Trusting the Data Analytics Process from the Perspective of Different Stakeholders

Sven Gehrke, Sandra Niemz, Johannes Ruhland
Chair of Business Informatics,
Friedrich-Schiller-University,
Jena, Germany
e-mail: sven.gehrke@uni-jena.de, sandra.niemz@uni-jena.de, johannes.ruhland@uni-jena.de

***Abstract*—The paper at hand shows different aspects of the concept of trust. Using the Cross Industry Standard Process for Data Mining (CRISP-DM) phase model, we stress the information asymmetry of the typical stakeholders in a data mining project. Based on the identified influencing factors in relation to trust, problematic aspects of the current approach are verified. We execute various interviews with the stakeholders and the results of the interviews confirm the theoretically identified weak points of the phase model with regard to trust. Based on the finding, we sketch amendments and future research areas.**

Keywords-trust; data mining; CRISP DM; stakeholder management.

I. Motivation

Big data analyses take up an ever-larger part of our lives or influence them indirectly. This can be derived from the number of scientific publications [1], which can be assessed as an indicator of the researchers' interest in the subject. Another indicator is the trend in Internet searches on the subject, e.g., for the term "Data Science" [2], which not only reflects academic interest, but also a broader public interest. In addition to the theoretical interest, the number of mass market products that are essentially based on big data analyses has been increasing for years [3]. The same interest can be seen for the term “social media” [4].

Data mining algorithms are largely based on heuristics, i.e., finding probable solutions with limited knowledge and time. This goes hand in hand with probabilities and trust. While there is an extensive literature canon on trust in general - differences and similarities in trust in general and in specific technology - the interactions between people who request, create, operate and use a specific data mining application with regard to trust and its elements have been little explored [5]. The present paper reports on various studies of the relationship between big data analyses and, in particular, their representation and trust depending on the stakeholders involved. Based on the interviews with major stakeholders of the data mining process, the paper points out open issues and challenges found during the survey.

The rest of the paper is structured as follows. In Section 1, we provide an overview on standard data-mining procedures and, based on a literature review, we examine different concepts on trust depending on the field of study. Both these concepts – data-mining methodology and the components influencing trust – are then linked together. In Section 4, we present the interview results from several identified main stakeholders. Since the concepts behind the survey were not equally known by all those involved, mainly semi-structured interviews were chosen. The results were analyzed qualitatively using Mayring's approach [6]. In Section 5, we give a conclusion and list the main findings.

II. Literature Review

A. *Big Data Analytics (BDA) / Data Mining (DM)*

Big data analyses are not just about the underlying data and the resulting analyses but, as with other information systems, about the organization of the processes and organizational views [7]. This requires the presentation in a holistic end-to-end model that connects and coherently maps the individual elements. Phase models are traditionally used in project management [8] [9], while process models, such as Business Process Model and Notation (BPMN) [10] have established themselves to map the various aspects involved.

Regardless of the learning type or the methods, the Framework Cross Industry Standard Process for Data Mining (CRISP DM) has established itself as the standard procedure in data mining projects [11] [12]. The model describes the basic sequence of individual phases, the relationships between one another (see Figure 1) and the tasks contained therein (see Figure 2).

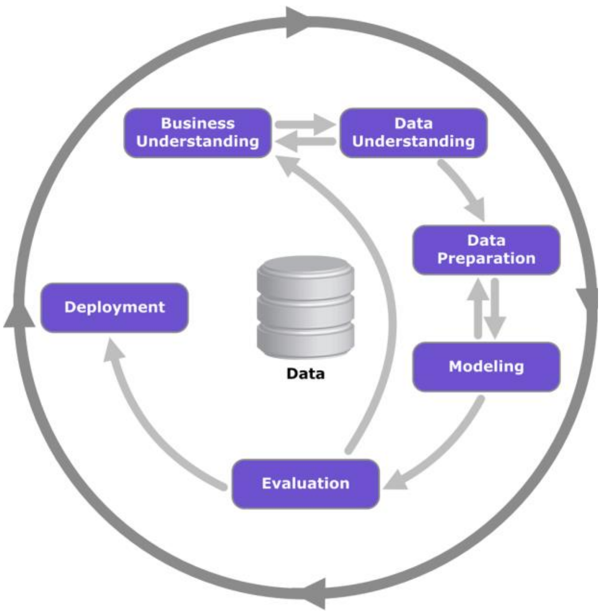

Figure 1. CRISP DM model [12 p. 5].

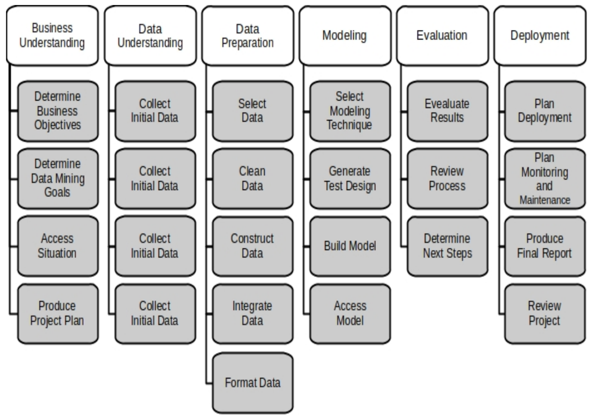

Figure 2. CRISP DM - detailed phases/ tasks [12 p. 6].

The organizational view deals, among other things, with the balancing of interests and information of the stakeholders involved. Only if all those involved find their needs taken into account and respected, they will use the results of the analysis. Obviously, an elementary component like trust has to be treated at every step of the process and cannot be added after the fact. A transfer of trust between different stakeholders is, therefore, necessary in many process steps for trust in the result of the data analysis

Interestingly, typical tasks are listed and described, but no stakeholder identification or explicit role assignment is provided for the tasks in the framework. Accordingly, there is also no consideration of the relationship of trust between the individual stakeholders. As part of this paper, in addition to the standard model, the phases within the framework of a stakeholder analysis, the typical roles involved in a RACI matrix and the information flow were analysed (see also [13] [14]). This is certainly open to discussion in detail, but several things become clear. In the individual phases, there is an information asymmetry to be overcome between the stakeholders and thus a situation of trust in the above sense. Thus, not only is the algorithm itself afflicted with probabilities and thus (trust) risks, the roles involved must also build trust among each other in order to resolve the information asymmetry - and this at different times and in different directions. If data mining is used purely in-house, the business user represents the users or consumers of data mining in the company and can manage the transfer of information and trust. If, however, the consumer of data mining is the general public, it is necessary for the “business user” to put themselves in the most varied of perspectives in the best possible way and to create the basis for the transfer of trust for all stakeholders not directly involved. To make matters worse, the general public is very heterogeneous - both in their personal attitudes and experience, as well as in their institutional environment.

The business user should take the perspective of all major representatives and anticipate and manage their expectations. Looking at the model, (see Figure 1) it becomes clear that this management has to be taken into account both when considering the origin of the data and when communicating the evaluation results / key figures.

TABLE 1. ROLE OF STAKEHOLDERS DESCRIBED BY RESPONSIBILITY ASSIGNMENT MATRIX (RACI)

	Project Sponsor (PS)	Business User/ Analyst (BA)	Data Analyst/ Scientist (DA)	Information Ownership/ Flow
Business Understanding	**a**	**r**	c	PS → BA → DA
Data Understanding	**a**	c	**r**	BA/ DA
Data Preparation/ Modeling	**a**	i	**r**	DA
Evaluation	**a**	**r**	c	DA → BA → PS
Deployment	a	r	c	PS

a = accountable; = responsible; c = to consult; i = to inform

B. Acceptance and Trust

Trust is a complex concept that is defined differently in numerous disciplines depending on the specific circumstances. An additional complicating factor is that trust is also used in everyday scenarios, which results in a multitude of meanings without any reference to a concept. In a much-cited 1964 standard by Kaplan, the author goes so far as to recommend that researchers focus on a specific component of trust rather than a generalized view [15]. If, in addition to the use of the term in one language, one also considers the translation into other languages, there is a large number of uses and synonyms with clear deviations in the connotation. Cooperation, confidence and predictability are closely related terms that Mayer et al. used to describe the term in English [16 p. 729]. Trust is the basis for accepting vulnerability from people, technology and, in our case, the use of data analysis results [17].

To measure acceptance of an application or technology in business informatics, the Technology Acceptance Model (TAM) is often used [17] [19].

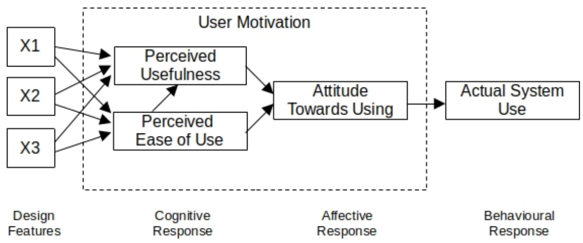

Figure 3. Technology Acceptance Model [41 p. 24].

In Figure 3, the "Attitude Towards Using" is the readiness for use, which is influenced by "Perceived Usefulness" and "Perceived Ease-of-Use". The “Perceived Usefulness” describes the expected benefit, while the “Perceived Ease-of-Use” describes the costs for the user to learn how to use the technology and thus indirectly the costs of building trust. Due to its simplicity, it is easy to use and popular. However, the model focuses on the user and the lack of consideration of the situation / structure is often criticized [20 p. 30]. Furthermore, it does not take time into account and, therefore, a separation of initial trust and continued trust is not explicitly described in the model.

Closely related concepts to trust come from psychology, sociology and social psychology, the latter being understood as a bridge between the former. While psychology focuses on the person-to-person relationship, sociology focuses on the organization-to-organization relationship. What they all have in common is to define trust as "the willingness to take risks" [21 p. 103] or the "intention to accept vulnerability" [22 p. 395]. Basically, Mayer, Davis and Schoorman show that both the personal influence and the organizational or institutional influence of taking risks can be characterized on the basis of competence, benevolence and integrity [16].

After an intensive comparison of the literature, McKnight and Chervany succeeded in bridging the gap between the above-mentioned disciplines and showing the interdependencies [23]. The authors separate between *trusting believes* as the extent to which a target is likely to behave in a way that is "benevolent, competent, honest, predictable in a situation" and the *trusting intentions* as the extent to which a person is willing to make himself vulnerable to another person's actions [23].

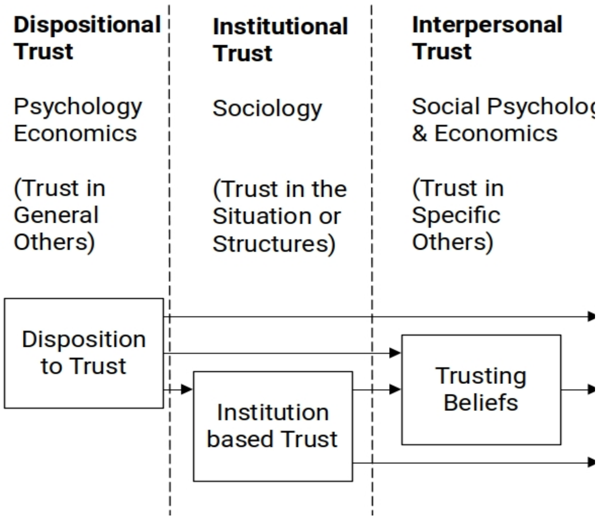

Figure 4. An interdisciplinary Model of High Level Trust Concepts [23]

In a further work, McKnight et al. show that the original characteristics (aka trusting beliefs) of trust in people can be transferred to trust in technologies (functionality, helpfulness, reliability) [5 p. 9]. Thus, the original concept is not limited to natural persons, but can be transferred to people, technologies / objects and even processes [24]. While Figure 4 links different research fields and explains influencing factors from the personal disposition to the trusting intentions, it does not take into account the dimension „time". While continued trust and trusting intentions result from experience and, therefore, the balance of incentives and penalties resulting from trusting, initial trust results from trust transfer - either from person, groups or places [25]–[27].

The microeconomic theory as a further subject area investigates trust in its own branch of research - information theory. It offers a complementary model to the explanatory approaches above. Here, the focus is on trust in goods and the costs of evaluating their properties, less on individual disposition. The underlying assumption is that the information market does not exhibit high degrees of transparency. That is, to evaluate the information, the information must be known, so one has to invest in learning it to evaluate it [28]. In principle, a distinction is made between three types of goods: search goods, experience goods and credence goods. Search goods can be evaluated before purchase or use and, therefore, trusted due to previous experience or easily available product information. Search goods are well known and represent continued trust. Experience goods can be evaluated only after purchase and, therefore, trusted after the purchase and need either a transfer of trust or reduced penalties (e.g., a „refund policy"). Credence goods cannot be evaluated due to prohibitive information retrieval costs or singularity and depend always on external trust transfer [29] [30]. This links the model nicely to the initial model: search goods have a strong linkage towards trust transfer and/ or previous experienced trust, experience goods need initial trust and a positive experience balance for continued usage, and credence goods cannot personally be evaluated over time at all and depend entirely on trust transfer.

From their perspective, all the models presented above are justified and complement each other. The technology acceptance model focuses heavily on the acceptance of a technology without addressing explicitly the wider trust aspect. However, McKnight's model explains the impact of trust in technology and links personal and environmental attitudes, while the microeconomic model deepens insights into how costs affect trust transfer.

C. Linking CRISP DM and Trust

There is only very limited amount of literature about what influences trust in data mining analyses. If one looks at the underlying knowledge and skills of the individual stakeholders in the individual phases of the CRISP DM model (see Figure 1), they will see a strong information asymmetry. To make matters worse, it is noticeable that the responsibility changes in the phases (see Table 1), and external information interests are not explicitly in the focus of the model. This is associated with relatively high costs for obtaining information. Thus, the next step is to evaluate the current practical approaches and problems by interviewing the stakeholders involved.

There are two basic approaches to understand data - compressing the information into key figures/ metrics and visualizing it in graphics. Measures have been around since the dawn of mathematics and are widely used in a wide variety of scientific fields. visualisations of data are just as old as key figures, but have been experiencing increased interest since the 1970s, beginning with Tukey's "Exploratory Data Analysis" [31] and the "Box Plots". Currently, the increased computing power enables complex representations of high-dimensional data and leads to innovative and highly complex forms of representation, e.g., t-distributed Stochastic Neighbor Embedding (t-SNE) [32]. While t-SNE or Uniform Manifold Approximation and Projection (UMAP) help to understand the data directly, graph based visualisation helps to understand hierarchies and dependencies between data or Key Performance Indicators (KPIs) [33]. As an independent specialist discipline, however, visualisation is quite young [34].

 ISBN: 978-1-61208-994-2

Both approaches are to be viewed critically: the key metric α (significance level) used most frequently in statistics and the corresponding p-value (probability) are so often misinterpreted in the scientific literature that the American Statistical Association opposed in 2018 another use of the term “significance” pronounced [35] [36]. On the other hand, visualisations are not problem-free either. A sub-goal of visualisation is to increase the perceived and cognitively processed amount of information and to capture interdependencies in the data structures through aggregation and emphasis. This is intended to support the correctness of decisions and confidence in the decision [37]. Thus, visualisations are not neutral either and depend on the ideas of the creators [38].

Recently, to overcome the aforementioned issues, there is a trend to combine key figures/ metrics with visualisations or allow using interaction in the visualisation. In order to develop a balanced strategy across all critical goals and their respective KPIs, it is particularly important to discover the inherent relationships between all KPIs. In this case, graph-based representations are particularly suitable [39] [40].

D. Identified Research Area

Freedom offered by modern societies, access to loads of sources of information and increased complexities force people to cope with the uncertainty of the global world by themselves. If one considers the factors presented that are important for the trust of the individual stakeholders and if one also considers the CRISP DM phase model, it becomes apparent that the currently leading framework does not explicitly pay attention to the transfer of trust between the stakeholders involved. It should also be clear that trust must be observed from the beginning to the end - interrupting the analysis process would disrupt the transmission of trust. It is important to consistently monitor the level of knowledge of all persons involved. Since the analysis process cannot be explained personally to everyone, it is important to create a framework that passively enables it. A process-oriented, graph-based visual representation as well as additional, generally understandable visualisations should help to show the connections and thus reduce the information asymmetry between the stakeholders. This should be practically substantiated in the following summarized interview with the stakeholders (see Table 2).

III. Practical Survey

As in the theoretical model, trust or the transfer of trust depends on the personal environment (“dispositional trust”), the environment (“institutional trust”) and specific influencers (“interpersonal trust”). In order to obtain a representative picture, identified stakeholders were interviewed using different survey methods . In addition to the typical stakeholders already identified (see Table 1), the list was expanded to include “normal consumers” (“consumer A”) and “informed consumers” (“consumer B”).

In the interviews conducted, the aim was to show which aspects - from the stakeholder's point of view - are necessary in order to develop or transfer trust in data analyses. The questionnaire was based on the identified trust influencing factors during all phases of CRISP DM. Consumer were asked to compare results from data mining analyses with previous expert analyses.

TABLE 2: OVERVIEW OF STAKEHOLDERS AND THEIR INTERVIEWS

Stakeholder	Interview Type	Interview Channel
S1: Data Analyst	semi-structured	face-to-face
S2: Business User	semi-structured	face-to-face
S3: Project Sponsor/ Management	semi-structured	face-to-face
S4A: Normal Consumer	semi-structured / closed	telephone
S4B: Informed Consumer	semi-structured	face-to-face

A. S1 Data Analyst Representative

The data analyst has specialist knowledge of the technical analysis of data, its consistent preparation and the use of appropriate statistical or data mining methods. He needs information about the data used and the business objective of the analysis.

Interview

For the interview, 3 data scientists were asked independently of each other which indicators they believe are relevant in order to trust the data and models. Then they were presented with various business KPIs and visualisations of their department together with the respective business representative and the similarities and differences in understanding were determined.

Main concerns and issues

In the business understanding, it was difficult for the stakeholders involved to interpret the specific KPIs. Concrete examples and the representation of the processes through graphics were essential for understanding.

The interviews showed that the KPIs used to justify the analysis results were rarely understood or misunderstood.

In principle, graphic representations were preferred by the other stakeholders involved. More complex representations were accepted, but required more detailed descriptions, and here again the data analysts often struggled with the business terms. As a compromise for understanding, several simple graphics that build on one another were used.

B. S2 Business Representative

The business analyst represents the business perspective of the departments and has special knowledge of his department. He alone can make sense of the data and explain their origin and meaning and practically validate the results.

Interview

For the interview, 3 representatives were interviewed independently of each other regarding their intentions during the phases in which they are responsible. Then they were presented with various KPIs and visualisations of their department together with the respective data analysts and the similarities and differences in understanding were determined.

Main concerns and issues

The concerns and issues of data analysts reflect the concerns and issues found among business users.

C. S3 Project Sponsor/ Management Perspective

In addition to the roles of data analyst and business user, which are involved operationally, the project sponsor is another relevant role that is more strategically oriented. His trust in the implementation and the results decides initially and finally on the resources used and the use of the results. As a managing role, which is not directly involved but is regularly informed about the analysis and the results, his trust can be seen as the first test of trustworthiness. This role is also responsible to a not inconsiderable extent for the design of the environment, ergo it exerts a great influence on the "institutional trust".

Interview

For the analysis, 10 senior IT managers were asked about their criteria for building trust in a guided interview. The following section summarizes the answers and the underlying intentions.

Main concerns and issues

Looking at the key considerations and underlying intentions, the focus is clearly on promoting institutional trust rather than understanding individual BDA and its metrics. The interviewees emphasized that building a high-quality and transparent data infrastructure is essential for trust in the results. There were different opinions as to whether this should be done step-by-step or with a "big bang". While the majority emphasizes the "step-by-step" approach and thus the step-by-step understanding of the data, a minority fears that too narrow a focus will limit the reference power of the data too much. All project sponsors emphasize that a common understanding is important. To achieve this goal, KPIs, a commonly understood language - which also includes visualisation, are used. For the most part, however, the internal stakeholders are taken into account, but the perspective of the external stakeholders is primarily included through reference to legal data sources.

D. S4 Consumer

When analysing the consumer, a distinction must be made between two stakeholders. The difference lies in the existing experience with analyses. One group are the consumers who have no experience with data mining analysis, data and procedures (see Table 2: "S4A: Normal Consumer"). On the other hand, there are consumers who were not directly involved in the analysis but have personal experience with similar data mining projects (see Table 2: "S4B: Informed Consumer").

Consumers use the results directly for their own purposes, e.g., fitness wearables or evaluations in magazines. However, consumers can also be indirectly affected by the results of the analysis, e.g., as bank customers who are subject to a risk classification when requesting a loan.

E. S4A: Normal Consumer

Interview

In a semi-structured interview, 23 people between the ages of 20 and 60 were asked which factors are relevant for them in different contexts in order to trust data mining analyses.

Main concerns and issues

The results of the data mining were accepted to a very limited extent. Without a well-founded justification for the refusal, it was doubted that the data were representative and reflected the personal circumstances. Although trust was positively influenced by the spread of the BDA (e.g., wearables/ web portals) and by certificates, the results are doubted by 70% - 80% of the respondents and used personally. When it comes to acceptance, the personal opinion of a specialist or friends prevails. If trust has arisen through the transfer of trust from third parties, the trust is not shaken by isolated negative examples or experiences. In principle, the respondents do not see themselves in a position to validate the data bases and functionalities and need support from their environment.

F. S4B: Informed Consumer

Interview

In the interviews, three scientists were questioned in a semi-structured interview. They were not actively involved in the analyses, but they were familiar with the environment.

Main concerns and issues

The expert survey revealed that these people generally trust the analyses, but inform themselves about the data collection, data processing and methods used on a random basis. A renowned environment of the BDA reduces the scope of own validations, but is not sufficient.

IV. CONCLUSION

In principle, the results of the data mining process are accepted by the stakeholders involved in the analysis, but trust in the results correlated strongly with the proximity to the process and the associated costs of information procurement.

In the business and data understanding, the business and data analyst representatives attached great importance to understanding the data and assessing its quality. Less value was placed on a detailed verbal description, the focus was more on use cases and easily understandable key figures.

While the specialists tend to orientate themselves towards the key figures of their specialist area during the evaluation, the other stakeholders involved prefer visual representations in addition to the key figures. Key figures are accepted without understanding their meaning in particular. In order to understand the statements and to trust the results, visual representations prevail. There, too, a trend towards rather simple, well-known representations was discernible. For example, a combination of histogram and box plot was preferred to a violin plot.

The preparation of the analysis process for later users was less of a focus. A fundamental desire was established among all stakeholders to present their findings or interests transparently. However, they often do not realize that the technical terms that describe their special field are not universally understandable. Thus, in data science projects, the focus should be on a more understandable language in advance. In particular, the visualisation seems to have a greater influence on the overall understanding than on specific key figures.

Based on the findings, it would certainly be helpful to add a stakeholder-oriented view to the CRISP DM framework. It should be essential to meet both the information needs of the specialists and to balance the information asymmetry among the stakeholders. In addition to specific, subject-related key figures, this view should be integrated into the CRISP DM process as well as simple visualized representations and generally accessible key figures. This view should also depict the chronological sequence and, so to speak, represent the information transformation nose to tail. In order to do justice to the different levels of knowledge, it should be able to depict different levels of detail.

References

[1] Y. Zelenkov and E. Anisichkina, "Trends in data mining research: A two-decade review using topic analysis," *Bus. Inform.*, vol. 15, no. 1, pp. 30–46, Mar. 2021, doi: 10.17323/2587-814X.2021.1.30.46.

[2] "Google Trends 'Data Science,'" *Google Trends*. https://trends.google.de/trends/explore?q=%2Fm%2F0jt3_q3&date=all (accessed Aug. 15, 2022).

[3] "Google Trends 'Fitness App,'" *Google Trends*. https://trends.google.de/trends/explore?date=all&q=%2Fg%2F11cny0zppc (accessed Aug. 15, 2022).

[4] "Google Trends 'Social Media,'" *Google Trends*. https://trends.google.de/trends/explore?date=all&q=social%20media (accessed Aug. 15, 2022).

[5] D. H. Mcknight, M. Carter, J. B. Thatcher, and P. F. Clay, "Trust in a specific technology: An investigation of its components and measures," *ACM Trans. Manag. Inf. Syst.*, vol. 2, no. 2, Art. no. 2, Jun. 2011, doi: 10.1145/1985347.1985353.

[6] P. Mayring, "Qualitative content analysis," *Companion Qual. Res.*, vol. 1, no. 2, pp. 159–176, 2004.

[7] A.-W. Scheer and K. Schneider, "ARIS — Architecture of Integrated Information Systems," in *Handbook on Architectures of Information Systems*, P. Bernus, K. Mertins, and G. Schmidt, Eds. Berlin/Heidelberg: Springer-Verlag, 1998, pp. 605–623. doi: 10.1007/3-540-26661-5_25.

[8] G. Garel, "A history of project management models: From pre-models to the standard models," *Int. J. Proj. Manag.*, vol. 31, no. 5, pp. 663–669, Jul. 2013, doi: 10.1016/j.ijproman.2012.12.011.

[9] P. D. de M. Sánchez, C. G. Gaya, and M. Á. S. Peréz, "Standardized Models for Project Management Processes to Product Design," *Procedia Eng.*, vol. 63, pp. 193–199, 2013, doi: 10.1016/j.proeng.2013.08.176.

[10] M. Chinosi and A. Trombetta, "BPMN: An introduction to the standard," *Comput. Stand. Interfaces*, vol. 34, no. 1, pp. 124–134, Jan. 2012, doi: 10.1016/j.csi.2011.06.002.

[11] P. Chapman *et al.*, "Step-by-step data mining guide," *SPSS*, p. 76, 2000.

[12] R. Wirth and J. Hipp, "CRISP-DM: Towards a standard process model for data mining," in *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, 2000, vol. 1, pp. 29–39.

[13] C. J. Costa and J. T. Aparicio, "POST-DS: A Methodology to Boost Data Science," in *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*, Sevilla, Spain, Jun. 2020, pp. 1–6. doi: 10.23919/CISTI49556.2020.9140932.

[14] L. Wehrstein, "CRISP-DM ready for Machine Learning Projects," *Medium*, Dec. 19, 2020. https://towardsdatascience.com/crisp-dm-ready-for-machine-learning-projects-2aad9172056a (accessed Jun. 11, 2021).

[15] A. Kaplan, *The conduct of inquiry: methodology for behavioral science*. Scranton (Pa.): Chandler, 1964.

[16] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An Integrative Model Of Organizational Trust," *Acad. Manage. Rev.*, vol. 20, no. 3, pp. 709–734, Jul. 1995, doi: 10.5465/amr.1995.9508080335.

[17] B. A. Misztal, "Trust: Acceptance of, Precaution Against and Cause of Vulnerability," *Comp. Sociol.*, vol. 10, no. 3, pp. 358–379, 2011, doi: 10.1163/156913311X578190.

[18] F. D. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Q.*, vol. 13, no. 3, p. 319, Sep. 1989, doi: 10.2307/249008.

[19] N. Marangunić and A. Granić, "Technology acceptance model: a literature review from 1986 to 2013," *Univers. Access Inf. Soc.*, vol. 14, no. 1, pp. 81–95, Mar. 2015, doi: 10.1007/s10209-014-0348-1.

[20] B. Lunceford, "Reconsidering technology adoption and resistance," *Explor. MEDIA Ecol.*, p. 20, 2009.

[21] N. Luhmann, "Familiarity, Confidence, Trust: Problems and Alternatives," p. 10.

[22] D. M. Rousseau, S. B. Sitkin, R. S. Burt, and C. Camerer, "Not So Different After All: A Cross-Discipline View Of Trust," *Acad. Manage. Rev.*, vol. 23, no. 3, pp. 393–404, Jul. 1998, doi: 10.5465/amr.1998.926617.

[23] D. H. McKnight and N. L. Chervany, "What is Trust? A Conceptual Analysis and an Interdisciplinary Model," *AMCIS 2000 Proc. P382*, p. 8, 2000.

[24] J. B. Thatcher, M. L. Loughry, J. Lim, and D. H. McKnight, "Internet anxiety: An empirical study of the effects of personality, beliefs, and social support," *Inf. Manage.*, vol. 44, no. 4, pp. 353–363, Jun. 2007, doi: 10.1016/j.im.2006.11.007.

[25] D. Belanche, L. V. Casaló, C. Flavián, and J. Schepers, "Trust transfer in the continued usage of public e-services," *Inf. Manage.*, vol. 51, no. 6, pp. 627–640, Sep. 2014, doi: 10.1016/j.im.2014.05.016.

[26] D. Harrison McKnight, V. Choudhury, and C. Kacmar, "The impact of initial consumer trust on intentions to transact with a web site: a trust building model," *J. Strateg. Inf. Syst.*, vol. 11, no. 3, Art. no. 3, Dec. 2002, doi: 10.1016/S0963-8687(02)00020-3.

[27] K. J. Stewart, "Trust Transfer on the World Wide Web," *Organ. Sci.*, vol. 14, no. 1, pp. 5–17, Feb. 2003, doi: 10.1287/orsc.14.1.5.12810.

[28] B. Kahin and H. R. Varian, Eds., *Internet publishing and beyond: the economics of digital information and intellectual property*. Cambridge, Mass: MIT Press, 2000.

[29] S. Fließ, "Qualitätsmanagement bei Vertrauensgütern," *Mark. ZFP*, vol. 26, no. Sonderheft 2004, pp. 33–44, Jan. 2019, doi: 10.15358/0344-1369-2004-Sonderheft-2004-33.

[30] K. Mitra, M. C. Reiss, and L. M. Capella, "An examination of perceived risk, information search and behavioral intentions in search, experience and credence services," *J.*

Serv. Mark., vol. 13, no. 3, pp. 208–228, Jun. 1999, doi: 10.1108/08876049910273763.

[31] J. W. Tukey, *Exploratory data analysis.* Reading, Mass., 1977.

[32] L. Van der Maaten and G. Hinton, “Visualizing data using t-SNE.,” *J. Mach. Learn. Res.*, vol. 9, no. 11, 2008.

[33] M. Graham, J. B. Kennedy, and C. Hand, “A comparison of set-based and graph-based visualisations of overlapping classification hierarchies,” in *Proceedings of the working conference on Advanced visual interfaces - AVI ’00*, Palermo, Italy, 2000, pp. 41–50. doi: 10.1145/345513.345243.

[34] H. Reiterer, T. M. Mann, G. Mussler, and U. Bleimann, “Visualisierung von entscheidungsrelevanten Daten für das Management. In: HMD, Praxis der Wirtschaftsinformatik,” vol. 212, 2000, pp. 71–83.

[35] V. Amrhein and S. Greenland, “Remove, rather than redefine, statistical significance,” *Nat. Hum. Behav.*, vol. 2, no. 1, pp. 4–4, Jan. 2018, doi: 10.1038/s41562-017-0224-0.

[36] R. L. Wasserstein and N. A. Lazar, “The ASA Statement on *p* -Values: Context, Process, and Purpose,” *Am. Stat.*, vol. 70, no. 2, pp. 129–133, Apr. 2016, doi: 10.1080/00031305.2016.1154108.

[37] J.-A. Meyer, *Visualisierung von informationen*. Place of publication not identified: Springer, 1999.

[38] E. Bussemas, “Mehr als Balken und Torten. Eine experimentelle Befragung zur Wahrnehmung von interaktiven Datenvisualisierungen im Journalismus,” *Medien Kommun.*, vol. 66, no. 2, pp. 188–216, 2018, doi: 10.5771/1615-634X-2018-2-188.

[39] M. P. Brundage, W. Z. Bernstein, K. C. Morris, and J. A. Horst, “Using Graph-based visualisations to Explore Key Performance Indicator Relationships for Manufacturing Production Systems,” *Procedia CIRP*, vol. 61, pp. 451–456, 2017, doi: 10.1016/j.procir.2016.11.176.

[40] N. Elmqvist *et al.*, “Fluid interaction for information visualisation,” *Inf. Vis.*, vol. 10, no. 4, pp. 327–340, Oct. 2011, doi: 10.1177/1473871611413180.

[41] F. D. Davis, “A technology acceptance model for empirically testing new end-user information systems: Theory and results (Doctoral dissertation).” Massachusetts Institute of Technology, 1985.

 ISBN: 978-1-61208-994-2