components that are suitable for visualizing the knowledge objects collected in the KMS pilot. Figure 4, Figure 5, and Figure 6 show screenshots, which comprise different D3 views – such as the collaboration and hotspot view – which are provided by the pilot. These views relate to the use cases described later in section 5.

ALLEGRO Follows the Model-Driven Development Approach

In very simple terms, the ALLEGRO development process involved three main steps. First, the engineering team models the application, the business processes, and the usage cases. Second, the development team transforms this model into a more technical model and into code. The test team uses the engineering team model to test the code produced by the development team. Innovator is a modeling suite widely used at the FEA.

Visual Analytics and Cloud Computing by our Pilot

The architecture of this KMS pilot consists of four main layers: a web-based frontend (1), a service layer containing RESTful Web Services (2), a data layer that is based on Cassandra NoSQL database (3), and a layer comprising backend processes that inserts data into the database (4).

This architecture complies with a cloud-based KMS architecture that is capable of handling big data. Figure 3 shows the architecture of the KMS pilot. The backend processes include connectors that are responsible for the data mining, and incrementally import data from the modeling tool Innovator, from

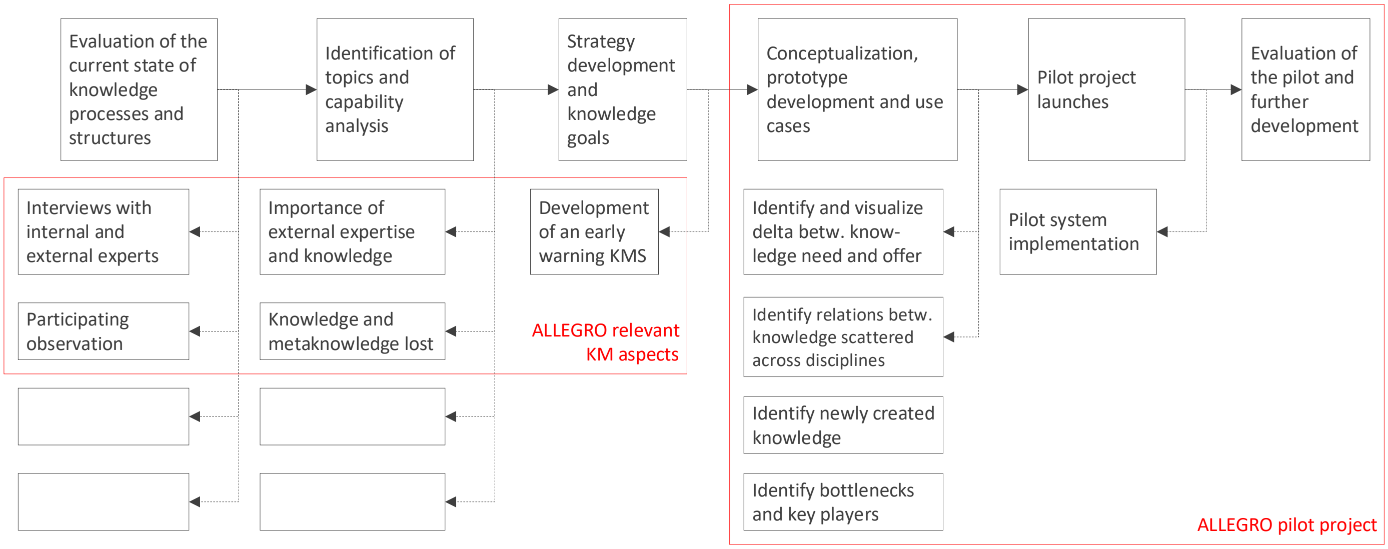


Figure 2: Schematic overview of the methodical procedure

the Confluence server, and from the Gerrit Server that is connected with a Git Repository. Gerrit is a temporary repository, into which developers commit their code for review loops. After the review process is complete, the code will be merged into the central Git repository.

The web-based frontend is running on an integrated Jetty server – a Servlet API 3.0 compliant web container. The frontend builds on the Angular 4 Architecture Stack including D3 graph components for Big Data visualization [42]. These graph components used by the frontend fit the architectural approach of using Cassandra in the backend, since Cassandra is able to return *ResultSets* in JSON Format directly – no conversions or Object-Relational (O/R)-Mapping are needed.

The service layer includes RESTful Web Services, which provide the knowledge-related data that is required for visualization to the frontend. To put the interaction of the components together, the knowledge data collected at the backend is imported into the Cassandra database. Then, the services select this data upon requests made at the frontend and provide it to the D3 components located in the Angular-based frontend.

Reasons for the Pilot's Architectural Design

Why was the pilot designed and developed in this specific way? In this context, the most relevant question might be "Why Cassandra, and why embedded in a container? Does this embedding conflict with cluster formation?" The data basis for the analytical KMS is large – knowledge to be discovered hides behind the facade of more than 740,000 LOC (lines of code), which developers have enriched with several million commits. New ones are added every day. The volume of data and heterogeneity of its sources (Innovator, JIRA, Confluence, and Wiki pages) speak clearly in favor of a NoSQL DB. But why a column store (and not a graph database or key value store)? Graph databases are primarily designed for complex data structures and column stores are a more generic approach suited to handling large volumes of data. For our purposes, the high-performance handling of a lot of data is important. The structuring by column-oriented databases is more meaningful than an irreversible fixation of our complete data layer by graph orientation. For our knowledge management purposes, we need analyses such as the summation of individual attributes. This includes, for example, the number of classes or commits in a particular knowledge area that can be defined at the package level. For such aggregates over many lines with single or few columns, column-oriented databases are not only suitable, but designed for purpose. Another feature favoring Cassandra is the advanced and powerful JDBC driver (DataStax Java Driver) for smooth use in Java. The Cassandra Query Language (CQL) query language is similar to SQL and works as a kind of "SQL for Cassandra".

But why embedded? The desired delivery model of the FEA requires a consistent implementation of the one-way-use principle. Embedding all layers into a container might complicate clustering, because the embedded Cassandra instances have to recognize and connect each other automatically. Apart from the fact that this can be done using discovery mechanisms at driver level, even if only for advanced users, the advantage of the one-way-use principle prevailed in this project.

The other components are synergistically matched. The NoSQL database can already return its results in JSON format, which the REST service simply dispatches. We use the Servlet 3.0 API and annotate the corresponding servlet method with @Produces (MediaType.APPLICATION_JSON). The servlet method directly returns the result obtained from Cassandra to the angular service that is implemented on the client side. This service on the client side takes the result without the developer having to make further conversions. It maps the result to the corresponding JSON entity class using automatic type casting (TypeScript offers a simple "as" syntax). This eliminates a lot of glue code. Data transfer and conversion are performed automatically according to the convention over configuration principle. Finally, D3 and Angular are also easy to integrate with one another.

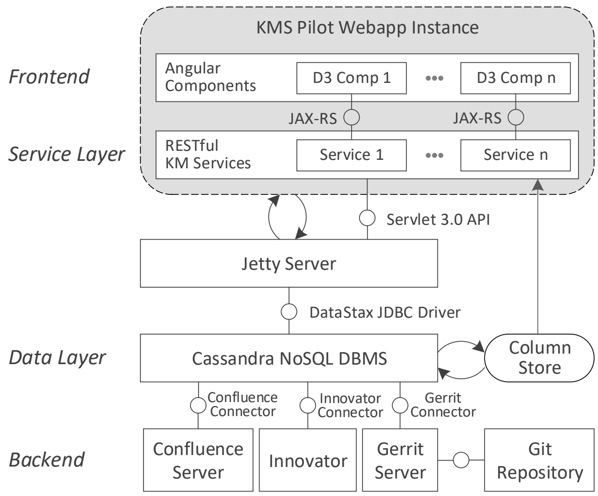


Figure 3: Architecture of the KMS Pilot

5 USAGE

The KMS pilot is used in many ways. Usage examples are described below, along with the reasons for bringing them into focus. The following describes the most essential cases for ALLEGRO covered by the KMS pilot:

Identifying and visualizing commonly developed knowledge in clustered knowledge areas: This usage relates to Figure 4, which visualizes how collective and relational knowledge has been commonly developed over three years. We used the D3 to create a “Chord Diagram”, which illustrates which developers have committed source code into the ALLEGRO source repository with which other developers/colleagues. The outside edge of the circle represents knowledge bearers – in this case committers of ALLEGRO source. A diagonal line between two or more committers means that these persons have worked (effectively together) at the same area and edited collocated code fragments (classes, methods, etc.).

The motivation for this illustration is to identify whether there are individuals solely committing in a certain area. This case is intended to answer the question of “who has unshared knowledge?”

Identifying and visualizing knowledge developed by an individual: Figure 5 shows the development of a person’s individual tacit and explicit knowledge. The identification and visualization happens by means of code ownership. We used the D3 to create a “Bubble Chart”, which incorporates a circle packing algorithm to illustrate code ownership. A developer who is committing in a certain knowledge area, where other developers do not commit anything, is deemed the sole owner of this code. Our pilot highlights this area red. When other developers start committing in this area, the color changes from red to green. The diameter represents the modified total lines of code in one code area. Circles located in the same encompassing circle are classes within the same package, which is represented by the encompassing circle. The color red means a code is owned by a single person. The color yellow means that more people are involved (knowledge is shared), and thus visualizes collective knowledge. The color green means that many people are involved (knowledge is sufficiently shared in a collective body).

This case intends to reveal in which areas (unshared) knowledge is bound to single person. This unshared knowledge is identified in the previous paragraph.

Identify hot spot areas of knowledge: This case is demonstrated by Figure 6, which displays the pilot system’s hot spot view. We used the D3 to create a “Bubble Chart”, which incorporates a circle packing algorithm to illustrate hot spots. The diameter represents the volume of lines of code in a knowledge area. The darker the color, the more commits exist in this area. It is possible to show how the focus on areas shifts by visualizing the timeline in a non-additive way– each year can be analyzed separately by not including commits from the prior year. Another hot spot metric would be a quotient involving the total number of commits and the number of commits in an area. This view represents a starting point for the investigation and visualization of causal knowledge.

ALLEGRO is complex in its business logic. In some areas, code is being changed frequently in a way that does not grow the code volume. Such knowledge areas indicate high requirements for quality. This case intends to ensure that such knowledge will be shared. It is highly problematic if an area is identified as unshared knowledge (e.g. during the second form of use described above).

Identifying experts, key players, and bottlenecks: The analysis of crawled documents, commit history, and reviews provides an overview of key individuals who have specific explicit knowledge in distinct areas, as well as tacit knowledge if the knowledge bearers have indicated their relevant knowledge domains in databases or available documents. Another view of this analysis depicts the degrees on how given knowledge areas are covered by individuals.

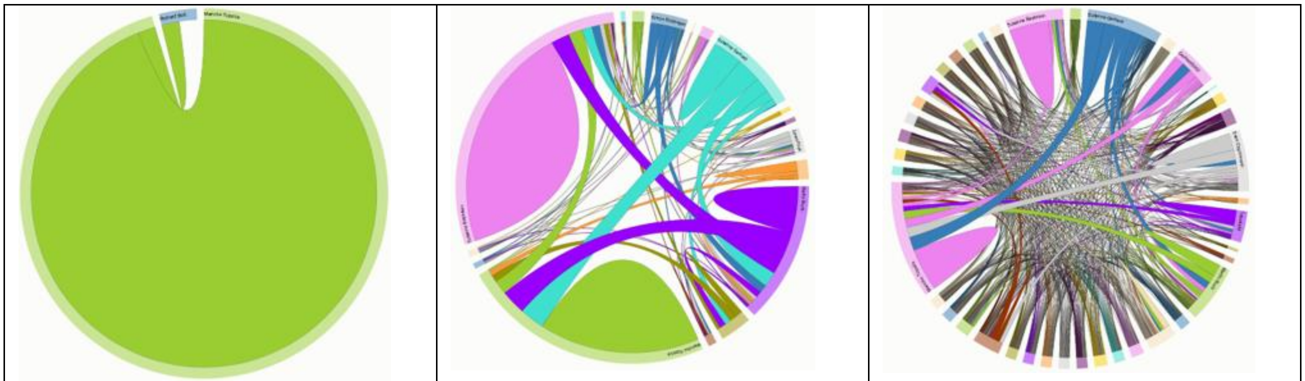
Identifying and visualizing the delta between knowledge need and supply automatically: Matching the search queries (sorted by frequency) in the organization-wide Wiki, Confluence and Knowledge Portal with the delivered search results gives a strong indication of a delta between knowledge needs and knowledge supply (between suppliers and demanders of knowledge).

Identifying relations between knowledge of one area that is scattered across disciplines: This is accomplished by crawling a document’s metadata in the file share that is commonly used by the requirements engineering team, test team, and design & implementation team, including matching this metadata with commit metadata. The result allows us to predict whether the design and implementation team will run into a required knowledge bottleneck during development of the next ALLEGRO release. The pilot extracts diverse knowledge areas from artifacts of requirements engineering and thus several taxonomical aspects of knowledge, focusing thereby on relational knowledge. The requirements engineering team is currently working on the next release and compares these knowledge areas with the available knowledge in the design and implementation team.

Identifying newly created knowledge: By analyzing the content creation and history in the organization-wide Wiki, Confluence, and Knowledge Portal, the pilot provides a big picture of knowledge development in existing areas and emerging new areas of knowledge that become relevant for teams. Such knowledge discovery is not limited to content analysis, but also includes the analysis of models that created by the requirements engineering team using the Innovator modeling tool.Hence, this function addresses the dynamics of the knowledge creation process.

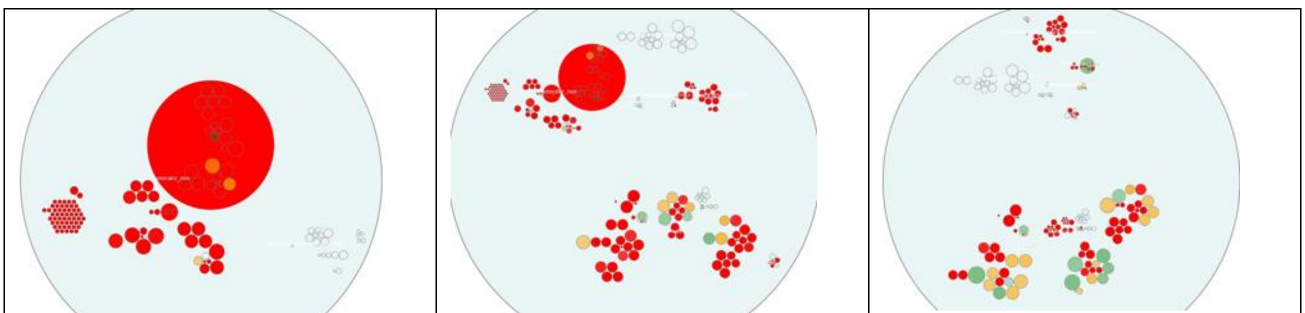
Enabling better staffing: Because the pilot can to visualize and highlight knowledge areas that are insufficiently covered, personnel can be recruited and/or skilled in these areas and staffed for specific tasks.

Graph based knowledge queries: As described, the database is a graph database storing the current knowledge available in ALLEGRO. This allows for the formulation of knowledge queries involving specific knowledge objects that are interconnected. For



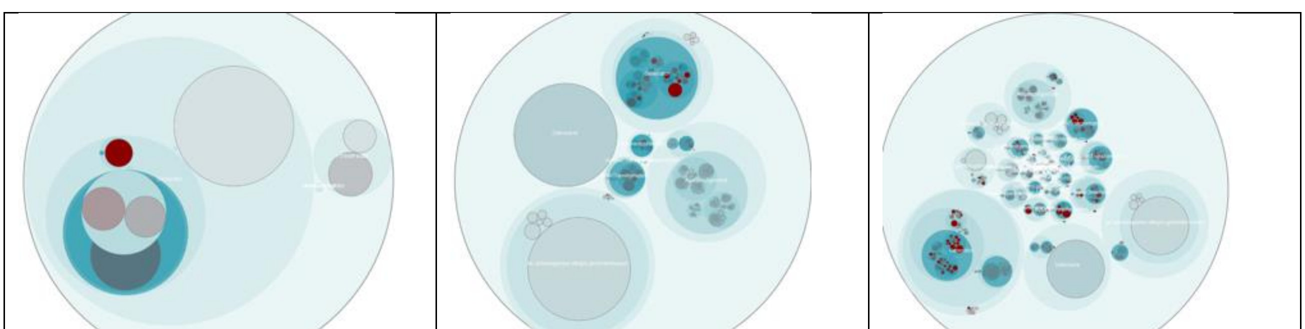
2010	2011	2012

Figure 4: Timeline of commonly developed knowledge



6/11	10/11	12/11

Figure 5: Timeline of individual-centric knowledge development from 06/11 until 12/11



2010	2011	2013

Figure 6: The Pilot's Hot Spot View displays knowledge areas (darker), in which knowledge is heavily developed

instance, it is possible to query knowledge holders and the connected people with which they share their knowledge.

6 EVALUATION

To validate its functionality, the pilot system was thoroughly tested and assessed by various members of the ALLEGRO team during the development and testing phases. Knowledge managers, project managers, and developers served as test subjects. The test trials started with a brief introduction on the intended usage of the system and its functionalities. Various tasks encompassing the usage cases described above were to the testers. The testers had to identify knowledge islands, staff a team for a potential new project, and unveil knowledge gaps in a current project. The system’s performance was assessed by comparing the results of the testers and by interviewing the testers on their impressions. The main findings of the interviews can be summarized as follows:

Knowledge Managers get a holistic and atomic overview of the organizational knowledge situation. They can trigger both the project management and the developers to take appropriate action, e.g. to share their knowledge or provide capacity for counter actions against a loss of knowledge. Furthermore, "this kind of identification of knowledge bottlenecks supports short-, mid-, and long-term planning and development of domain experts and strategic knowledge areas". Knowledge managers also found the efficient identification of newly created knowledge a major benefit, given the effort formerly required to gather and systemize new knowledge and its bearer. They also mentioned that "the identification of similar knowledge across disciplines enables the implementation of interdisciplinary expert groups". In turn, this may lead to synergetic effects between diverse disciplines and departments. Lastly, the test group emphasized the pilot’s' potential to transform individual knowledge to collective knowledge.

Project Managers gain a basis for deciding on measures such as budgets for knowledge acquisition and transfer, or investments in externalizing, sharing, or renewing knowledge. The pilot also enables them to efficiently identify key knowledge bearers without having to contact the domain knowledge manager. Project managers considered the pilot "a useful way of unveiling knowledge deficits in teams and for staffing, since domain experts are easily to identify". The pilot helps to “counter a lack of causal and relational knowledge in project teams” by supporting the incorporation of relevant roles in project teams.

Developers can see what kind of knowledge will be required from them and can prepare themselves through training or by skilling themselves in knowledge areas that are or will become relevant for them. Most of the developers appreciated the pilot's architecture, which enables a fast and relatively effortless integration of new functionalities into the system. They frequently stated that, “Now we are aware of which technology know-how (e.g. rich client or server) they need from us to develop the subsequent release.”

The pilot evaluation identified various benefits for the diverse test groups. In summary, all participants were positive about the handling and the functionalities of the system. However, further benefits will emerge from day-to-day use of the pilot system, especially when it becomes more sophisticated and established.

7 LIMITATIONS

Although we have made great progress with our analytical knowledge management pilot, we have more to do. The pilot in its current phase is solely developed and used within the design and development team. Expansion to the whole ALLEGRO team remains pending. Although the pilot crawls different sources (Git, Gerrit, Innovator, and Confluence), relationship matching and identification can be optimized. For example, it is not easy to detect whether a Confluence article belongs to a certain area of knowledge or is related to a piece of code that has been committed by an individual. The fact that two artifacts have been authored by the same person within the same period of time does not provide sufficient evidence to conclude that these two artifacts belong to the same knowledge area.

Currently, the pilot cannot distinguish whether knowledge remains tacit at one individual. Additionally, our pilot does not capture knowledge that has been shared face-to-face. Currently, we indicate that a person potentially contains a lot of unshared tacit knowledge if his/her code ownership is high within a certain knowledge area.

8 DISCUSSION AND OUTLOOK

One of the most influential and charismatic knowledge management researchers describes the essence of knowledge creation as an "endless innovation" [15]. Enabling and supporting the creation of knowledge on individual and organizational levels within an organization remains a major challenge, despite awareness of its hidden innovation potential.

The development of an early warning KMS that incorporates visual analytics has been initiated bottom-up as an answer to concrete challenges within the highly knowledge-intensive ALLEGRO project. However, the well-established and structured knowledge management framework of the FEA builds the basis for immediate support and establishment of appropriate activities in order to meet these challenges.

Given the specifics of procedural knowledge as implicitly embodied in individuals, the key source for the identification of critical process success factors has been the internal and external experts involved in the ALLEGRO project. The usage examples of the pilot system described above address all relevant taxonomical aspects of knowledge in the present context – tacit, explicit, individual, collective, procedural, causal, and relational – by supporting the handling with both knowledge objects (such as documents and data) and knowledge subjects (collaboration and communication, learning processes, and competence development). Furthermore, the pilot project and its results are one appropriate response to the actual knowledge management challenges within the public sector.

Due to the size of the ALLEGRO software system, the number of team members, and the knowledge involved, this project provides an appropriate volume of data to be processed and visualized by the KMS. The pilot currently provides very useful insights into the FEA's situation of available, emerging, fluctuating, and required knowledge. However it has not yet reached its full potential. The system is not yet (but should be) available as a mobile application for use on a device. This would fit the FEA's "Mobile First" strategy and take advantage of the REST and Angular 4 technology stack. In its current state, the pilot is internally used for investigating knowledge developments within the ALLEGRO software development. In the near future, the pilot should be available and applicable to all FEA projects. There are also some pattern recognition development possibilities. In future, the pilot should represent a Knowledge-as-a-Service platform that incorporates machine learning and In-Memory technologies in order to provide stronger insights into the FEA's knowledge situation.

REFERENCES

[1] Gronau N, Thim C, Ullrich A, Vladova G, Weber E (2016) A Proposal to Model Knowledge in Knowledge-Intensive Business Processes. BMSD 2016 - Proceedings of the Sixth International Symposium on Business Modeling and Software Design, SCITEPRESS, pp. 98-105. DOI: 10.5220/0006222600980103

[2] McQueen, R. "Four Views of Knowledge and Knowledge Management" in Proceedings of the Fourth Americas Conference on Information Systems, E. Hoadley and I. Benbasat (eds.), August 1998, pp. 609-611.

[3] Zack, M. "An Architecture for Managing Explicated Knowledge," Sloan Management Review, September 1998.

[4] Haas, M. R. (2006). Knowledge gathering, team capabilities, and project performance in challenging work environments. Management Science, 52(8), 1170-1184.

[5] Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. MIS Quarterly, 107–136.

[6] McEvoy, P. J., Arisha, A. and Ragab, M. (2015) A Review of Knowledge Management in the Public Sector: A Taxonomy. International, Forum on Knowledge Asset Dynamics. Bari, Italy, 2015.

[7] Garlatti, A., Massaro, M., Dumay, J., & Zanin, L. (2014). Intellectual Capital and Knowledge Management within the PS. A systematic literature review and future developments. International Conference on Intellectual Capital and Knowledge Management, 175–185.

[8] Sandhu, M. S., Jain, K. K., and Ahmad, I. U. K. B. (2011). Knowledge sharing among PS employees: evidence from Malaysia. International Journal of PS Management, 24, 206–226.

[9] Willem, A. and Buelens, M. (2007). Knowledge sharing in PS organizations: The effect of organizational characteristics on interdepartmental knowledge sharing. Journal of Public Administration Research and Theory, 17, 581–606.

[10] Seba, I., Rowley, J., and Delbridge, R. (2012). Knowledge sharing in the Dubai Police Force. Journal of Knowledge Management, 16(1), 114–128.

[11] O'Riordan, J. (2005). A Review of Knowledge Management in the Irish Civil Service. Institute of Public Administration.

[12] Ihringer, S. (2002) Knowledge Management: A Key Factor of Future Sustainability of State and Administration (in German), electronically published: https://www.itdz-berlin.de/dokumente/splitter/splitter_2002_3.pdf, retrieved: 20.05.2017.

[13] Al-Adaileh, R. M., & Al-Atawi, M. S. (2011). Organizational culture impact on knowledge exchange: Saudi Telecom context. Journal of Knowledge Management, 15(2), 212–230.

[14] Bullinger, H. J., Wörner K., Prieto J. (1998). Knowledge Management – Models and Strategies for Practice (in German), in: B.H.D., Wissensmanagement – Schritte zum intelligenten Unternehmen, S. 21-39.

[15] Nonaka I., Takeuchi H. (1995). The knowledge creating company: how Japanese companies create the dynamics of innovation. New York: Oxford University Press, p 72; 96-104.

[16] Nonaka, I. (1994) "A Dynamic Theory of Organizational Knowledge Creation," Organization Science(5:1), pp. 14-37.

[17] Davenport, T.H., De Long, D.W. and Beers, M.C., 1998. Successful knowledge management projects. Sloan management review, 39(2), p.43-57

[18] Koskinen, K. U., Pihlanto, P., & Vanharanta, H. (2003). Tacit knowledge acquisition and sharing in a project work context. International journal of project management, 21(4), 281-290.

[19] Williamson, T. (1995). Is knowing a state of mind? Mind, New Series, pp. 533-565 104(415), 533-565.

[20] Nonaka, I., & Toyama, R. (2003). The knowledge-creating theory revisited: knowledge creation as a synthesizing process. Knowledge management research & practice, 1(1), 2-10.

[21] Carlsson, S. A., El Sawy, O., Eriksson, I. V., & Raven, A. (1996, July). Gaining Competitive Advantage Through Shared Knowledge Creation: In Search of a New Design Theory for Strategic Information Systems. In ECIS (pp. 1067-1076).

[22] Willke H. (2001). Systemic Knowledge Management (in German). Stuttgart. 2. edition.

[23] Franken R, Franken S, 2011. Integrated knowledge and innovation management (in German). Wiesbaden: Gabler.

[24] Voigt, H. (2004) Particularities of Knowledge Management im the Public Sector (in German): http://www.community-of-knowledge.de/beitrag/besonderheiten-von-wissensmanagement-im-oeffentlichen-sektor/, retrieved: 20.05.2017

[25] Lenk, K. and Wengelowski (2002) Knowledge Management for Admistratative Action (in German) in: Edeling, T., Jann,W. und Wagner, D. (Hrsg.), Wissenssteuerung und Wissensmanagement in Politik, Wirtschaft und Verwaltung, Opladen: Leske und Budrich, 2002.

[26] Barachini, F. (2009). Cultural and social issues for knowledge sharing. Journal of Knowledge Management, 13(1), 98–110. doi:10.1108/13673270910931198

[27] Salleh, K., Chong, S. C., Syed Ahmad, S. N., & Syed Ikhsan, S. O. S. (2012). Learning and knowledge transfer performance among PS accountants: an empirical survey. Knowledge Management Research & Practice, 10(2), 164–174. doi:10.1057/kmrp.2011.46

[28] Parker, R., and Bradley, L. (2000). Organisational culture in the PS: evidence from six organisations. International Journal of PS Management, 13, 125–141. doi:10.1108/09513550010338773

[29] Bhatt, G. D. (2001). Knowledge management in organizations: examining the interaction between technologies, techniques, and people. Journal of knowledge management, 5(1), 68-75.

[30] Gronau, N. (2009). Process-oriented Knowledge Management: Methods and Tools for Utilizing the Competitive Factor Knowledge in Enterprises (in German). Oldenbourg Verlag.

[31] Tsui, E., Cheong, R. K., & Sabetzadeh, F. (2011). Cloud-based personal knowledge management as a service (PKMaaS). In Computer Science and Service System (CSSS), 2011 International Conference on (pp. 2152-2155). IEEE.

[32] Li, J., Tao, F., Cheng, Y., & Zhao, L. (2015). Big data in product lifecycle management. The International Journal of Advanced Manufacturing Technology, 81(1-4), 667-684.

[33] Antonova, A., Gourova, E., & Nikolov, R. (2009, September). Extended architecture of knowledge management system with Web 2.0 technologies. 10th European Conference on Knowledge Management ECKM09.

[34] Sultan, N. (2010). Cloud computing for education: A new dawn? International Journal of Information Management, 30(2), 109-116.

[35] Liao, C. N., Chih, I., & Fu, Y. K. (2011). Cloud computing: A conceptual framework for knowledge management system. Human Systems Management, 30(3), 137-143.

[36] D. A. Keim, F. Mansmann, J. Schneidewind, J. Thomas, and H. Ziegler: Visual Analytics: Scope and Challenges. Lecture notes in computer science, No. 4404 (2008), pp. 76-90

[37] Q. Zhang, R. S. Segall, and M. Cao (2011). Visual Analytics and Interactive Technologies: Data, Text and Web Mining Applications. Hershey, PA: IGI Global

[38] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, et al. (2008). Visual Analytics: Definition, Process and Challenges. In: A. Kerren and J. T. Stasko and J.-D. Fekete and C. North. Information Visualization - Human-Centered Issues and Perspectives, Springer, pp.154-175, 2008, LNCS.

[39] Karagiannis D., Telesko R., (2000) "The EU-Project PROMOTE: A Process-oriented Approach for Knowledge Management", In Proceedings of the Third International Conference on Practical Aspects of Knowledge Management (PAKM 2000), October 30-31 2000,Basel, Switzerland

[40] Maier, R. and Remus, U. (2003) "Implementing process-oriented knowledge management strategies", Journal of Knowledge Management, Vol. 7 Issue: 4, pp.62-74, doi: 10.1108/13673270310492958

[41] Sivri, S. and Krallmann, H. (2014): Process-Oriented Knowledge Management within the Product Change Systems of the Automotive Industry. In proceedings of the 25th DAAAM International Symposium on Intelligent Manufacturing and Automation, DAAAM 2014.

[42] A. Lerner and V. Powell (2014). D3 on AngularJS: Create Dynamic Visualizations with AngularJS. Leanpub, 15.04.2014.