In this course unit [CM-W-WEB] a structured software development process of web applications is illustrated. The application development makes use of current software development concepts, such as Behavior-Driven Development (BDD), Domain-Driven Design (DDD), microservice architectures including a systematic design of the web Application Programming Interfaces (API) of the microservices.

(1) An overview of the relevant development concepts is given and the application development process using these concepts is introduced. The core concepts and technologies are illustrated with the example of a web application, by name TodoListManagement (TLM), is introduced.

(2) The development of advanced web application requires a complex tool environment. The organization of software development projects is supported by a specific set of project management and version control tools which are described in more detail in this chapter.

(3) This chapter deals with the analysis, the first phase of the development process, in which the requirements are specified. In the approach described in this course unit, so-called Gherkin features are used. Those features are the central artifact of the concept of Behavior-Driven Development (BDD).

(4) The domain model and the APIs are central artifacts of the subsequent design phase. In the domain model established according to the principles of Domain-Driven Design (DDD), the (structural and functional) knowledge needed to implement the software system is formally captured.

(5) In the implementation phase the functionality as is was specified in the analysis and design phase is coded based on a microservice architecture and frameworks supporting the frontend and backend implementation.

(6) In the deployment phase, the tested application is deployed in a Docker-based execution environment.
The motivation behind Behavior-Driven Design (BDD) is to build and deliver better software [Sm15:3].

(1) According to a number of studies, nearly half of all software projects fail to deliver in some significant way.
(1.1) In a CHAOS Report published by the Standish Group, in 42% of the software projects one or more of these problems occurred.
(1.2) This results in billions of dollars in wasted effort.

(2) The two goals center around the questions of WHAT to develop and HOW to develop.
(2.1) This goal describing the WHAT is shown by the vertical (y) axis in the figure.
(2.2) This goal describing the HOW is shown by the horizontal (x) axis in the figure.

(3) BDD encourages business analysts, software developers, and testers to collaborate more closely by enabling them to express requirements in a more testable way, in a form that both the development team and business stakeholders can easily understand.

BDD Behavior-Driven Development

Throughout this course unit, the TodoListManagement (TLM) web application by which a user can manage his/her todos in lists, is used to demonstrate the technologies and methods of web application development. This application is also in the focus of the initial phase of the practical and seminar course offered in parallel to the lecture. A detailed technical description can be found in [CM-W-TLM].

(1) To be able to develop complex web applications it is not only necessary to (theoretically) understand the quite complex concepts, but also to (practically) apply them. With the TLM example the practical use of the concepts is demonstrated.

(<<client>> TLM Frontend, HTTP. <<server>> TLM Backend) Since TLM is a web application it consists of a web client and a web server which communicate via the Internet application protocol HyperText Transfer Protocol (HTTP). The diagram describes the physical architecture of the TLM application and uses modeling elements specified by the Unified Modeling Language (UML). It is a so-called deployment diagram by which the physical systems (in this case, TLM Frontend and TLM Backend) are modeled as specific UML symbol called nodes.

(Web Technologies) Technologies play an important role in web application development. According to the physical architecture of a web application, frontend and backend technologies as well as technologies concerning the interface (Application Programming Interface, API) can be distinguished)

(A, TypeScript, JS) The "A" logo stands for Angular which is a leading frontend technology. Angular uses the script language TypeScript which is based on JavaScript (see JS logo).

(TLM GUI) This is a cutout of the TLM Graphical User Interface (GUI) which is implemented in Angular.

(Spring, Java) Spring is one of the leading Java-based backend technologies.

(OpenAPI Specification) The OpenAPI Specification is a manufacturer-independent defacto standard for the definition of REST APIs.

(REST API) (TLM REST API) REST stands for Representational State Transfer which is the most widely used concept for the specification of web APIs. (POST /todo-lists/{id}/todos, GET /todo-lists/..., PUT ...) This is an excerpt of the REST API which is provided by the todo-list microservice. Each REST operation is a specific HTTP method, such as GET or POST. For example, the REST operation "POST /todo-lists/{id}/todos" has a parameter which describes the todo which is added to the todo list identified by "id".

<table>
<thead>
<tr>
<th>API</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Application Programming Interface</td>
<td>HTTP HyperText Transfer Protocol</td>
</tr>
<tr>
<td>JS</td>
<td>JavaScript</td>
</tr>
<tr>
<td>REST</td>
<td>Representational State Transfer</td>
</tr>
<tr>
<td>TLM</td>
<td>TodoListManagement</td>
</tr>
<tr>
<td>UML</td>
<td>Unified Modeling Language</td>
</tr>
</tbody>
</table>

The TodoListManagement (TLM) web application serves as an example to demonstrate the concepts of web application development introduced in this course. All artifacts and the code of the web application is available in the GitLab [CM-G-TLM].

(1) A concept that is used to specify and test the functionality is the Behavior-Driven Development (BDD). The main artifact of BDD are the so-called features which are described in the language Gherkin.

(2) The sample project is a simple todo administration software.
(2.1) It should be possible to create a new todo list and add it to the existing todo lists. Lists can have a title and an unlimited number of todos.
(2.2) Existing todo lists can be edited.
(2.2.1) The title can be changed or a new todo can be added to the list.
(2.2.2) It is also possible to mark a todo as done.
(2.3) Another requirement is to delete existing lists.

(1. Feature: Creating a todo list) Most of the feature description is written in a natural language. In order to allow interpretation with the computer, Gherkin prescribes some keywords (in the example: Feature, Background, Scenario, Given, When, Then, And) which must be considered. Gherkin is used by the Cucumber framework [Cuc-Ger].
(2. As a ...) (3. I want ...) (4. So that ...) Each feature starts with a story which informally introduces the role and what this role wants to do why.
(6.) (11.) The feature consists of two scenarios. Each scenario describes a specific part of the behavior of the application which is specified by preconditions (Given), actions or events (When) and the testable outcomes (Then).

(edu.kit.cm.springBootDemo.cucumber.RunCucumberTests) In all later phases of the software development process the definition can be used to test the software. In this way, it is possible to test whether a software meets the requirements and, in case of changes to the code, to check whether all requirements are still being met. The tests are started as Junit tests and can be integrated into all existing tools [Jun-Ju4].

BDD Behavior-Driven Development
TLM TodoListManagement

The figure introduces the different phases that have to be carried out to develop software systems, esp. web applications. Besides the programming language, natural language and several description languages are needed to develop a web application in a sound and systematic way.

(1) Iterative in this context means that the phases are not completely passed in a sequential order but jumps to a the preceding phase are the usual case.

(Feature, Domain Model) Features are the main artifacts from Behavior-Driven Development whereas the domain model is the main artifact introduced by the Domain-Driven Design (DDD).

(Unified Modeling Language, eXtensible Markup Language, Hypertext Markup Language) are widely used description languages in the development process of web applications. Gherkin is a language which is used in the BDD approach.

(OpenAPI) This is a standardized language by which an API (Application Programming Interface) of a microservice based on REST (REpresentational State Transfer) can be specified.

<table>
<thead>
<tr>
<th>API</th>
<th>Application Programming Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDD</td>
<td>Behavior-Driven Development</td>
</tr>
<tr>
<td>DDD</td>
<td>Domain-Driven Design</td>
</tr>
<tr>
<td>REST</td>
<td>REpresentational State Transfer</td>
</tr>
<tr>
<td>UML</td>
<td>Unified Modeling Language</td>
</tr>
<tr>
<td>XML</td>
<td>eXtensible Markup Language</td>
</tr>
</tbody>
</table>
This figure illustrates the various architectural concepts and their relationships. It should be noted that no standardized definitions exist. For example, [CMU-SA] lists more than 30 definitions for software architecture.

(Software Architecture) The software architecture takes a logical view on the system by dividing it into logical components. (System Architecture) The system architecture takes a physical view on a system and describes its structure which consists of network and hardware components [VA+08]. Furthermore, their properties and relationships to each other as well as to their environment and other systems are shown. A possible modeling type is the deployment diagram. (Implementation Architecture) The implementation architecture describes the structure of the system from a technical point of view [RM+06]. This includes packages, libraries and frameworks. (Macro Architecture, Miocro Architecture) The macro architecture is used to describe the principles of the subsystems, e.g. microservices or modules, on a black box level. The micro architecture deals with the internal structure of a single subsystem [Ot13].

(Architectural Style) Architectural styles [RH08] are application-independent solution principles that are used throughout the project. These can be assigned to the following categories: communication (e.g. message bus oriented), deployment (e.g. N-tier architecture) and structure (e.g. layer architecture). (Design Pattern) Design patterns [RH08] provide program code-specific patterns (e.g. visitor pattern) for recurring problems. (Architectural Pattern) Architectural patterns [RH08] are solutions to recurring problems which, unlike design patterns, affect several architectural elements.

(Component / Deployment / Class Diagram) Deployment View) The Unified Modeling Language (UML) provides different types of diagrams to specify the different types of architecture as models.

Challenges in Software Development++

- Goal 1: To develop the right software (WHAT should be developed?)
- Goal 2: To develop the software right (HOW should the software be developed?)

Example: Web Application Demo TodoListManagement (TLM)++

- Physical systems: Web client (frontend), web server (backend)
- HTTP (HyperText Transfer Protocol) which is an Internet application protocol.

Example: Web Application Demo TodoListManagement (TLM)++

- Angular: A leading frontend technology
- TypeScript: Used by Angular and based on Java Script
- Spring, Java: Spring is one of the leading Java-based backend technologies.
- OpenAPI Specification: A manufacturer-independent defacto standard for the definition of REST APIs.

Software Development Process++

- In the analysis phase
- Remark: Features are described by using the language Gherkin. They are the central artifact of the Behavior-Driven Development (BDD).

Types of Architectures and Relationships++

- Software architecture: logical view, UML component diagram
- System architecture: physical view, UML deployment diagram
(1) The fundamental principles of each plane (e.g. layering) are described in the following.

(Network Plane) Web applications run on several independent systems which heavily use Internet technologies for communication purposes. The network plane consists of communication layers. In each such layer the communication is organized by standardized protocols (e.g. IP, HTTP).

(System Plane) Internet protocols (such as the web protocol HTTP) are based on a client-server principle. The system plane is characterized by different operating systems leading to heterogeneous IT infrastructures.

(Application Plane) Web-based applications are located on the application plane. The term "web-based" especially means that the presentation layer is implemented by a standard web browser.

(DATA, BUSINESS LOGIC, PRESENTATION) The logical separation concerns the different (logical, conceptual, architectural) aspects a distributed application has to cover. There are three aspects – data, function (or business) and presentation – which can be found in every application.

(2) The Unified Modeling Language (UML) is an adequate language to graphically describe the architecture of a software application. Modeling is necessary in order to develop the application in a structured and systematic way.

(Stereotype describing a component) A component in UML is modeled as a rectangle with a specific icon in the right upper corner. This icon is a graphical representation of a stereotype which is further defined in the UML meta model. An equivalent textual representation of this stereotype defining a component is <<component>>.

(Ball-Socket symbol describing an interface) The ball represents the provided part and the socket represents the required part of the component interface.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB</td>
<td>Data Base</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical User Interface</td>
</tr>
<tr>
<td>HTTP</td>
<td>HyperText Transfer Protocol</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>UML</td>
<td>Unified Modeling Language</td>
</tr>
</tbody>
</table>
There are so-called communication reference models that define the functionality of the layers that build the network plane. Two important models are shown on this page [KR05, Ta06].

(HTTP, Request Content, Provide Content) The HyperText Transfer Protocol (HTTP) is an Internet application protocol which allows a HTTP client to request and also change content residing at an HTTP server by offering different HTTP operation, such as GET or POST.

(1) (2) The merging of the two lower layers in the Internet architecture has "organizational" reasons whereas the missing layering in the application system is conceptually motivated.

HTTP HyperText Transfer Protocol  
IEEE Institute of Electrical and Electronics Engineers  
IP Internet Protocol  
ISO International Standards Organization  
OSI Open Systems Interconnection  
TCP Transmission Control Protocol

On the system plane of a web application the involved systems take the roles of clients and servers.

(1) Processing is usually organized according to a client-server principle: The client asks the server to do some processing.

(2) Central tasks the server systems listed in the slide have to provide, are:
- File server: Upload and download of files
- Database server: SQL-based search queries
- Print server: offers functionality specific to the print service, such as queuing or prioritizing of print orders
- Application server: contains executable code specific to one or more applications

SQL Standard Query Language
A microservice architecture is located on the application plane as it is introduced in ++Planes and Three-Layer Application Architecture+++. It consists of three parts (backend, backend-for-frontend BFF, frontend) and two types of application programming interfaces (BFF) via which these parts are interacting (backend microservices API, BFF microservices API).

(Presentation Layer) This layer renders the UI elements in the browser. Technologies that support the implementation are Angular and Bootstrap.

(Presentation Logic Layer) The presentation includes a logic which controls the interaction with the BFF microservice API.

(Application Logic Layer) This layer realizes the orchestration of the backend microservices in order to provide the BFF microservices required by the frontend. A technology that supports the implementation of this functionality is Spring.

(Domain Logic Layer, Infrastructure Layer) These layers implement the backend services which are CRUD operations on the domain objects.

In contrast to a traditional three-layer application architecture the business logic layer in a microservice architecture is split into two layers, the domain logic layer and the application logic layer. The reason for that is to promote the reuse of business logic functionality by distinguishing between application-agnostic (= domain logic) and application-specific (= application logic) functionality.

<table>
<thead>
<tr>
<th>API</th>
<th>Application Programming Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>BFF</td>
<td>Backend For Frontend</td>
</tr>
<tr>
<td>CRUD</td>
<td>Create, Read, Update, Delete</td>
</tr>
</tbody>
</table>
The software development process applied by C&M combines the concepts of Behavior-Driven Development (BDD) and Domain-Driven Design (DDD). Both concepts provide complementary contributions to the layered microservice architecture as the figure illustrates.

(1) An implemented feature can be seen as a deployable increment of the software system. (Feature1, Feature 2, ...) The ordering of the features in the figure implies that the first feature should cover the core functionality of the software system.

(2) The domain model makes sure that the static and dynamic domain knowledge is consistently used by each feature. This ensures that the features build a consistent whole although each feature is developed and deployed independently from other features.

BDD
Behavior-Driven Development

DDD
Domain-Driven Design
The activity diagram gives an overview of the whole engineering process which in fact is divided into two sub-processes: According to our understanding of DDD, a domain model is the foundation of all software applications belonging to this domain. We regard the domain model as the knowledge of a domain expert he or she makes available to the software developer by the model. This knowledge is the application-agnostic part of the application meaning that each application of the domain implements this knowledge in its domain logic layer. We understand this domain knowledge as the "physics" that each application must adhere to. Evans calls this part of functionality the heart of the software system [Ev03].

(Starting Points) (Strategic Modeling of the Initial Context Map) The starting point of the domain modeling process is set earlier as the starting point of the application development process. The reason is that it makes sense to have an initial context map of the domain before the development of the first application starts. In the following we describe the main characteristics of the application development process.

(Context Map) The main artifact of the domain modeling process is the context map in which all micro-service-based applications belonging to the domain must fit into.

(Strategic Modeling of the Bounded Contexts as Part of the Domain’s Context Map) In the application development process, the integration of the application under development into the context map of the domain is done in the second step after the first step in which the the BDD-based requirements analysis is carried out.

(Tactical Modeling of a Bounded Context) A bounded context is part of a context map. It represents a candidate for a microservice which can be developed by an independent team [Ne15]. DDD considers the modeling of the content (i.e. the functionality) of a bounded context as tactical modeling.

(Relation View) The most relevant artifact of the tactical modeling is the relation view which is built in the third step. By the relation view central DDD concepts (such as entity and value objects) including their relationships and operations are represented.

(Implementation of each Bounded Context as a Microservice Accessible via its Web API) The functionality modeled by the relation view is implemented in the domain logic layer of the microservice architecture. A systematic approach to implement the bounded context as a microservice (fourth step) is a core characteristic of the proposed development process. One of the key demands of DDD is to keep the model and its implementation in sync in order to assure that the model and the implementation is the same.

The two remaining steps concern development activities around the core implementation of the microservice, namely the implementation and test and the deployment of the application.

The figure shows that concepts specific to the microservice field do not replace, but complement, existing software engineering concepts.

(Analysis Techniques) In the first phase of the engineering process, the analysis phase, different techniques to elicit and specify the requirements exist.

(Ubiquitous Language (BDD, DDD)) In the process we apply for the development of microservice-based software, the requirements are specified by Gherkin features which are part of Behavior-Driven Development (BDD). There is a match between BDD and microservices since BDD provides a ubiquitous language which is needed to design the APIs of a microservice.

(Domain Modeling (DDD)) BDD fits well to the Domain-Driven Design (DDD) which we estimate to be one of the most relevant concepts for the design of microservices.

(Design Patterns) DDD is a set of design patterns which are a well-known engineering concept to capture the knowledge of the structure (architecture, code) of a software system.

(Microservice Architecture) One of these DDD patterns, called Layered Architecture, provides the foundation of the microservice architecture. By the Layered Architecture pattern, the functional layer from the well-known three-layer application architecture is divided into a domain logic layer and an application logic layer. The functionality captured in the domain logic layer has to be application agnostic. This is exactly the functionality a microservice API should provide. By using the DDD patterns, a domain model of the domain, the software system belongs to, is built.

(API Design) The domain model supports both the implementation of the microservice and the design of the web API.

(Frontend/Backend Frameworks, Microservice-Supporting Frameworks) For the implementation of a micro-service-based web application, powerful frameworks for the implementation of the frontend and the backend are available. Certain frontends provide a specific support for the development of microservices, especially related to the web API via which the functionality of the microservice can be accessed.

(CI/CD) One of the driving forces of microservices is the continuous integration (CI) and continuous deployment (CD) of implemented software systems. In contrast to monolithic architectures, microservices are loosely coupled. Changes in the software system relate to single micro-services which can be tested and deployed independently from the other part of the software system. In practice, each microservice is deployed in a so-called container which provides the computing infrastructure to run the microservice.
(1) ++Planes and Three-Layer Architecture++
- Application plane
- System plane, network plane

(2) ++Planes and Three-Layer Architecture++, ++Microservice Architecture++
- Separation of the business logic layer (middle layer) into two layers, the application layer and the domain layer
- Explicit introduction of two APIs, the Backend Microservice API and the Backend for Frontend Microservice API

(3) ++Features and the Domain Model in a Microservice Architecture++
- BDD: Each feature is a vertical cut through all layers of the microservice architecture
- DDD: The domain model is implemented in the domain logic layer and provides the semantical foundation for all features

(4) ++BDD/DDD-Based Software Development++
- BDD: Influences the analysis by providing the specification based on features. In addition, a feature defines the software increment which is implemented in an agile way.
- DDD: The domain model is the central design artifact of the development process.
(1) On the one hand, the tools support the software engineers and make software development easier and more efficient. On the other hand, the broad spectrum of different tools makes software development more complex since the correct use of the tools is a time-consuming challenge for each developer.

1.1 The phases are: analysis, design, implementation, testing, deployment. Management tools cannot be assigned to one specific phase are which are used by the teams for communication purposes and for the organization of the software project.

(2) This type of development tools is indispensable for the development of software.

2.1 Construction tools are used in the implementation phase.

(3) An example of a software requirements tool is Cucumber which supports the approach of behavior-driven development (BDD). Examples of software design tools are UML modeling tools or tools to specify the APIs (application programming interface) of a software architecture.

(4) Software testing tools include test generators, test execution frameworks, test evaluation, test management and performance analysis. Tools which support the deployment are called software configuration management tools [Ca01].

(5) The main support provided by this type of development tools is project planning and tracking of tasks.

Software engineering tools cover two major aspects which must be handled in every software project: the management aspect dealing with organizational and project-related problems and the development aspect dealing with technical and (software) system-related problems.

(Communication and Collaboration Platform) The tools for solving either organizational or technical problems of software engineering are not decoupled. If a technical problem occurs this has influence on the overall planning. In order to manage these relationships a common communication and collaboration platform for management and development is useful.

(Software Repository) A software repository enables the storage of different versions of the software in a way that each software engineer can communicate the changes he has made in a systematic way to the other team members.

(Project Management and Organization) The relevance of this group of software engineering tools rises with the size of the projects and the software engineers involved.

(Project Planning and Tracking) This type of tool is used by a project manager (in Scrum this role is called project owner) to plan the resources (esp. the developers) that are available in the project. In [Ca01] the project planning and tracking is part of the so-called software engineering management tools.

(Issue and Problem Tracking) A team of software engineers uses this type of tool to organize technical, i.e. software-related tasks.

(Software Development) In this group of tools the development process starting with the analysis and ending with the deployment of the software is supported. The list of tool types mentioned in the figure is not complete. Tool types which are missing include the software quality tools or software maintenance tools.
The main dimension according which the development tools can be ordered are the development phases (from analysis to deployment). The tools supporting the analysis and design and the implementation and testing can be grouped according to the software architecture (which in our case is a microservice architecture).

(Project Management and Version Control) These tools support the overall organization of the software project and the communication between the project members. At C&M two different tools sets are used: the Atlassian toolset extended by Microsoft tools, esp. SharePoint on which the C&M Teamserver is based.

(Analysis and Design) Analysis requirements at C&M is done by taking the approach of behavior-driven development (BDD) based on the tool Cucumber. For the design the two most relevant tools are Enterprise Architect for the domain model and Swagger for the API specification. In addition to these tools the Microsoft Office tools (Word, PowerPoint) and the Bitbucket Wiki are applied for documentation purposes.

(Implementation and Testing) In this phase the frontend and backend of the web applications are constructed. The integrated development environment (IDE) used for frontend development is Visual Studio Code and frontend frameworks are Angular and Bootstrap. As IDE for backend development Eclipse is used and Apache Spring (esp. Spring Boot for the microservice implementation) is used as backend framework.

(Build and Deployment) The build and deployment of the microservices is carried out via a build pipeline by which the concept of continuous integration and continuous deployment is provided. The result of the build pipeline is a Docker image (= application container image) since Docker is used as the container environment at C&M. Kubernetes is used to manage the Docker containers in order to reach a high scalability and robustness of the service landscape.
Analysis and design comprises the development work that is needed to implement the right software in the right way.

(1) Analysis is concerned with the question which software the user requires [CM-W-BEH].
(1.1) Gherkin is a language which introduces a minimal structure into the largely informal description of the requirement description.
(1.2) Cucumber allows to test an implemented software based on the specified features. The test are called user acceptance tests.

(2) Design is concerned with the static and dynamic structures of the software system that is to be implemented [CM-W-DOM].
(2.1) The concept of domain-driven design (DDD, [Ev04]) does not include a way how to specify the parts of a domain model. The Unified Modeling Language (UML) is a widely accepted modeling language which provides the flexibility to express the needed DDD modeling elements. Enterprise Architect is one of the leading professional UML tools to implement the DDD-to-UML approach.
(2.2) Based on a formalized UML-based domain model, the web application programming interfaces (API) of the microservice architecture can be systematically derived. The specification of the APIs is based on the OpenAPI defacto standard. Swagger is a well-known tool to specify and test the web APIs.

[CM-W-BEH] Cooperation & Management: BEHAVIOR-DRIVEN DEVELOPMENT, WASA Course Unit. https://team.kit.edu/sites/cm-tm/Mitglieder/2-1.WASA
[CM-W-DOM] Cooperation & Management: DOMAIN MODELING, WASA Course Unit. https://team.kit.edu/sites/cm-tm/Mitglieder/2-1.WASA
(1) **Spectrum and Types of Tools**
- Software construction tools: Program editors, compilers / interpreters, debuggers

(2) **Classification of Software Engineering Tools**
- Project Management and Organization: The relevance of this group of software engineering tools rises with the size of the projects and the software engineers involved.
- Software Development: In this group of tools the development process starting with the analysis and ending with the deployment of the software is supported.

(3) **Example: Tool Environment**
(3.1) SharePoint, Confluence, Jira, Bitbucket, Trello
(3.2) Git, Bitbucket
(3.3) Cucumber
(3.4) Enterprise Architect
(3.5) Angular and Bootstrap
(3.6) Spring
(3.7) Docker, Nexus, Portainer
In this chapter the software engineering tools are introduced that are used to manage and organize the project and the people involved in the whole software engineering process.

(1) Software systems tend to become complex which means that a whole team of software engineers is needed to develop the system. That is why software engineering means teamwork.
   (1.1) The project teams are built in the beginning of each semester.
   (1.2) The coach takes care that all team members contribute to the teamwork.

(2) In this context, project management comprises all tasks that should be carried out to organize the development tasks in a systematic and clear way.
   (2.1) Features specify the requirements the software system must fulfill. They are described in a language called Gherkin. Since the agile development process is organized along the features with highest priority this artifact plays an important role for the project planning.
   (2.2) The division into frontend-related and backend-related tasks make sense since they demand different competences.
   (2.3) At C&M, the Bitbucket wiki is used to store the artifact descriptions which are edited collaboratively by the members of the development team.

(3) The development of Git started in 2005. It is currently the most widely used version control system.
A version control system (VCS) is one of the most relevant software tools to develop complex software in a team [Atl-Wha].

(1.1) A local VCS is a database that keeps all the changes to files under revision control. The differences from one file version to the next file version usually is stored in a certain format. An example of a local VCS is RCS (Revision Control System). The central and distributed VCS are explained in what follows [CS14].

(2) A disadvantage of local VCS is that a collaboration of developers working on distributed systems is not supported which has led to the central VCS. In a central VCS, clients running on different systems check out files from the central place.

(2.1) Subversion is the successor of CVS which stands for Concurrent Versions System.

(3) In a central VCS, clients only check out single files and not the repository (i.e. the whole file system).

(3.1) The full mirror includes the full history of all the files in the repository. This leads to a high stability.

(3.2) BitKeeper today is an open-source distributed VCS which was used to develop the Linux kernel. Some conflicts caused by the fact that BitKeeper at that time was proprietary (owned by a company called BitMover) has led to the development of Git.

Mercurial is a free software which was announced and developed at the same time as Git in 2005.

(Computer A <---> Computer B) Note: There is only an "indirect" relationship between the participating computers A and B (leading via the server computer).

CVS Concurrent Versions System
RCS Revision Control System
VCS Version Control System

Git was developed by Linus Torvalds, the initiator of the Linux kernel. Development started in 2005 [HP16]. The reason for the development of Git was the need of a version control system for the Linux kernel development.

1.1 Git allows to record and manage the changes of files.
1.2 Git is the appropriate tool both for single persons and for large projects in which hundreds of developers cooperate (such as the Linux kernel development).
1.3 Examples of other VCS include CVS (Concurrent Versions System) or Subversion (successor of CVS) which store data as a list of file-based changes (since the changes of files are stored as deltas, they are named delta-based version control).

2 A snapshot in Git is a picture of all the files of the project. For efficiency reasons for all files that did not change from version to the other only a link to the previous identical file is stored. (Version 3, A1, B, C2) The snapshot Git takes for Version 3 are links for A1 and B since they did not change (illustrated by the dotted lined ellipses). Only C2 is stored since it is different from C1 (which is file C in Version 2).

This page gives an overview of the most relevant commands provided by the distributed version control system Git [Atl-Get].

1. A Git repository (also called Git repo) is a virtual storage of the versions of code from a project. 
   1.1. `git init` creates a new .git subdirectory in the current working directory. It will also create a new master branch. 
   `git clone <repo url>` creates a copy of the indicated remote repository. The URL format depends on the network protocol that is used. 
   `git config` sets Git configuration values (e.g. email or username) on a global or local project level. It also allows to create shortcuts for frequently used Git operations in order to work with Git more efficiently.

2. In Git, each change of the local repository is done via a so-called staging area which can be seen as a buffer between the working directory and the project history. 
   2.1. `git add` and `git commit` compose the fundamental Git workflow. They are the means to record versions of a project into the repository's history. The stage is used by Git to group related changes into highly focused snapshots before actually committing it to the project history (when `git commit` is executed). 
   `git stash` (germ. bunkern) temporarily saves uncommitted changes and reverts them when a developer later wants to continue his work. 
   `.gitignore` is a file which contains all names of files that should be ignored by Git (e.g. build artifacts or machine generated files, such as compiled code).

3.1. `git status` displays the state of the working directory and the staging area. 
   `git log` provides information regarding the committed project history.

4. These commands allow to explore old commits and undo the changes in different ways.

5. By using these commands the Git history can be rewritten or altered.

6. These commands are used by a developer to share a series of commits with other developers.

GitLab is an open source software solution which uses the functionality of the version control system Git. Examples of solutions comparable to GitLab are: GitHub (owned by Microsoft), Bitbucket (owned by Atlassian).

(1) Several large tech companies (e.g. NASA, CERN, Boeing, IBM, LRZ, KIT) use GitLab.
(1.1) The management includes the documentation by a wiki, workflow support by an issue tracker and automatic build function by a CI/CD pipeline.
(1.2) Go (often referred to as Golang) is a statically typed, compiled programming language syntactically similar to C and designed at Google. Ruby on Rails, or Rails, is a server-side web application framework written in Ruby. Vue.js (commonly referred to as Vue; pronounced like view) is an open-source JavaScript framework for building user interfaces and single-page applications and it is inspired by Angular.

(2) All KIT members (students, employees) who have a KIT Active Directory (AD) account can use KIT GitLab [KIT-GitLab].
(2.1) The Steinbuch Centre of Computing (SCC) is the IT service provider at the KIT.
(2.2) Each KIT AD account has the structure "xy1234".
(2.3) The profile can be added by clicking on the entry "Settings" from the menu which can be opened by clicking on the arrow on the right hand side of the public avatar.

(3) C&M uses GitLab for the documentation of all artifacts and the versioning of all the code that are developed during the structured development process (see also page ++BDD/DDD-Based Software Development Process++).
(3.1) The origin of and main concepts behind the repository stem from [Zi19].

AD
Active Directory

SCC
Steinbuch Centre for Computing


[KIT-GitLab] Karlsruhe Institute of Technology: GitLab am KIT SCC. https://git.scc.kit.edu

The goal is to make all relevant information on the software, developed by C&M, available in GitLab. The documentation especially includes the analysis and design artifacts specified during the development.

(1) It should be taken into account that the documentation is carried out in a “normal” Git repository (and not in the Wiki as one might expect). The reason for this is the missing branch concept in a Wiki which is needed by C&M to implement the reviewing as shown in the process on the right-hand side.

((content change)) (Commit Changes to New Branch) As shown by the documentation process, each major change which concerns the content (i.e. the semantics) leads to the generation of a new branch in which the changes are committed.

(Merge Request) (Accept / Merge Changes) By this GitLab operation the changes are provided to other persons involved in the documentation. Usually, the main responsible person or owner of this part of the documentation is merging (i.e. adopting) the changes.

(2) GitLab uses an extension, called GitLab Flavored Markdown (GFM) [GL-Mark] of the standard Markdown, which is a lightweight markup language with plain text formatting syntax.

(2.1) Alternatives are Typora, which is a WYSIWIG editor or GitLab's Web IDE (but not GitLab's single-file online code editor).

(1 # C&M – Documentation Guidelines ...) The figure on the lower right-hand side is a screen dump taken from Visual Studio Code. It shows the README Markdown file from the CM-Documentation repository in the left-hand window and its Markdown preview in the right-hand window.

Note: In all C&M documents screen dumps with black background color should be avoided. Therefore, the default setting of Visual Studio Code should be changed (File | Preferences | ColorTheme | Light+) when screen dumps are taken.

GFM GitLab Flavored Markdown

The structure and the content of a GitLab repository is shown with the repositories of the TLM web application.

1. The TLM demo provided by C&M is accessible via GitLab [CM-G-TLM].
   (CMTTeam) This is the main C&M Group in GitLab. It contains guidelines for the correct usage of GitLab and subgroups for the different projects of C&M.
   (TodoListManagement) This is a subgroup of the "CMTTeam" group and it contains the repositories of the TLM project. Every project of C&M has its own subgroup in the "CMTTeam" group.
   (0Doc_TodoListManagement) For each project a repository "0Doc_<project name>" should exist in which the artifacts that are valid for the whole project (i.e. frontend and backend) are stored.
   (1UI_TodoListManagement) This is the repository in which the frontend code (UI stands for user interface) of the TLM application is stored.
   (2BFF_TodoListManagement) In this repository the backend for frontend (BFF) code of the TLM application is stored.
   (3MS_TodoListManagement) This is the repository in which the backend microservices (MS stands for micro service) of the TLM application is stored.
   (4Test_TodoListManagement) Here the end-to-end tests (Cucumber tests) are stored.

2. Every C&M project subgroup has a "0Doc_TodoListManagement" repository that documents the projects and contains the analysis and design artifacts. The structure of this repository is defined in the C&M documentation guidelines [CM-G-Doc]. The guidelines also define the process that is used to make contributions to the documentation repository.

3. All the information stored in the source folder is loaded into the development environment by using the appropriate Git command (which is "git clone https://git.scc.kit.edu/cm-tm/todolistmanagement/3ms_todolistmanagement.git" in the case that the backend source code is loaded).
   (green check mark) Pipeline status for the last commit. A green check mark means the pipeline was finished successfully, a red cross means the pipeline failed.
   (master) This field is used to select the branch that is currently shown.
   (src) The source code files are part of the folder "src". An example of a source code file which describes the todo domain object is: "src/main/java/edu/kit/tml/springdemo/domain/model/Todo.java".
   (.gitignore, Dockerfile, .gitlab-ci.yml, pom.xml) These are the control files needed for building and deploying the software.
   (README.md) In this file the most relevant information can be found which describes what the software provides, how it can be used, and what the current status and open tasks are.

[CM-G-TLM] Cooperation & Management: GitLab TLM group. https://git.scc.kit.edu/cm-tm/todolistmanagement
The name Jira stems from the Japanese name for Godzilla (the name of a monster) which itself is a reference to Jira's main competitor, Bugzilla (an open source software released by Netscape Communications). It is pronounced "Jee-rah". The software is developed by Atlassian since 2002 and it is written in Java [Wik-Jir].

(1) Jira is used for Project Management, Bug Tracking, Issue Tracking and Workflow [tut-Jir].
(1.1) This functionality is needed in help desk, support and customer services.
(1.2) This includes task tracking, requirement management and process management.
(1.3) Such tools include source control programs such as Git or Version Control Systems (CVS).

(2) Issues are the central concept since a project is a collection of issues and a workflow tracks the lifecycle of an issue.
(2.1) Examples for projects are: software development project, marketing project, employee performance system.
(2.2) Examples of issues are: task or sub-task of a story, bug or defect, helpdesk ticket can be logged as an issue.
(2.3) Examples of statuses an issue are: open, in progress, resolved, reopened, close

(3) The Jira dashboard is the first page that shows up (https://www.tutorialspoint.com/jira/jira_dashboard.htm).
(3.1) These links are projects, issues, boards and create.
(3.2) In this section the issues list assigned to users is displayed. Another example of a main section is "Activity Stream" which display activities done by the user.

EXERCISES: TOOL ENVIRONMENT (II)

(1) ++Overview++
- Chronological planning
- Distribution of tasks
- Documentation

(2) ++Version Control System++
- Local: VCS Client and VCS server run on one system. Disadvantage: collaboration of developers working on distributed systems is not supported.
- Central: Clients running on different systems check out files from the central place.
- Distributed: Client fully mirrors the repository. Each system holds a full backup of all data.

(3) ++Git++
- Git stores snapshots in each client. A snapshot is a picture of all the files of the project. For efficiency reasons for all files that did not change from version to the other only a link to the previous identical file is stored.
- Subversion (successor of the Concurrent Versions System CVS) stores data centrally on the server as a list of file-based changes (since the changes of files are stored as deltas, they are named delta-based version control).

(4) ++Bitbucket / Git Workflow @ C&M++
- Branch: represents an independent line of development for the repository. A branch includes a new working directory, staging area, and project history.
  - Types
    -- Master (origin/master): Before any new branches can be created, the main branch called master is automatically created.
    -- Stable (origin/stable): This branch is optional and contains the last stable version which was deployed.
    -- Feature (feature/featurename): In this branch new functionality is developed
    -- Bug (bug/bugname) In this branch faults in the software are solved. There is not difference how to work with these branches by using Bitbucket and Git operations.

(5) ++Jira++
- Jira: Professional incident management tool which allows to track incidents based on flexible workflows
The right software can only be built when the problem the software should solve is well understood.

(1) At project start the level of ignorance is at its maximum.
(1.1) The learning path is neither linear nor predictable.

(2) It is hard to predict what the development team will learn as the project progresses.
(2.1) It is not advisable trying to force reality to fit into your plan according to a Swiss Army proverb: "When the terrain disagrees with the map, trust the terrain."

https://team.kit.edu/sites/cm-tm/Mitglieder/2-4.Literatur/BDD
A main characteristic of the process behind the Behavior-Driven Development (BDD) is the collaboration of three central roles (business analyst, developer, tester) in the development process. A fourth role is the business owner using the software.

1. The business owner ... business needs.) The first step is similar to the traditional development process.

2. The business analyst ... requirements together.) Two aspects in BDD are new and different compared to the traditional process:
   (i) The collaboration of the business analyst, developer and tester.
   (ii) The structured specification all three roles agree on.

3. The scenarios guide the developer ...) For the developer the specification defines what should be implemented.

4. The tester ... for the tests) For the tester the specification defines what should be tested.

5. The automated tests ... document the application) For the business analyst (and the two other roles) the specification and the outcome of the tests document the software and its current state of implementation. The scenarios can be seen as a low-level technical documentation and provide up-to-date examples of how the system works.

6. The structured collaboration of the roles taking part in the development process leads to a software which does the right thing and which does the thing right.

1. The advantage of BDD compared to the traditional development process is the common specification which prevents misunderstanding and miscommunication, especially between the business analyst and the developer. This specification which is structured according to a defined grammar and keywords is the core of BDD. It consists of a so-called story and scenarios described in natural language.

(Extended, Adapted) The text "6 The tested ..." was extended. The role of the business analyst (i.e., a green instead of a grey stickman) was adapted in step 5.
BDD provides a language Gherkin consisting of specific keywords used to define the features (and the scenarios to test these features) which the software system must fulfill. The proposed template is taken from two sources: The upper feature description part is taken from [DNA-Story] and the lower scenario description part is taken from [WH12].

(1) In the description the business value of the feature should be recognizable. Although Gherkin does not prescribe any structure and use of specific keyword in this part, it is recommended to use the structure "As a – I want – So that" proposed by Dan North.

(2) Behind BDD's Given-When-Then concept a finite state machine can be seen. The structure and the keywords "Given – When – Then" are prescribed by Cucumber.

Developers use lots of different words to describe what they want to build. The following terminology taken from [Sm15:88] is used to provide a clear understanding in this confusing part of software development.

(1) The project vision provides a high-level guiding direction for the project. The problem must be clarified to understand which software system is needed by which users to solve the problem.

(2) The software system must have a measurable, positive impact on the business of the customer the software is built for.

(3) Capabilities give users or stakeholders the ability to realize some business goal or perform some useful task. A capability represents the ability to do something; it does not depend on a particular implementation. For example, "the ability to book a flight" is a capability.

(4) A feature is a piece of functionality that is delivered to the end users or to other stakeholders to support a capability that they need in order to achieve their business goals [:98]. A feature is something that users can test and use in isolation. A feature can deliver business value in itself; once a feature is completed, it could theoretically be deployed into production immediately, without having to wait for any other features to be finished first. The description of the features is generally more effective than describing what the application does.

(5) The typical format of a story is "As a – I want – So that".

(6) Scenarios written in the format "Given – When – Then" make up the core of an executable specification.
By defining the vision, business goals and capabilities, the software system to be developed can be described on a high abstraction level.

(1) The vision should make clear for which purpose the software system is used. In this case the minimal core functionality of the todo list management is to organize the todos in the easiest and quickest way possible. By this vision a Minimum Viable Product (MVP) is described.

(2.1) More customers are attracted by a modern application which is easy to use and which provides intelligent functionality.
(2.2) Offering such an application, customer loyalty will increase and additional products can be sold to the customer.
(2.3) Revenue can also be generated from the advertising shown in the application.
Remark: The business goal pursued by C&M is to practically demonstrate the students the BDD/DDD-based software development with an easy example.

(3) The formulation of the capabilities in this example is close to the features that the (minimal) todo list management system should provide.

MVP Minimum Viable Product
The example is taken from the TodoListManagement (TLM) which allows a user to organize todos in different lists.

1. As the example shows (user acceptance) tests are written in a way that they can be understood by everyone, especially also by non-technical experts who know the domain and users who should use the system.

2. Each scenario describes one example case.

3. This is the reason for Dan North to use the term Behavior-Driven Design instead of Test-Driven Design since the specification of the behavior is the goal and the test is only a means to reach this goal.

4. Since everyone can read and write a test like this a common understanding between stakeholder and developer of what the system should do can grow. Stakeholders are motivated to think about and write down further scenarios that are to be considered.

5. Given I have the possibility to create a todo list The steps should be formulated in a way that no assumptions are made with respect to the GUI (Graphical User Interface). Visualization can be further elaborated by additional UI prototypes illustrating the interaction of the user with the software.
The artifact which describes the Gherkin features according to the BDD is the main outcome of the requirements analysis. Since all further development steps, esp. the modeling of the domain, depend on this artifact it should be documented in an adequate and defined way.

(1) The Gherkin features provide a requirements specification which is an integral artifact of the software development process. Therefore, developers must have access to the features in their IDE (Integrated Development Environment) since they run through the features in order to carry out user acceptance tests of their software implementation. Nevertheless, features are not only relevant for the developers, but also for the stakeholders for which the software is developed. Therefore, an adequate way of documentation of the feature collection must be made available not only for the developers, but also for the stakeholders.

(1.1) The

(2) The specification of features is a creative process which makes it impossible to state complete rules how to go through this process. Nevertheless, best practices can be formulated which give some valuable hints for the process.

(2.1) The first (one to three) features should specify the functionality of a so-called Minimum Viable Product (MVP).

(2.2) A business value will usually not be provided by one single functionality but by a set of coherent functions by which a user can fulfill a specific task.

(2.3) The terms of the ubiquitous language should be used. These terms should be understood by the software developer and the business owner/analyst.

Another best practice is to use a declarative and not an imperative way to express the scenario steps.

MVP Minimum Viable Product
EXERCISES: ANALYSIS

(1) How to Build the Right Software
The proverb concerns the analysis of the requirements the software should fulfill. It means that the plan must always be adapted to reality – and not the other way round.

(2) Behavior-Driven Development (BDD) Process, Roles and Main Artifact
- Three central roles: business analyst, developer, tester
- Scenario description: This specification which is structured according to a defined grammar and keywords is the core of BDD.
  -- Hint: Feature description including stories and scenarios would be the correct term.

(3) Systematic Analysis of Requirements
- Vision: states what the project wants to achieve
- Business goal: describes what the business will get out of it
- Capability: expresses what users and stakeholders should be able to do to deliver these goals
- Feature: represents software functionality to support capabilities
- Story: makes clear which role wants to carry out what for which purpose
- Scenario: expresses a concrete example in a defined format by which a feature is illustrated

(4) Gherkin Feature and Including Scenarios
A story consists of two parts
- Narrative description of a requirement structured according to "As a – I want – So that"
- A set of acceptance criteria presented as scenarios structured according to "Given – When – Then"

(5) Example: TLM Feature
- Scenario: Successful creation of a todo list
  Given I have the possibility to create a todo list
  When I create a todo list with the title "WASA course"
  Then the todo list with the title "WASA course" is created

(6) Dealing With Gherkin Features
- The developer and the stakeholder.
  This leads to the requirement that features should not only be documented in the IDE of the developer but also in external format (e.g. Word or Bitbucket Wiki).
The motivation of Domain-Driven Design (DDD) is taken from Martin Fowler's foreword of Eric Evans's book "Domain-Driven Design – Tackling Complexity in the Heart of Software" [Ev04].

1. A general definition of a domain is: A sphere of knowledge, influence, or activity.
   1.1 A model itself can be defined as a system of abstractions.
   1.2 By this language domain experts and technologists are tight together.

2. There are many things that make software development complex. A deep understanding of the problem domain is in the heart of this complexity.
   2.1 An effective domain modeler can both use the whiteboard during his talk with the customer and work together with a programmer on a Java program.
   2.2 Only few people can do it well – and the competence to build good domain models is not easy to teach.
   2.3 Domain models are not first modelled and then implemented. Instead, an experienced modeler gets splendid ideas from earlier releases of the domain model.
   2.4 DDD teaches a lot of people how to apply modeling concepts to add structure and cohesion to the software development.

3. Based on this vocabulary the activity and the hard-to-learn skill of domain modelling can be explained.

DDD  Domain-Driven Design

To understand how domain modeling can be practically applied it is important to understand where it can be located in the overall software development process.

(1) The domain model is located in the design phase of a software development process. Since the domain model is used to clarify specific aspects of the domain with a domain expert (not with the user) it also contributes to the analysis phase.

(1.1) The domain model concentrates on the domain knowledge that is needed to provide the functionality of the software system. A precise and semi-formal description as a domain model is the main challenge of the design of the software system.

(1.2) Only one layer of the overall software architecture, the so-called domain logic layer is covered by the domain model. Since this layer can be seen as the heart of a software system it provides the kernel of the system. A major goal of domain modeling is to keep the model (of the domain) and implementation (of the domain logic layer) close together.
According to the overall BDD/DDD-based development process in the subsequent step the domain logic contained in features should be added to the domain model.

Note: In what follows the term feature is synonymous to a Gherkin feature.

(1) A feature specifies a specific part of the business logic from the viewpoint of a user who applies the functionality in a certain role ("As a ..." in the story) and in certain situations ("Given ..." in the scenario).

(1.1) The other part is the application logic. It is important to notice that in the DDD approach the business logic is divided into the domain logic and the application logic. The domain logic is application agnostic which means that this logic is relevant for other applications of this domain. Therefore, the domain logic is the re-usable part of the business logic appearing in a feature.

(2) Like the process of feature specification the derivation of domain knowledge from the features is a highly creative process. Nevertheless, some best practices do exist.

(2.1) By answering this question those terms of the feature should be identified that are relevant for the understanding of the user requirement.

(2.2) Only those aspects identified by Question 1 should be added to the domain model which are not specific to the regarded application.

(2.3) Three types of relationships can be distinguished: general associations, containment (composition, aggregation), inheritance
The derivation of domain knowledge is shown with a part of the TLM feature that was introduced on a previous page.

(1) The approach is to go through the lines of the feature and identify those terms which are presumably relevant for the understanding of the domain and the functionality related to the domain, i.e. the domain logic.

(2) (3) The list should be seen as a candidate of terms that might become part of the ubiquitous language of the domain. For those terms which certainly will be part of the language, a definition should be formulated.

(4) The relation view is specified as a UML class diagram as illustrated in what follows.
In DDD, entities and value objects are the fundamental domain objects.
(1.1) The graphical notation is a rectangle which expresses the stereotype "<<class>>." 
(1.2) This notation is also used for all other modeling elements, such as the structuring element of a package that has been introduced before.
(Hazard Warning, PointOfInterest, Thesis) These are examples of domain objects from different domain models and software systems.

(2) The term collection is an artifact that is created in the analysis phase of the C&M software development process.

(3) The following types of relationships are standardized by UML which can be used in a class diagram.
(3.1) The association is the most general relationship between two domain objects DO1 and DO2. The association can be named (e.g. defines, is responsible for, creates). An arrow indicates the direction of the association.
(3.2) The aggregation is a stronger version of association and it is used to indicate that a domain object DO1 contains another domain object DO2 [MH06:86]. The lifetimes of the containing domain object (DO1) and the contained domain object (DO2) do not depend on each other. An aggregation is modeled by using an empty diamond arrowhead.
(3.3) The composition is as the aggregation a containment relationship, i.e. domain object DO1 contains another domain object DO2 [MH06:86]. In contrast to an aggregation, the lifetime of the containing domain object (DO1) depends on the lifetimes of the contained domain object (DO2). This means that the contained domain object (DO2) cannot exist without the containing domain object (DO1). A composition is modeled by using a filled diamond arrowhead.
(3.4) Generalization which is otherwise known as inheritance is used to describe that a domain object DO1 is a generalization of a domain object DO2. Therefore, the sub domain object DO2 inherits all properties from the super domain object DO1. The generalization relationship can be denoted as "is a" or "is a type of". It is modeled by an empty generalization arrow.

DO Domain Object

So far, the domain objects have been identified. In a next step, domain modeling concepts are applied in order to build the domain model. On this page, this step is shown with the example of a part of the TodoListManagement (TLM) domain model.

Note: This part belongs to the so-called "bounded context" TodoManagement which is introduced in one of the next pages.

1. The stereotypes (e.g. <<shared entity>>, <<entity>>, and <<value object>>) are defined by the Domain-Driven Design (DDD) concepts. Each domain object belongs to one or more such stereotype.

2. Entities are domain objects that can be clearly identified. They can contain attributes and methods.
   (TodoList, Todo) These are the two main domain objects of the example.
   (2.1) If two Todo domain objects have the same status and contain the same information, they are the same Todo domain object.
   (2.2) Example of an attribute and a method of the Todo entity are content and changeContent().
   (2.3) If objects are shared between bounded contexts, the stereotype is annotated with "shared".
   (<<shared entity>> TodoList) Since a TodoList entity is shared between the bounded contexts "TodoManagement" and "UserDirectory" it is modeled by using the stereotype <<shared entity>>.
   (<<shared entity User) The entity User is not part of the domainTodoListManagement but it is provided by the domain IAM (Identity and Access Management).

3. Value objects represent a descriptive aspect of the domain for which is only important what they are, but not who or which they are.
   (3.1) This lack of constraint gives design freedom and simplifies to optimize performance (e.g. by copying and not sharing a value object).
   (<<value object>> Color) Since it is not important which "red" was selected this information of the domain model is modeled as a value object.

Remark: Further stereotypes like aggregates or domain functions do exist but they are not explained here.

IAM Identity and Access Management
The domain model emerges from two types of modeling, the tactical modeling and the strategic modeling.

(1) The tactical modeling corresponds to the well-known proceeding of modeling knowledge as conceptual class diagrams added by dynamic diagrams to model certain time-related aspects. The result from the tactical modeling can be seen as the core content of the domain model.

(1.1) Four types of views are distinguished. Only the process view allows to model dynamic aspects, whereas the relation view, domain object view, and interaction view focus on the static aspects.

(1.2) The relation view corresponds to conceptual class diagrams as will be shown with an example.

(2) The strategic modeling is the real new part of domain modeling and the result, a so-called context map, provides the link to the microservice architecture.

(2.1) The structuring of the domain knowledge (as it is described by tactical modeling based on the four view types) is carried out in two steps: The whole domain is structured into subdomains and the subdomains are structured into so-called bounded contexts. To find an adequate structure coupling that domain knowledge which is coherently connected "from its nature" is one of the main challenges in domain modeling. The term "bounded context" is derived from the approach of a Domain-Driven Design (DDD, [Ev04]). A bounded context form the architectural perspective can be seen as a microservice candidate. This is the reason why a domain model serves as a blueprint for the microservice architecture which implements the model in the domain logic layer of the layered architecture. Domain objects can be shared by bounded contexts. These objects are called shared entities.

(2.2) DDD does not prescribe a concrete representation of the model. At C&M, the modeling elements from the Unified Modeling Language (UML, (CM-W-UNI)) are used to model the results from the strategic (and also from the tactical) modeling.

(<<subdomain>> SubDomain1) The subdomain is modeled as a standard UML package. "SubDomain1" stands for the name of the name of the subdomain formulated in UpperCamelCase.

(<<bounded context MicroserviceCandidate1>> MicroserviceCandidate1) To be able to model a bounded context in UML a new stereotype <<bounded context>> is introduced. The UML modeling symbol of a package is used as graphical representation. "MicroserviceCandidate1" stands for the name of the bounded context formulated in UpperCamelCase.

(<<shared>> SharedEntity, uses) This is the way how domain models shared between bounded contexts. In the example, the bounded context MicroserviceCandidate1 shares the domain object "SharedEntity" with a bounded context being part of SubDomain2.

**DDD** Domain-Driven Design
**UML** Unified Modeling Language

[CM-W-DOM] Cooperation & Management: DOMAIN MODELING, WASA Course Unit. https://team.kit.edu/sites/cm-tm/Mitglieder/2-1.WASA


The context map is a diagram which can be seen as the starting point of domain modeling since it provides the domain structure in which the tactical domain models are placed in a systematic and consistent way. A context map does not focus the domain content (i.e. domain objects and its relationships), but a coherent structure of the business functionality.

(1) The modeling elements that can be used in a context map are introduced in what follows.

(1.1) By introducing subdomains the overview of a larger domain can be kept. Each subdomain should have a size which can be easily overlooked.

(1.1.1) Core subdomains are in the center of the context map since they are the building blocks of the domain which is to be described. The TodoManagement subdomain is part of the core. Supporting and generic subdomains provide some additional functionality that are needed. An example of such a subdomain in the TodoListManagement domain is the IAM subdomain.

(1.2) (1.2.1) A bounded context packages coherent functionality in a component. In the software architecture derived from the context map a bounded context is a candidate for a microservice.

(1.1.1) Core subdomains are in the center of the context map since they are the building blocks of the domain which is to be described. The TodoManagement subdomain is part of the core. Supporting and generic subdomains provide some additional functionality that are needed. An example of such a subdomain in the TodoListManagement domain is the IAM subdomain.

(1.2) (1.2.1) A bounded context packages coherent functionality in a component. In the software architecture derived from the context map a bounded context is a candidate for a microservice.

(1.3) The modeling elements of a foreign bounded contexts and an anti-corruption layer allow to model functionality from foreign domains.

(1.4) If there is an overlap of a part of the domain model in two or more bounded contexts it might make sense to model this overlap in a shared kernel.

Since a shared kernel can be seen as a kind of directory and not as a running software component, it is modeled a a package. In the example, the shared kernel TodoInformation describes the domain logic with respect to the todos which are shared by two bounded contexts, TodoManagement and Notification.
(1) Not only the modeling elements (e.g. subdomain or bounded context) but also the relationships between these modeling elements are described by a context map.

(1.1) As the modeling elements itself the relationships between them are specific to a context map. In the following each of the four specific relationships are described with the exemplary context map.

(<<conformist>>) This type of relationship is a directed association between two bounded contexts. The bounded context building the source of the directed association is called upstream (in this case, UserDirectory) and the destination (UserDirectoryACL) is called downstream. In a conformist relationship the downstream must conform to (i.e. strictly obey) any changes made by the upstream even if the model and the implementation from the upstream does not fit to the downstream.

(<<customer/supplier>>) A directed association between two bounded contexts in which the downstream is the supplier (in the example, TodoManagement) and the upstream is the customer (Calendar). The association means that the upstream team which develops the Calendar must coordinate planned changes with the downstream team which develops the TodoManagement.

(<<partnership>>) An undirected association between two bounded contexts. A partnership association means that the two teams responsible for the bounded contexts depend on each other and must, therefore, cooperate closely. In the example, this is the case for the bounded contexts TodoManagement and Notification although the shared kernel TodoInformation might lead to a looser coupling.

(<<import>>) This directed association is used between a bounded context and a shared kernel which needs this part of the domain model described by the shared kernel. In the example, the shared kernel is imported – and in this way, shared – by the two bounded contexts TodoManagement and Notification.

(Packet Grouping) This type of association allows to model packet groupings by defining subpackages of a package. In the example, User Directory ACL is a subpackage of TodoManagement [Spa-Nes].

(1.1) ++Domain Modeling and Software Development++
- The domain model provides the functionality of the domain logic layer.

(1.2) ++From Features to the Domain Model++
- Features contain domain knowledge which should be formalized in a domain model according to DDD
- Features specify that part of the business logic of the software system that should be implemented in the domain logic layer and which should be separated from the application logic layer.

(2) ++Example: TLM Feature and Including Domain Knowledge++
- The approach is to go through the lines of the feature and identify those terms which are presumably relevant for the understanding of the domain and the functionality provided by the domain, i.e. the domain logic.
- In the TLM examples these are
  -- Objects: user, todos, todo list
  -- Activities: manage todos, create a todo list, add a todo

(3) ++Modeling the Domain Content++
- Composition: filled diamond arrowhead
- Aggregation: white diamond arrowhead
- Difference: In contrast to an aggregation, the lifetime of the containing domain object (DO1) depends on the lifetimes of the contained domain object (DO2)
- Example: TodoList is a composition of Todos -> Deletion of the TodoList leads to the deletion of all contained todos
  TodoList is an aggregation of Todos -> Deletion of the TodoList does NOT delete the contained Todos

(4) ++Tactical and Strategic Modeling++
- Strategic modeling provides the overall structure of the domain model by defining subdomains and bounded contexts.
- Relationship: Bounded contexts are the candidates for microservices.

(5) ++Example: TLM Context Map++
- <<conformist>>: This type of relationship is a directed association between two bounded contexts. In a conformist relationship the downstream must conform to (i.e. strictly obey) any changes made by the upstream even if the model and the implementation from the upstream does not fit to the downstream.
- <<customer/supplier>>: The association means that the supplier must coordinate planned changes with the customer.
- <<partnership>>: A partnership association means that the two teams responsible for the bounded contexts depend on each other and must, therefore, cooperate closely.
Besides the domain model the Application Programming Interface (API) is an important design artifact. The API is specified in the second phase of the design phase. Both artifacts, the domain model and the API, are strongly related.

(1) As has been shown, the features as the central analysis artifact are a good basis to determine the domain objects.

(1.1) For instance, an explorative prototype, based on the user requirements, shows which interactions are needed.

(1.2) The output of this step is called "candidates". For each candidate, the software architect has to choose if the domain object is exposed or not. They are transformed into "real" resources by a transformation described in the following steps.

(2) Based on the resource candidates the resource types are determined.

(2.1) The semantics of each resource type are described in more detail in [Gi18]. The resource types used in this approach are based on the current research results in this field:

- PrimaryResource: A resource which represents a domain object and which is directly addressable. A primary resource can be derived from an object of the domain model that has no inbound associations.
- SubResource: A resource which is a logical part of an other resource and which is directly addressable. These are resources which have one or more inbound associations.
- ListResource: A resource which contains more resources as list elements. Domain object with multiple instances or which is in a 1..n relationship and which represents the parent.
- InformationResource: A resource that can not dereference itself but is associated with an identity. A domain object which is represented as a value object and which has an inbound association.
- ActivityResource: A resource that represents a business process. Domain object which is modeled as a domain service.

(3) The semantics of each resource type as well as the domain objects and their associations allow to derive heuristics and transformation rules.

(3.1) The heuristics identify a resource type based on the domain model. But only the resource candidates have to be considered for identification. Each heuristic is linked with a transformation rule that transforms a resource candidate into a "real" resource by assigning a stereotype from a dedicated UML profile called resource orientation and placing it into a so-called resource model. For example, the heuristic for a primary resource is the following: A domain object is an entity and has no outgoing associations. The transformation rule for this heuristic consists of these steps: The first step is to copy the entity into the resource model and remove all existing stereotypes, except the identifier. The next step is adding the stereotype "PrimaryResource" to the copied entity. The resource model applies the previously mentioned UML profile and extends the UML with concepts of REpresentational State Transfer (REST). It is important to note that also behavioral aspects can be modeled within the resource model. To keep the model lightweight, CRUD operations were not modeled and instead considered as implicitly given.

API Application Programming Interface
REST REpresentational State Transfer

The initial resource model (called RM1 in the refinement process) is created based on the underlying domain model, as showed on ++Example: NavSG Resource Model++. In addition to the structural aspects, the behavior-specific aspects of the domain are also transferred in a suitable form so that the behavior is also reflected in the resource model. For a complete and detailed description of these steps see [Gi18].

1.1) Each resource is supplemented by a Uniform Resource Locator (URL) component based on its assigned resource type. By following links between resources, the complete URL can be derived. For the creation of corresponding components, linguistic patterns and best practices are used to increase the usability and ultimately the reusability. An example is showed on slide ++Example: NavSG Resource Model++. 

1.2) After adding addressing information to RM2, the interactions with a resource from the client is considered. The offered behavior by the domain has to be mapped onto a uniform interface which is a constraint of the architectural style REST. To simplify the mapping, existing patterns were identified and assigned to the individual resource types. Regarding the resulting influence, each pattern was extended by their influences on quality characteristics. Since the web API is the link between a client and a resulting microservice, the patterns also show their effect on different interaction participants. After this steps, RM3 has been created.

1.3) The creation of representation schemas is the next logical step. The schemas define the structure of the request/response body. In addition to listing the required and optional attributes, they can also be used to define valid attribute values or to show sample values of some attributes. Based on these schemas, mock values can be generated that may be suitable for web API testing.

1.4) The last step represents the versioning of the resource model and finally the web API. Version information using Semantic Versioning (SemVer) is added to RM3. Since this is only a small addition from the technical perspective it is still an RM3. The representation of this version information in the web API is treated by patterns. Besides, a process is introduced on how to change the version information. According to the versioning process more than one resource model may exist at the same time. RM3 is the final result and an example is showed on ++Example: NavSG Web APIs (Excerpt)++.

2) Each refinement step can be found in many of today’s web APIs. The refinement process combines the expert knowledge and forms a systematical and traceable process to build well-designed web APIs for the modeled domain. Web APIs are an important asset of digital transformation and have to be designed with care.

API Application Programming Interface  
RM Resource Model  
REST REpresentation State Transfer  
URL Uniform Resource Locator
The TLM resource model is derived from the TLM domain model which was described on a previous page ++Example: TLM Domain Model++.

(1) Not every domain object is a resource candidate. That is because not every domain object can or should have public access. There is a variety of reasons for that.

(1.1) In the TLM application Todo might be completely accessed via methods from TodoList. In this case the domain object Todo must be made public and it would not become a resource.

Remark: In the API introduced in one of the next pages Todo is handled as a resource. This means that a different design decision has been taken.

(2) The resource candidates are assigned stereotypes. Those types can be PrimaryResource, SubResource, ListResource, InformationResource or ActivityResource.

(2.1) In TLM User is a PrimaryResource since it is a domain object without inbound relations.

(2.2) Every object which is associated with a PrimaryResource is a SubResource. Further objects with inbound relationships are also SubResources. TodoList is a SubResource of User.

(2.3) A resource which can have more than one instance is categorized as a ListResource. That is the case for both resources, User and TodoList, in the example.

(3) The path by which the resource can be accessed is built using an algorithm described in [Gi18, Hü18].

(3.1) The starting point of the path is a PrimaryResource. Afterwards all SubResources of that resource are handled.

(3.2) If it is a List Resource a plural "s" is extended. Otherwise the singular is concatenated.

(3.3) If the resource was an entity in the Domain Model, it can be identified by an ID. The ID is appended to the path as well. It is denoted with curly braces which means that its value is defined at runtime.

++PrimaryResource ...<identitifier>++ The class diagram shows the resource model of TLM. The resources have a stereotype (like PrimaryResource), attributes, methods and a path tag which denotes the path under which the resource can be accessed.

(1) The resource model itself is the starting point for deriving a web API. The resource model does not contain any design decisions regarding the technology or platform. Therefore the developer can choose any technology during the further development. The technology decisions can be made independently by the respective development teams.

(2) The choice of the appropriate protocol is based on their characteristics which have to comply with the needs of the resource model. In doing so the application level protocol should fit to the service consumers and if necessary their non-functional requirements. Some issues for this choice are the connection type (synchronous or asynchronous), the required exchange format and required methods.

(2.1) Safe methods are methods that do not modify resources. Idempotent methods are methods which outcome won’t change no matter how often they are called. HTTP safe methods are GET, OPTIONS and HEAD, while idempotent methods are GET, PUT, DELETE, OPTIONS and HEAD.

(2.2) Examples of protocols for accessing resources are HTTP and COAP (Constrained Application Protocol). While COAP is mainly used for machine to machine communication in unreliable environment, HTTP is the most used application layer protocol.

(3) To make further use of the web API it can be transferred into a machine interpretable specification format. By this independent testing of the backend and frontend, automated building of the API and an automated code template creation can be enabled. Examples for machine interpretable API specifications are Open API (formerly Swagger) or RAML. For practical use at least the following has to be done.

(3.1) The meta information includes its version, general information (title and description) as well as contact information for support queries.

(3.2) Furthermore, reusable components have to be described. These components could be schemas, responses and request parameter. For instance, a query parameter description consists of a name, a short explanation, and a data type.

(3.3) An interaction can be composed of the reusable components.

(4) If the API is used in a large environment, it is very important that it is as stable as possible and meets a certain quality standard. Otherwise the API becomes less useful and generates more costs as actual benefit. Therefore an review can be performed to check its quality criteria. Possible reviewers are the API consumers as well as dedicated experts in the API design field. If there are required actions, it is necessary to populate it in the resource model and propagate it through the entire process again.

COAP Constrained Application Protocol
HTTP Hypertext Transfer Protocol
RAML RESTful API Modeling Language
OpenAPI is one of the most broadly accepted API description formats.

(1) The OpenAPI Specification is a manufacturer-independent defacto standard. Examples of other description formats are API Blueprint or RESTful API Modeling Language (RAML).

(1.1) The OpenAPI Initiative was founded in 2015 as an open source project under the Linux Foundation [OAI-FAQ]. A major driver was SmartBear which donated the Swagger Specification to the OpenAPI Initiative

(1.2) Examples of such companies are Google, Microsoft, IBM, Apigee.

(2) JSON and YAML are both markup language which are based on the same concept (lists and scalars). Each JSON document is a valid YAML document. The first line of the document (i.e. swagger: '2.0') defines that this is an OpenAPI 2.0 specification.

(2.1) In the "info" JSON/YAML object, metadata of the OpenAPI specification can be found, such as the "title" or the "contact" person of the service. All information is described by a cascaded list.

(2.2) In "paths" the (relative) URLs of the endpoints and the HTTP methods are indicated. In "responses" the results including the HTTP status codes are described.

(2.3) Reusable parts are objects that consist of a set of properties. To be able to reuse the object a name is defined which is use as a reference.

<table>
<thead>
<tr>
<th>JSON</th>
<th>JavaScript Object Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAML</td>
<td>RESTful API Modeling Language</td>
</tr>
<tr>
<td>YAML</td>
<td>YAML Ain’t Markup Language</td>
</tr>
</tbody>
</table>

The YAML document of the TLM OpenAPI specification is available on the Bitbucket repository [CM-B-swa].

(2. info) The “info” part contains general information (such as title, description, contact) that describe the TLM API.

(1) (5. host ...) This is a central part of the API path [SWA-API] which is built as follows:
<operation> (e.g. GET) <scheme> (HTTP) <host> (todo-management.cm.tm.kit.edu) <path> (todo-lists)

(2) (7. paths:) An endpoint corresponds to a resource that the API exposes. An examples of endpoint or resource of the TLM API is "/todo-lists" (line 8.).

(3) (9. get:) (22. post:) As can be seen in the example, a single path can support multiple operations. In this case, GET /todo-lists and POST /todo-lists.

(4) (10. tags:) (23. tags:) For example, SwaggerUI (see also page ++Example: TLM API++) uses "tags" to group the displayed operations. Often the endpoint names are used as tags which is also done in this example.

(5) (13. operationId: getTodoListUsingGet) (26. operationId: createtodoListUsingPOST) These are the names of the corresponding methods in the backend code. Usually, the (optional) operationIds are automatically generated, i.e. when the OpenAPI specification is generated from the code).

(6) (27 parameters:) If an operation (such as POST or PUT) has parameters in its request these are described as body parameters ("- in: body", line 28). In the example, the parameter "newTodoList" is specified as body parameter of the operation POST /todo-lists by which a new todo list is created.

(32. schema:) (33. Sref: '#/definitions/TodoList-GenericRequest') The schema describes the ... The "Sref" indicates that for modularity purposes this description is part of the "definitions" section of the OpenAPI specification.

(7) (14. responses) (34. responses) In this part of the "paths" object the possible results of the get operation are described. In the example a successful (200, line 15) is defined.

(Sref ...) This is a reference to the "definitions" part of the OpenAPI specification where the different data structure related to the domain objects (e.g. Todo or TodoList) are defined.

For the specification of the REST-based TLM API the tool Swagger was used [CM-S-TLM]. The page shows screen dumps taken from the tool SwaggerUI.

(GET /todo-lists) This HTTP GET operation of the TLM API returns all todo lists and all todos for each todo list. The complete request URL is: "https://todo-management.cm.tm.kit.edu/todo-lists"

When clicking on a list element representing an operation, details of this operation is presented. The screen dump on the right hand side shows the details of the TLM API operation "GET /todo-lists".

(Parameters) Since is TLM API operation is based on a HTTP GET it has no input parameters.

(Responses) The response of a successful request (i.e. "Code 200" meaning "OK") is listed in the text field underneath "Example Value". 

("id", "title", "color", "todos") These are the elements which define a todo list. 

("number", "position", "content", "done", "description") A todo is defined by these five elements.

(POST /todo-lists/ ...) (GET /todo-lists/{id} ...) In what follows all REST-based operations of the TLM API are listed. For each operation a short text describes what is performed by the operation. For example, by sending a request "POST/todo-lists" a new todo list is created and a request "GET /todo-lists/{id}" finds the todo list with the specific "id" (or provides an error response if a todo list with the "id" does not exist).

1. In which way do domain objects contribute to the design of an API?
2. Which resource types can be distinguished and what are examples taken from the TLM application?
3. What are the different steps of the refinement process?
4. What is OpenAPI and what are its main parts?
5. What is an example of a REST operation from the TLM API?

++Derivation of the API from the Domain Model++
- Domain objects are candidates for the resources which are the main building blocks of an API (= resource-oriented web API)
- Those domain objects which should be accessed via the service interface become API resources

++Example: TLM Resource Model++
- PrimaryResource: A resource which represents a domain object and which is directly addressable. A primary resource can be derived from an object of the domain model that has no inbound associations.
  -- TLM example: User
- SubResource: A resource which is a logical part of another resource and which is directly addressable. These resources which have one or more inbound associations.
  -- TLM example: TodoList
- ListResource: A resource which contains more resources as list elements. Domain object with multiple instances or which is in a 1..n relationship and which represents the parent.
  -- TLM example: User, ListResource
- InformationResource: A resource that can not dereference itself but is associated with an identity. A domain object which is represented as a value object and which has an inbound association.
- ActivityResource: A resource that represents a business process. Domain object which is modeled as a domain service.

++Refinement of the Initial Resource Model++
- Assigning addressing schema heuristics and transformation rules for addressing
- Adding interactions by using a pattern language
- Creating representation schemas for resources
- Adding version information for subsequent changes

++API Description Format OpenAPI++
- OpenAPI is an API description format published by the OpenAPI Initiative
- JSON or YAML documents
- Main parts
  -- info: metadata such as the "title" or the "contact" person of the service. All information is described by a cascaded list.
  -- paths: the (relative) URLs of the endpoints and the HTTP methods are indicated. In "responses" the results including the HTTP status codes are described.
  -- definitions: reusable parts are objects that consist of a set of properties. To be able to reuse the object a name is defined which is use as a reference.

++Example: TLM API++
- GET /todo-lists: returns all todo lists and all todos for each todo list
- POST /todo-lists: a new todo list is created
- GET /todo-lists/[id]: finds the todo list with the specific "id" (or provides an error response if a todo list with the "id" does not exist).
In the subsequent implementation phase the functionality as is was analyzed and designed is coded based on a microservice architecture and frameworks supporting the frontend and backend implementation. All implementation work is carried out in a test-driven manner based on unit tests and user acceptance tests which are defined by the scenarios being part of the Gherkin features. In the deployment phase the tested application is deployed in a Docker-based execution environment. As the microservice approach is strictly followed, the concept of Continuous Integration/Deployment/Delivery (CI/CD) can be supported.

(1) The BDD/DDD-based development process (Behavior-Driven Development / Domain-Driven Design) is an agile approach which means that implementation of the software should start as soon as possible.

(1.1) A (Gherkin) feature specified according to the BDD approach usually is the part of the software which is incrementally designed, implemented and tested.

(1.2) This part of the implementation concentrates on the domain logic being part of the domain logic layer of a microservice architecture. It can be carried before the APIs of the architecture are designed.

(1.3) Unit tests check if the implementation of the domain model fulfills all domain-specific assertions included in the domain model.

(2) This is the main part of the implementation work. A prerequisite for the start of this implementation is the existence of the API specifications of the microservice architecture.

(2.1) User acceptance tests are carried out by testing the scenarios occurring in the feature. In order to be able to test a feature the step definitions occurring in the features must be implemented.
The implementation of a web application can be separated into two parts, the frontend implementation (client side) and the backend implementation (server side).

1. Since the tasks that have to be carried out during frontend and backend application require different competences and the use of different tools it makes sense to build two separate teams.

2. The backend for frontend (BFF) service bridges the gap between the requirements of the frontend to the backend and the available backend services.

3. The domain model and the ubiquitous language define the common language of both teams. If changes of the domain model become necessary during the development a clear agreement between the two teams is necessary.
The figure illustrates the overall implementation architecture which is based on the MVC pattern (Model-View-Controller). The focus of the illustrated architecture is on the communication of a Backend-For-Frontend (BFF) microservice with Backend (B) microservices (Backend Microservice 1 and Backend Microservice 2) using the Spring technology.

Note: Each of these microservices (BFF and B) run on different web servers which communicate via HTTP.

(View) The view in a microservice architecture is realized by the frontend. Technologies that can be used for the frontend implementation are Angular and Bootstrap.

There are two types of controllers in the architecture:
(BFF Controller) This is the controller directed to the frontend. It supports the frontend by transforming data for the view.
(B Controller) This is the controller related to a microservice. It starts the request and orchestrates the services.

There are also two types of models in the architecture:
(BFF Model) This is the data model that is underlying the BFF microservice.
(B Model) Each B microservice has its own data model and its own database to persist the data specific to the one microservice.

B Backend
BFF Backend For Frontend
MVC Model View Controller
POJO Plain Old Java Object

The example taken from [Ir18] shows how the TodoList which is a central TLM domain object is systematically transferred into its implementation. Several JavaX annotations (e.g. @Entity, @Column, @OneToMany [Ora-JEE]) are used to solve the persistency issues of the implementation.

<<shared entity>> TodoList) The TodoList is a <<shared entity>> which means that it is owned by one bounded context and used by at least one other bounded context. This domain object is defined by three (alphabetically ordered) attributes, color, nextTodoNumber, and title, and a number of methods (e.g. createTodo).

(1 @Entity) This JavaX annotation is used for persistency purposes. The class is marked as an entity when it is written into the database.

(4 @Column ...) Specifies the mapped column for this property.

(11 @OneToMany ...) The class TodoList can be mapped to many Todos. The annotation @ManyToOne works the other way round.

(14 @OrderBy ...) The mapped objects are ordered by ascending positions.

(20 public TodoList (...) The constructor has no parameters. New created TodoLists are empty.

(26 createTodo(String content) ...) This is the first method provided by the class TodoList. In contrast to the model of the domain object the methods are not alphabetically, but logically ordered.


A core part of the software development according to the Domain-Driven Design (DDD) is to make sure that the model and the implementation are the "same". A change of one of the two artifacts must lead to an equalizing change of the other artifact. (1.1) The domain logic is only found in the domain logic layer. The other layers of the microservice architecture are separated from the domain logic as to promote reusability. The application logic ensures the access to the domain layer.

<<shared entity>> TodoList …) TodoList is a central domain object of the TLM domain model as described on previous pages.

(43 public Todo createTodo ...) The method createTodo takes the parameter content that is required to create a new Todo.

(46 throwExceptionIfDuplicate ...) Two identical Todos in a TodoList are not allowed. Hence the new Todo is checked against duplicates.

(47 Todo todo = new Todo...) A new Todo is created and added to the todos list. Then the newly created Todo is returned.

This page gives an overview of the modular architecture the Angular framework is based on.

(1) An Angular module is a class decorated with NgModule. Decorators (such as @NgModule) are functions that modify JavaScript classes.

(1.1) Each Angular App has at least one such Angular module class, the root module (conventionally called AppModule). The Angular module system is complementary to the JavaScript module system in which each file is a module and all objects defined in the file belong to that JavaScript module. The two types of module systems must be well separated which is not that easy since both use the same vocabulary of "imports" and "exports".

(Library Module) A library module is a collection of JavaScript modules. Angular ships with several such library modules, called Angular libraries each beginning with the @angular prefix. Examples are: @angular/core, @angular/platform-browser.

(Metadata) Describe how a class needs to be processed. They are inserted into classes with decorators.

(2) A component controls parts of the screen called view and each defined by a template. The component supports the view by providing presentation logic which is implemented as a class.

(3) A template is a form of HTML that tells Angular how to render the component.

(3.1) Directives give instructions on how templates should be rendered. There are three types of directives in Angular. First, there are components that are directives with templates. Second, there are structural directives, such as ngIf and ngFor, that manipulate the DOM layout by adding and removing DOM elements. There are also attribute directives, such as ngStyle, that change the appearance or behavior of elements, components or directives.

(Property Binding, Event Binding) The mechanism for coordinating parts of a template and parts of a component is supported by the data binding concept of Angular. Data binding is conducted by adding markup bindings (specific forms) to the HTML template to tell Angular how to connect the component with the DOM.

(4) A service can be nearly everything, such as a logging service or data service. Or it can be a tax calculator or an application configuration. There is nothing specifically Angular about services. Angular has no definition of a service. There is no service base class, and no place to register a service.

(4.1) (Injector) Angular uses dependency injection to provide new components with the services they need. When Angular creates a component, it first asks an injector for the services that the component requires.

The main structure of the Angular architecture is demonstrated with the example of the part of the TLM frontend which presents the data and functions related to a todo list.

(TodoList) This component handles everything a todo list offers.

(HTML 5 & Bootstrap) The HTML defines which elements are shown in the browser. It includes Angular syntax like "*NgFor" which is used for addressing each list element. The todolist.component.html uses the Bootstrap element "card".

(CSS) Elements can use custom style sheets. These are defines in the "*.css" classes.

(TodoListComponent) The file todolist.component.ts handles the actions from the HTML elements. For example, if a button on the web page is pressed, it is forwarded to this file. If an action requires a request, this is forwarded to the service.

(TodoService) The TodoService is the interface to the BFF. It forwards every request.
(1) ++Overview++
++Separation into Frontend and Backend Implementation++
- A prerequisite for the start of this implementation is the existence of the API specifications of the microservice architecture.

(2) ++Spring-Based Backend Architecture++
- The pattern is called MVC (Model-View-Controller).
- Backend components: BFF Controller, BFF Model, Microservice Controller, Microservice Model

(3) ++Example: TLM Domain Model and Implementation++
- @OneToMany: The class TodoList is related to many Todos since a todo list cant contain any number of todos.
- @ManyToOne: A user can have many, i.e. any number of todo lists.

(4) ++Architecture of an Angular Application++
- Component: controls parts of the screen called view which are defined by a template. The component supports the view by providing application/presentation logic which is implemented as a class.
- Template: is a form of HTML that tells Angular how to render the component.
- Relation: The mechanism for coordinating parts of a template and parts of a component is supported by the data binding concept of Angular. Data binding is conducted by adding markup bindings (specific forms) to the HTML template to tell Angular how to connect the component with the DOM (https://angular.io/guide/architecture#data-binding).
The microservice system architecture consists of a frontend, backend-for-frontend (BFF) and a backend. Angular is used as implementation technology for the frontend and Spring as implementation technology for the BFF and the backend microservices.

(1) A microservice architecture divides the software system into three distributed subsystem (or system parts) frontend, BFF, and backend. These systems communicate via resource-oriented web APIs.

(Diagram) The Unified Modeling Language (UML) provides the deployment diagram to model this physical view on the software application [MH06] as shown on the right hand side of the slide. The three-dimensional boxes represent the UML modeling element of a so-called node (which describes a computer system). On a node contain so-called artifacts are located which describe software programs running on that node.

(2) Angular is a popular open-source framework to implement the frontend based on HTML (HyperText Markup Language) and TypeScript which is an extension of JavaScript.

(.angular-cli.json) A JSON file (JavaScript Object Notation) which contains configuration information of the whole project setup.

(app.module.ts) A Transcript file which defines the modules that are bootstrapped by the framework at its start.

(3) Spring is one of the most popular open-source development framework for Java-based application servers. Many available open JavaEE technologies, such as Java Persistence API (JPA) are used and extended by Spring.

(4) The three system parts of the microservice architecture (usually) run on distributed client and server systems which communicate via the Web. The way how specification of the web APIs are specified is described in a previous chapter.

<table>
<thead>
<tr>
<th>HTML</th>
<th>HyperText Markup Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>JPA</td>
<td>Java Persistence API</td>
</tr>
<tr>
<td>JSON</td>
<td>JavaScript Object Notation</td>
</tr>
<tr>
<td>UML</td>
<td>Unified Modeling Language</td>
</tr>
</tbody>
</table>

Continuous Deployment and Build Pipelines

(1) Continuous Deployment (CD) treats each and every check-in as a release candidate.
(1) The software has to go through multiple stages from check-in to production.
(2) The multiple stages inside a build are modeled as build pipeline.

(2) A CD tool provides specific functionality to support the build pipeline.
(1) Define and visualize the pipeline.
(2) Support manual actions when stage transitions are not fully automated.

(3) One pipeline per microservice should be the goal.
(1) An exception might be acceptable in the initial project phase.

(1) Continuous Deployment (CD) gives constant feedback on the production readiness of each and every check-in [Ne15].

(1.1) The reason for the stages is that there is a sequence of the needed activities. An example is to conduct fast tests before slow tests (more general: to conduct different types of tests in a specific sequence).

(1.2) The figure on the right side shows the different stages of a build pipeline on an abstract level.

(2) The complexity of such a CD tool is often underestimated. It does not make sense to try to hack and extend Continuous Integration (CI) tools to make them do CD.

(2.1) The tool supports the developer to keep a good overview of the process and the current status.

(2.2) In the case of a manually executed "User Acceptance Test" (UAT) the tool supports this process by showing the next available build ready to be deployed in the UAT environment, deploy it and show the test result in the visualized build pipeline.

(3) The service is an artifact (coming with different, platform-specific shapes) that is created and moved through the pipeline.

(3.1) In the initial phase, the service boundaries are often not clear and changes across service boundaries are more likely. As soon as a service API stabilizes it should be moved into an own build.

---

CD  Continuous Deployment
CI  Continuous Integration
UAT User Acceptance Test

After the software has been implemented it must be built and deployed which requires another specific set of tools.

1) DevOps contains the two terms development and operations. It summarizes all practices that bring together software development and software operations in a way that the development cycles are shortened and the deployment frequency is increased. Microservice architectures support the DevOps concept since microservices have the property that they can be tested and deployed independently from the whole software system. This is not the case for monolithic systems.

1.1) Continuous integration (CI) is a DevOps practice by which the code from the involved developers are built and tested in short time intervals (e.g. several times a day) in order to discover and solve integration problems as soon as possible.

1.2) Continuous delivery (CD) extends the CI practice by releasing the software after it was built and tested (i.e. continuously integrated).

1.3) Continuous deployment (also abbreviated by CD) which is often confused with continuous delivery, means that every change is automatically deployed to production. Since the software must be released before it can be deployed, continuous deployment is an extension of the DevOps practice of continuous delivery.

(Maven) A Java tool by which a project's build, reporting and documentation can be managed.
(Gradle) A Java-based build management tool (comparable to Apache Maven) which uses a domain-specific language
(Cucumber) A library by which user acceptance tests based on Gherkin features can be carried out.
(Junit) A framework to test Java programs based on unit tests.
(Nexus) A repository for build artifacts produced by Docker, Maven, npm, and others.
(Docker) A software system which supports container virtualization in order to deploy applications in isolated containers.
(Portainer) An open-source management user interface to manage a Docker host or swarm cluster.

Docker is a software product developed in Go by Docker Inc. The first version of Docker was published in 2013.

(1) Docker extends the existing container technology from Linux in a way that it provides a complete and easy-to-use solution for the creation and the operation of containers.

(2) Container virtualization offer applications the isolated usage of resources (CPU, RAM, network) without use of heavyweight virtual machines. Each container runs as an isolated process on the host operating system. This offers important advantages compared to virtual machine virtualizations, such as lightweight resource consumption (esp. CPU, storage), shorter start time, easier distribution.

(3) A container is a runnable instance of an image while an image is a read-only template with instructions for creating a Docker container. The standardized description leads to a portability of containers. In addition, deployment becomes easier and more stable.

(4) The instructions are formulated in a simple syntax. An example of a command that can be used in a Dockerfile is "docker run".

"$ docker run -i -t ubuntu /bin/bash" (i) creates a new container containing an ubuntu image, (ii) allocates a read-write filesystem as its final layer, (iii) creates a network interface including the assignment of an IP address to the container, (iv) starts the container and executes the /bin/bash command.

(4.1) One important property of layers is that they are shared between all installed images. In the example, the (operating system) layer "Ubuntu" and the (programming language) layer Java are shared by the applications A,B, and C.


[CM-W-MIC] Cooperation & Management: MICROSERVICES, WASA Course Unit. https://team.kit.edu/sites/cm-tm/Mitglieder/2.-1.WASA


This page illustrates an example of a microservice environment as it is established at C&M.

**(Humane Service Registry)** This registry is used during the development to guarantee a consistent design of the architecture. The humane service registry visualizes the service landscape by offering a user-defined list which is stored in the C&M Teamserver.

**(Version Control)** The source code of the services is organized with the version control system Git and stored at GitLab.

**(CI/CD Pipeline)** The CI/CD pipeline is a customized pipeline executed whenever code is pushed to the repository. Hereby a step initiates the next one, when its run was successful. Each step runs in its own Linux Docker container as defined by the gitlab.ci.yml which is stored in the version control system.

**(Runner Server, PaaS Server)** The pipeline includes two servers: (i) the runner server executing the build, and (ii) the PaaS (Platform as a Service) server which is responsible for storing and deploying the images. A GitLab runner executes the tests and builds a runnable artifact. For example, in the case of a Java application, a JAR is the resulting artifact.

**(Tested Code)** The test results are produced by running Maven Verify [Mav-Int]. It includes to validate whether all information, like dependencies, are available, compile the source code and to test the code. The project information and dependencies which are needed are given by the pom.xml.

**(Runnable Artifact)** Furthermore Maven Verify packages the compiled code to an executable. In our case it is an lightweight jar file. This jar is then stored as a pipeline artifact to be used in the next step.

**(Docker Image, Running Container)** Next, the runnable artifact is copied into a bare image, the result is the so called Docker Image. This Docker image is pushed to the registry on the PaaS server. The registered Docker image is deployed in a running container which can be monitored with Portainer.

---

CD Continuous Deployment  
CI Continuous Integration  
PaaS Platform as a Service

GitLab provides a CI/CD pipeline for every repository. Every repository that wants to use the CI/CD pipeline needs to configure the pipeline with the .gitlab-ci.yml file. This file is stored in the root directory of the repository and it contains the definitions of the jobs and the order of their execution. A complete configuration reference can be found in the GitLab documentation [GL-CIC].

Each pipeline can contain multiple jobs which can be divided into several stages. Jobs of the same stage can be executed in parallel. Jobs of the next stage can only be executed when all jobs of the previous stage have been finished successfully.

The code of the GitLab Runner is open source and is available in the GitLab Runner repository on Gitlab.com [GL-git]. It is written in Go and is designed to run on GNU/Linux, macOS and Windows operating systems. To execute the pipeline and its jobs, a GitLab Runner is needed. The GitLab Runner can be run on a server or a desktop computer.

The GitLab Runner supports multiple ways to execute the jobs in so-called executors:
- SSH: The SSH executor can be used to connect to a remote machine and execute the commands of the pipeline job on this machine over SSH.
- Shell: The Shell executor is used to execute the commands of pipeline jobs locally. Depending on the operating system that is used, different shells are used.
- Docker: GitLab Runner can execute jobs on Docker images which are specified in the .gitlab-ci.yml by using the Docker executor. To run Docker, commands inside the Docker container needs to be in privileged mode. Further information can be found in the GitLab documentation [GL-Use].

More executors for VirtualBox, Kubernetes and more are also available.

A GitLab Runner can be assigned to a group or a single repository. If a GitLab Runner is assigned to a group, all child groups and repositories of that group can use it.

The image shows the currently used C&M CI/CD pipeline with three stages: Build, Package, Deploy. Each stage contains one job and the execution of the jobs was finished successfully, as shown by the green checkmarks.

The screen dump on the right shows an excerpt of the .gitlab-ci.yml that is used in the microservice repository of the ElectricCarCharger project.

(1) The .gitlab-ci.yml file contains the configuration of the GitLab CI/CD pipeline for a GitLab repository. The main elements of the pipeline configuration are the jobs. A job has no predefined name or prefix, it can have any name possible. The only limitations are already existing keywords like image, services, stages, and a couple more. The full list can be found in the GitLab documentation [GL-CIC].

(1.1) (25. script) Every job requires the script keyword which defines the commands that are executed in this job. These commands are executed by the GitLab Runner that runs the script in a shell.

(1.2) (21. cache) The files that are used by a job can be cached. This way the files do not need to be downloaded or created for every run. As an example, the build-app job (15.) caches the Maven repository (24.) so it does not have to be created in every run.

(1.3) (18. artifacts) With the artifacts keyword the artifacts that were created by running the job can be copied and attached to the job on GitLab after the job finished successfully. The artifacts of a job can also be used by jobs which are part of a following stage.

(2) Jobs can be assigned to stages. Jobs of the same stage can be executed in parallel. Jobs of the next stage can only be executed if all jobs of the previous stage have been finished successfully.

(10. stages) The stages are defined outside the jobs by using the stages keyword, followed by the stages that should be used. The stages are ordered by their order in the pipeline config.

(3) (3. variables) With the variables keyword it is possible to define environment variables that are available in all jobs of the pipeline.

(4) (1. image) The GitLab CI/CD pipeline allows the use of Docker images. The default Docker image for the pipeline can be set with the image keyword on the same layer in the pipeline config as the stages and jobs. Inside each job the image keyword can be used to assign another Docker image to a job.

Kubernetes (short name K8s, Greek for governor) was first announced by Google in 2014 and its design is heavily influenced by Google’s Borg system. While Borg is entirely written in C++, the rewritten Kubernetes is implemented in Go. The main goal of Kubernetes is to automate the deployment, scaling, and management of applications that run in a container-virtualized infrastructure. One of the most relevant supported container tools is Docker. Since many cloud services offer a Kubernetes-based platform or infrastructure as a service (PaaS or IaaS) it is a key technology in the field of cloud computing.

Google describes Kubernetes as a platform for the orchestration of containers. Orchestration in this context means that all resources are made available (in a highly elastic and efficient way) in order to provide the services.

(kubectl) Administration interface (kubernetes control) that serves for the communication between client and master.
(Master) The central management unit of a Kubernetes cluster consisting of the four components API server, etcd, controller manager, and scheduler.
(API Server) Provides the whole Kubernetes API.
(ETCD) Stores all metadata and configuration data of Kubernetes in a key-value store.
(Scheduler) Schedules the pods by placing them in a queue from where they are assigned to a node. Opposed to schedulers in operating systems, it is not responsible for executing the pods.
(Kubelet) Fulfills the task of executing a pod. Since kubelets run on a node an overload of the master is prevented.
(Kube-Proxy) Distributes the requests to the container based on a load balancing algorithm.

(Pod) This is a running process on the Kubernetes cluster. Each pod contains one or more containers. The one-container-per-pod concept is used in most situations. Containers are able to intercommunicate between pods which have their own internal IP address.
(Docker) A container runtime which can be used in Kubernetes since it supports the Container Runtime Interface (CRI). Other examples of container technologies supported by Kubernetes are XEN and Rocket.

The Git repository is responsible for the version management of the source code. Any commit to the versioning system also triggers the CI/CD pipeline.

The continuous integration and continuous deployment pipeline links development and operations together. It consists of six steps which will be explained in the following.

The first step of the pipeline performs a static code analysis.

After that, manually written unit tests confirm if the application still behaves as expected.

To review the application, it is deployed in a Kubernetes environment.

The penetration tests are automatically carried out by the OWASP Zed Attack Proxy Project (ZAP) and the report is checked by the Neural Evaluator.

At this stage the pipeline evaluates the behavior of the application under load.

In the last step the tested application will be pushed to a container registry to be available for deployment and also deployed in the production cluster.

The container registry is necessary to manage different container images of an application. This allows different application versions to be reused. It is used by the CI/CD pipeline to upload new images and by the deployment cluster to download the required images.

The development cluster is a Kubernetes-based environment which is used to deploy the applications. This cluster is part of the pipeline and ensures both application and deployment functionality.

The production cluster is the final step of the CI/CD pipeline and serves to release a tested and functioning application.

As attackers always adapt and improve their approaches, limiting security testing to the development process is insufficient. There are several approaches to implement a self-learning security system which evolves alongside the attackers.

OWASP  Open Web Application Security Project
ZAD    Zed Attack Proxy Project
(Scaler) The scaler scales the cluster horizontally and vertically according to the current load. To make a decision about whether scaling makes sense, it uses Prometheus to obtain metric data.

(Gateway) All microservices in the Kubernetes cluster are protected by a gateway. Its tasks are to regulate the consumption of the services regarding authentication and billing. Further tasks of the gateway could include enforcing security policies, possibly determined through machine learning. Apigee by Google contains a gateway, which can work together with the PingIntelligence API Security Enforcer (ASE) to provide such security feature.

(Neuronal DP) The Neuronal Decision Point learns from incoming requests and derives policies to determine if a new request represents an attack or not. One possible implementation of this concept is the API Behavioral Security component of the PingIntelligence solution provided by Ping.

(FaaS) The Function-as-a-Service concept can be integrated into Kubernetes using a serverless framework. The serverless framework allows to deploy individual functions within the Kubernetes cluster. Two suitable use cases are extensions and asynchronous jobs.

(Extensions) Functions to extend existing software environments with additional functionality. If the software product to be extended is closed, which means that it is not possible to make your own changes to the source code and the development of your own service for the extension is too much effort, then this functionality is suitable to be developed as a function. As a result, the development time and thus the development costs are reduced.

(Async Jobs) FaaS is suitable for complex calculations that can be executed asynchronously, i.e. for which it is not relevant that they deliver a result as quickly as possible. If no calculations are pending, the associated function can be scaled to zero, so that no resources are consumed during this time. When calculations are due, the function is scaled accordingly. Since the result does not have to be available immediately, there are no disadvantages from a cold start.

<table>
<thead>
<tr>
<th>ASE</th>
<th>API Security Enforcer</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP</td>
<td>Decision Point</td>
</tr>
<tr>
<td>FaaS</td>
<td>Function-as-a-Service</td>
</tr>
</tbody>
</table>
EXERCISES: DEPLOYMENT

(1) ++Deployment Diagram of the Implementation Architecture++
- Nodes: Frontend, BFF, Backend system
- Interfaces: Web APIs, e.g. specified by using OpenAPI

(2) ++Continuous Delivery (CD) and Build Pipelines++
- CD treats each and every check-in as a release candidate
- CD requires a build pipeline which consists of several stages (resulting from different fast/slow/user acceptance/performance tests) which the software must pass.

(3) ++Docker++
- A Docker container runs as an isolated process on the host operating system and provides a virtualized environment to run applications.
- A Docker container is a runnable instance of an image.
- Created by instructions in a Docker file

(4) ++C&M Microservice Environment++
- The result is a application container image (Docker image) which runs in a container (Docker container).