(note that according to some authors such as Kerzner [Ker09]), staffing is part of directing).



Figure 2.1.: Project Management Process Groups according to [Soc04]. Ellipses represent the project management process groups and arrows denote the **output dependencies (results, documents)** between the process groups.

According to [Soc04], project management is accomplished through the application of initiating, planning, controlling and closing processes. Those processes are tightly interrelated, as shown on figure 2.1. The links among the process groups denote the result dependencies between them. For example, a project plan, which is a result of the planning process, is an input for the execution. This example does not claim that process groups are executed sequentially. On the contrary, process groups are overlapping activities that occur at varying levels of intensity throughout each phase of the project [Soc04]. Each process comprises of multiple actions. The actions of those processes depend highly on one another, and a failure in one action can lead to problems in other areas as well due to the integrative nature of project management.

A project or a phase of the project needs to be authorized (see Initiating processes on figure 2.1). Planning describes the project objectives and the selection of the course of action. The coordination of resources is the main task of the executing processes. The closing processes encompass the actions that need to be taken to complete a phase or a project, including the preparation of the formal documentation. The fifth of the project management process groups is controlling.

Project controlling is of high interest for this work, as it is its problem domain. Controlling is a sub-field of project management (see figure 2.2). It is the measuring of the project progress towards the planned goals, evaluating the remainder of the work and taking corrective actions [Ker09]. Project performance should be monitored regularly to discover deviations from the plan. The reasons for those deviations should be identified and the corresponding preventive actions should be decided upon.

The core sub-processes of project controlling are, according to [Soc04], perfor-



Figure 2.2.: Project Controlling. The main and the facilitating processes of controlling are presented as well as the relation of project controlling to project management.

mance reporting and integrated change control. The controlling process group also contains facilitating processes, such as cost control, quality control, risk monitoring and control, etc.

Performance reporting is a core process of project controlling. It consists of the collection of project performance information, reporting the status and progress of the project, as well as forecasting. Variance, trend and earned-value (EV) analysis (see section 2.2) is performed in order to accomplish the goals of the process. Status reports describe the current state of the project, while progress reports relate the status with the plan. The two types of reports can be combined under the term *performance reports*, hence the name of the process. They include information on the cost and quality, among others, and analysis of this information when needed. The output of the process are performance reports and change requests, as can be seen from figure 2.3. Change requests are created in case that an aspect of a project, for example, scope, schedule or cost, needs to be changed.

Integrated change control is the second core process of project controlling and is tightly connected with performance reporting. It is concerned primarily with the detection of changes and their management. The outputs of this process are updates to the project plan, corrective actions and lessons learned

2. Theoretical Background

Core Controlling Processes	Output
Performance Reporting	Performance reportsProject aspects change requests
Integrated Change Control	 Project plan updates Corrective actions Lessons learned

Figure 2.3.: Outputs of Project Controlling

(see figure 2.3). Corrective actions are those, which try to influence the performance, so that it corresponds to the plan. Lessons learned are the reasons for taking an action. They are important for the project at hand as well as other projects, because they give hints on how to act in certain situations or how not to. Project control is often realized with the assistance of tools. Therefore project controlling systems are briefly discussed.

A project controlling system is one that supports the user in the planning of performance, the observation of the actual performance, its evaluation and the process of decision-making [Lew00]. Such a system could offer assistance in determining deviations and corrective actions. A differentiation and comparison of available solutions that could be used in the project controlling domain is presented in section 2.3.

2.2. Progress Measurement and Analysis Techniques

Project progress needs to be measured in order to realize project control, as already discussed in section 2.1. The measured progress needs to be analyzed so that corrective actions can be taken. A variety of measurement and analysis techniques are available to project managers. They are discussed in this section after the notion of *metrics* is introduced.

Metrics are a source of important data for the measurement of progress and project control. According to the Institute of Electrical and Electronics Engineers (IEEE), a metric is a *quantitative measure of the degree to which a system, component, or process possesses a given attribute* [CB90]. A project metric can therefore be defined as a quantification of the degree to which a project aspect is fulfilled. The central aspects of a project are its cost, time and scope. They are also called project constraints or project objectives. Those aspects can be refined in many different ways. Figure 2.4 offers a common refinement of the so-called triple constraint. The refinement adds the risk, resource and quality aspects next to cost, time and scope. The choice of project metrics heavily depends on the project at hand. Other types of metrics, based on different project

attributes can be devised.



Figure 2.4.: Project Constraints: the sextuple constraint

Common basic project metrics are schedule, number of defects, number of risks, to-do items, number of items that need rework, cost, resource capacity, etc. Some of those metrics might make sense on their own, but are capable of offering much better insight when refined with the help of variance analysis.

2.2.1. Variance Analysis

Variance analysis is a method of control analysis, relying on simple subtraction. Once the only common method, it remains an important tool [HL04] in the hands of project managers. Variance analysis can be used to show the difference between the planned and the actual value of a metric. More rarely it is used to relate estimated to planned values or estimated to actual values. Among the popular metrics are "Scheduled start versus actual", "Planned budget versus actual cost", "Planned man-hour versus actual man-hours", "Estimated finish versus planned finish", etc.

Although those metrics are helpful, they could also be misleading. Consider the following situation: At the end of the first month of a project, the spent budget is 80 000 \in . The planned budget for the first month was 100 000 \in . Though it seems that the project is doing well, it might actually be delayed and over budget. This example shows that variance analysis could sometimes be ineffective for analysis and reporting when used on its own. Variance analysis can be supplemented by more reliable assessment and prediction methods, such as Earned-Value Analysis.

2.2.2. Earned-Value Analysis

Earned-Value Analysis (EVA) provides metrics, measuring the divergence from the initial expectations as well as cost and schedule prediction metrics. It is based on metrics, such as Planned Value, Earned Value and Actual Cost.

Planned Value (PV) is the planned budget of an item of work. Actual Cost (AC) is a further EVA metric, measuring the factual cost of an item of work. **Earned Value (EV)** is the summed planned value (PV) of the work actually performed between the date of the measurement and the project start. EV of a completed project is equal to the planned value. EV represents the value of the completed work expressed in terms of the budget assigned to that work.

Figure 2.5 shows an example for the three metrics that were addressed above. The vertical line at the sixth week indicates the moment of measurement. By the sixth week, the actual cost is above the planned value. This means that for the money that has been spent, more work should have been done. Therefore the earned value at the sixth week is below the planned.



Figure 2.5.: Example for Earned-Value Analysis [rnHL]. The leftmost tilted line represents AC, the rightmost represents EV and the middle one – PV.

The most useful EVA metrics are the Cost Performance Index (CPI) and the Schedule Performance Index (SPI) [HL04]. **CPI** measures the efficiency of a project and is calculated as the ratio of the earned value (EV) to the actual cost (AC). According to the example in figure 2.5, by the sixth week of a project 80 000 \in have been spent, which have generated only 40 000 \in of value. The CPI value is therefore 1/2. This implies that the performed work is twice more expensive than planned. Values of CPI<1 indicate that the project is over budget and CPI>1 means that the project is under budget.

SPI measures the time progress of a project. It is the ratio of the earned value (EV) to the planned value (PV). In the example of figure 2.5, SPI = 40 000/ 60 000 = 2/3. This value indicates that the work that has been performed in the six weeks since the start of the project should have been performed in four weeks according to the plan. Values of SPI<1 indicate that the project is behind schedule, SPI>1 means that the project is ahead of schedule.

For more information on Earned-Value Analysis refer to [Web03] and [HL04]. Another technique, widely applied in project controlling is milestone trend analysis.

2.2.3. Milestone Trend Analysis

Milestone Trend Analysis (MTA) tracks the progress of the project milestones and provide a quick overview of the deviations from the planned schedule. An MTA chart has two dimensions – report dates are mapped on the horizontal axis and the milestone dates according to the plan are placed on the vertical axis (see figure 2.6).



Figure 2.6.: Example for Milestone Trend Analysis [rnHL]. The curves indicate the forecasted completion dates of the milestones.

Three milestones are included in the example of figure 2.6. The planned date for the architecture milestone is 30.06.2009. The dotted line indicates the forecast for the accomplishment of the milestone. A horizontal line indicates that the milestone date is kept. The curves for the acceptance test specification milestone indicates a forecasted delay and the requirements line indicates that the milestone is reached prior to the planned date.

As already mentioned in section 2.1, project controlling systems support the user in the observation of the actual performance, its evaluation and the process of decision-making as well as offer assistance in determining deviations and corrective actions. A differentiation and comparison of available solutions that could be used in the project controlling domain is presented in the next section.

2.3. Terminology: Cockpit vs. Dashboard vs. BI

This section aims to establish a common understanding of terms, such as cockpit, dashboard and business intelligence solution among others. A classification is offered. Each term of the classification is described and compared to the other terms. The provided definitions are used throughout the paper.



Figure 2.7.: Nomenclature fuzz, a mindmap

Figure 2.7 shows a mind map of some commonly used terms in the field of project controlling that are sometimes used interchangeably. The terms *cockpit* and *dashboard* are very overloaded and often (incorrectly) used as synonyms. Special attention is therefore given to the differentiation of those two terms in section 2.3.6. *Mashup* and *scorecard* are also used, though much rarely, to denote dashboards or cockpits thus increasing the terminological confusion. The similarities and differences between business intelligence solutions/suits, cockpits, dashboards are often disregarded.

Figure 2.8 offers a classification of the terms with respect to decision-making support. Depending on the interpretation of the terms, different classifications are possible. In the following, the terms depicted on the figure are defined and their interrelations are discussed.



Figure 2.8.: A terminology classification according to decision-making support

2.3.1. Performance measurement versus performance management system

Often a system is referred to as a performance measurement or as a performance management solution. In the following, the two are also used to support the classification of the other notions on figure 2.8.

Performance measurement is the regular collection and reporting of data to track work produced and results achieved [Vir98]. Performance management, on the other hand, is taking action in response to actual performance to make outcomes better than they would otherwise be [Loc11]. Performance management systems therefore offer more support to the user in the process of decision making.

2.3.2. Report

A report is regarded in this work as the most basic tool with respect to decisionmaking support. It is static, because it refers to only one particular point in time. A report contains data, which is retrieved once at the creation of the report and is never updated later. It visualizes data with a fixed granularity, targeted for a concrete audience, which makes a report unsuitable for other audiences, requiring less or more details. Reports themselves do not offer means to influence the development of a project. They are backward-looking and are therefore mostly passive with respect to decision-making support. Their usage is most appropriate when raw data needs to be viewed in a readable format.

2.3.3. Mashup

A mashup is a web application that combines data from more than one source into a single integrated tool, providing a new functionality that was not originally provided by either source [DBBRB08]. Sometimes the term *Mashboard* is also used to denote the same concept. A prominent example for a mashup, though not from the project controlling domain, is a website, combining location data from Google Maps with real estate data to provide a new service for people in search of real estate.

Comparison A mashup performs the following three functions: data extraction, processing of the extracted data and data visualization. It is therefore more powerful than a report as far as decision-making support is concerned (see section 2.3.2). The term mashup assumes much higher degree of automation as compared to a report thus better supporting the user. Data in a mashup generally does not get outdated as it is regularly extracted from the data sources and not only at the creation of the mashup.

2.3.4. Dashboard

Dashboard is a semantically overloaded term, used to refer to car dashboards, desktop widgets, weather gadgets and software management systems. In the latter field, specialized types of dashboards exist for tracking many corporate functions. Dashboards for controlling the operations of human resources and recruiting, sales, security, information technology, project management, customer relationship management and other areas exist.

Project dashboards are of special interest for this work. A project dashboard view is presented on figure 4.2 in chapter 4. They support the users in managing project performance (see also section 2.1 Project Management and Project Controlling). What the multiple definitions for a dashboard have in common is that dashboards collect, summarize, display and manage information using highly-tuned sets of performance metrics [Fit08]. Metrics are an integral part of a dashboard and perhaps the most important success factor. Typical project metrics are presented in section 2.2. From the user's point of view, a dashboard is a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance [Few06]. For further reference regarding other aspects of dashboards, such as design, implementation and deployment, consult [Eck10].

Comparison Dashboards and mashups offer a very similar set of functionalities with respect to decision-making support and are therefore placed at the same level on figure 2.8. Dashboards can be realized as stand-alone, web or desktop applications, whereas mashups are generally associated with the web. Another difference is that the term *mashup* does not necessarily imply the usage of metrics.

2.3.5. Scorecard

Scorecards and dashboards are two very similar concepts. Lots of ongoing discussions regarding their similarities and differences exist in the literature and in research.

Comparison Some authors claim that there is no difference to dashboards, other see the difference in the type of metrics and the target users of the two solutions. In the author's opinion, scorecards are dashboards, which relate the metrics to the strategic goals. Scorecards therefore provide a further dimension and are therefore considered superior to dashboards, as depicted on figure 2.8. This view is supported by Eckerson [Eck10], who notes that a dashboard is a performance monitoring system, whereas a scorecard is a performance management system.

2.3.6. Cockpit

The term "cockpit" originates from the pits dug in the ground, where cock fights took place. In the 18th century, it has been associated with battlefields and places of combat in general. It was later adopted by World War I pilots to refer to the control room of fighter planes, equipped with measurement instruments, monitors and controls. Besides for airplane control rooms (see figure 2.9), the metaphor is currently used to refer to software products that assist their users in making decisions, based on specific information and metrics. A project cockpit provides means for collecting, interpreting, and visualizing measurement data in order to provide purpose- and role-oriented information to all involved parties during the execution of a project [JM02].

Based on the analysis conducted in this work it can be concluded that the functionality of a *Software Project Control Centre* correspond to that of a Cockpit. Therefore the two terms are seen as synonymous with respect to decision-making support are placed on figure 2.8 at the same level.

Comparison: Dashboard vs. Cockpit The term cockpit is often wrongly used as a synonym for dashboard. The following addresses this wide-spread



Figure 2.9.: A Boeing 737 Cockpit (source: [Lan])

terminology confusion. The definition of a cockpit is detailed and opposed to that of a dashboard. The general characteristics and applications of a cockpit are discussed after establishing a common understanding on what a cockpit is and what it is not.

The main difference between dashboards and cockpits is *control*. Dashboards comprise of gauges and status indicators representing performance metrics, but there are no control knobs. This essential difference is sometimes disregarded and dashboards are referred to as cockpits. A cockpit with no controls is just a dashboard monitoring the impending crash, as noted by John Fitch [Fit08].

Cockpits support their user in making future-oriented decisions by providing the control knobs next to the monitoring tools. In addition to the features of a dashboard, cockpits compare figures (for example, the values of a metric between one month and the next) and discover trends. Cockpits assist the analyst by providing projected figures. The offered assumptions and suggestions are usually based on experience data and information about previous decisions in similar situations. Besides suggesting decisions, cockpits could provide an analysis of the consequences of those decisions.

Important to note is that cockpits are not an universal solution to all problems. Cockpits can support the decision-making process for operational decisions, where project performance needs to be improved. Cockpits do not claim to be able to "make" strategic decisions, i.e. decisions, which are long-termed and influence an organization's mission and vision.

2.3.7. Business Intelligence

Business intelligence (BI) is an architecture and collection of integrated operational as well as decision-support applications and databases that provide the business community easy access to business data [MA03]. It is a broad field, containing technologies such as Dashboards, Cockpits, Decision Support Systems (DSS), Executive Information Systems (EIS), On-Line Analytical Processing (OLAP), relational and multi-dimensional analytical processing. According to Michael Reed [MR11], BI can be broken down to the following four components: multi-dimensional analysis tools, query tools, data mining tools and visualization tools. This classification implies some central aspects of a BI solution:

- data can be viewed and analyzed from different angles (see section 2.4),
- it supports advanced SQL search queries,
- automatic search for data patterns and
- advanced visualization.

Business Intelligence is very often associated with data warehousing (see section 2.4) as data warehouses are widely used in such solutions. For this reason, the term data warehouse is sometimes wrongly used as a synonym for business intelligence.

BI and Project Performance Management BI originates from the business field, but it offers many techniques and solutions that can be reused in project controlling as well. This observation is a result of the market research and the conducted requirements engineering process for cockpits (refer to chapters 3 and 4). The features of a cockpit solution are often covered by the functionality of a Business Intelligence Suite and therefore business intelligence solutions are also included in the classification shown in figure 2.8. Most BI solutions offer an extended set of functionalities, such as finance and resource management, to name a few. Business Intelligence solutions may be built modularly, enabling the reuse of components for project cockpit applications.

An analysis of the software development life cycle at a big software development company, such as the Generali Deutschland Informatik Services (see 1.5) shows that large amounts of data are generated by a variety of tools. Important project status information, such as the number of implemented requirements and test cases, productivity, schedule, etc. can be derived from this large data massif with the help of query, data mining and analysis tools. Making use of a business intelligence, this data can be summarized and visualized to assist project managers and other roles from the software development field in the process of progress monitoring and decision-making. Work has been done in this direction by Rong Ou [Ou07], who built a real-time data warehouse to host project-related data from different systems and a web-based data visualization system using business intelligence techniques.

Comparison The differences between business intelligence and cockpit products are primarily in the historical evolution of the two concepts and their implementation. Although data warehouses can be used with both, data warehouses are generally more often related to business intelligence. Business intelligence suites offer a variety of metrics for various areas, including human resources, sales revenues, etc. and thus extensively support the analyst or manager. With respect to decision-making support in the domain of project controlling and performance management, however, the two solutions have very similar features.

2.4. Data Warehouse

Data warehouses often are a component of dashboard, cockpit and business intelligence solutions. The following describes the properties of a data warehouse and compares data warehousing to other similar technologies.

Ralph Kimball, one of the pioneers of the data warehousing field, defines a data warehouse as *a copy of transaction data* specifically structured for query and analysis [Kim96]. Data warehouses are not used as a primary source of data. Instead, subsets of data from multiple heterogeneous sources are copied to a data warehouse to serve a new purpose. Unlike standard databases, which are designed for efficient data storage, data warehouses are **designed for optimal data retrieval**.

Bill Inmon extends the set of properties of data warehouses by looking at them from a different and more detailed perspective: "A warehouse is a subjectoriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process" [Inm95]. A data warehouse provides insight about a **particular subject** and does not simply provide all the information about a company's ongoing operations. Data is **integrated** from a variety of sources into a coherent whole. All data in the data warehouse is **time-variant** as it is associated with a particular time period. Data warehouses are often regarded as "read-most" data containers because the **data is stable** in a data warehouse - more data is added, but data is "never" removed (i.e., data is removed only if it gets too old to serve the particular subject of the data warehouse).

2.4.1. Benefits

Data warehouses have many advantages, determined by the above-stated properties. They include the following, according to Limaye [Lim09]:

- All data of interest, regardless of the location and type of its source, is represented in a data warehouse as a single common data model. Data warehouses eliminate the complex handling of inter-system dependencies for versioning of data. As a consequence, retrieval of data is simplified.
- Complex queries can be executed easily and efficiently, because no data mapping is performed and no communication with multiple remote data sources is needed during query execution. Data is pre-aggregated to answer expected queries when the data warehouse is built.
- Reporting and analysis of data is simplified as inconsistencies are identified and resolved prior to loading the data into the data warehouse.
- Data warehouses can record historical information for data source tables. The historical development of the data can be used for trend reports, exception reports, and reports that show actual performance versus goals.
- Information in the data warehouse is under the control of data warehouse users. They decide how long to store the data, because they work on a copy of the source data. Even if the information sources are unreliable or purge data, the information in the warehouse can be stored safely for the necessary period of time.
- Data warehouses provide retrieval of data without slowing down operational systems. When, for example, a test report needs to be created, the required data is loaded from the data warehouse and not from the database of the test system.

In order to obtain those benefits, data has to be organized in a way, different from the way it is organized in a relational database. The following addresses the design of data in data warehouses and thus highlights the differences between relational databases and data warehouses. Data design is a key milestone in understanding what a data warehouse really is.

2.4.2. Data warehouse design

As already mentioned, data in a data warehouse has to be organized in such a way, so that quick retrieval is possible. Traditional relational OnLine Transaction Processing (OLTP) databases are optimized for gathering and storing data instead. Their design allows them to add, delete and update data quickly, but at the same time hard to analyze for several reasons. OLTP databases hold a large number of tables, usually connected by many relations. Stored procedures, sometimes taking long to execute, are usually used to access data. The scheme of relational databases makes it easy to work with individual records. For dashboards, cockpits and BI solutions this is a disadvantage, because they operate mainly on aggregated data.

The star schema (see figure 2.10) is an established solution for granting quick data retrieval. The star schema structures data in less tables compared to



Figure 2.10.: Sales data warehouse: Star schema. The database model diagram shows related tables with defined foreign and private keys.

the schema of a relational database and thus minimizes the number of table joins. The star schema contains a central fact table and multiple dimension tables. Besides facts and dimensions, attributes and attribute hierarchies are key components of the star schema.

The **fact table** is built up of key values and raw data. Non-key attributes are often called measures. To achieve fast retrieval of data, the fact table contains many redundancies and is therefore generally long and has millions of rows. For this reason, data warehouses need much more disk space as compared to relational databases. The fact table is connected to multiple dimension tables.

A dimension table contains a primary key and all attributes of a dimension. Attributes are arranged into hierarchies, if supported by the dimension. The time dimension on figure 2.10, for example, contains the parent/child attribute hierarchy "Year - Quarter - Month - Order Date".

The process of filling the data warehouse is called Extract-Transform-Load (ETL). It includes copying the data from the sources, merging it, checking the consistency of the data and finally filling the tables of the star schema. What is missing at the end of the ETL process is the aggregated information. It is inevitably required for the analysis of data and reporting and therefore data warehouses have to offer it. Data warehouses realize the quick retrieval of aggregated data by using OLAP (OnLine Analytical Processing) cubes.

2.4.3. OLAP cube

An OnLine Analytical Processing (OLAP) cube (see figure 2.11) is a data structure, used by data warehouses to facilitate analysis and visualization. The term implies that the OLAP cube has three dimensions, though this is only a special case. It is also generally not a cube, as the items of its dimensions need not be the same number. Although the term is strictly speaking incorrect, it is a suitable metaphor, which has become an industry standard. OLAP (or data) cubes



Figure 2.11.: Sales data warehouse: OLAP Cube (source: [Tek02])

aggregate the facts in each level of the attribute hierarchy of each dimension for the star schema at hand. One can drill-up or down into an OLAP cube to view the aggregated data on a different level. Other common operations that can be applied to data cubes are slicing, dicing, roll-up and rotating (pivoting).

Figure 2.11 shows the OLAP cube, corresponding to the discussed sales warehouse star schema. Time, Product and Location were chosen as dimensions. The front upper left cell contains a concrete "fact" - the number of cell phones, sold in North America in the year 2000. The sum of the Nokia cellphones, sold in the first quarter of the year 2000 in Canada, for example, can be quickly retrieved, as it is pre-calculated and already contained in the cube. By storing data in an aggregated form, the data warehouse has the answers to the user's questions in advance.

2.5. Project Collaboration

A collaboration software can be considered every technology or software environment that supports people to work together on a task in order to achieve a common goal. Collaboration could be helpful in the context of project management due to the benefits listed below. The list is based on Seth Bates' observations regarding project challenges [SB05].

Improved communication Project collaboration software improves the communication in a project by introducing a further communication channel beside the verbal one. It can be used to broadcast updates to the project objectives and the decisions made and to identify scope changes. In turn, this can reduce unproductive time of team members and the project risk.

Management of project-related information Such information may be scattered throughout the organization. Multiple versions of data or documentation might be available due to exchange over e-mail. Project collaboration addresses those challenges.

Crossing geographical boundaries Organizations create teams primarily based on individual skills. Putting the best team together more and more frequently means involving people from different locations. A benefit of collaboration software is that it helps the members of geographically dispersed teams to be in close contact by making all project information and materials available any time of the day.

Visibility into the project Project managers most often need to create reports on the project status at certain intervals or on demand. This involves the calculation of required metrics and the compilation of the project information. Collaboration software saves time and effort by letting stakeholders, team members, customers, etc. view the project status, be aware of its progress, scope and objectives any time without involving the extra effort of the project manager.

Easier management of changing resources The composition of teams usually does not change during the course of small or short projects. This is not the case with bigger projects, where the required skills of the team members might be different at different points of the project. Especially for long-running projects, the availability of the team members may change. New project members need to know the decisions that have been made in the past. They should be informed of the project goals and objectives. Relocated tasks need to be assigned to the new project members. The know-how of the people, leaving a project should also be preserved within the project. Those issues can be addressed by project collaboration software.