Quō vādis, Data Business?

A Study for Understanding Maturity of Embedded System Companies in Data Economy

Sami Hyrynsalmi[1(✉)], Helena Holmström Olsson[2], Jan Bosch[3], and Sonja Hyrynsalmi[1]

[1] LUT University, Mukkulankatu 19, 15210 Lahti, Finland
`sami.hyrynsalmi@lut.fi`
[2] Malmö University, Nordenskiöldsgatan 1, 211 19 Malmö, Sweden
[3] Chalmers University of Technology, Hörselgången 11, 412 96 Göteborg, Sweden

Abstract. Data has been claimed to be the new oil of the 21st century as it has seen to be able both to improve the existing products and services as well as to create new revenue streams for its utilizing company with a secondary customers base. However, while there is active streams of research for developing machine learning and data science methods, considerably less has been done to understand and characterize data business activities in the software-intensive companies. This study uses a multiple case study approach in the software-intensive embedded system domain. Four large international embedded system companies were selected as the case study subjects. The objective is to understand how the case companies are developing their activities for successful utilization of the data. The study identifies six distinct stages with their own challenges. In addition, this study serves as a starting for further work for supporting software-intensive embedded system companies to start data business.

Keywords: Data business · Embedded system · Data economy · Multiple case study

1 Introduction

A modern day idiom—credited to be said by Clive Humbly in 2006 as well as by various other authors—states that *data is the new oil* [7], a new kind of a primary material which can be then processed and transformed either into new products and services or smarter ways of doing business. Despite the early skepticism whether, for instance, big data is a fad or a fashion in management [c.f. 1], a large body of literature have emerged to support various aspects related to data science [4], ranging from machine learning an AI engineering methods and tools [3] to discussion on data-driven economy [8].

Despite the interests towards the data, its processing and business, a less has been done to understand data business activities in the software-intensive embedded system companies [2,5]. The software-intensive embedded system company here refers to an organization which produces and sells hardware-based devices which include a remarkable software components, including the embedded software as well as the possible software services offered in the top of the product. For instance, a GPS tracking smart watch has both software services (i.e., the portal where the location is mapped) as well as the hardware part. Modern embedded devices are often well-connected and monitored, they are held in various locations and they are often logging data about their own performance as well as the environment where they are operated.

This research was motivated by the question presented in the title of the paper–where the data business is currently going in the embedded system companies. This research focuses on the following questions:

RQ How software-intensive embedded system companies are maturing in their data business activities?

That is, the objective is to gain understanding how and where the business related activities of data are evolving. To achieve this, we carried out a multiple-case study.

2   Research Process

The researchers selected eight large embedded system companies which were known to either have already activities in data business or have publicly discussed on the possibility to improve data usage. From the contacted companies, four agreed to take part in the case study. The data for this study was collected via in-depth interviews with the case companies as well as documentation, website and other public materials such as news and press releases. For the case interviews, the key informant technique [9] was used. The empirical inquiries were carried out during the second quarter of the year 2022 and the interviews were hold in May–June, 2022.

All selected cases are well-established, large international companies working, broadly defined, in the field of embedded systems. That is, all of the selected companies are producing physical devices or components. In addition, all of the companies have remarkable and experienced software divisions, which are responsible for programming the embedded systems as well as creating other support systems ranging from back-end software to online portal services.

All of the selected companies have offices in Nordic countries, where the on-presence interviews were held. As agreed with the companies, their identifies will be kept confidential and details, which could help to identify a company, will be abstracted. The companies are:

CASE A. The company is manufacturing electronics products for both business-to-business and business-to-consumer markets. The devices collect various information regarding its own functioning as well as the environment where it has been installed. As the collected data might be sensitive for the devices' users, the company has set strict technical and organizational culture barriers to prevent the misuse of the data. Nevertheless, the company has initiatives where data—cleared from sensitive information in the customers' own servers—can be enriched and used for valuable new innovations.

CASE B. The company is working in an automotive and transportation industry. Also their devices collect various information regarding the device itself as well as the environment where it is operated. The company is well-aware of the business potential of their data sets and have been sharing some of its data with their partners as well as with a non-profit organization.

CASE C. The company is manufacturing electronics products for mainly business-to-business markets. As a side product of their main operations, the devices manufactured by the company collected number of different information regarding its environment and activities. The company is well aware of the business potential of its data assets, yet it is limited by the contracts between its customers, legislation and conservative business culture in its field.

CASE D. The company is working in an automotive and transportation industry. The devices manufactured by the company have a number of sensors, which collect a wide-range of information. The company is well aware of the business potential of data economy and has set a target to reach new revenue streams out of digital services.

Overall, the selected companies represent relatively exemplar cases of using the collected data to support business. In the following section, more details of these companies data business activities and experiences are viewed.

3   Results

We identified six distinct stages where the companies are in their journey to benefit from data economy. Two of these stages are close to each other and thus we organized the identified stages into a model incorporating five steps where one step is divided into two sub-stages. The overall model is illustrated in Fig. 1 and the stages are explained with details in the following.

(1) Aspiring User. In this stage, a company typically has acknowledged the new business possibilities—both for improving the existing ones as well as founding new opportunities—of the data. Yet, the actions and activities are limited due to various reasons such as lack of organizational and technical capabilities or due to, e.g., legal and contractual issues. For instance, the European GDPR regulation is perceived to restrict business opportunities.
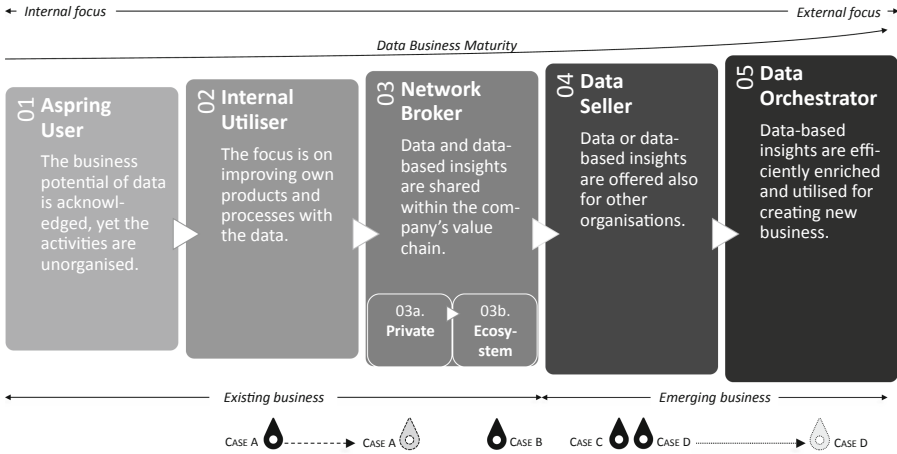
Fig. 1. An illustration of the distinct stages in the model of the embedded system data-business journey. The picture illustrates the focus on the usage of the data (either to improve internal processes or create new business for external stakeholders) as well as the distinction of using data to support existing business or using data to create new business. Each case study subject's location during the study in the model is depicted with markers. The dashed transition in the CASE A illustrates their moving to the new maturity stage. The dashed transition in the CASE D depicts their vision of the desirable maturity position.

Challenges. The main challenges related to the data business activities lay in organization of the data collection, its warehousing, and processing activities in the company to support effective and comprehensive use of data. In addition, the objective is to avoid or dismantle the silo effect—where each unit or department locally optimizes the data collection and using, preventing cross-organizational use—inside the organization.

(2) Internal Utilizer. In this stage, a company is focusing on improving its own products and processes with data analytic. The utilization might be for, e.g., identifying technical and quality problems in the products, or measuring customers' usage for improving the product or product development. The usage of the data is, however, limited to the borders of the company.

Challenges. The main challenges relates to identify and organize efficient utilization of the data for improving the products and services, as well as find usable, valuable information from the collected data.

(3) Network Broker. In this stage, a company is already efficiently using the collected data for improving their internal activities and is focusing distributing value captured from the data to their partners. In this stage, we observed

two distinctive steps how companies were maturing; therefore, also this stage is divided into two sub-stages accordingly. The sub-stages are:

(3a) *Private broker*. In this sub-stage, a company is sharing its data and data-based insights only with a close organization in upstream and downstream in their value chain. A partner who have, e.g., participated in the collection of this data or can use the collected data to improve its part in the value chain. The ownership of the data might be either the embedded device manufacturer or its operator, yet the shared data is strictly restricted into this particular actor.

(3b) *Ecosystem broker*. In this sub-stage, a company is sharing its data-based insights with all of its customers. The main distinctive steps to the previous sub-stage is that the company can combine information from various customers and offer insight from the data conglomerate. For instance, a company can offer an information how a single customer is performing compared to his domains average.

While among the studied organizations the path to mature towards a full-scale data-driven business has gone through including both sub-stages of the third step, it is likely that organizations can move also directly from Stage 2 into Stage 3b of the model. However, the challenges and questions faced in these two sub-stages are perceived by interviewees to be some extent different. For instance, concerns regarding quality and correctness of the data where seen to be larger when the data and data-based insights are shared with other ecosystem partners (Stage 3b) than just with the closest companies in the up- or downstream in the value chain (Stage 3a).

(4) Data Seller. In this stage, a company has started to monetize the collected data with third-party organizations or open market. The main difference to the previous stage is that the company might have been able to derive valuable information for some organization or a customer who is completely unrelated to the core business area where the data was initially collected. Alternatively, the company is selling data or data-bases insights into the third-party organizations working in the same industrial area, yet which are not part of their own ecosystem (for discussion on ecosystems and their borders, see [6]).

Challenges: Whereas previous stages were mainly related to supporting existing business of the companies, in this stage a company would entering into emerging new businesses. This raises also challenges that are not present in the previous stages. For instance, the suitable monetization models of the data as well as the value sharing with the initial data collectors become more focal questions. From the technical side, guaranteeing high data quality as well as the privacy of the data was seen as an important issue. In addition, as the customer base becomes wider and more variable, also the concerns of data theft (e.g., unagreed copying and storing of the data) and mischief are more common than when working with trusted customers and suppliers.

Also in this stage, the question of what is the role of the data business in the company at hand need to be addressed. Being able to become an efficient data seller requires that the company invest in building technological and organizational capabilities for data business. Often this might happen at the price investing less in the core business or the core products of the company. Thus, at least in the case interviews in the domain of embedded systems, it was seen that it is a strategical question for the board and the top-level management whether or not to invest into becoming a data-driven business company.

(5) Data Orchestrator. In this stage, a company is capturing value from the gathered data in broad and efficient ways. The company is capable of enriching different data sources together and offering data-based insights and services as well as raw data for selected customers when seen reasonable. Technical solutions as well as organizational culture and business development support agile and flexible operations in data economy, thus allowing swiftly actions. As illustrated in the interviews, the aim is reach so high maturity in the actions that offering access for ecosystem partners either to some raw data values or data-based insight would be matter of mere hours of implementation work. Overall—including both technical and organizational—maturity of the company also provides capabilities for secure handling of data and following regulations such as GDPR.

Challenges. The main challenge perceived to be linked in the fifth stage relates the creation and orchestration of a business ecosystem around the data. As pointed out by the key informant in CASE D, a challenge would be to get cross-company efforts to solve real customer problems in the ecosystem instead of each ecosystem members solving their small, more local problems.

4   Discussion

The title of this paper asked *quo vadis, data business*–where are you going, data business? To answer that, as well as the set research question, this study carried out a multiple-case study with embedded system companies.

The key observation is that the studied companies seem to follow rather similar paths when their actions and activities in the data business are maturing. To some extent, the journeys have been surprisingly similar as both CASE B and CASE C have ended up sharing their collected data for free with non-profit organizations.

However, we were able to recognize a set of distinct stages through which the companies have matured. In addition, we identified and characterized a goal stage for a few of the studied companies. Nevertheless, it is worth to note that not all companies are aiming towards that direction. For instance, due to the limitations put to the company CASE A by the market where they are operating,

the company seemed not to be focusing to became a full-scale data business company.

The identified stages and the illustrated model might serve as a starting point for the future work in understanding and supporting software-intensive embedded system companies in their data business journey. However, it is worth to note that the presented model is not a maturity model—remarkably due to the lack of alternative journeys and small number of cases—and it should not be taken as an optimal journey for the practitioners. However, it represents a data business journey that some companies have successful being able to start. Thus, the model and described cases might be able to help practitioners to identify problem areas and how to move forward.

In addition, the technical issues were not seen as major challenges for the case companies in the in-depth interviews. Contrarily, the interviewees more often referred to legal, contractual, and organizational issues. When explicitly asked, one interviewee stated that they are easier to solve. However, as it is visible in the review by Cao [4], less work has been done on those non-technical issues. As discussed in the review of the related work, more focus has been put on solving technical issues.

Moreover, as pointed out by one of the key informants, companies are *"almost like starting to try to run, before they can crawl. It won't work."* Often there seems to be defined high overall goals for the data business activities, yet the organizational capabilities and competences might not yet support going into full-scale data business. Therefore, improving through small distinct step might be a way forward for many companies.

While data business and its opportunities are intriguing avenue for embedded system companies, entangling a company into data business activities in full-scale also raises the question of rethinking the core business of a company [c.f. 10]. Yet, by following the decade old discussion whether all companies will become software companies or will they be replaced by ones, it also can be asked whether there is an option for a software-intensive embedded system company of not becoming into data business.

Finally, to answer to the researcg question where the data business of software-intensive embedded system companies is going, the answer is towards an ecosystem-like business where data is shared and enriched. However, the unanswered question is whether all software-intensive embedded system companies are heading there and what alternative business strategies are left for them if they do not.

Naturally, this study has limitations. First, the generalization of the results is limited due to the focus on a small number on quite well performing companies. The used sampling strategy favored large companies who are active in marketing their actions. As this is a consequence of the selected research approach, further work is needed to study other kinds of data business journeys and to larger support for this study. Second, while the selected four companies are rather distinct in their products and services, their experiences in the data business are rather similar. While the selected research approach aimed for greater diversity

of the data business experiences, a more generalizable result could perhaps be achieve with a larger number of cases.

5   Conclusions

This study used a multiple-case study approach to understand how software-intensive embedded system companies are maturing in their data business activities. From the collected data, six distinct stages were identified. The studied companies were maturing through these stages in their data business journeys. In addition, challenges and main activities in each stages were discussed.

References

1. Abrahamson, E.: Managerial fads and fashions: the diffusion and rejection of innovations. Acad. Manag. Rev. **16**(3), 586–612 (1991)
2. Bosch, J., Holmström Olsson, H.: Digital for real: a multicase study on the digital transformation of companies in the embedded systems domain. J. Softw.: Evol. Process **33**(5), e2333 (2021)
3. Bosch, J., Holmström Olsson, H., Crnkovic, I.: Engineering AI systems: a research agenda. In: Artificial Intelligence Paradigms for Smart Cyber-Physical Systems, pp. 1–19. IGI Global (2021)
4. Cao, L.: Data science: a comprehensive overview. ACM Comput. Surv. **50**(3), 43:1–43:42 (2017)
5. Holmström Olsson, H., Bosch, J.: Going digital: Disruption and transformation in software-intensive embedded systems ecosystems. J. Softw.: Evol. Process **32**(6), e2249 (2020)
6. Hyrynsalmi, S.: Letters from the War of Ecosystems–An Analysis of Independent Software Vendors in Mobile Application Marketplaces. TUCS Dissertations No 188, University of Turku, Finland (2014)
7. Lahtiranta, J., Hyrynsalmi, S.: Crude and rude? an old way in a new oil business. Commun. ACM **61**(11), 34–35 (2018)
8. Loebbecke, C., Picot, A.: Reflections on societal and business model transformation arising from digitization and big data analytics: a research agenda. J. Strateg. Inf. Syst. **24**(3), 149–157 (2015)
9. Marshall, M.N.: The key informant technique. Fam. Pract. **13**, 92–97 (1996)
10. Zook, C.: Finding your next core business. Harvard Bus. Rev. **85**(4), 66–75 (2007)