Goal-Oriented Data Storytelling from User Requirements: An LLM-Assisted Method for Industrial Analytics

Vânia Sousa[1,2,*], Cristiana Vieira[1,2], Pedro Guimarães[1,2], Fernando Pereira[3], Marco Sá[3], António Vieira[2,1] and Maribel Y. Santos[2,1]

[1]CCG/ZGDV Institute, University of Minho, Campus de Azurém, 4800-058 Guimarães, Portugal
[2]ALGORITMI Research Center, University of Minho, Campus de Azurém, 4800-058 Guimarães, Portugal
[3]CEI by Zipor, Rua dos Açores 278, Zona Industrial das Travessas, 3700-018 São João da Madeira, Portugal

Abstract

The Digital Transformation of manufacturing, driven by Industry 4.0 and the green transition, has led to a surge in sensor deployment and real-time data collection. This shift gives rise to contexts in which large volumes of complex multivariate time series data present valuable opportunities for advanced analytics, while simultaneously posing significant challenges for effective interpretation. Analytical dashboards are widely used to support decision-making, yet their impact is often limited when design choices do not align with users' analytical goals or cognitive workflows. A promising response to this challenge is data storytelling, which combines data visualization with narrative structures to enhance comprehension, especially in high-pressure, multi-stakeholder environments. However, in complex industrial contexts, the task of identifying and preparing relevant data for analysis presents considerable challenges due to the massive data volume constantly generated. Recent advances in Artificial Intelligence, particularly Large Language Models (LLMs), present new opportunities to automate and enhance the development of such goal-oriented dashboards. It is therefore necessary to investigate how they can be incorporated into a method that applies them for data storytelling in data-intensive contexts. In light of this, this paper proposes a method for designing analytical dashboards that integrate multivariate sensor data with goal-based storytelling techniques, supported by LLMs to accelerate and guide the development process. The proposed method is instantiated in a real-world industrial case, within the PRODUTECH R3 "Industry-UP" project, in the CEI use case for anomaly detection and operational optimization in sensorized stone-cutting machines. The results show that the method reduces manual intervention, identifies data gaps in earlier stages, and delivers dashboards directly traceable to strategic goals, improving both development efficiency and decision-support quality.

Keywords

Analytical Dashboards, Analytical Requirements, Data Storytelling, Large Language Models, User Requirements

1. Introduction

The digital transformation of manufacturing environments, driven by the advent of Industry 4.0, has led to the widespread adoption of sensors and real-time data collection systems [1]. This transformation is increasingly intertwined with the green transition, as industries seek to enhance not only operational efficiency but also environmental sustainability through data-driven strategies [2]. In this highly automated and data-driven context, industrial processes now generate large amounts of multivariate time series data, offering valuable opportunities to monitor equipment performance, detect anomalies, and optimize operations. However, transforming these complex high-frequency data into clear and actionable insights remains a significant challenge.

Analytical dashboards have become key tools for supporting decision-making in industrial contexts. However, their effectiveness often falls short when the design is not aligned with the analytical goals

ER2025: Companion Proceedings of the 44th International Conference on Conceptual Modeling: Industrial Track, ER Forum, 8th SCME, Doctoral Consortium, Tutorials, Project Exhibitions, Posters and Demos, October 20-23, 2025, Poitiers, France

*Corresponding author.
vania.sousa@ccg.pt (V. Sousa); cristiana.vieira@ccg.pt (C. Vieira); pedro.guimaraes@ccg.pt (P. Guimarães); fjp@inocam.com (F. Pereira); m.sa@zipor.com (M. Sá); avieira@dsi.uminho.pt (A. Vieira); maribel@dsi.uminho.pt (M. Y. Santos)
0009-0002-1279-6651 (V. Sousa); 0009-0000-5128-3571 (C. Vieira); 0000-0003-3390-8528 (P. Guimarães); 0000-0002-1059-8902 (A. Vieira); 0000-0002-3249-6229 (M. Y. Santos)

and reasoning processes of users [3]. Conventional dashboards may present raw data in a visually appealing way, but fail to guide users toward meaningful conclusions, especially when decision makers are not experts in data analysis. Moreover, the way dashboards are organized and the elements included in their content also highly influence their efficacy and value.

A promising approach to address this challenge is data storytelling, which combines data visualization with narrative techniques to enhance comprehension and engagement. Rather than merely showing data, storytelling structures information aligned with the cognitive processes of users, helping them to make sense of complex patterns and draw actionable conclusions [4]. Therefore, in industrial contexts, where decision-making often involves multiple stakeholders and fast-paced environments, the use of storytelling could significantly improve communication, insight discovery, and response time.

Recent advances in Artificial Intelligence, particularly Large Language Models (LLMs), have opened new opportunities to transform how dashboards are conceived and developed. LLMs can automate parts of the analytical design process, such as structuring user requirements, mapping data to analytical goals, and suggesting visualizations tailored to decision-making needs. This automation can reduce the manual effort required, accelerate iteration cycles, and make goal-oriented dashboards more accessible even to users without advanced technical expertise. Beyond simple assistance, LLMs can act as intelligent mediators, embedding domain-specific knowledge into the design process and facilitating the creation of coherent, actionable narratives from complex datasets [5].

Building on these developments, the purpose of this paper is to extend the methodology proposed by [6] for the systematic design of storytelling dashboards, by integrating LLMs in it, to automate specific steps. This integration enables the refinement of analytical requirements, the mapping of tasks to available data, and the proposal of goal-aligned visualizations with reduced development overhead. By aligning dashboard design with user needs and decision-making goals, and by leveraging LLMs to accelerate and guide the development process, the proposed method ensures that visualizations not only present relevant data but also support interpretation and action in complex contexts.

The main contribution of this work lies in demonstrating how LLMs can operationalize and enhance a well-established goal-oriented methodology, showing that they can reduce the number of steps requiring manual intervention, streamline requirements-to-visualization pipelines, and improve the traceability between decision goals and dashboard elements. The proposed method is instantiated in an industrial use case for the reindustrialization of production technologies.

The paper is organized as follows. Section 2 reviews the relevant literature on visual analytics, dashboard design, and LLMs. Section 3 introduces the industrial use case and its specific challenges. Section 4 presents the proposed method, detailing each phase from requirements elicitation to visualization design. Section 5 discusses the results and insights derived from the instantiation of the proposed method in the industrial context. Finally, Section 6 concludes the paper and outlines directions for future research.

2. Related Work

The use of dashboards in industrial and organizational contexts is crucial for data-driven decision-making, with studies exploring their adoption and usage in manufacturing and production environments. Works such as [7, 8] and [9] emphasize the importance of aligning visual analytics tools with user roles, decision needs, and organizational goals. However, they also highlight persistent challenges in transforming complex data into useful insights, especially when dashboards are designed without considering users' reasoning processes or analytical goals.

The effectiveness of dashboards, or their limited adoption when they are not well-aligned with users' mental models or tasks, is a recurring issue across the studies. Walchshofer et al. [10], for instance, explore the socio-technical barriers encountered during the adoption of a visualization tool in a traditional manufacturing company. The study highlights difficulties related to training, dashboard creation, and adaptation to digital tools, illustrating the importance of more user-centric approaches. In the same way, Musleh et al. [11] and Mahmoodpour et al. [12] recommend involving stakeholders throughout the dashboard development process to increase usability and trust, but note a lack of structured methodologies that link dashboard features to user goals in a clear and traceable way.

To address these challenges, several scholars have proposed goal-driven or intent-driven dashboard design methodologies. Studies such as [13, 8] advocate for the organization of dashboards according to decision-making goals and analytical purpose, often using intentional modeling approaches. These frameworks aim to guide the selection and presentation of data according to the user's comprehension or decision-making requirements. While promising, such approaches are rarely extended to support storytelling or enhanced interaction with the dashboard content.

Storytelling has been examined as a method to enhance analytical reasoning and improve data interpretation. Studies such as [9, 14, 15] illustrate that storytelling functions not just as a narrative layer but also as a systematic approach to communicate analytical intent and support users in building explanations and making decisions. For instance, Hutchinson et al. [15] show how narrative principles might guide users through coordinated visualizations, facilitating the connection between visual patterns and interpretative insights. Nonetheless, several systems need manual writing or significant design effort, limiting their applicability in fast-paced or dynamic analytical environments.

Recently, LLMs have emerged as effective tools for assisting users in data exploration and interpretation. The works [7] and [16] demonstrate the capability of LLMs to automate the generation of narratives, assist in visualization specification, and provide contextual explanations. The LEVA framework [16], in particular, introduces a novel structure that employs LLMs to enhance visual analytics workflows during the on-boarding, exploration, and summarization stages, serving as an intelligent mediator to facilitate users' interactions with data and visual analytics systems, making them more accessible, intuitive, and efficient. The recent survey by Hutchinson et al. [15] emphasizes the integration of foundation models, including LLMs and multimodal LLMs, into visual analytics workflows, outlining opportunities for multimodal interaction, automated insight creation, and user guidance. These advances are rapidly converting dashboards into collaborative analytical partners.

Notwithstanding these advances, significant gaps remain. Existing methods often regard LLMs as standalone tools - such as suggesting charts or summarizing data - without integrating them into a comprehensive method that considers decision-making purpose, user reasoning, and interaction design. For this reason, this paper presents a method that uses an LLM-supported process to accelerate a storytelling methodology for dashboard design. In our approach, the role of users is limited to providing analytical requirements, while the method is operationalized by a Data Engineer (DE) who validates, refines, and contextualizes the outputs suggested by the LLM. This profile corresponds to a professional with expertise in data management and analytics, capable of interpreting domain-specific requirements and ensuring technical correctness. By placing the DE as mediator, the method focuses on aligning visualizations with decision goals and creating structured, explainable narratives. Consequently, the approach improves the comprehensibility, applicability, and overall use of analytical tools for decision makers, even when they lack advanced data expertise, since the DE bridges the gap between requirements and implementation.

3. Industry-UP: The CEI Use Case

The PRODUTECH R3 agenda[1], funded by the Portuguese Recovery and Resilience Plan (PRR), is an Innovation Pact aimed at transforming the Production Technologies Sector into a driving force for national economic growth, promoting resilience, climate transition, digital transformation, and innovation in industry. The agenda includes 15 transformative programs, grouped into 5 areas. The Industry-UP project, in the specific area of promoting the *"efficiency in the use of resources and direct integration of renewable energies in production processes"*, aims to develop a holistic operational and retrofit structure for both new and existing industrial equipments, maximizing their efficiency, extending their useful life, and increasing their return on investment. This project is demonstrated in five specific industrial case studies, and the work of this paper is related to CEI's Use Case.

[1]https://r3.produtech.org/en

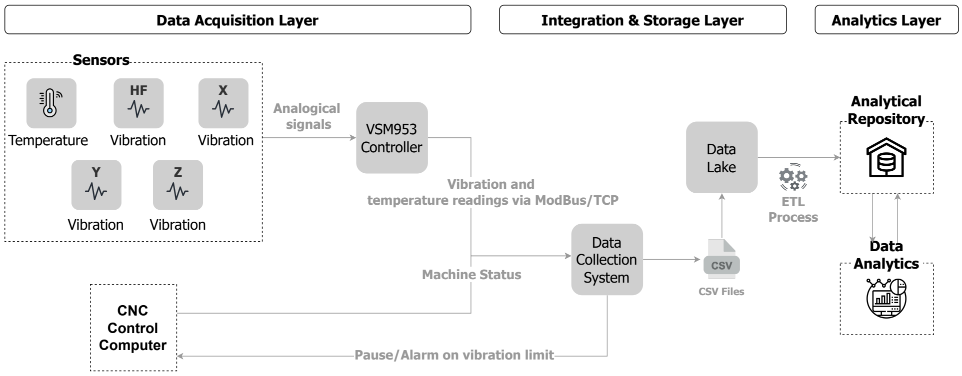

Figure 1: Industry-UP's CEI Use Case Data Architecture.

CEI by Zipor[2], is a group established in 1995, focused on innovative, intelligent, and flexible cutting solutions for the footwear industry. It has expanded internationally, introducing waterjet cutting technologies and spin-offs for advanced software and hardware. In 2003, it expanded to the ornamental stone sector and, today, with brands like Pegasil[3] and more than 2,500 pieces of equipment produced, it is a benchmark in industrial cutting and testing solutions [17]. In the Industry-UP project, CEI provides sensorized equipment for stone-cut machines, making available a vast amount of data to detect and decrease the number of anomalies. This is the context of the CEI's use case.

For data collection and analysis, it is essential a multi-layer data architecture. Figure 1 depicts the data architecture that delineates the information flow for condition monitoring of a Computer Numerical Control (CNC) machine using piezoelectric vibration and temperature `Sensors`. These sensors capture high-frequency (HF) and multi-axis (X, Y, Z) vibration signals, along with thermal measurements. The `VSM953 Controller` processes these analogical signals, calculating key statistical indicators like Root Mean Square (RMS) speed and RMS acceleration. The data is transmitted via the ModBus/TCP protocol, while the `CNC Control Computer` communicates the machine's operational status via TCP/IP protocol. Both streams of data are integrated within the `Data Collection System`, which performs data sampling at two-second intervals and stores the data in `CSV files` for analysis. The Extract, Transform, and Load (ETL) process is carried out weekly when the data is made available in the `Data Lake`, to feed the analytical repository for advanced data analytics.

4. Method for LLM-Assisted Data Storytelling

Lavalle et al. [6] propose a user-centered methodology for designing storytelling dashboards that aligns with decision-makers' analytical requirements and mental models. This approach ensures that the visualization organization is aligned with these cognitive processes, improving data interpretation and reducing the risk of misinterpretation, and optimizing the user experience compared to traditional dashboards. The developed dashboards help users find relevant information for decision support, make the analytical process more intuitive, and reduce the number of requirements answered incorrectly.

As highlighted by the authors of [6], in complex contexts, such as the industrial one, the task of identifying and preparing relevant data for analysis presents considerable challenges due to the massive data volume constantly generated. They further argue that to effectively use their methodology in such contexts, it is necessary to support the dashboard design process with tools that help automate parts of the visualization creation process. Moreover, they advocate that this automation must always be grounded in a rigorous analytical structure, meaning a systematic framework that links strategic, decision, and information goals to the supporting data and visualizations. Using this methodology as a basis, we propose a method for LLM-assisted data storytelling (Figure 2).

[2] https://www.ceigroup.net/
[3] http://www.pegasil.pt

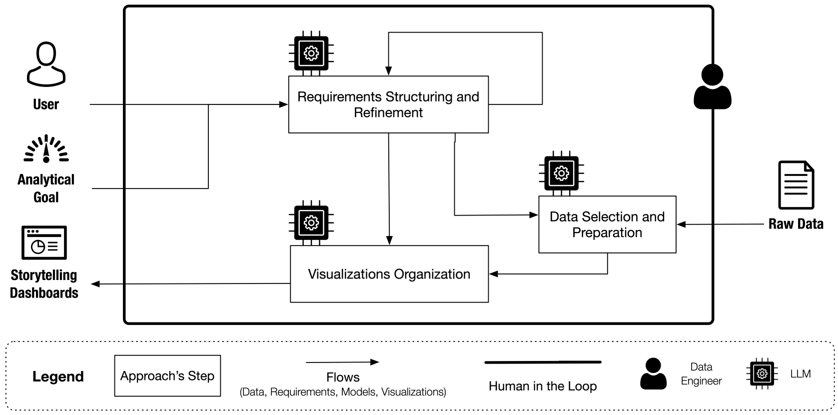

Figure 2: Method for LLM-assisted Data Storytelling.

As can be seen in Figure 2, the proposed method is composed of three steps:

- **Requirements Structuring and Refinement**: The analytical requirements of the users are gathered, prioritized based on the MoSCoW approach [18], and structured in an iStar model [19] with the support of an LLM to align strategic, decision, and information goals.
- **Data Selection and Preparation**: The mapping between analytical tasks and available data is done in this step, as is the identification of relevant hierarchies and filters, and the design of the data analytical model with the support of the LLM suggesting relationships, metrics, and mitigation strategies for data gaps.
- **Visualizations Organization**: The LLM is used to suggest suitable visualizations, and groups of visualizations, for each strategic goal, taking into account the dimensionality and purpose of the analysis.

These three steps are interconnected in an iterative flow. Step 1 defines the strategic, decision, and information goals that guide all subsequent activities. Step 2 operationalizes these requirements by mapping them to available data sources and designing the analytical model. Finally, Step 3 translates the goals into concrete dashboard elements. The connection between Step 1 and Step 3 reflects the fact that visualizations must remain aligned with the original decision goals, ensuring traceability between requirements and visual outputs, even when mediated by the data selection and preparation stage.

In the following subsections, the method's steps are presented along with information of the Industry-UP CEI use case, highlighting their instantiation in real industrial data. For reproducibility purposes, the detailed information can be found in [20].

4.1. Requirements Structuring and Refinement

The first step of the method considers as input the analytical requirements of the users, which can be specified as a set of analytical goals that the users want to be met. As different users have different goals, and as in complex scenarios the number of users and goals is usually high, the analytical goals need to be prioritized. For that, this work adopts the **MoSCoW** approach used in software engineering and developed by Clegg and Barker [18]. This approach prioritizes the requirements according to four distinct groups: i) **M**ust have: requirements with high priority, critical to meeting delivery deadlines; ii) **S**hould have: requirements considered as important, but not critical to meeting delivery deadlines; iii) **C**ould have: desirable requirements, but not critical, being incorporated only if there are time and resources for that; and iv) **W**on't have: requirements to be considered in the future.

To support the development of storytelling dashboards, the MoSCoW prioritization framework offers a structured approach for capturing and contextualizing analytical requirements. It organizes each requirement by ID, description, associated analytical question, required attributes, computed metrics, and priority level, ensuring stakeholders clearly understand both the need and its intended objective.

Users should define their requirements as clear, goal-oriented statements and assign priorities using the structure outlined in Table 1. This template, complemented by illustrative examples, helps ensure consistent interpretation and facilitates a shared understanding among stakeholders.

Table 1
Template for the Specification of Analytical Goals/Questions.

ID	Analytical Goal/Question	Priority
AR1	Monitor the evolution of vibrations by sensor and when they pass their threshold	Should
AR2	Monitor the evolution of spindle electrical current and speed, and when they pass their threshold	Should
...	...	...

In the CEI's use case, 21 analytical requirements were gathered from the user needs and were made available to the LLM as a numbered list format, in this particular case, ChatGPT[4], to structure the requirements in an iStar model that must consider Strategic Goals (SG), Decision Goals (DG), and Information Goals (IGs) [21]. SGs reflect the overarching objectives of the business process being improved, representing a transition from the current state to a desired future outcome. DGs translate these strategic intentions into actionable decisions that leverage information to benefit the organization. They address the question: *"How can a strategic goal be achieved?"*. IGs specify the data requirements necessary to fulfill a DG, answering: *"How can decision goals be achieved in terms of information required?"*. IGs define the data to be collected, typically through analytical processes, and can be expressed either as specific goals or as descriptions of the required analysis.

To consider the evolving nature of a development process, the LLM was provided with the MoSCoW classification (in list format) to aid in the iStar model, allowing DEs to plan development cycles. Additionally, metadata about available raw data was made available to the LLM as textual descriptions, enhancing the efficiency of the method and aligning requirements and data while also ensuring awareness of the application domain.

The specific prompt used to interact with the LLM in this step is as follows:

> Consider the following set of analytical requirements. Consider that Strategic Goals (SGs) are related to the main objectives of the business process that are being enhanced, representing a desired change from a current situation to a future one; Decision Goals (DGs) represent decisions that use information to provide benefits for the organization, operationalizing the SGs into actions by answering the question, "How can a strategic goal be achieved?"; Information Goals guide the information needed to achieve a DG by responding to the question, "How can decision goals be achieved in terms of information required?" IGs outline the data that must be gathered, usually through analysis. As a result, they can be described in terms of goals or in terms of the analysis process. Consider that IGs are decomposed into Tasks (T). Based on this, suggest an iStar model to represent the analytical requirements. For the derived SGs, DGs and IGs, classify them as Must have, Should have or Could have, considering the classification available in the set of analytical requirements. The analytical requirements are: *«to be defined»*.

The iStar model generated by ChatGPT was returned in a structured list format, which was subsequently converted by the DE into a graphical representation for analysis and visualization. Figure 3 depicts the final iStar model after the refinements introduced from the interaction between the DE and the LLM. In particular, the DE suggested a revision to IG3.1.3, highlighting that a threshold was incorrectly assigned to the *speed rate* variable. This feedback led to the reformulation of IG3.1.3 in accordance with the provided inputs. In general, the identification of the iStar model with the support of the LLM was very effective, requiring only a limited number of refinement iterations. The process was relatively fast, with the DE mainly focusing on validating domain-specific aspects (e.g., correcting thresholds or clarifying variable definitions), while the overall structure proposed by the LLM proved consistent with the analytical requirements.

[4]ChatGPT-4, released by OpenAI in May 2025, accessed via the web interface at https://chat.openai.com

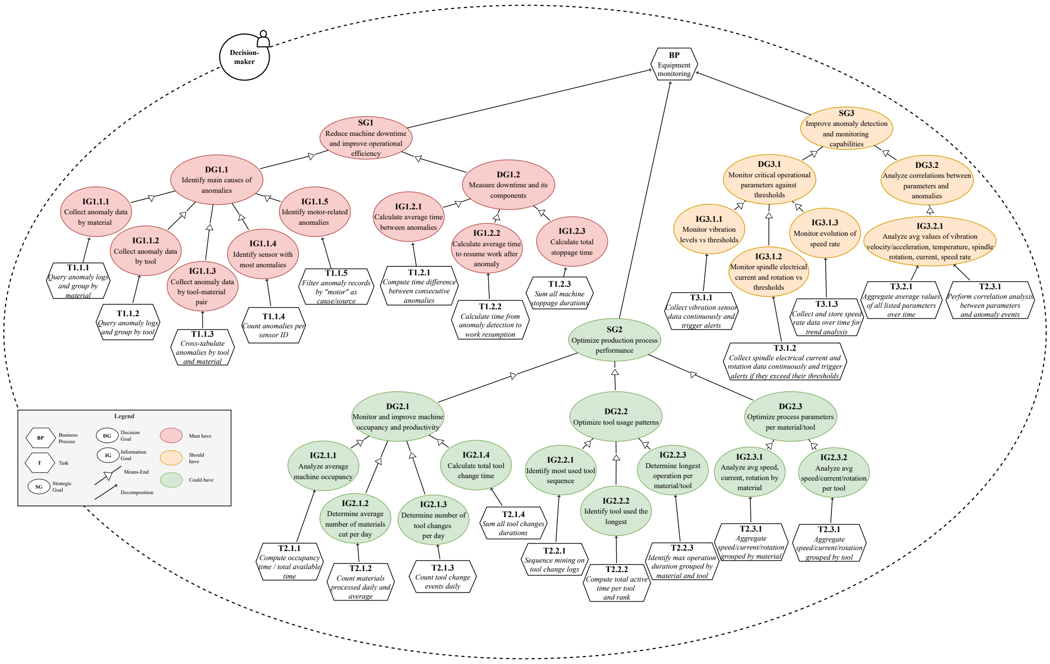

Figure 3: iStar Model after the Requirements Structuring and Refinement Step.

4.2. Data Selection and Preparation

Once the iStar model is identified, including the SGs, DGs and IGs, the selection and preparation of the relevant data must be started to design and implement the analytical repository that will integrate the relevant data and support all visualizations. The input of this step includes the iStar model (image format) and the metadata of the available data as textual descriptions. With this information, the LLM must: i) map the tasks integrated in the IGs to data sources, attributes and metrics (indicators) (subsection 4.2.1); ii) identify the hierarchies and filters to be considered in the decision-making process (subsection 4.2.2); and iii) propose the model of a data warehouse system [22], the analytical system considered in the presented use case (subsection 4.2.3), complementing the Data Lake presented in Figure 1.

In addition to organizing the information needed to propose the supporting data model, this second step plays a crucial role in uncovering potential data gaps that may hinder the fulfillment of the analytical requirements.

4.2.1. Mapping Tasks/IGs to Data Sources, Attributes, and Metrics

For this mapping, the next prompt should guide the LLM in this mapping process. The prompt used to interact with the LLM in this step is as follows:

> Consider the previous iStar model and the following metadata. Map the tasks/IGs to data sources, attributes, and metrics to better select and prepare the data needed for analyses and detect potential data gaps. The resulting table must include the task ID, the task description, the data source(s), the attribute(s), the metric, and, if applicable, the calculation formula. In another table, for each identified data gap, suggest mitigation strategies. The metadata are: *«to be defined»*.

For the available iStar model and data sources, the resulting mapping for SG1 is depicted in Table 2. For clarity, the resulting mapping for the other SGs is not shown. However, it can be found in [20].

For this instantiation of the method with the presented use case, five data sources were available: i) **vibr_monitor_YYYY_MMDD.csv**: stores data about the sensors' measurements, including the anomaly values. On a weekly basis, these files are made available containing data from the previous week; ii) **components.csv**: contains contextual information about components (sensors or motor); iii) **variables.csv**: contains contextual information about the variables measured by the sensors, in particular their threshold, important to detect anomalies; iv) **mat_types.csv**: contains contextual information about the materials cut by the sensorized machine; and v) **tool_description.csv**: contains contextual information about the tools used to cut the materials. While Fig. 1 emphasizes the sensor streams as the primary source of information, these readings are complemented by contextual CSV files. These files do not represent additional independent data sources but rather provide metadata that enriches and contextualizes the sensor measurements, enabling a more complete analytical model.

Table 2
Mapping Tasks/IGs to Data Sources, Attributes, and Metrics supported by the LLM for SG1.

Task	Task Description	Data Source(s)	Attribute(s)	Metric	Formula / Notes
T1.1.1	Query anomaly logs and group by material	vibr_monitor_ YYYYDDMM.csv, mat_types.csv	Time, MType	Count of anomalies per material	Filter anomaly events → group by MType
T1.1.2	Query anomaly logs and group by tool	vibr_monitor_ YYYYDDMM.csv, tool_description.csv	Time, Tool	Count of anomalies per tool	Filter anomaly events → group by Tool
T1.1.3	Cross-tab anomalies by tool and material	vibr_monitor_ YYYYDDMM.csv, tool_description.csv, mat_types.csv	Time, Tool, MType	Count of anomalies per tool-material pair	Filter anomaly events → group by both Tool and MType
T1.1.4	Count anomalies per sensor component	vibr_monitor_ YYYYDDMM.csv, variables.csv	Time, V1–V12, Temp	Count of anomalies by sensor	Map V1–V12 + Temp to sensor components; group
T1.1.5	Filter anomalies caused by motor	vibr_monitor_ YYYYDDMM.csv, components.csv	Time, SCurr, SRpm	Count of anomalies where component = motor	Compare SCurr, SRpm to thresholds
T1.2.1	Compute time difference between consecutive anomalies	vibr_monitor_ YYYYDDMM.csv	Time	Avg time between anomalies	Sort anomaly timestamps, take time deltas, average
T1.2.2	Calculate time from anomaly detection to resumption	vibr_monitor_ YYYYDDMM.csv	Time, Status/Operation flag (*gap*)	Avg resume time	Need indicator for when work resumed
T1.2.3	Sum all machine stoppage durations	vibr_monitor_ YYYYDDMM.csv	Time, Status/Operation flag (*gap*)	Total downtime	Requires event logs for stop/start states

As mentioned, identifying data gaps is key for formulating effective mitigation strategies. Table 3 highlights several gaps detected by the LLM, along with proposed solutions to address them. It is the DE's responsibility to validate and, if necessary, refine these strategies. In this instantiation, refinement involved selecting the most appropriate suggestion based on the specific context of the use case. For instance, in G2, the LLM initially suggested the following mitigation strategy: "*Add record of machine status (running/inactive) or infer from parameter changes*". However, the best option is to infer from the parameter, since we know from the process rules that when an anomaly is detected, the machine automatically stops working and the speed attribute is equal to zero. When the machine resumes working, the speed attribute will be non-zero.

Despite these refinements, another adjustment was necessary. The LLM identified a gap in the variable.csv dataset: it expected a direct mapping between sensors and variables, but in this dataset the mapping is implicit, since each variable entry already includes an attribute that indicates the corresponding component. In other words, although the LLM did not recognize the sensor as a component, the dataset contained the information needed to establish this relationship.

Table 3
Data Gaps and Mitigation Strategies suggested by the LLM and refined by the DE.

Gap ID	Description	Mitigation Strategy
G1	No explicit anomaly flag/event in vibr_monitor_YYYDDMM.csv	Define anomaly rules based on thresholds in variables.csv and mark events
G2	No "work resumption" indicator for T1.2.2	Infer indicator from speed parameter changes
G3	No explicit "machine stop/start" events for T1.2.3 and T2.1.1	Derive from zero-speed/zero-current periods
G4	No explicit tool change event logs for T2.1.3, T2.1.4, T2.2.1	Infer when Tool changes between records
G5	No start/end operation markers for T2.2.3	Infer operation by pair material-tool
G6	Tool active time not directly logged for T2.2.2	Track active intervals per tool using continuous readings

4.2.2. Hierarchies and Filters

For the identification of hierarchies and filters, the following prompt is used:

> Taking into account the previous suggested mapping, suggest hierarchies and potential filters that can be applied to the data. The hierarchy definition will allow decision-makers to explore the storytelling dashboard at different levels of detail. The filters serve as a tool to identify the reusable data. The dashboard filters will reduce the dimensionality of the data being visualized while also removing their graphical representation. Please, identify the filters taking into account the DGs.

Identifying hierarchies and filters has several benefits. Hierarchies allow "drill-down" in the storytelling elements (such as charts, tables or maps, for instance), while filters let users isolate relevant slices of data depending on the decision goal, reducing dimensionality and improving performance and clarity. When combined, both help to reuse data for different goals, since filters standardize how subsets of data are selected.

Regarding the hierarchies, after refinements of the DE, the LLM identified the following:

- **H1 - Time based hierarchy:** Year →Month →Day →Hour →Minute →Second.
- **H2 - Equipment hierarchy:** Component →Variable.
- **H3 - Material hierarchy:** Material Type →Hardness →Material ID.
- **H4 - Operational state hierarchy:** Operation Classification →Derived Event Details.

It is important to clarify that *H4 - Operational state hierarchy* relies on derived data through the data gap mitigation strategies and, therefore, cannot be inferred directly from the original metadata.

For the filters, the suggested and refined filters are grouped by DG to match the intended analytical exploration purposes:

- **DG1.1:** Time, Material, Tool, Component, Variable.
- **DG1.2:** Time, Operation Classification, Derived Event Details.
- **DG2.1:** Time, Operation Classification, Material, Tool.
- **DG2.2:** Time, Tool, Operation Classification, Derived Event Details.
- **DG2.3:** Time, Material, Tool, Component, Variable.
- **DG3.1:** Time, Component, Variable.
- **DG3.2:** Time, Component, Variable, Material, Tool.

In this instantiation, the hierarchy and filter proposals were refined to reflect the real availability of data and analytical requirements. The final version presented here already incorporates these refinements: some hierarchy levels were removed since they did not exist in the metadata, incorrect relationships were corrected, and inconsistent concepts eliminated. These adjustments were mostly punctual but ensured the hierarchies and filters remained feasible and aligned with requirements. For instance, in the Material hierarchy, levels Material Type, Material Hardness, and Material ID were consolidated under a single broader concept while still allowing drill-down exploration when required.

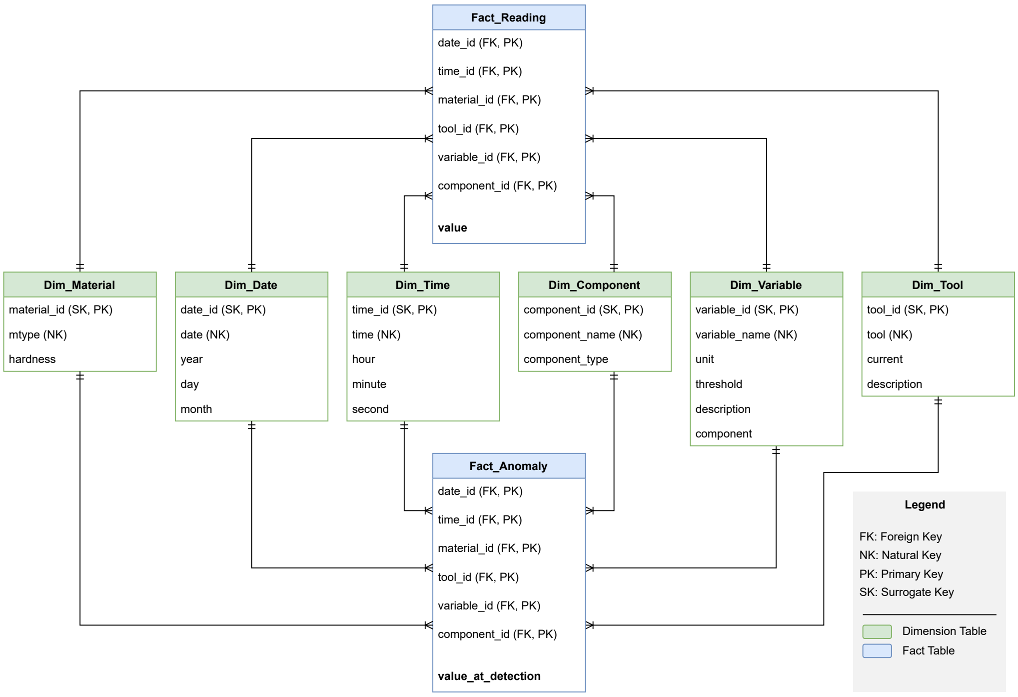

Figure 4: Analytical Data Model proposed by the LLM and refined by the DE.

4.2.3. Analytical Data Model

With all previous information, the LLM can suggest the model of the data warehouse to be implemented to support the visualizations. The DE is in charge of verifying and refining the model, if needed. The prompt used in this task is as follows:

> Considering the suggested mapping, hierarchies, and filters, propose an analytical data model, based on a Data Warehouse system, that responds to the analytical requirements.

Figure 4 depicts the model verified by the DE, a constellation schema that includes two fact tables and six dimension tables. The `Fact_Reading` table allows the storage and analysis of continuous readings from the sensorized equipment, supporting time-series analyses, calculation of performance indicators and monitoring of operational parameters over time. The `Fact_Anomaly` table supports recording and analysis of threshold-exceeding occurrences, enabling the evaluation of anomaly frequency, distribution, and impact. The `Dim_Date` and `Dim_Time` dimensions provide temporal granularity for trend analysis and comparisons, while the remaining dimensions give the necessary context to link readings and anomalies to materials, tools, components, and monitored variables.

Identifying the final version of the analytical data model required the highest level of intervention from the DE compared to other steps. The LLM produced an initial constellation schema that captured the main entities and relationships, but several refinements were needed to adapt it to best practices in data warehousing [22] and to the specifications of the industrial case. For instance, the original `Dim_Time` was split into two separate dimensions (`Dim_Time` and `Dim_Date`), and discussions were necessary regarding the optimal number of fact tables and the definition of materialized views to ensure performance. Despite this higher level of intervention, the process was still more efficient than manual design from scratch, as the LLM provided a solid starting point that accelerated the overall development.

4.3. Visualizations Organization

To structure the storytelling dashboards, the LLM is tasked with organizing storytelling dashboards by SG, considering the type and dimensionality of data when proposing visualizations. This involves identifying the best visualization type (line chart, bar chart, table, map, etc.) for each element in a dashboard, ensuring alignment between data and targeted analysis. The used prompt is as follows:

> Considering the refined analytical data model and the refined iStar model, suggest the type of visualizations (line chart, bar chart, table, map, etc.) that best fits the targeting analysis to include in each dashboard. Organize each dashboard by SG and consider the type and dimensionality of the data when proposing the visualizations to adopt. Provide the output in a table. Suggest the best way to organize the visualizations in the dashboards, taking into account the best practices of information visualization.

The proposed story includes 3 dashboards, each related to a specific SG, covering the 21 analytical tasks. Table 4 presents the proposals of visualizations for *SG1 - Reduce machine downtime and improve operational efficiency*, which was refined in the interaction of the DE with the LLM. Due to space constraints, only the proposed visualizations for SG1 are presented, as the corresponding requirements were classified as "Must have" in the MoSCoW approach. In this step, one example of the refinements made can be seen in T1.2.2. Initially, the LLM suggested a gauge chart to represent the average resume time, but there is no indication of a targeting value for this time. Therefore, it was necessary to suggest a change in the visualization type, to a KPI card, in order to better meet the analytical requirement.

Table 4
Visualizations proposed by the LLM and refined by the DE for SG1.

Strategic Goal	Task	Visualization
SG1 - Reduce machine downtime and improve operational efficiency (Must)	T1.1.1 Query anomaly logs and group by material	Horizontal bar chart
	T1.1.2 Query anomaly logs and group by tool	Horizontal bar chart
	T1.1.3 Cross-tabulate anomalies by tool and material	Heatmap
	T1.1.4 Count anomalies per sensor ID	Horizontal bar chart
	T1.1.5 Filter anomaly records by “motor” as cause/source	Side-by-side bar chart
	T1.2.1 Compute time difference between consecutive anomalies	KPI card
	T1.2.2 Calculate time from anomaly detection to work resumption	KPI card
	T1.2.3 Sum all machine stoppage durations	KPI card

Regarding the visualization proposals, the LLM produced a mockup that organized how the data should be displayed across the dashboard. This mockup served as a starting point for the DE, who refined the layout, adjusted visualization types, and ensured consistency with analytical requirements. The refined version was then implemented in the dashboarding tool. This process reduced design effort, as the LLM accelerated the structuring of the dashboards while leaving final validation and improvements to the DE. The complete set of refined visualization proposals is available in the shared repository [20], and the dashboard presented in Section 5 corresponds to this final validated version.

5. Results and Discussion

Figure 5 presents the dashboard developed for SG1, which aggregates all analytical tasks derived from the iStar model into a coherent and interactive visualization environment. The dashboard was implemented in Grafana[5], fed by a refined Data Warehouse schema and materialized views that ensure performance and responsiveness during user interaction.

The dashboard offers a comprehensive view of anomaly occurrences and downtime, providing both diagnostic and descriptive analytics. The top-left panels show horizontal bar charts ranking materials

[5]https://grafana.com/

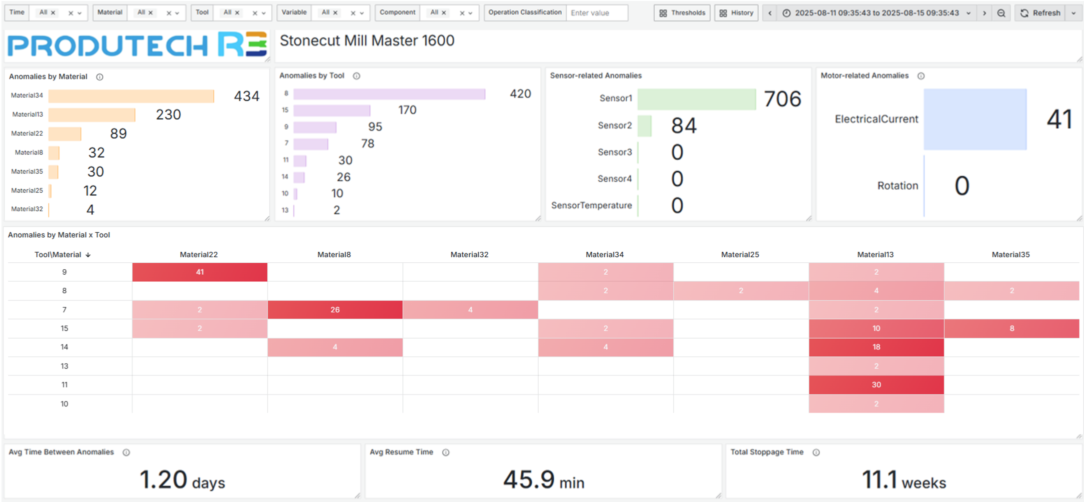

Figure 5: Dashboard of SG1 - Reduce machine downtime and improve operational efficiency.

and tools by the total number of anomalies detected. In the example shown, *Material34* and *Tool 8* emerge as the most frequent sources of anomalies. This direct ranking allows operators to prioritize inspection and maintenance efforts on the most problematic resources. The central heatmap reveals the interaction between materials and tools, highlighting specific combinations that produce disproportionately high anomaly counts (e.g., *Tool 9* with *Material22*). This supports targeted corrective actions rather than general adjustments. Separate panels provide breakdowns of anomalies by sensor readings and by motor-related variables. This distinction helps identify whether anomalies originate from mechanical were captured by the sensors or from deviations in motor performance (e.g., abnormal current or rotation values). The lower section of the dashboard presents KPIs for operational continuity: average time between anomalies, average resume time after an anomaly, and total stoppage time. For the available data, the monitored CNC stone-cutting machine experienced an average of 1.2 days between anomalies, a 45.9-minute average resume time, and an accumulated downtime of 11.1 weeks, which are useful insights for production planning and maintenance scheduling.

The SG1 dashboard is just one example of the method's output. The same approach was applied to SG2 and SG3, proving that the process is replicable, scalable, and adaptable to different industrial goals while maintaining strong alignment between high-level strategy and operational data analytics.

Although the dashboards were not formally validated by the industrial stakeholder at this stage, they demonstrate the feasibility of the method by producing coherent, goal-oriented visualizations aligned with the defined analytical requirements. The process also confirmed that using the LLM accelerated development: instead of designing dashboards from scratch, the DE build on structured proposals and mockups generated by the LLM, which significantly reduced design effort and iteration time.

The proposed method offers several distinctive advantages. The main identified advantage is the reduction of the number of steps in the methodology that serves as the basis for the proposed method [6]. This demonstrates that LLMs can effectively automate parts of visualization creation - such as requirements structuring and refinement, data selection and preparation, and visualization organization - which is crucial for adopting a storytelling methodology in complex industrial contexts. The SG-DG-IG-T chain ensures a direct link between strategic goals and the final visualizations (tasks, T), providing transparency for stakeholders, facilitating validation, and guaranteeing alignment with organizational priorities. Concretely, the method allowed the identification of potential data gaps at an early stage and the definition of corresponding mitigation strategies, as well as the selection of visualization types that remained aligned with the analytical requirements. This alignment reduced the risk of inconsistencies between requirements, data, and dashboards, supporting a more reliable and efficient design process.

An important observation from the case study is that the interaction between the DE and the LLM proved to be efficient. While the DE retained responsibility for validating and refining the outputs, the majority of suggestions generated by the LLM required minor adjustments. This reduced the number of manual interventions and corrections, showing that the LLM acted as a valuable accelerator rather than an additional source of overhead.

6. Conclusion and Future Work

This paper introduces an LLM-assisted method for designing analytical dashboards that integrates multivariate sensor data with a goal-based storytelling approach. By extending an established methodology with automation capabilities, the method uses LLMs to support the structuring of analytical requirements, mapping of tasks to available data, and selection of visualizations aligned with user goals.

The results showed that, while human validation remains necessary, LLMs can anticipate a significant portion of the design process, accelerating dashboard development and enhancing the traceability between analytical requirements and final visualizations. In practice, the DE plays a critical role in reviewing and validating the outputs at every stage, ensuring that proposed mappings, hierarchies, and visualizations are both technically correct and contextually relevant. This validation step is particularly important given the possibility of LLM hallucinations, where outputs may appear plausible but lack factual accuracy or alignment with the available data. Acknowledging these limitations does not diminish the contribution of the LLM; rather, it underscores the value of combining human expertise with automation. When guided and verified by domain experts, the LLM becomes a powerful accelerator, reducing repetitive work and enabling a more efficient and traceable design process.

Future work should expand the application and evaluation of the method in a wider variety of industrial and non-industrial contexts, identifying strengths and limitations and assessing its practical impact. It should also focus on quantifying improvements in decision-making performance, development time, and user satisfaction, as well as exploring ways to refine the interaction between LLMs and human designers to maximize the benefits of automation while preserving contextual accuracy and relevance.

Acknowledgments

This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the R&D Unit Project Scope UID/00319/Centro ALGORITMI (ALGORITMI/UM) and the European Union under the Next Generation EU, through a grant of the Portuguese Republic's Recovery and Resilience Plan (RRP) Partnership Agreement, within the scope of the project PRODUTECH R3 – "Agenda Mobilizadora da Fileira das Tecnologias de Produção para a Reindustrialização". This paper uses icons made available by www.flaticon.com.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT and Grammarly for sentence polishing and rephrasing, and ChatGPT for supporting the proposed approach. All generated content was reviewed and edited by the authors, who take full responsibility for the final text.

References

[1] T. Zheng, M. Ardolino, A. Bacchetti, M. Perona, The applications of industry 4.0 technologies in manufacturing context: a systematic literature review, International Journal of Production Research 59 (2021) 1922–1954.

[2] F. Shrouf, J. Ordieres-Meré, A. García-Sánchez, Digital transformation and green transition in manufacturing: A review of the role of industry 4.0 technologies, Journal of Manufacturing Systems 62 (2022) 801–815.

[3] G. Mazurek, B. Małgorzata, P. Gawrysiak, Designing dashboards for strategic decision-making: A goal-oriented approach, Information Systems Frontiers (2023).

[4] J. D. Boy, E. Segel, J. Heer, Data storytelling: A literature review and an agenda for research, Information Visualization 22 (2023) 142–163.

[5] J.-H. Wang, R. Chen, C.-M. Chang, C.-Y. Cheng, C.-L. Lai, Y.-C. Su, H.-Y. Chou, H.-W. Lin, C.-W. Chang, W.-C. Lin, S.-H. Chen, M.-F. Chiang, S.-H. Wu, C.-H. Chen, Dashchat: Interactive authoring of industrial dashboard design prototypes through conversation with llm-powered agents, 2025.

[6] A. Lavalle, A. Maté, M. Y. Santos, P. Guimarães, J. Trujillo, A. Santos, A methodology for the systematic design of storytelling dashboards applied to industry 4.0, Data Knowledge Engineering 156 (2025) 102410.

[7] Z. Keskin, D. Joosten, N. Klasen, M. Huber, C. Liu, B. Drescher, R. H. Schmitt, Llm-enhanced human-machine interaction for adaptive decision making in dynamic manufacturing process environments, IEEE access (2025).

[8] M. Musleh, A. Chatzimparmpas, I. Jusufi, Visual analysis of industrial multivariate time series, in: Proceedings of the 14th International Symposium on Visual Information Communication and Interaction, Association for Computing Machinery, 2021, pp. 1–5.

[9] T. Langer, T. Meisen, Visual analytics for industrial sensor data analysis., in: ICEIS (1), 2021, pp. 584–593.

[10] C. Walchshofer, V. Dhanoa, M. Streit, M. Meyer, Transitioning to a commercial dashboarding system: Socio-technical observations and opportunities, IEEE Transactions on Visualization and Computer Graphics 30 (2023) 381–391.

[11] M. Musleh, A. Chatzimparmpas, I. Jusufi, Visual analysis of blow molding machine multivariate time series data, Journal of Visualization 25 (2022) 1329–1342.

[12] M. Mahmoodpour, A. Lobov, M. Lanz, P. Mäkelä, N. Rundas, Role-based visualization of industrial iot-based systems, in: 2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA), IEEE, 2018, pp. 1–8.

[13] J.-S. Jwo, C.-S. Lin, C.-H. Lee, An interactive dashboard using a virtual assistant for visualizing smart manufacturing, Mobile Information Systems 2021 (2021) 5578239.

[14] G. A. Topalian-Rivas, J. Wassermann, M. Severengiz, J. Krüger, Automated dashboard generation for machine tools with opc ua compatible sensors, in: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), volume 1, IEEE, 2020, pp. 1009–1012.

[15] M. Hutchinson, R. Jianu, A. Slingsby, P. Madhyastha, Foundation model assisted visual analytics: Opportunities and challenges, Computers & Graphics (2025) 104246.

[16] Y. Zhao, Y. Zhang, Y. Zhang, X. Zhao, J. Wang, Z. Shao, C. Turkay, S. Chen, Leva: Using large language models to enhance visual analytics, IEEE transactions on visualization and computer graphics 31 (2024) 1830–1847.

[17] C. Zipor, Cei group timeline, 2025. URL: https://www.ceigroup.net/geral-2.

[18] D. Clegg, R. Barker, CASE Method Fast-track: A RAD Approach, CASE method, Addison-Wesley Publishing Company, 1994.

[19] F. Dalpiaz, X. Franch, J. Horkoff, istar 2.0 language guide, CoRR abs/1605.07767 (2016).

[20] V. Sousa, C. Vieira, P. Guimarães, F. Pereira, M. Sá, A. Vieira, M. Y. Santos, Goal-oriented data storytelling from user requirements: An llm-assisted method for industrial analytics, 2025. URL: https://doi.org/10.5281/zenodo.17258832.

[21] A. Maté, J. Trujillo, X. Franch, Adding semantic modules to improve goal-oriented analysis of data warehouses using i-star, Journal of Systems and Software 88 (2014) 102–111.

[22] R. Kimball, M. Ross, The Data Warehouse Toolkit: The Defi nitive Guide to Dimensional Modeling, Third Edition, third ed., John Wiley & Sons, 2013.