# Deeper Insights into Product Development Through Data Visualization Techniques

Jens Michael Hopf, Jivka Ovtcharova

## HAL Id: hal-01377475
## https://inria.hal.science/hal-01377475

Submitted on 7 Oct 2016

# Deeper Insights into Product Development through Data Visualization Techniques

Jens Michael Hopf, Jivka Ovtcharova,

Institute for Information Management in Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
{ michael.hopf@partner.kit.edu, jivka.ovtcharova@kit.edu }

**Abstract.** Product Data Management (PDM) system as primary component of the PLM concept has the goal of supporting product development by appropriate methods. Interactive visualization methods support various user groups (e.g. designers, project leaders, managers) in the interpretation and analysis of large and complex data structures in PDM systems which lead to improved decision making. This paper presents a novel approach for the application of visualization methods in PDM systems and a generic architecture for efficient data interpretation.

**Keywords:** PDM, Data Visualization, Architecture, Product Development.

## 1    Introduction

PDM systems are complex and structured data repositories that store and represent diverse information and business processes have grown steadily in complexity over time and thus this makes it difficult to prepare and present complex data structures in a short time period. Users should not spend a great deal of time interpreting and analyzing data. Tufte [1] describes principles as follows: *Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space*. Due to the constantly growing amounts of data, interactive visualizations are necessary to take new perspectives on complex data structures and to gain new insights. The perception of an elaborate understanding of activities and processes across all stages of the product life cycles is also supported by Stark [2] for PDM systems. The presentation of fragmented information in a single view from various sources allows the user to monitor and recognize relevant business-related correlations in a limited way. However, information about objects and relationships cross-context is rarely linked, so that coherent information structures cannot be recognized by the user. The user obtains only a piece of information displayed without perceiving the underlying structures and relationships of the object. The user may retrieve further information only through additional system queries, which reduces the recognition of data correlations by cross-context and combinatorial information. To ensure that information is not only textually represented, but also closely linked, interdisciplinary approaches such as Visual Analytics (VA) can support the combinatorial information processing in PLM in order to gain new

business insights. The interactive visualization of PLM information with various methods such as timelines, graphs, maps, and matrices supports the user in rapid interpretation of information. Eppler and Lengler [3] describe as follows: *Visualization methods can help the user to articulate implicit knowledge and to stimulate new thinking*. This paper presents visualization methods applied for PDM systems. A generic architecture that extracts data from the PDM system and prepares them for realistic visualization frontends is presented. This paper is organized as follows: Section 2 analyzes and evaluates currently used visualization techniques within the PLM field. Sections 3 and 4 compare data views and present an architecture to access PDM data for various visualizations. Section 5 introduces visualization methods and describes the insights gained from visualizations for the PLM context. Section 6 discusses the limitations. Finally, in Section 7, conclusions and suggestions for further research are formulated.

## 2    Related Work

The majority of PLM manufacturers use visualizations only to display 3D models using 3D viewers as well as present basic diagrams, for example, cost structures in pie charts and line graphs. Interactive visualizations are rarely found in PDM systems, although the demand for evaluation of ever-growing amounts of data exist. Parametric Technology Corporation (PTC) offers *Windchill Visualization Services* (WVS) as integrated visualization for 2D and 3D data in PTC products [4]. Siemens offers the NX-tool *HD3D Visual Reporting* for product analysis. The tool allows the identification of problems by creating predefined visual reports based on various data sources [5]. Various *Teamcenter Lifecycle Visualization* solutions by Siemens allow the visualization of 2D and 3D product data during various product life cycle phases [6]. PLM vendors can visualize not only 2D and 3D data, but also metadata from PDM systems, ERP systems, and software systems supporting the Product Development Process (PEP). For example, *Dassault Systèmes* [7] offers configurable dashboards to organize data more clearly through multiple frames (called widgets). Widgets retrieve, prepare, display, and partially visualize data from different sources for a desired user perspective. Thereby, the user obtains a richer and cross-system information view with a limited set of interactive functionalities on the data set. Other PLM providers use bar charts and pie diagrams to supply visual information. However, a more in-depth interaction with data sets is not provided. A timeline-based data view for visualizing information from PDM systems was investigated by Hopf and Ovtcharova [8]. An overview about visualization methods for PLM has been presented in a case study in [9]. A taxonomy of graph visualization techniques has been presented in [10]. A visualization platform dedicated for aircraft product development information has been developed by [11] covering early product stages from requirements analysis to product design. A UML-RUP-based visualization model for enterprise processes in PLM systems was presented in [12]. Other researchers have developed an interactive graph visualization system for graph structure investigations [13], an interactive tool for analysis of millions of nodes and edges in an adjacency matrix was presented in [14], and visual queries in a UI was

presented in [15] in order to display semantic information. Since most PDM providers follow a web-based approach, several visualization technologies can be used for gaining knowledge in various phases of the product development process (PDP) and downstream stages. While 3D models use file formats such as WebGL, X3D, 3DXML, U3D and VRML for product design, product service, and sales, a series of web-based tools and libraries such as D3.js, Datacopia, Sigmajs, and arbor.js exists for process-supporting and product-related PLM data. Visualization techniques are rarely used for data sets in PDM systems which are not directly related to 3D models. This particularly refers to the PLM data such as product-related discussions, meetings, documents from requirements management and product design, project tasks, and material costs.

## 3  Data Views

Today's PDM systems present information either through predefined OOTB data views or views tailored to specific customer requirements. Data views usually present structured tables with objects of a certain type. Object-related functions perform appropriate changes on the object or navigate to other objects and object types using the associated relationship. Data views of an object type and related relationships are limited to the context in which the object is located. A cross-functional interlinkages of objects in a unified view which contains various object types and relationships from different life cycle phases does not exist. Figure 1 schematically shows the function-driven representation of a tabular data view for the component *Program and Project Management* of PDM systems. Data views are often nested and are separately accessed and viewed by users. Objects in the current view are only visible based on the object that has been selected in the previous view. All other objects that are associated to other clusters are sorted out by the underlying logic and thus not considered in the view and therefore invisible for the user. Graph-based views have a different perspective on objects. The user can immediately select the desired objects of the view and does not need to start navigation through various views to reach the desired view output. However, the number of objects displayed in a graph view can be substantial, and therefore has to be restricted on certain criteria. Dependencies and relationships of an object to other objects are visible to the user through graph-based views. Thus, the user can investigate related objects that were hidden in tabular views. Thereby, the user is able to identify correlations of various object-clusters without calling and comparing different views. To ensure that preselected and small data amounts not only are taken into account in a graph-based view, but also allow new user perspectives on data sets, the existing user view must be expanded for the integration of new interactive visualization methods. The extents to which the visualization method is appropriate for a specific context depends on the concrete question to the accessible PLM data set. The most known visualization types are collected by Zoss [16]. The majority of visualization methods differ in their representation of spatial, temporal, structural, or multidimensional data. Moreover, there is a variety of other visualizations which use techniques of different categories and can be described as hybrid visualizations.
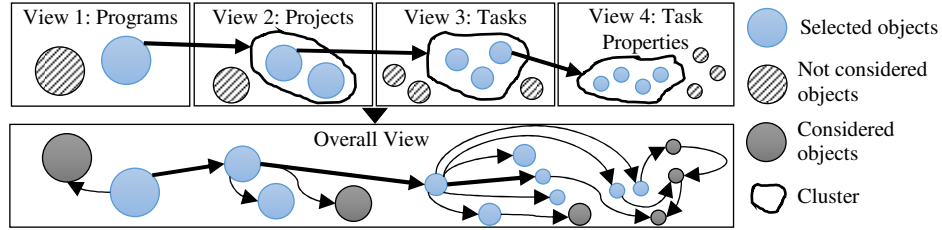
View 1: Programs  View 2: Projects  View 3: Tasks  View 4: Task Properties

● Selected objects

▨ Not considered objects

● Considered objects

⬡ Cluster

Overall View

**Fig. 1.** Tabular object approach (see *View 1 to 4*) vs. Graph View (see *Overall View*)

## 4 Architecture Design

To ensure that all necessary PLM information is retrieved by the visualization via a standard interface, an architecture has been developed and is responsible for querying the data from the PDM system as well as preparing the data sets in the appropriate format for the various visualization methods. The components of the visualization architecture have not been deeply integrated into the PDM system to keep the integration flexibility for other software systems. Proprietary Application Programming Interface (API) supplied by the PLM provider allows access to the vendor-specific data structure and has to be transformed for various visualization methods. Figure 2 shows schematically the developed architecture. The multi-tier architecture consists of three components (see Figure 3): (1) The *Visualization Frontend* (VF) for the information presentation, (2) *Information Collector* (IC) that provides the logic and data structure, and (3) *PDM Connector* (PC) to build queries and retrieve data from the PDM system.
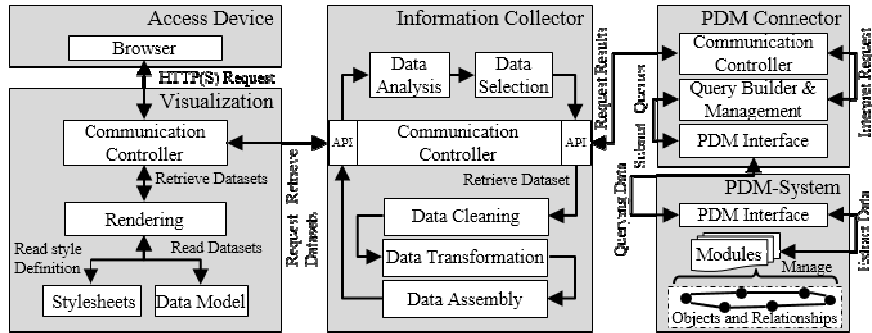


**Fig. 2.** Architecture View

### 4.1 Components

The *Visualization Frontend* (VF) is responsible for the presentation of complex information structures. All relevant information for the visualization is provided in the required data format, so that the visualization frontend can focus on the user experience. A request concerning mandatory information required for the presentation

is sent to the *Information Collector* (IC). The response contains the data in the appropriate format that can be processed by the visualization which concentrates on the content presentation in an interactive context through the web browser, smart user interaction and navigation through the data sets, and loading of provided data objects in the visualization. The IC receives incoming requests from the VF and interprets the parameters to select the data sets. The specific request is subsequently sent to the *PDM Connector* (PC) by which a response is carried out with the corresponding data and contains the requested data in the PDM specific data format. The IC filters out irrelevant information and converts the data into the format that is required for the visualization frontend. Afterwards, the data structure is loaded and processed by the VF. The PC has the role of data broker to receive incoming request from the IC, formulate the database query, and return the results. A request already contains all contextual information to generate the query which is subsequently submitted to the database. The PC is also responsible for performing the authentication to the PDM system and uses proprietary libraries for the communication to the PDM systems provided by the respective manufacturers.

## 4.2 Data Exchange and Data Filtering

JSON (Java Standard Object Notation) is used for the data exchange which is a compact and simple structured data format in a readable text format. The data can be nested and data types such as objects and arrays are supported. Moreover, the syntax of JSON has a lower overhead compared to XML examined in [17]. This characteristic of JSON is especially an advantage for mobile users, because less data needs to be transferred to the mobile device and thus the mobile device battery is less impacted. The communication between the VF and IC depends on the data format required by the specific web-based VF. In addition to the type and amounts of data objects, visualization configuration parameters are declared for user interaction behavior and styling. The reduction of the enormous data volume through temporal selection of data objects may be advisable to ensure the performance of real-time visualization and thus the interactive user experience.

## 4.3 Data Access, Security, and Conversion

User views on data sets are usually defined based on user roles. Therefore, the PDM system defines access control on a subset of data by user roles and context parameters without providing access to the entire database. The current user role concept of PDM systems is sufficient for interactive visualizations, as long as only contextual data need to be visualized. However, access to the entire PLM database is necessary to gain new insights from an overall perspective, which requires the expansion of the current user role concept. Otherwise, limited data sets could lead to erroneous conclusions drawn by the user. Additional user roles could be added to the user account to allow access to a broader amount of data, but the administrative effort appears to be inefficient. A new user role that allows access to the entire database, but with reduced data granularity appears to be an appropriate approach to ensure the

balance between enterprise security and gained insights by users. The data from the PDM system received by the *PC* is in a raw state, so that a data conversion is necessary for the *VF*. The conversion process aims to transform the raw data into the data format required by the *VF*.

# 5 Visualization Methods

The components of a PDM system are responsible for various engineering disciplines. Data sets vary enormously according to the type of object and relation, attributes, and dimensions and therefore not every visualization method is suitable for different user contexts. Before a visualization method is applied, the data structure must be known and an idea about the significance of the visual representation must exist that facilitates the interpretability of the underlying PLM data by the user. When both conditions are fulfilled, a suitable visualization method can be selected to represent data, for example in object relations, comparison to each other, partitions or compositions. In recent years, a variety of visualization methods and tools have been developed such as Tableau Software [18], matplotlib [19], and Highcharts [20]. One of the most common libraries is the JavaScript library D3.js (Data-Driven Documents) [21], which allows the creation and manipulation of interactive visualizations for data sets that increase dynamically. The following visualizations realized with D3.js library have been adapted and integrated for the architecture.

## 5.1 Tree Maps for Cost Analysis of Product Components

Tree Maps are suitable for hierarchical structures to assess and interpret ratios in size. Therefore, a rectangle represents exactly one item and variety of rectangles are proportional to the size as well as can be grouped by color. Tree maps are useful to visualize the cost of product components and cost blocks from various life cycle phases for staff, prototyping, production, service, maintenance, and disposal. Figure 3a shows the visualized component costs of a vehicle. The user can assess the proportions of the color-coded cluster immediately and determines the cluster representing the largest cost as well as comprehends the relations among the rectangles and clusters. The Tree Map supports the user to obtain an overview of the cost structure and allows him to implement appropriate measures for cost optimization.

## 5.2 Chord Diagrams for Flow Analysis of PLM Activities and Dependencies

Chord diagrams are useful for data matrices with interactions to represent relationships between objects and object groups (also known as nodes). The objects are arranged in a circle which represents individual objects that are classified into different areas. The objects are connected within the circle with curved bands to represent the relationship type. A source object can have various relationships to a couple of target objects. It can be recognized what type of object influences other

object types in the PDM system. As a result, the flow pattern of PLM activities can be identified (see Figure 3b) when a new meeting object triggers new task objects, which in turn produce a new collection of document objects. This presupposes that necessary data are stored and accessible in the PDM system.

## 5.3    Bundle Visualization for Dependencies Verification and Traceability

Bundle visualizes relationships wherein connections between objects are represented by lines drawn within the circle. This visualization type is especially interesting for PLM with regard to traceability to track vendor codes, serial numbers, and used part and assembly configurations. Moreover, various examples exist in the field of Change Management, Configuration Management, and Compliance Management. In Figure 3c, the user can examine the dependencies between the test cases (TC-*n*) and customer requirements (R-*n*). Moreover, it can be seen that test case TC_*A40* has a significant relevance to cover a range of customer requirements. Based on these insights, a PLM user responsible for validation and testing could thus intensify the focus on the review of test cases with special dependencies.
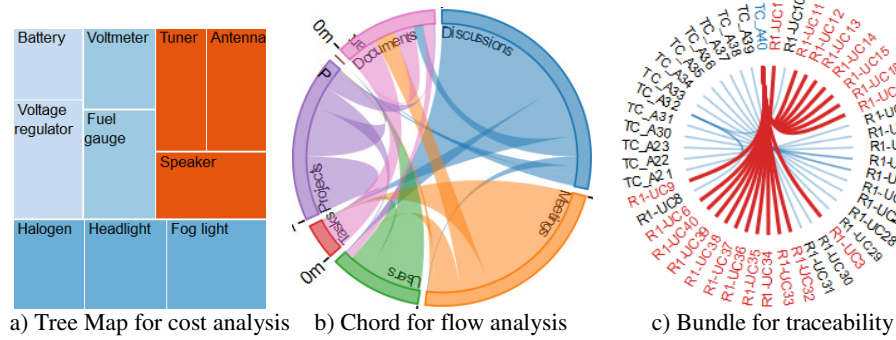


a) Tree Map for cost analysis    b) Chord for flow analysis    c) Bundle for traceability
**Fig. 3.** Visualization methods

## 5.4    Timeline of Historical Results and Events

Decisions concerning product requirements, cost and material of products belong to the everyday work routine of decision makers and can be based upon a variety of reasons which trigger certain events and initiate related activities. All object activities are logged by the PDM system to fulfill the demands of security audits. A lot of object-related event data sets that are accessible in a text-based form have been accumulated over an extended period of time. A chain of events and activities in a historical sequence of an object supports the interpretation of PLM data sets in a specific context. To ensure that such events remain traceable in a chronological order and can be perceived by users, interactive visualizations are suitable not only to select specific event types through the timeline but also to navigate through a series of object events. For example, the navigation can take place through the temporal events of components and assemblies or user's activities. Figure 4a shows the simultaneous navigation through a horizontal time axis scale for object type people, parts and tasks.

It can be seen, which activities have been performed in a given period of time and which events occur in a time-staggered time frame. Each object event is represented by a point. If several events occur within a time period, the points are displayed overlapped. The user has the ability to zoom into the time period in order to break down the temporally overlapping events, and thus obtain a better overview.

### 5.5 Comparison of Parallel Coordinates Visualization for Product Feature

The visualization of multidimensional data allows the analysis of various attributes in different units. An individual filter can be applied for each attribute to reduce the count of data sets. This method is suitable to compare existing parts and assemblies which may be reused in a new product, product requirements, project characteristics, and engineering changes. Depending on the context and assumed user perspective, different object types and their attributes can be compared and analyzed. Figure 4b shows the product data and attributes of various displays. Each display is represented by a horizontal line. Filters can be used for each axis by specifying the desired range of the attribute values (blue lines) to determine the optimal display. The user is fully aware that alternative displays exist (gray lines) and can adapt the requirement and expectations accordingly in case that the optimal display components do not exist.
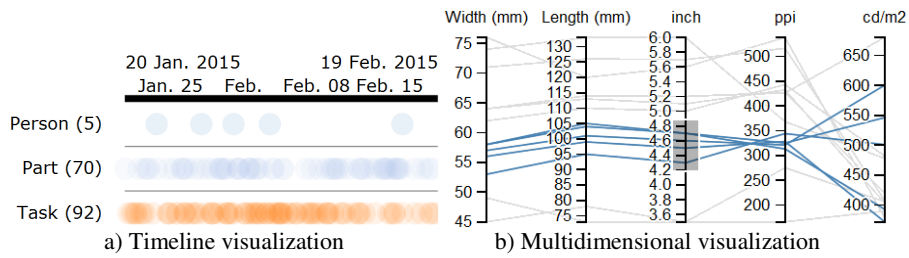


a) Timeline visualization      b) Multidimensional visualization

**Fig. 4.** Comparison of PLM object attributes by interactive visualizations

Several other methods have been adapted for the virtualization system. For example, *Object Maps* are used to visualize complex object-based networks of relationships and dependencies in lifecycles. The *Bullet chart* as variation of the bar chart, but with a richer and smarter view has been used to compare a primarily focused measurement with other secondary measurements which are set in the desired context. The *Calendar Chart* is used to present and analyze data in a temporal context (e.g. shows in what period additional tasks have been redefined and in which months the project tasks have progressed quite slowly caused by absence of project participants).

## 6 Limitations

*Historical data sets*: Interactive visualizations allow users to gain insights based on historical and current data. Any changes occurring in the future can only be predicted if the future behavior with the same logic corresponds to the past behavior. However, this is rarely the case, because external influences (e.g. changes in market conditions)

can induce user behavior strongly and finally lead to changes in established methods, defined PLM processes, and social communication structures.

*Lack of question clarity and unsuitable visualization method*: Without a specifically formulated question, no insights can be gained by the user from the visualization, because required data sets and visualization methods are derived from the question which may lead to the results that existing PLM data sets are insufficient and additional external data sources have to be included. There is also the risk that an inappropriate visualization method is selected that either new insights are not allowed or the visualized data are too complex to be interpreted by users. In addition, too much information highlighting in the visualization can overstrain users.

*Loss of data sets significance*: The total amount of data cannot be represented in every interactive visualization, because data structures can be arbitrarily complex such as found in the Engineering Change Management (ECM) investigated in [22]. Therefore, the selection and compression of the amount of data is necessary to ensure the clarity of the data and efficiently recognize correlations. A falsified perspective on the data can arise by incorrect data selection and compression which mislead users to draw wrong conclusions. The challenges from a business analytics perspective for PLM data in context of innovation management have been summarized in [23].

## 7    Conclusions and Future Works

Interactive visualizations provide an added value in gaining new insights for PLM users. This paper has presented interactive visualizations for PDM systems and a generic architecture that has been implemented as an integrative developed visualization system. The presented visualizations have focused mainly on PLM data generated during the PDP, but not on the representation of 3D models. The visual representations of context-specific data which are provided by the developed architecture enables user groups to derive new insights and make conclusions. Thus, design engineers obtain a deeper insight into various aspects of product design, whereas project leaders obtain a deeper understanding of meeting customer requirements and purchasers about cost optimizations.

A generic and component-based PLM visualization framework for individual applications that cover frequently occurring patterns from PLM field would be a future step in this research field to consolidate individual-based visualization approaches. We see great potential in the development of interactive PLM visualizations that are tailored for different user perspectives and product life cycle stages. Therefore, the expertise of each domain expert is required to understand industry-specific subjects and different work processes and requirements.

# References

1. Tufte, E.R.: The visual display of quantitative information, 7. print., Cheshire, Conn: Graphics Pr. (1983)
2. Stark, J.: Product lifecycle management: 21st century paradigm for product realisation. 2nd, Springer (Decision engineering), London, New York (2011)
3. Eppler, M.J., Lengler, R.: Towards a periodic table of visualization methods: GVE (2007)
4. PTC Inc., PTC Creo - Visualization - Product Development Lifecycle, http://www.ptc.com/product/creo/visualization (visited on 01/18/2015)
5. Siemens PLM Software, Visual Reporting - Produktanalysen im 3D-Konstruktionsprozess, http://www.plm.automation.siemens.com/de_de/products/nx/for-design/visual-analytics/reporting.shtml (visited on 01/18/2015)
6. Siemens PLM Software, Lifecycle Visualization, http://www.plm.automation.siemens.com/en_us/products/teamcenter/lifecycle-visualization/ (visited on 01/18/2015)
7. Dassault Systèmes, http://www.3ds.com/products-services/netvibes/ (visited on 01/18/2015)
8. Hopf, J.M., Ovtcharova, J.: Data View Patterns for an Object-Based PLM Timeline Approach under Consideration of the User Perspective. In: Informatics & IT Today (2), Publishing Society Ltd, Slovakia (2014)
9. Guo, C., Chen, Y.V., Miller, C.L., Hartman, N.W., Mueller, A.B., Connolly, P.E.: Information Visualization for Product Lifecycle Management (PLM) Data. In: 121st ASEE Annual Conference & Exposition, Indianapolis (2014)
10. Nazemi, K., Breyer, M., Kuijper, A.: User-Oriented Graph Visualization Taxonomy: A Data-Oriented Examination of Visual Features. In: Human Centered Design, pp. 576--585. Springer, (2011)
11. Wang, H., Zhao, G., Wang, W., Chen, C.: The design and implementation of the platform of interactive information visualization on aircraft product development. In: 2012 3rd International Conference on System Science, Engineering Design and Manufacturing Informatization, pp. 180--184. IEEE (2012)
12. Chiabert, P., Lombardi, F., Sauza Bedolla, J., Martinez Gomez, J.: Visualization Model for Product Lifecycle Management. In: Annals of Faculty Engineering Hunedoara 11 (1), (2013)
13. Tominski, C., Abello, J., Schumann, H.: CGV—An interactive graph visualization system. In: Computers & Graphics 33 (6), pp. 660--678., (2009)
14. Elmqvist, N., Do, T.N., Goodell, H., Henry, N., Fekete, J.D.: ZAME: Interactive Large-Scale Graph Visualization. In: Pacific Visualization Symposium 2008. PacificVIS'08, pp. 215--222. IEEE, (2008)
15. Shamir, A., Stolpnik, A.: Interactive visual queries for multivariate graphs exploration. In: Computers & Graphics 36 (4), pp. 257--264., (2012)
16. Zoss, A.: Introduction to Data Visualization: Visualization Types, Hg. v. Duke University, http://guides.library.duke.edu/vis_types (visited on 01/14/2015)
17. Ying, M., Miller, J.: Refactoring legacy AJAX applications to improve the efficiency of the data exchange component. In: Journal of Systems and Software 86 (1), pp. 72--88. (2013)
18. Tableau Software, http://www.tableau.com/ (visited on 01/18/2015)
19. matplotlib, http://matplotlib.org/ (visited on 01/18/2015)
20. Highcharts, http://www.highcharts.com/ (visited on 01/18/2015)
21. D3.js (Data-Driven Documents), http://d3js.org/ (visited on 01/18/2015)
22. Jarratt, T., Eckert, C., Caldwell, N., Clarkson, P.: Engineering change: an overview and perspective on the literature. In: Research in Engineering Design 22 (2), pp. 103--124., (2011)
23. Rohleder, C., Lin, J., Kusuma, I., Ozkan, G.: Business analytics in innovation and product lifecycle management: Poster paper. In: IEEE 7th International Conference on Research Challenges in Information Science (RCIS), pp. 1--2. IEEE, (2013)