Product Catalogue Process Improvement and Advanced Monitorization

José Luís Faria de Sá

Master's Dissertation

Supervisor: Prof. Bernardo Almada Lobo

Integrated Master in Industrial Engineering and Management

2019-07-01

Abstract

An accurate and well managed product catalogue is crucial for a company with an online luxury marketplace business model. Firstly, luxury clients are incredibly demanding on the quality of information of the product they are acquiring; secondly, the company's partners who sell their products in the marketplace expect nothing less than a premium quality service on what accounts to the management of their products' information. This service not only relates to the accuracy of information that is displayed to the customer, but also to guaranteeing trusted visibility on the status of the items a partner wants to sell at the marketplace.

As an online marketplace's product offering grows wider, so does the complexity on managing all those items' information. This creates a problem for the company, that struggles with presenting the customer with the highest information quality possible, in order to drive sales, but also to the sellers that commercialize their products on the marketplace - being unable to tell these partners why their products are not selling will ultimately create friction in the partnership, putting it at risk.

Heavily supported by advanced process re-engineering and monitoring techniques, this dissertation provides an end-to-end approach to product catalogue management monitoring for an e-commerce marketplace. With a strong emphasis on how crucial it is for everyone involved in a process execution to contribute to its modelling, documentation and improvement, this project presents the several stages a product needs to flow through in order to be available to sell at the marketplace, as well as the appropriate performance indicators to closely monitor aforementioned flow. The developed concepts, processes and respective performance metrics are then consumed by a set of monitoring tools that provide partners and e-tailer with the so needed visibility on the online and offline product catalogue, as well as powerful data insights that support the evaluation of the product catalogue's health.

Finally, backed by a deep research work on the state-of-the-art of product information management tools and on the developed concepts and processes of product workflow management, this thesis provides a scrutinous evaluation of three different providers of tools of this sort.

Resumo

Um catálogo de produtos com informação precisa e bem gerida é crucial para uma empresa cujo modelo de negócio assente num *marketplace* de luxo *online*. Em primeiro lugar, os clientes de luxo são extremamente exigentes quanto à qualidade da informação dos produtos que pretendem adquirir; em segundo lugar, os parceiros da empresa que vendem os seus produtos no *marketplace* esperam nada menos que um serviço de qualidade *premium*, no que diz respeito à gestão da informação desses itens. Este tipo de serviço não se refere apenas à precisão da informação exibida ao cliente, mas também à garantia de uma visibilidade confiável do estado dos itens que um parceiro deseja vender no *marketplace*.

À medida que a oferta de produtos de um *marketplace online* se torna mais ampla, aumenta também a complexidade da gestão de toda a informação relativa a esses itens. Isto cria um problema para a empresa, que se esforça para apresentar ao cliente a melhor qualidade de informação possível, de modo a impulsionar as vendas, mas também para os vendedores que comercializam os seus produtos no *marketplace* - a falta de capacidade de justificar por que é que os artigos de um parceiro não estão a ser vendidos acabará por criar atrito na parceria, colocando-a em risco.

Fortemente apoiada por técnicas avançadas de reengenharia e monitorização de processos, esta dissertação fornece uma abordagem de ponta a ponta para a monitorização da gestão do catálogo de produtos de um *marketplace online*. Enfatizando o quão crucial é, para todos os envolvidos na execução de um processo, participarem e contribuírem para a sua modelação, documentação e melhoria, este projeto apresenta as várias etapas pelas quais um produto flui para estar disponível para venda no *marketplace*, bem como os indicadores de desempenho apropriados para acompanhar de perto esse fluxo. Os conceitos, processos e métricas de desempenho desenvolvidos são consumidos por um conjunto de ferramentas de monitorização que fornecem aos parceiros e retalhista *online* a visibilidade necessária do catálogo de produtos *online* e *offline*, além de conclusões que suportam a avaliação da saúde do catálogo de produtos.

Finalmente, apoiada por um profundo trabalho de pesquisa sobre o estado-da-arte das ferramentas de gestão da informação de produto e sobre os conceitos e processos desenvolvidos acerca da gestão do fluxo de um produto para ser vendido *online*, esta tese fornece uma avaliação minuciosa de três fornecedores de ferramentas deste tipo.

Acknowledgments

First of all, I would like to thank Farfetch for giving me the opportunity to participate in and develop this project. To the Operations Strategy team, thank you for the way you integrated me as one of your own. I found in you, not only colleagues, but also great friends.

In particular, I owe my gratitude to my project team - Ana Souteiro, David Santos and Pedro Justo - for their unrelenting support throughout this project. A special thank you to my mentor, Inês Pereira, not only for her guidance throughout this dissertation, but also throughout the start of my professional career.

I would also like to thank my dissertation supervisor, Prof. Bernardo Almada Lobo, for his help and counsel throughout the development of this thesis. A huge thank you, not only to all the MIEGI professors for all the knowledge provided through the years, but also to the entire course's administration for the amazing opportunities they have given me the luck to take part in.

Throughout the five years I have spent as a MIEGI student, I have made great friends with whom I have been through incredible experiences and learned a lot. A very special thank you to them, for their company, for their knowledge, and most importantly, for their friendship.

Finally, but not less important, I would like to thank my parents and my sister for their constant support. There is not doubt that without them I would not be what I am, nor where I am today.

"Ideas are like rabbits. You get a couple and learn how to handle them, and pretty soon you have a dozen."

John Steinbeck

Contents

Acronyms and Symbols

KPI	Key Performance Indicator
PCM	Product Catalogue Management
PIM	Product Information Management
PLM	Product Lifecycle Management
PXM	Product Experience Management
FTP	File Transfer Protocol
HENRY	High Earner Not Rich Yet
PDP	Product Detail Page
PLP	Product Listing Page
CTO	Chief Technology Officer
BPM	Business Process Management
SS	Spring-Summer
AW	Autumn-Winter
IS	Information Systems
IQ	Information Quality
CRISP-DM	Cross Industry Standard Process for Data Mining
SMART	Specific, Measurable, Attainable, Realistic and Time-bound
CLEAR	Collaborative, Limited, Emotional, Appreciable, Refinable
YoY	Year-on-Year
GUI	Graphics User Interface
MDM	Master Data Management
BO	Back-Office
FO	Front-Office
SO	Season Original
CS	Current Season
MS	Main Season

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Picture yourself as a clothes' store manager. To drive sales, you need to have items on the shelves, and in case there are articles missing, you can replenish them, since you can physically see which ones are missing. Now picture that your store has hundreds of thousands of different items on the shelves and you have no visibility on which ones are missing. With items missing, you will have lower sales, and you will not have any idea on why those items are not selling, since you do not know whether they are available for purchase or not.

This is an analogy for the businesses that experience exponential growth within a short time period, in particular, highly successful online marketplaces. The abrupt increase on the number of sellers commercializing their products in the marketplace, leads to a boom in the number of online items available for sale, and an even bigger number of items that are requested by those partners to be sold at the marketplace. Now, the problem with such rapid growth is that the technology, people and procedures do not accompany it, making it unsustainable. At a certain point, it becomes very difficult for these businesses to know which of the requested items to be sold are not available for sale, and neither the reasons for those items to be offline.

The problem affects these enterprises in an end-to-end scale, as it impacts all of a product's lifecycle since the moment a partner requests it to be sold until it shows up online in the website. It is a type of problem which is very hard to put a monetary value on, since we have no idea of knowing if an item that was offline could have been sold, but it directly affects the company's results and final customer and partners' satisfaction.

Being unable to tell partners why their products are not online puts those partnerships at risk and could even be the cause of their termination. With that said, providing the company and partners with an accurate visibility on online and offline products is of extreme importance to the enterprise (Hunter et al., 2018).

1.2 Company Description

Farfetch started as an e-commerce marketplace with the mission of connecting boutiques spread out around the globe to online buyers. By using Farfetch's services, small boutiques that could only sell to local markets, are enabled to sell internationally, allowing them to increase their customer base, while the customers themselves have access to unique items that would otherwise be unavailable for purchase. The company not only allows these boutiques to sell their products in the marketplace, it also provides a full production service – sellers send the items they want to sell at the marketplace to the production center to be photographed by the company's specialists.

Being the middleman between boutiques and end customers also eliminates the big risk associated with fashion players – seasonality, short lifecycles and, consequently, fast obsoleteness. By focusing on connecting them, Farfetch, unlike its competitors, has no need in acquiring stock.

The company naturally grew from the marketplace mindset to an online luxury retail platform. The expertise that the organization acquired throughout the years allowed it to diversify its product offering. The partners can now make use of not only Farfetch's marketplace and logistical network, but also of its white label service – Black and White – which offers custom made websites and a dedicated production service, so that partners can have their own dedicated online store.

With double digit Year-on-Year (YoY) growth, 10 million website visits per month, customers in almost 190 countries and world-renowned partners, Farfetch prides itself of providing a wide range of products, making it the global leader in the online luxury fashion industry. At the end of the day, the company's main mission is to "move fashion forwards with technology" (Neves (2018)).

1.3 Project Goals

The goal of this project is to provide Farfetch and its partners with the adequate concepts and processes to monitor their catalogue status and product workflow. To do so, two different tools will be developed - a dashboard that provides the company with a snapshot of the current status of each item, and a second dashboard that includes a historical view of the health of a partner's catalogue. To guarantee a scalable and robust catalogue monitoring, the project will also support the creation of a mid-term vision which will include the support of the entire organization in gathering the needed requirements that a product information management tool needs to have. Furthermore, the recommendation of an already existing tool developed by a third-party provider will be made, in case the company decides to acquire one.

1.4 Main Stakeholders

This project is important to the entire organization, and indeed, it can affect its operations and consequentely, it is also important to define who the main stakeholders are.

Product workflow starts since the moment a partner asks the company for a product to be sold online until that product is available for purchase. This means that catalogue monitoring starts at the partner side, making them one of the most important stakeholders. The commercial department is the connection point between partner and organization, which means that this project will also affect how they communicate with the partner.

The product will then be shipped to Farfetch's installations so that it can be photographed, described and receive other specific attributes needed for it to be sold, which takes us to the Creative Operations Department – the other main stakeholder. Since the product spends most of its time in this department, it is very clear that catalogue monitoring will directly affect how it operates.

All data generated during the product's lifecycle needs to be stored in an organized way for it to be used for future projects and to have a history on product catalogue, meaning that the Business Intelligence department is another big stakeholder in this project. Additionally, as mentioned in section 1.3, the team will also support the Engineering Department in the development of a tool to accurately manage product information, making them another key stakeholder.

In order to solve the inefficiencies in managing catalogue, this project was proposed by the Operations Strategy Department. A project team was established with members from the mentioned departments - Creative Operations, Commercial, Business Intelligence and Engineering. All the people involved played key roles in the project, and without them no achievements would have been possible.

1.5 Methodology

The adopted methodology for this project will follow two separate, not necessarily subsequent, workstreams. The first stream will be short-term focused and with the goal of providing the project stakeholders with the adequate tools and processes to allow for a proactive catalogue management monitoring.

In order to provide the company with these tools, we will first need to perform a detailed analysis on the path a product takes until it reaches the marketplace and is available for sale. The outcome of this stage will be a process map, or product workflow. To do this we will need to analyze all different sorts of requirements a product needs to have to show up online, and at what stage of its workflow those requirements are assigned to the product. This stage of the project will involve dealing with large amounts of data, treating it and analyzing it. To do so, the team will make use of the Cross-Industry Standard Process for Data Mining (CRISP-DM, for short).

With a developed and mature product workflow, we will then move on to analyzing its different stages, by creating several metrics with the goal of providing advanced performance evaluation and linking it to the partners' and the organization's performance. The newly developed metrics will then be gathered in an easy to use tool, that will assist the catalogue management stakeholders in closely monitoring its status.

The second workstream will be mid to long-term focused. The end goal of this stream will be guaranteeing that the concepts, processes and tools developed during the first stage are robust

and scalable. To do so, the project team will act as a support team and work alongside with the engineering and technology development teams to gather the needed requirements and provide a recommendation from a business perspective for a tool that will guarantee that all the work performed thus far will be applied to the entire organization.

1.6 Thesis Outline

This dissertation is organized in 6 chapters. Chapter 1 presents the company description, the motivation of this thesis and the respective goals. Chapter 2 includes the theoretical background analysis that will support the decisions made throughout the project. This chapter includes a deep analysis on best practices regarding product information management. Chapter 3 provides a detailed description of the problem, by focusing on how catalogue management monitoring is currently performed in the company and pointing out its flaws. Based on the investigation performed in Chapters 2 and 3, Chapter 4 details the plan that was put in motion to solve the problem. It explains how the work was divided in different stages. The discussion of the obtained results is performed in Chapter 5. This chapter discusses the several deliverables this project produced, along with the proposal of a midterm vision that will guarantee these deliverables will still be used in the years to come. Finally, Chapter 6 will summarize the findings throughout this project and focus on possible future works on the subject.

Chapter 2

Literature Review

This chapter will deep-dive on the theoretical background that will support the decisions made throughout the project. We will first analyze the environment which this project takes part in, followed by a study on the several concepts that constitute business process management. Afterwards, we will focus on studying the theoretical baseline that supports data modelling and visualization. Additionally, an investigation on the importance of information quality and how it can affect both customer and organization will be performed. Finally, this chapter will include a research on what constitutes a product information tool, along with a revision on the state-of-the-art of this matter.

2.1 Environment

This project takes part in Farfetch, more specifically in its marketplace business unit. It is then crucial to study the e-commerce luxury market, along with the specificities of the luxury client.

2.1.1 E-commerce Luxury Market

Okonkwo (2009) argues that placing the Internet in the context of the unique relationship luxury has with its clients would be atypical. Buying a luxury item is not only about the product, but the experience of purchase itself, which is hard to replicate on the Internet. Being unable to replicate some of these aspects for an e-commerce experience, players have turned to other approaches to improve this process. Ease of purchase and return, speed of sending, among others, are elements that can heavily impact the online shopping experience. Adams (2013) adds that nowadays customers value an easy shopping experience, which is the strong side of online retailing and is reinforced by how convenient, fast and available at any time it is. (Dauriz et al. (2014); Harris and Dennis (2002)).

Furthermore, the industry is moving to a consumer-centric era, meaning that customers now expect a more personal shopping experience. As identified on McKinsey's report, "Getting Personal" is one of the key trends for 2019's fashion business (Amed et al. (2018)). There's abundance of data on the customer, which allied to technologies such as artificial intelligence will allow the

industry to achieve or even surpass the level of care and personal experience associated with luxury brick and mortar stores.

Additionally, besides the difficulties with conveying a luxury experience through an online platform, from a business perspective, the problem with online shopping is how easy it is for a customer to buy from a competitor if the desired product is not available in the first option (Bhatti et al. (2000)). This acts as an incentive for retailers to focus on improving their time to market and on optimizing the type of information that is displayed to the buyer, and the way it is displayed – a topic we will focus on ahead.

Finally, it is hard to tell what the future holds for the luxury fashion business, but online luxury shopping is here to stay. Consumers are more tech savvy, and fear of online purchases is getting lower by the day. Achille et al. (2018) also adds that by 2025 almost 20% of luxury sales will be made online (values are given in Figure 2.1).

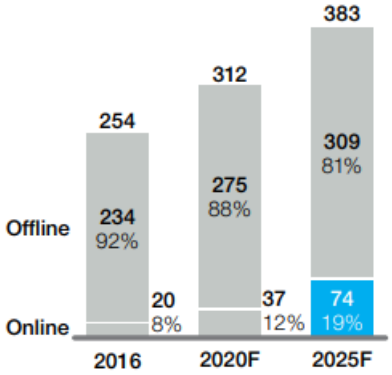

Figure 2.1: Luxury goods sales progression. Source: Achille et al. (2018)

2.1.2 Luxury Client

Having defined the luxury market, its current state and how it connects with e-commerce, it is also important to characterize the typical luxury buyer. According to *Delloite*'s report (Arienti (2019)), a new luxury consumer class is rising nowadays and is going to become increasingly relevant in the future - the HENRYs (High-Earners-Not-Rich-Yet). These customers are smart and tech-savvy. With a smartphone on their hands they have access to every luxury brand and online retailer in the world, which, according to Okonkwo (2010), makes it harder for brands to retain the informed client's attention. Since this type of customers is strongly influenced by technology and social media, big brands' marketing investment focus has shifted to a digital approach with a strong focus on using social media to engage with them.

Okonkwo (2010) also defines the luxury client as "international", "empowered" and "insatiable". The consumer is used to shop in major luxury cities around the world and consequently expects a seamless and coherent brand experience between offline and online channels. The luxury client is empowered due to their influence - with a simple comment on a social network, they can effectively influence and shape the opinions of millions of people. Okonkwo (2010) adds that the

online luxury client has become less brand loyal and more website loyal, meaning that no matter the brand, a customer will avoid buying its products if its website does not deliver.

2.2 Business Process Management

Since this project will have a strong focus on process management and key performance indicator (KPI, for short) elaboration, it is critical to first analyze the underlying theoretical baselines that support the methodology used throughout the dissertation.

Before we dive into what comprises business process management (BPM), it is important to clearly define what a process is. Zairi (1997) argues that a clearly defined process is composed of predictable and definable inputs and outputs, clearly defined tasks and a linear flow and logic.

Business process modelling can be defined, in broad terms, as a systematic methodology to analyze and continuously improve critical processes in an organization which will support companies in sustaining a competitive advantage. It is an all-encompassing approach to organizational improvement that requires the usage of strategic and operational elements, modern tools and techniques and the involvement and cooperation of everyone in the organization (Zairi (1997), Hung (2006)).

Zairi (1997) argues that to achieve its goals, a BPM approach to improve organizational performance needs to guarantee that all the major activities are mapped and documented; there is a customer focus through horizontal linkages between activities; the implemented processes are continuously analyzed in order to perpetually identify possible opportunities for improvement; and finally, proper change management will be conducted, since this is what will ultimately allow a process to be successfully implemented.

2.2.1 Business Process Re-engineering

As Zairi (1997) mentioned, an ideal approach to BPM should have all major activities mapped and documented. This means that the organization needs a systematic approach to perform such documentation. This is where Business Process Re-engineering comes into effect.

Grover et al. (1995) propose a systematic approach to document and map the current state of a process, followed by an improvement and implementation stage. This technique can be divided in four stages (Figure 2.2). The first stage focuses on analyzing how current operations are executed to the utmost detail, which means understanding who is involved in the process, what their tasks are and describe how they are executed. Having gathered all the needed process knowledge during the first stage, the as-is model of the operation is developed. This will involve using standardized visual representations of business models in order to convey the clearest, most understandable idea possible. The third stage will focus on analyzing the as-is model and identifying every opportunity for improvement, which can range from the introduction or removal of small tasks to a complete restructuring of the organization. Finally, having identified the achievable opportunities for improvement, the to-be model of the operation is elaborated. This final model needs to be validated

by everyone in the organization and introduce new, significant changes that culminate in a more efficient process.

Figure 2.2: The Process Re-engineering Methodology

2.2.2 Process Implementation

The to-be model then needs to be implemented in the organization. This stage will most likely involve telling people to do their work in a different way, which can be argued as one of the most difficult tasks with the BPM methodology, known as change management. Moitra (2002) adds that putting the to-be model in motion may involve changes in people's attitudes, behaviors and abilities to improve their performance, which will most likely cause some friction in the organization, delaying the implementation process, or in some occasions, rendering the whole project useless.

Rozenfeld and Costa (2007) argue that identifying the needed changes in an organization is as important as implementing them efficiently. In fact, the benefits of a new, more efficient process are as good as the implementation procedure.

There are several change management methods to implement organizational transformations. Kotter (1995) proposes an eight step framework to successfully transform an organization. This framework starts by establishing a company-wide sense of urgency for initiating a transformation process by identifying potential crisis or major opportunities for improvement. Followed by this, a group of people with enough power to lead the transformation process needs to be assembled. Creating and communicating a vision of the desired outcomes is paramount to the success of the project. This will allow the involved people to focus on the identified objectives and work towards achieving them. This vision will not be immutable, however. It is important to empower others to act on the vision and improve it by encouraging creative ideas and actions in order to guarantee an even more successful transformation process. Additionally, establishing milestones throughout the project with the goal of creating quick wins will allow the management team to leverage on the increased credibility from the achieved quick wins to create even more change and consequently, an even better outcome. Finally, part of the implementation process also includes thinking about the future and guaranteeing that the introduced changes will be carried on and further improved in future projects.

2.2.3 Performance Evaluation

Having implemented the new process model, it is now important to define and structure how its performance will be evaluated. This is a critical stage, since it will effectively tell the organization

whether the transformation process introduced new benefits. Kueng (2000), Enoma and Allen (2007) and Maté et al. (2017) argue that assessing an organization's performance is essential to detect potential areas for improvement, or even recognize problems before they occur, allowing the enterprise to take preventive actions. Metrics used to evaluate a process or an organization's performance are denominated key performance indicators (KPIs). According to Doran (1981) the created KPIs to monitor performance should be SMART, which stands for Specific, Measurable, Attainable, Realistic and Time-bound. However, performance metrics' creation should also be an iterative and continuous process, which is supported by Economy (2015), stating the KPIs should also be Collaborative, Limited, Emotional, Appreciable and Refinable (CLEAR).

Kaplan and Norton (1992) propose a categorization of the different performance indicators dividing them in four groups of perspective: Customer, Business, Innovation and Financial. Customer perspective indicators translate if a business is delivering the right products or services to the client (e.g., percentage of sales from new products). Internal business perspective indicators convey what the company must do to meet the customers expectations (e.g., cycle time). Innovation and learning perspective indicators express the organization's ability to improve and innovate (e.g., new products introduction versus competition). Finally, financial perspective indicators translate if a company's strategy is contributing for overall business improvement (e.g., cash flow). These four groups of perspective are represented in Kaplan and Norton's Balanced Scorecard - Figure 2.3.

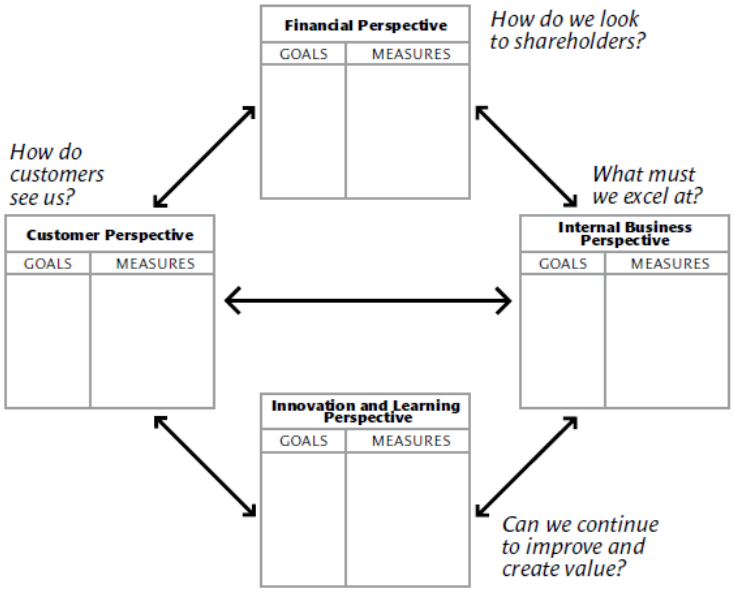

Figure 2.3: The Balanced Scorecard. Source: Kaplan and Norton (1992)

Attached to clearly defined KPIs are target values. These are typically specified in Service Level Agreements (SLAs) (del Rıo-Ortega et al. (2009)) - informal contracts between a service provider and the user of that service (Parish (1997)) - which , according to Larson (1998), clarify

the service responsibilities of both provider and buyer. The SLAs will be defined with the end goal of meeting business requirements and strategy objectives.

2.2.4 Process Owners

Braganza and Lambert (2000) argue that organizations need governance mechanisms that enable them to pursue and achieve their objectives, and add that many enterprises fail to implement new or restructured processes due to low or even lack of these mechanisms. This means that "to implement process-oriented change effectively, business leaders need to develop governance mechanisms that align process and business functions" (Braganza and Lambert (2000)).

Turetken and Demirors (2008) support the need of these governance mechanisms and propose a decentralized approach based on the idea that process modeling and monitoring should be performed by process owners - individuals that actually execute the processes. Process owners are people that will have ownership/governance over a specific process, as well as understand how it is executed in an end-to-end scale to the utmost detail.

2.3 Data Modelling and Visualization

This project will have a strong focus in data gathering and modelling, as well as on the utilization of powerful data visualization techniques. With this said, this section will deep dive on the CRISP-DM methodology, an industry standard technique for data gathering and modelling, followed by a literature review on the importance of data visualization.

2.3.1 CRISP-DM Methodology

According to Hand (2013), the underlying baseline behind the concept of data mining is that organizations produce and store large amounts of information, some of which is not necessarily relevant for the business. Data mining is the process of sifting through large amounts of data to identify patterns and establish relationships with the end goal of solving problems through data analysis. The problem with this is that in most organizations data can be scattered across different databases (such as SQL Databases, spreadsheets and email, among others) and retrieving and relating that data needs various tools and different people. The Cross Industry Standard Process for Data Mining (CRISP-DM) methodology proposes an approach to connect those tools and people in a more efficient way.

CRISP-DM consists of a framework aiming at standardizing the retrieval and processing of business data, making the whole process less costly, more reliable, more repeatable, more manageable and faster (Wirth (1995)). Figure 2.4 represents how the information flows in the CRISP-DM framework.

It consists of six different phases. The first phase – business understanding – focuses on understanding the several business objectives associated with the data mining project. This stage mainly involves meetings with project stakeholders to better understand what the implications of

Figure 2.4: CRISP-DM Methodology - Process Diagram. Source: Jensen (2012)

the project are. The second phase – data understanding – focuses on figuring out where to find the data, collect it, identify possible data quality problems and provide preliminary data insights. The first and second stages are mutually dependent. Identifying possible data quality problems requires a revamp on business understanding to better understand what the correct data is. Data preparation follows the second stage. The output of this phase will be the final dataset. The fourth stage – modeling – consists on applying the appropriate data modelling technique to the identified problem, whose data is stored in the created dataset in the data preparation stage. This can be, for instance, an optimization algorithm or a business process restructure. Evaluation is the fifth stage. In this phase, the applied model will be evaluated before being deployed. This is a critical point to understand if a key business objective was not identified during the business understanding phase. If that is the case, the process restarts. The final stage – deployment – consists on organizing and presenting the knowledge gained throughout the process. It is very important, at the beginning of the project, to understand how to make use of the created models (Wirth (1995)). Lastly, the outer circle symbolizes the cyclic nature of a data mining project, which is not finished once a solution is deployed, but it can trigger several other, more complex projects.

2.3.2 Data Visualization

Data visualization leverages the human brain's capabilities of quickly detecting unique visual properties to allow people to process data in a much faster fashion. By facilitating the recognition of patterns and relationships, data visualization allows to communicate a message in a much more compelling and faster way (Archambault et al. (2015)).

Eckerson and Hammond (2011) provide a definition on the two different categories of data visualization technology, visual reporting and visual analysis - the former uses charts and graphics to characterize business performance, while the latter empowers the users to visually explore data

through filtering, modelling and statistical navigation options. It is also important to mention that a data visualization tool is never an out-of-the-box solution and should always be adapted to the user's requirements. This means that prior to develop one, the users' needs should be gathered and implemented in the tool development process.

Fry (2008) argues that data display requires unique treatment and adequate planning. To support this, he proposes a seven step methodology to understand, model and display data. The first step consists on acquiring the data, followed by ordering it into meaningful categories. Afterwards, all data that is not of interest for the project should be removed. The fourth step focuses on applying statistical or data mining concepts, such as the CRISP-DM methodology described in Subsection 2.3.1 in order to place the obtained data into a logical context. Finally, a basic visual model to represent the data needs to be chosen and further improved, making the visualization clearer. To this model, additional methods to manipulate data, such as filtering capabilities, should be added.

2.4 Information Quality

This project has a strong focus on how information is handled within the company and presented to partners and end-customers, so it is important to analyze the theoretical background under what is known as information quality (IQ).

2.4.1 Impact on the customer

DeLone and McLean (2003) identify, in their success model proposal for an Information System (IS), information quality as one of the key factors that impact customer satisfaction. IQ is defined by McKinney et al. (2002) as the "users' perception of the quality of information presented (...)". This means that an e-commerce retailer with high information quality will allow users to make better, more informed purchasing decisions, which will ultimately drive sales and increase revenue, putting high information quality as one of the top, if not the first priority for an online retailer. Presenting a potential client with wrongful information will almost surely drive the customer away since there are practically no barriers to switch to another provider (Bhatti et al. (2000)).

Furthermore, it is not only the content that drives information quality, but also how that information is presented to the customer. According to Bryan Eisenberg - Buyer Legends co-founder and co-author of the Wall Street Journal Business Week - there are four different types of buyers – competitive, methodical, spontaneous and humanistic (Eisenberg (2013)) – and each one of them needs to be presented with product data in different ways. A methodical buyer, for example, will look for more technical and detailed information on a product, while a spontaneous buyer, driven by speed and emotion, needs to be presented with the most important and easy to understand details about the item as fast as possible. This means that in order to maximize conversion[1], the same product needs to presented in different ways to different types of customers.

[1]The number of users that effectively buy the product, out of those that have accessed the PDP.

2.4.2 Impact on the organization

Having trustful data presented to your customer will directly impact your sales, and consequently your business, but that is only one end of the spectrum. Having poor data quality flowing through your organization leads to hidden costs associated with dealing with that information. IBM states that in the US alone bad data cost $3.1 trillion in 2016 (IBM (2016)). This happens due to a phenomenon that author Thomas C. Redman entitled “The Hidden Data Factory”, as explained in Figure 2.5. This hidden factory is composed of people dealing with wrong data, such as sales-people having wrong forecasts, data scientists cleaning databases and senior executives cancelling their plans because they do not trust the numbers finance provided them with.

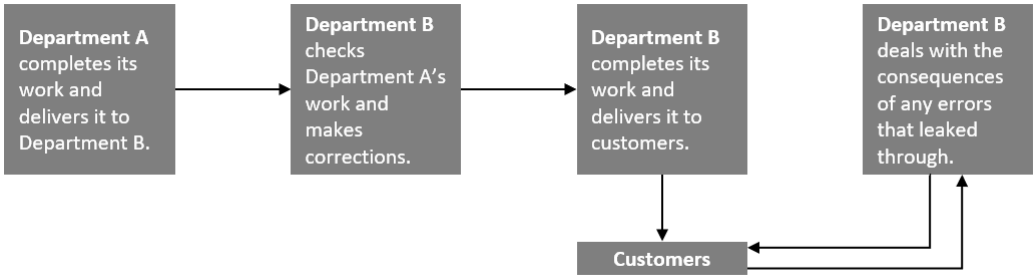

Figure 2.5: The Hidden Data Factory. Source: Redman (2016)

2.5 Product Information Management

As previously mentioned, this project will include the support of the entire organization in gathering the needed requirements that a product information management tool needs to have in order to provide a recommendation of an already existing tool developed by a third-party provider. To do so, this section will firstly focus on studying the concepts that constitute product information management, and culminate on an analysis of the state-of-the-art of this kind of product.

Jorij (2004) defines product information management (PIM) as the set of processes and tools used to centrally manage all information regarding a product, focusing on the essential data to market and sell the item across different channels (e.g., marketplaces such as Amazon or Ebay, brand-specific websites or social media advertising). Jorij also adds that "the end goal of PIM is to have one shared source of product information" and enforces the idea that information should enter and be stored in a company’s databases once. Gartner’s specialists emphasize the sense of trustful information that a PIM tool should handle by referring to it as the "Golden Record", to which all other copies of that data refer to for final validation (Radcliffe and O’Kane (2012)).

Forrester research specialists argue, in the Forrester Wave Q2 2018 Report for PIM solutions (Eppinger et al. (2018)), that enterprises need a product information management system to manage the several sales channels they distribute their products across as well as create, manage and deliver relevant content to the different resellers and market segments they sell in.

Not having a PIM tool places a company at a great competitive disadvantage, since it will struggle to distribute relevant content across different channels and markets. This will, in turn, make a market expansion a very slow and inefficient process, which will substantially reduce the speed at which a company can increase its customer base and, ultimately, its revenue. According to G2 Crowd's specialists, an enterprise has the need of a product information management tool if the following conditions apply (G2 Crowd (2019)):

1. There' is a great amount of product combinations, such as color and size, which cannot be efficiently managed using other systems;
2. The company focuses on regularly launching new products or iterations of existing ones;
3. The company distributes products from many producers into different markets;
4. There is great demand fluctuation based on seasonality.

2.5.1 Core Processes and Features

As mentioned above, having a PIM tool will make market expansions a much easier and efficient process. To do so, Jorij (2004) defines four core processes that such system should have, that are described in figure 2.6.

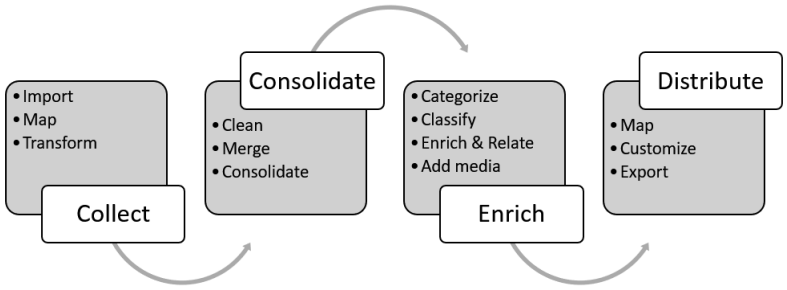

Figure 2.6: A PIM tool's four core processes. Source: Jorij (2004)

Data Collection

For a PIM tool to create the "Golden Record" it first needs to gather the required data. This means that the system needs to be prepared to import data from various sources, such as spreadsheets or SQL databases, map that data to the specific product attributes, and if necessary, even transform the data to comply with the company's standards.

Data Consolidation

As mentioned before, data for a specific product should enter and be stored in the organization once. However, if the company works with large assortments of items there will be situations where the same product data enters the system multiple times. One of the core functions of a PIM

tool is to detect and prevent this duplicated data from jeopardizing the single source of truth for each product.

Data Enrichment

Whenever a product is created in the PIM tool it will only contain basic information, such as unique identifier, color, material type, price and so on. In order to increase the chances for the products to be sold in the different channels they will be commercialized at, the product data needs to be enriched with unique qualities about the item such as more detailed descriptions, dimensions and photos.

Data Distribution

Despite not serving as publication channels themselves, PIM tools can export the enriched data in a ready to use format that will be fed into the different distribution channels.

2.5.2 State-of-the-art of Product Information Management Tools

A PIM solution can either be developed in-house or acquired from a vendor. The former has the advantage of completely satisfying all of the organization's requirements regarding information management and available features. On the other hand, it is a very lengthy, resource intensive process that needs the collaboration of everyone in the enterprise.

There are, however, PIM tool vendors that offer different features and respond to different requirements. With the increasing need of product information management solutions, vendors have come up with different iterations of PIM tools, such as Product Content Management (PCM), Digital Asset Management (DAM), Product Experience Management (PXM), among others. Some authors, such as G2 Crowd's specialists argue that there are significant differences between these products and consider them as different tools (G2 Crowd (2019)). Other authors and experts on the matter, such as Forrester Research's specialists, consider that all of them fall under the PIM's umbrella, but satisfy different needs (Eppinger et al. (2018)) - some of the tools have a stronger focus in managing different types of content, such as video, audio and images, while others may place a higher priority in performance monitoring, such as analyzing how efficient the company is being with content onboarding. For this project, we will share the same vision as Forrester Research's specialists, and consider that all of the above mentioned tools fall under the scope of a product information management tool, but offer different functionalities.

The PIM industry has evolved from offering simple data management software to what is now being called a Product Experience Management tool. This evolution is being fueled by the customer centric era we currently live in. A PIM tool should no longer focus solely on centralizing and distributing all product data, it should also optimize the content delivery to the customer analyzing the information and provide actionable insights on the consumer experience. Additionally, cloud computing and access to machine learning are some of the observable trends that are currently taking place in the PIM market. Cloud computing will allow these tools to better focus on improving their business functions instead of focusing on improving performance, and access to

machine learning will allow for an even further optimization of the customer experience (Eppinger et al. (2018)).

2.6 Final Considerations

It is also essential to pinpoint which of the above-mentioned concepts and methodologies supported each stage of this project.

In an initial stage, the CRISP-DM methodology was used to better understand how catalogue-related data was stored and organized in the company. Next, the studied concepts and frameworks for an efficient and effective business process re-engineering and management were put into practice to model the process that allows users to check the reasons for a product to be offline, as well as metrics that assess the catalogue's general health. Additionally, the concepts introduced in Subsection 2.3.2 helped the team in developing a set of dashboards that include and allow users to interact with all the contents that were developed throughout this project.

Finally, the knowledge gathered from the literature review performed around product information management (PIM) tools supported the investigation process to understand what type of features a tool of that sort needs to have which in turn, aided the team on the research for the best PIM vendor.

Chapter 3

Problem Description

As mentioned before, Farfetch has been showing double digit year-on-year (YoY) percentage growth for the past 10 years. This kind of growth is reflected in the number of brands and different products sold at the marketplace. In March 2018 Farfetch had more than 2200 different brands for sale with an estimated shelf value of € 140 million. The number of online products available for sale surrounded the 500 thousand items for the Spring Summer 2018 (SS18) collection alone. Throughout the entire duration of the SS19 season, there were almost 800 thousand different products from 3000 brands for sale at Farfetch.com, corresponding to a 60% catalogue size increase in comparison to last year's figures. The difficulty with such a fast growth is matching its speed with the evolution of company resources – people, processes and technology – which are what allow it to be sustainable.

That being said, this project aims to create a set of tools that support a sustainable business evolution regarding catalogue management monitoring. This section features an in-depth description of the problem. It will start with an explanation of the different concepts in play, such as brand tier, geopricing, blocking rules and slot management, followed by an analysis on how catalogue management monitoring was performed before the project deliverables were put in motion and how it affected the different stakeholders. Finally, a review of the shortcomings of the current solution is performed.

3.1 Key Concepts for Catalogue Management

Before deep-diving into the problem this project aims to solve, we should first define some of the key concepts and processes that will be used throughout this dissertation. This section will then feature an explanation of the different types of season that an item can be simultaneously assigned to - season original, current season and main season -, followed by an introduction to the different partner's hierarchical levels that Farfetch works with. Afterwards, a simple explanation on the specific place where a partner stores its stock to sell at Farfetch's marketplace will be provided. We will then explain how the company deals with the common situation where two or more different partners want to sell the same item - a duplicate - simultaneously. Additionally, we

will provide an explanation on how the company or its partners can use blocking rules to block an item from being sold online, and finally, we will explain how a partner can manage the stock it provides Farfetch to sell.

Product Season

In the fashion world, every item is associated with a season - Autumn/Winter(AW) and Spring/-Summer(SS). The styles for each one of these seasons vary from year to year, depending on observable customer trends. This means that it is also necessary to associate a year to a product. For example, the trends for Autumn/Winter 2018 (AW18) will differ from the ones in AW19. If a brand deems an AW18 product's style inadequate for the AW19 season, the item will be discontinued and not sold in the following year. However, if the style is still appropriate, the brand may decide to change its season from AW18 to AW19, turning the item into a carryover.

Farfetch's databases store product information in a way that tries to capture the season during which an item was originally created by its brand - Season Original (SO) - but also in a way that captures the action of updating an item's season to a new one, turning it into a carryover - Current Season (CS). This means that in Farfetch's databases a product is associated, at any point in time, with two different seasons - Original, which will remain unchanged throughout a product's lifecycle, and Current Season, which will be updated every time there is a season change - for example, if the item was created while the trend season was AW18, but it has been carried to the AW19 as per the partner's request, then its original season will be SO-AW18, while the current season will be CS-AW19.

Additionally, a third season concept will be introduced in this project - the main season (MS). This is the season in scope for a certain year. For example, in 2019 the main seasons will be MS-SS19 and MS-AW19.

Finally, it is also important to clarify which period of the year each season lasts through. This project will include several analyses throughout a season's duration. Each season lasts for 6 months, and its duration is weeks 22 to 47 for the AW season, and weeks 48 to 21 of the following year for the SS season. For example, an item created in March 2019 will have AW19 as its original, current and main seasons.

Brand Tier

It is important to understand what type of partners Farfetch deals with before we enter into product catalogue details. The company has two groups of partners: one, boutiques are identified as tier "T" and two, entire brands as tier "B". The main difference between these two categories is that boutiques sell several brands while the "B" partners sell their own brand. Within each of these tiers there are hierarchies that translate the size and importance of the partnership. These hierarchies range from level 0, 1 and X, with B0 brands and T0 boutiques being the most important partners for the company.

Stockpoint

This project will deal with large B0 partners with multiple stockpoints, so it is important to understand what this concept means. In simple terms, a stockpoint is the physical place where a partner stores all its stock. This means that for a partner to sell a specific item at a stockpoint, he needs to have that product created and with available stock at that stockpoint.

Duplicates

As explained before, Farfetch marketplace's business model consists, in very broad terms, in connecting partners spread out around the globe to consumers that would otherwise have no access to those stores. With the increase on the number of integrated boutiques and the integration of brands that commercialize their whole product catalogue at Farfetch's marketplace, it is only natural that at one point two or more partners want to sell the same product referenced as duplicate.

Partners have the freedom to commercialize any product they want at Farfetch's marketplace – albeit with some underlying rules (e.g., a sweater with explicit content will not be allowed to be sold) – meaning that there is no impediment if they want to sell the same product. However, it is in everyone's interest to eliminate the number of duplicates that are produced at Farfetch's production centers, which means that the company wants to avoid dedicating resources to the same product more than once and the partner wants to avoid having one less item in their inventory.

There are some mechanisms that prevent this situation. Duplicates can be detected at three different stages from the moment a partner tries to create the item in Farfetch's database until the product reaches the production facilities. A duplicate can be detected through its Designer ID – the brand's style code for each item -, meaning that two different items will never have the same Designer ID. If an item is detected as a duplicate, the partner can automatically associate its product to the already existing one, allowing it to be easily put online. If the item is not detected as a duplicate, it follows along the production process. This means that eventually a customer might find the same item twice or even more while navigating Farfetch's web-page – an undesirable situation that has a negative impact on the client.

Blocking Rules

Blocking rules are one of the main ways of managing the online availability/visibility of a product. They are rules that can be applied to specific products, an attribute or even a country, that will block items that do not obey those rules from being sold. For example, a product may be blocked from being sold overseas due to a certain material composition.

Stock Management

There are two possible options for a partner to manage its stock, one, Barcodes and two, Stock Sync. If a partner manages its stock via barcodes, every time a purchase is made, be it online or offline, the item will need to be scanned out by the partner himself using STORM - a partner-dedicated platform where the stock, along with other important information, can be managed. The same principle is applied whenever the partner wants to add stock to a specific item. For

this type of stock management, the partner will have to associate its store barcodes to the product identification code in Farfetch's database.

On the other hand, if the partner manages its stock by stock sync, inventory updates will be automatically sent on a regular basis to Farfetch. For this type of stock management, the partner will need to have its database integrated with Farfetch's one.

3.2 Key Processes for Catalogue Management

A product needs to undergo four critical processes in order to gather all the attributes it needs to be available for purchase at Farfetch's marketplace. These processes are the creation process, followed by the slot allocation, production process and culminate in the addition of remaining data. This sequence is better depicted in Figure 3.1. Each of these processes will be introduced in the following sections.

Figure 3.1: Key Processes for Catalogue Management

3.2.1 Creation Process

Before a partner ships an item to be photographed at Farfetch's production facilities, the product needs to be created in Farfetch's databases, hence the creation process. There are three alternatives that a partner has for creating an item: one, manual creation, two, file upload, and three, integrated creation. These options are further scrutinized below.

The creation process culminates in the generation of a unique product identification code – the Item ID – that will be stored in Farfetch's databases alongside all the product's attributes. Figure 3.2 provides a high level view on the process.

Manual Creation

Partners that undergo a manual product creation process create the items themselves one by one by filling in an online form in the STORM platform. During this process, the system might identify duplicate items, since one of the mandatory attributes for product creation is the Designer ID. If that situation arises, the items are automatically associated with the original product and the duplicate is not created. This is one of the three above-mentioned stages where duplicates are detected.

File Upload

For the creation of multiple items at a time, an Excel template is provided to the partner, where the column names represent the mandatory and optional attributes that a product needs to possess

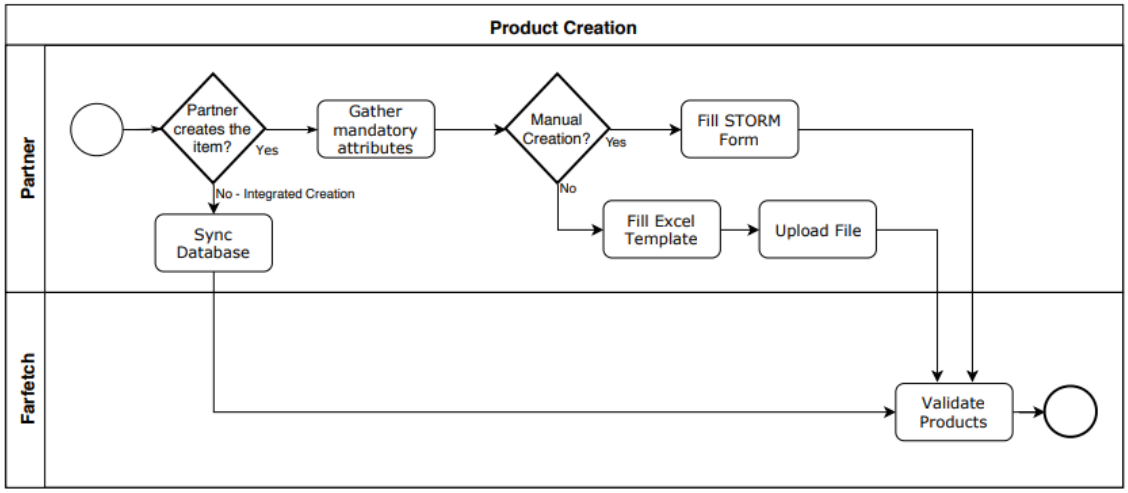

Figure 3.2: Product Creation Process

to be created in Farfetch's database. These are the same attributes the partner would need to fill in for the manual creation process. The difference here is that after filling in the Excel template, the partner uploads the file to the STORM platform and the items are automatically created in Farfetch's database.

Integrated Creation

The integrated alternative for product creation is only possible if the partner is integrated within Farfetch's systems. An integrated partner's database automatically connects with Farfetch's one. Needless to say that for a partner to be integrated, his data structure needs to be mapped in order to match Farfetch's data structure. There is a whole another process that describes how a partner integrates with Farfetch's systems, but that is not a topic of this project. This means that when a partner creates a product in its own system, the item is automatically created within Farfetch's system, which speeds up the creation process.

After creating the item, the partner will then assign it to a slot - a process that will be further scrutinized in the following subsection.

3.2.2 Slot Allocation

Before sending a product for production at Farfetch's facilities, the partner must first assign that product to a slot. A slot is basically a set of products. Usually a partner sends more than one item for production and allocates up to 50 items to the same slot. For a visual representation purpose, picture the slot as a box that can contain up to 50 items. For these products to be shipped for production, they first need to be fitted inside the box. The goal of this stage is to incentive partners to send multiple items for production simultaneously, and not one at a time.

3.2.3 Production Process

There are two types of production processes – shooting and FTP. Shooting – the most common production process - starts the moment the slot containing the item arrives at the production center. After opening the slot, the product's barcode is scanned in. This is known as the scan-in task. The last duplicate verification occurs at the scan in station, where a collaborator checks Farfetch's databases for identical items. After being scanned in, the product follows along a complex production line where it gets extensively quality checked to make sure it is not damaged until it is photographed. After the images are uploaded, a dedicated team will describe the product with a user friendly text that describes the essential product information such as the material type, in a more easy to read manner, while simultaneously, a separate team will quality check the photographs. After this, the item will be approved for scan out and can be shipped back to the partner. Figure 3.3 portrays a high-level view of the shooting production process.

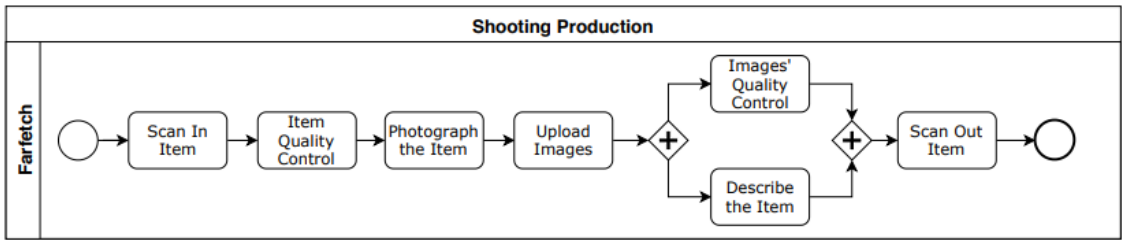

Figure 3.3: Shooting Production Process

FTP Production - meaning File Transfer Protocol Production – consists on a partner, after creating the item, uploading images that the partner himself produced for the item in question. These images are then quality checked for compliance with Farfetch's standards and follow the same path as with the shooting production. This type of production is relatively recent in Farfetch and only selected partners - majoritarily B0 partners - are allowed to use it. There are currently no mapped activities for this type of production, which could indicate that the process is not being executed in the most efficient way. Its management and utilization is done entirely via shared media folders and spreadsheets scattered across the organization. Figure 3.4 displays a high level view of how the FTP production process is currently executed in the organization.

3.2.4 Additional Data

After being produced, a product still needs some additional data in order to have all the needed attributes to be available for purchase at Farfetch's marketplace. This data includes defining the blocking rules for the product (if it has any), inserting the stock value, associating the product's barcodes (if the partner has integrated stock management), as well as the geoprice, better defined in the following section.

Geopricing

Geopricing is the practice of stipulating the price of a product according to the geographical location of the customer. The geoprice is defined by the brand of the product on a seasonal basis. For that, brands define a series of geographical areas and provide a price for each product that they want to sell in each of those regions.

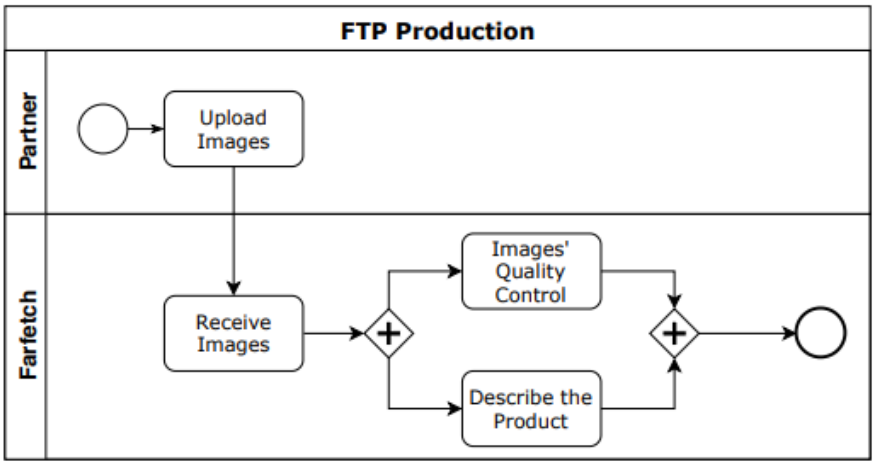

Figure 3.4: FTP Production Process

There are two types of geoprice – mandatory and suggested. The mandatory prices are, as the name says, mandatory attributes for a product to be available online. For example, if product A has a mandatory geoprice of 1000$ in the US, but boutique X wants to sell that product for 1050$, the item will be unavailable for purchase at that particular boutique. On the other hand, the suggested geoprices are not mandatory for a product to go online. They act as a price recommendation for that product to be sold in a specific area.

3.3 Catalogue Management

Finally, we must understand what catalogue management entails in the context of this problem and why it is so important for a company as Farfetch, or even any other e-commerce-related enterprise to have an end-to-end vision on its catalogue. A company's catalogue comprises all the products that it manages. From an e-commerce perspective, this means not only understanding which products are online and available for purchase, but also those that are offline and the reasons that justify that situation. From a customer perspective, catalogue management is critical to find the desired product with accurate and inspirational content, while from a partner perspective, catalogue management acts as a sales enabler – it is only possible to sell what is online. Finally, managing the product catalogue is also an important type of Farfetch – Partner interaction that can shape a partner's product offering for Farfetch to sell in the marketplace.

3.3.1 State of the Art of Catalogue Management

Now that the key concepts and processes are defined, we can dive on the state-of-the-art of catalogue management before this project started. Throughout this section we will also highlight the shortcomings of the above-mentioned processes.

To put it simply, Farfetch had no vision on the importance of catalogue management monitoring before the start of this project. Checking why an item was offline was a reactive process - it would only happen whenever a partner requested so. Additionally, there was no standardized process model to check the reasons for an item to be offline. This was a manual and slow process, based on private messages and emails. This means that the whole foundations for a stable catalogue management monitoring needed to be built, but first a baseline needed to be constructed to identify what constitutes the catalogue workflow. To achieve this, a discovery task-force composed of more than 20 people was assembled. From this study, the six main modules that constitute what will be the catalogue workflow were identified: (1) product creation; (2) slot allocation; (3) production; (4) geo-price management; (5) barcode management; (6) blocking rules management.

Supported by the literature review performed in Chapter 2, to guarantee a well-structured management of each one of these modules, four main components were identified: (1) processes; (2) process owners; (3) tools; (4) process monitoring. These individual components are further scrutinized in the next topics. The main outcomes of this discovery process are summarized in Table 3.1.

Table 3.1: Project Discovery Results

Catalogue Workflow	Product Creation	Slot Allocation	Production	Geo-Price Management	Barcode Management	Blocking Rules	End-to-End
Standardized Process Model	No	Yes	Yes	No	No	No	No
Owner	Yes	Yes	Yes	No	No	No	No
Adequate Tools	No	No	No	No	No	No	No
Monitoring	No	Yes	Yes	No	No	No	No

Processes

Each of the identified main modules has predictable and definable outputs, a linear and logic flow, clearly defined tasks and a predictable outcome, which according to the definition introduced in Chapter 2 is what characterizes a process. With this said, for every main module, the company needs a clearly defined, non-ambiguous process model that accurately translates the workflow inside these stages. Currently, only the slot allocation and production activities have standardized processes, with the remaining six modules' process models still being created.

Process Owners

As mentioned in Chapter 2, process owners are people that actually perform the process and have an end-to-end understanding of it to the utmost detail. They are crucial to better understand,

analyze and continuously improve the process they own. Currently, geo-price and barcode management, as well as blocking rules have no defined owners. This means, for example, that anyone in the company can create a blocking rule that stops a certain product from being sold.

Tools

For an efficient workflow in each one of the identified main modules, Farfetch needs dedicated tools. The currently existing ones were thought out for "Boutique Partners" and now the requirements are different, as these partners get larger and Farfetch onboards bigger players. There are no adequate tools for most of the stages. This has a strong impact on our monitoring abilities. Not having the needed tools, leads to having no centralized source of truth, which creates the already discussed hidden data factories.

Monitoring

Having dedicated owners, process models and tools are what allow catalogue monitoring to be performed, however, dedicated performance indicators will allow the company to identify potential areas for improvement and even hazardous situations before they occur. Currently, only the slot allocation and production stages are monitored. For example, Farfetch is monitoring the time an item spends in the production stage - the production lead time. This is, however, not enough. The company needs to understand, for example, how long does a product take from the moment of its creation until its first appearance online.

3.3.2 End-to-End Monitoring

In addition to the aforementioned six modules, there is a seventh hidden one: the end-to-end vision. Having an end-to-end vision on a complex process is quite challenging when it involves a common effort from several departments in the company, which is the case for catalogue management at Farfetch.

As companies grow, departments specialize in specific functions. This is beneficial to achieve high efficiency in repetitive tasks, but it also affects companies by creating organizational silos. The silo mentality is defined as a reluctance to share information across departments within the company. Silos cause problems primarily because they focus on fulfilling a function rather than an outcome (Dell (2005)).

This discovery process concluded that there was no end-to-end vision on catalogue management monitoring (e.g., there was no vision on what connected the creation and the production processes). Furthermore, it was discovered that more than 40% of the variables needed to perform an accurate end-to-end catalogue monitoring were not stored by the company, and that there were no common principles' definition on what is required to go online. (e.g., some people in the company saw product descriptions as a mandatory attribute for a product to be available online, while others saw it as optional and that descriptions could be added after a product went online).

If this project focused solely on creating the needed tools, processes and monitoring requirements for the six main catalogue workflow modules, without first creating an end-to-end vision on the whole process, the deliverables would simply be rendered useless.

3.4 Summary

Not being able to explain to a partner why some of their products were not online could have terrible consequences for the company. As an example, at one point Farfetch could not explain to a partner why its whole catalogue was offline, which lead to the partner threatening to terminate the partnership. In a different occasion, Farfetch noticed an almost 30% drop in a partner's number of produced items compared to the previous year, and neither the company, nor the partner could understand why that was happening, simply because there was no visibility on the offline items.

Farfetch currently has a reactive catalogue monitoring approach, meaning that the company only investigates the reasons that justify a product's online status if a partner requests so. By that time, the partnership is already at risk. The goal of this project is to first develop the adequate tools and processes that allow a proactive monitoring of a product's workflow, and after that support the organization by guaranteeing that the developed concepts, processes and tools are scalable and future-proof.

Chapter 4

Methodology

This chapter will focus on describing the methodology followed during the short and mid-term workstreams, mentioned in Section 1.5. As mentioned in Chapter 3 the discovery phase of the project allowed the team to understand that the company lacked an end-to-end vision on catalogue management monitoring. Consequently, building the needed processes and tools to support that vision was set as the first project priority and is what will constitute the End-to-End Monitoring section of this chapter.

After developing the adequate processes and tools to perform a proactive catalogue management monitoring, the project will focus on guaranteeing that what was developed will be sustainable and scalable. To do so, the project team will provide the company with a recommendation on the best tools for product information management.

4.1 Project Team Creation

Having identified what the critical problems with catalogue management monitoring were, it was needed to select a team responsible for tackling and providing solutions to the identified shortcomings. As mentioned in Section 1.4, this project affects the organization in an end-to-end scale, meaning that there are several separate teams working on the subject. However, this dissertation will focus on the work developed by the supporting business operations team, that collaborated directly with the content master data management team.

With this said, the main team structure includes three different departments. The Content Master Data Management department will provide insights and guide the project team on catalogue-related processes, data structure and storage side. The Engineering and Technology department will build the needed technological tools, based on the work developed throughout this project, which will be used by the entire organization for product catalogue management. Finally, the Business Operations department (part of Operations Strategy) will support the entire project team by developing the adequate concepts, tools and processes for catalogue management monitoring.

4.2 End-to-End Monitoring

As mentioned in Chapter 3, the discovery phase of the project allowed the team to understand that the company lacked an end-to-end vision on catalogue management monitoring. Consequently, building that vision was set as the first project priority. This means not only creating the product workflow - the process a product goes through to show up as online at Farfetch's marketplace - but also developing key metrics to monitor that flow and bridge the gap between commercial and operational performance indicators, leading to an holistic view of product catalogue. The end goal of this section is to define processes and their owners, as well as provide the company with the adequate tools and metrics for an end-to-end catalogue monitoring, which as discussed in Chapter 2 are the four pillars that support a well structured business process management.

During this stage the team was fighting against time, since there was pressure from B0 partners to provide them with visibility on the status of the products they requested to be commercialized at Farfetch's marketplace. Additionally, the team was dealing with enormous amounts of data, since the project required analyses of day to day variations at an item level, which caused long processing times on the databases. Consequently, to speed up the process, thirteen B0 partners were selected to be the first monitored brands. Despite this, all the developed concepts, processes and tools were created with a focus on scalability and sustainability.

4.2.1 Processes

Using the CRISP-DM methodology as support, a deep business understanding was gathered through a series of workshops. The key goal of this stage of the project was to gather the needed information to map the stages that compose the end-to-end catalogue management monitoring process, which, as mentioned before, there was no previous documentation on.

Having developed the needed documentation and standardized concepts will allow for two positive outcomes: (1) standardization of the journey a product makes since the moment a partner requests it to be sold online until it shows up as available for purchase at Farfetch's marketplace; (2) mapping the process to verify the reasons that justify a product being offline.

4.2.1.1 Product Workflow Stages and Leakages

Having no developed documentation and model meant that all the created concepts and terminology needed to be company-wide accepted and adopted. To guarantee this homogeneous vision on catalogue-related concepts, the organized workshops gathered people from different areas of the company - Product (Front and Back-Office), Business Intelligence, Black and White business unit, Creative operations, Global Brands and Operations Strategy.

These workshops resulted in a unified vision of the different stages a product needs to flow through since it is requested by a partner to be sold until it finally shows up as online at Farfetch's marketplace, as well as the set of information requirements that the product needs in order to reach

each one of those stages. It is also worth mentioning that a product can never be at two different stages at the same time.

Finally, it is also important to have a unified understanding of the several reasons that might prevent an item from flowing from one stage to the next one - the leakages. These are the reasons that will ultimately tell the company what is causing a product to be offline. With this in mind, the team created the concept of leakages and sub-leakages. The several stages and leakages a product can go through are described below and portrayed in Figure 4.1.

Figure 4.1: Product Stages and Leakages

Stage 1 - Item Requested

Requested items are products that a partner has asked the company to create, produce and sell at the marketplace. A product can only go through this stage if the partner has product creation by integration - explained in subsection 3.2.1. If this is the case, the requested items are the ones that the partner has created in its database and are going to be integrated in Farfetch's product catalogue.

Stage 2 - Item Created

As described in Subsection 3.2.1, an item is created only if its mandatory attributes (e.g., DesignerID, Price, Season, Gender) are correctly filled by the partner. As mentioned before, this can either be a manual or automatic process, depending whether the partner is integrated or not in Farfetch's system, and culminates in the creation of a unique product code - the ItemID - in the company's database.

Not every requested item gets created, though. This happens when for example, an error occurs when mapping the partner's information to a format that is accepted and understood by Farfetch's database.

Stage 3 - Item Created and Ready for Production

The data requirements for a product to be considered created and ready for production differ depending on the type of production process it has to go through - FTP or shooting.

On the one hand, if the item is to be produced via shooting, after being created, the partner assigns the product to a slot to be picked by a carrier and shipped to Farfetch's production center. The item is considered ready for production when the scan in task - better detailed in Subsection

3.2.3 - is performed. The slot allocation task is critical for a product to flow from the "created" stage to the "created and ready for production" stage, since it is the only way for it to be shipped to one of the multiple production facilities.

On the other hand, if the item is to be produced via FTP, the partner must upload the photographs to a shared folder. Since there is no need for the product to be shipped and photographed, there will be no slot allocation. However, Farfetch's information system is built in a way that does not allow an item to be sold at the marketplace if it does not have an allocated slot. To bypass this obstacle, a virtual slot is manually created and allocated to the item by Farfetch. Similarly, the scan in task will never be performed for this type of production and is replaced by the date of upload of the photographs by the partner. After having at least two uploaded media files, the FTP produced item is considered ready for production.

Stage 4 - Item Produced

Independently of the type of production an item goes through, after having uploaded photographs, associated descriptions, additional data entry and a specified category (e.g., men shoes), it is considered as produced.

Several situations might arise that stop an item from flowing from the ready for production to the produced stage, for example, if there is an error with the shooting process, the product will need to be photographed again.

Stage 5 - Item Ready to Go Online

For an item to be ready to go online, it needs to have a validated geoprice, in case it is mandatory, product barcodes and its associated blocking rules need to allow the product to be sold in at least one country.

Stages 6 and 7 - Item Online Back Office and Online Front Office

Finally, if the item meets all the aforementioned requirements, and the partner confirms that stock is available to be sold, the item will move to the "Online Back-Office (BO)" stage. The difference between the "Online Back-Office" and the "Online Front-Office (FO)" stages is a simple binary flag. If the item meets all the requirements and has the binary flag marked as true, then it is considered to be in the "Online FO" stage, however, there are situations where the item meets all the requirements to be available online, but is not marked as so in the database. Having a number of products in the "Online BO" stage different from the one in the "Online FO" stage is a rare situation, but might happen due to synchronization issues with the database.

It is important to mention that all the requirements are cumulative, meaning that for an item to be in a certain stage, it needs to have all the mandatory attributes from the previous stages, but the inverse situation is not true, meaning that a partner can supply the company with mandatory information for several subsequent stages at the same time. For example, at the creation stage, the partner can also upload the amount of stock that will be available for purchase for the item being created. This means, for example, that a product may have stock but still be in the "Created" stage

4.2.1.2 Leakages and Sub leakages

Besides having a common vision on which different stages a product can flow through during its lifecycle, it is also important to have a unified understanding of the several reasons that might prevent an item from flowing from one stage to the next one. These are the reasons that will ultimately inform the company on what is causing a product to be offline. With this in mind, the team created the concept of leakages and sub-leakages. Numbers 1 through 6 in Figure 4.1 represent the several leakages a product can flow through.

To put it simply, a leakage bridges the gap between two stages, and a sub-leakage is a deep dive on the leakage that aims to further explain the reasons that stopped a product from flowing from a stage to the next one. This means that in between two stages there is a leakage - (1) Not created; (2) Not ready for production; (3) Not produced; (4) Missing mandatory information; (5) Missing stock; (6) Lag between BO and FO -, and a plenitude of sub-leakages, for each one of those leakages. For example, if an item is already created in Farfetch's database, but the slot that contains it has not yet arrived to the production center, the product is considered to be in the "Created" stage, but at the same time, in the "Not Ready for Production Leakage", and the "Slot not arrived" sub-leakage is what explains this product status.

It is worth of note that each one of these sub-leakages is associated with a "Partner Action", "Farfetch Action", or a "No Action" if the product status still meets the SLAs. Using the same example as above, if a slot has not yet arrived, then we are facing a "No Action" type of sub-leakage, since neither the partner nor Farfetch can take any measure to make the slot arrive. However, if a slot is taking too long to arrive to the production center, the sub-leakage becomes a Farfetch action, since the company needs to take measures to ensure the slot will eventually arrive. The several sub-leakages of the "Not ready for production" leakage are described Table 4.1. For a more in-depth analysis on the different leakages and sub-leakages, the reader is referred to Appendix A.

Table 4.1: Product not Ready for Production Leakage and respective sub-leakages

Product not Ready for Production Leakage		
	Sub-Leakage	**Definition**
Shooting	**Not Added to Slot**	Items that are not ready for Production because they were not added to a slot.
	Added & Not Ready to Send	Items that are not ready for Production because they haven't a Ready to Send Date (RTS) defined, despite having already been added to a slot.
	Ready to Send & Not Allocated	Items that are not ready for Production because they do not have an Estimated Time Arrival (ETA) defined, despite having already been added to a slot and had a RTS.
	Slot not Arrived	Items that are not ready for Production because their slot has not arrived yet.
FTP	**Partner did not Send Media**	Items that are not ready for Production because the Partner did not send any images yet.
	Error on Media Integrations Process	Items that are not ready for Production because the Media Integrations Process is with an error, despite the partner has already sent the media.
	Media not in FTP Folder	Items that are not ready for Production because there are not any images in the FTP shared folder, despite the partner has already sent them and there is no error in Media Integrations Process.
	Missing Views in Folder	Items that are not ready for Production because there is only one image in the FTP shared Folder.
	Not Added to Slot	Items that are not ready for Production because they were not added to any slot.

4.2.1.3 Product Availability

Having clearly defined product stages and respective requirements meant that the team could finally create a process that would allow the project stakeholders to check more efficiently the reasons that are causing an item to be offline - the product availability process.

This process provides a set of clearly defined instructions to sequentially check which one of the mandatory requirements for a product to be available online is missing and what actions to take whenever the missing mandatory field is found (e.g., if the geoprice is not validated, the partner should be contacted and asked to validate it). Figure 4.2 portrays how this process is executed.

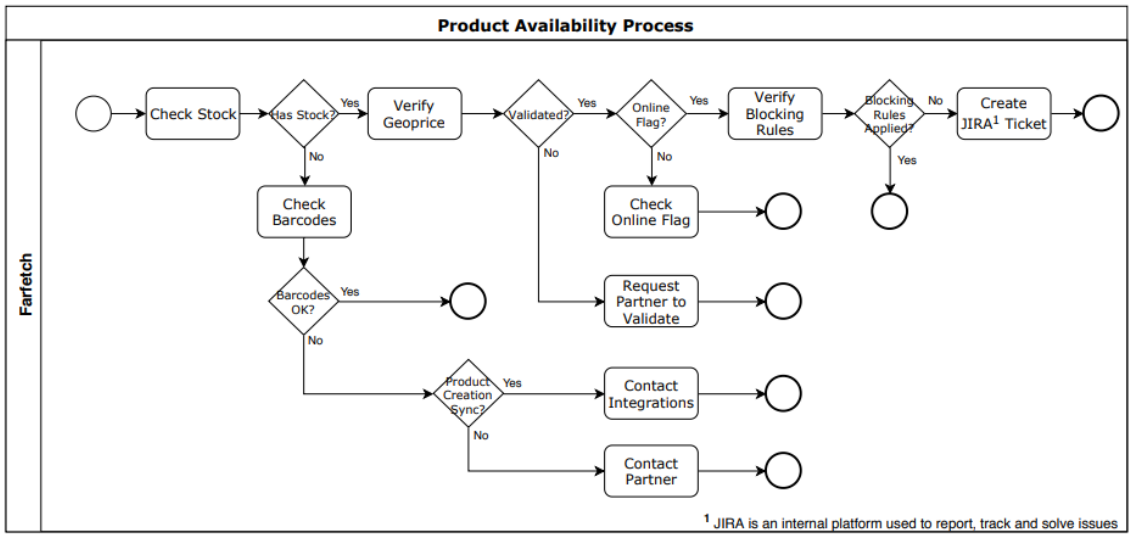

Figure 4.2: The Product Availability Process

In case a JIRA ticket is created, the process flows to the engineering team to further investigate what might be causing the item being offline, a process that is out of scope of this dissertation.

4.2.2 Process Owners

As discussed in Chapter 2, Turetken and Demirors (2008) define process owners as individuals that actually execute and have an end-to-end understanding of the process they own. With this said, it is important to define who will be the owners of the end-to-end catalogue monitoring process, who will be referred, from here on out, as catalogue specialists.

Each catalogue specialist will be responsible for monitoring at least a B0 partner's entire catalogue. To do so, they will make use of the developed concepts and modelled process introduced in Subsection 4.2.1. The end goal of a catalogue specialist's functions is to provide the monitored partners with the so needed visibility on the requirements a product needs to be available for purchase at Farfetch's marketplace. These individuals will not, however, report directly to the partners. They will, instead, report to the company's delivery managers - Farfetch's collaborators that have direct contact with the partners.

4.2.3 Tools

The one shortcoming with the modeled process - product availability - is that it is far too complex, meaning that it takes a long time to go through all the possible reasons that may be causing a product to be offline. This is a big obstacle for the project, since the end goal of the company is to become proactive in catalogue monitoring. This means that the product availability process will not be applied to a single product, but to the entire catalogue of the monitored B0 partners, and eventually to all the company's partners.

To surpass this obstacle, the team will need to create tools that automatically run through all the reasons that cause a product to be offline as well as provide catalogue health related metrics that allow, for example, to evaluate how the number of online products has been progressing through time.

Requirements

As with any other type of tool, its development needs to be done taking into account the users' requirements. To gather this information, several meetings were conducted to understand what the catalogue specialists needed. The identified requirements are as follows:

- Automatic data refreshment will allow users to have access to the most recent product availability information;
- Drill-down capabilities are essential to perform analyses for the monitored partners as a whole or for a single partner;
- Filtering capabilities are also a feature that will allow users, for example, to analyze only SS or AW items;
- Ease of data extraction - As we will discuss in Subsection 4.2.5, catalogue specialists will elaborate a weekly report containing all the information gathered in the catalogue management monitoring tool. This means that one of the most important features is the ability to extract data onto an excel spreadsheet to allow further data manipulation;
- Online publishing will allow several users to handle the tool at the same time.

Taking these requirements into consideration as well as the several software platforms available for Farfetch collaborators, *Tableau* was the chosen solution to support the development of these tools. It combines SQL server integration, which will possibilitate the connection with the company's data sources, with powerful data visualization features.

4.2.4 Performance Evaluation

Supported by the literature review performed in Chapter 2, evaluating the performance of the newly introduced processes is critical for three main reasons:

1. Understand if the new processes introduced new benefits to the organization;
2. Detect potential areas for improvement;
3. Identify problems before they occur.

With this said, the developed processes and tools were done so with a continuous improvement objective in mind. Creating SMART and CLEAR metrics to monitor the introduced transformations will allow the team to closely supervise the impact that catalogue management monitoring will have across the organization, which will, in turn, support the identification of potential areas for improvement and process re-engineering. Having operational perception of catalogue management monitoring metrics will allow not only partner facing teams to answer brand concerns, but also business oriented ones to connect the impact that the project might have on the business (e.g., number of sales).

As with the product availability process model, every developed metric was created from scratch. It is also relevant to mention that all the metrics were created with a historical baseline in mind, meaning that every KPI should allow a comparison between the day of analysis and previous time periods in order to take conclusions regarding the progression of the monitored partners. This requirement proved to be harder to satisfy, however, since there was no data stored in the organization related to the progression of an item through its workflow, limiting the scope for historical comparisons.

Having developed the needed metrics to evaluate product catalogue management monitoring performance, it was then needed to understand the frequency with which they will be monitored and reported. One of the short term deliverables for the project team was to develop a weekly routine with the goal of delivering a report detailing the stages, leakages and sub-leakages for each product to each one of the monitored partners - this routine is better detailed in Subsection 4.2.5. With this said, the team concluded that, when appropriate, the developed metrics should also be included in this weekly report.

Some metrics however, are only useful to analyze and report at the end of a season, that is, with a 6 months frequency. The "*produced items that have been online*" metric is an example of this kind of metrics. It only makes sense to compare this KPI at the end of a season, since there could be situations where a produced item is offline for the whole duration of a season and only goes online by the end of it.

Online Items Related Metrics

It is critical for both monitored partner and Farfetch to know how many different items are available for sale in the company's marketplace. The first group of created metrics focuses on providing an overview of the evolution of the number of online items with the ability of filtering by season

original - the season during which the product is created, as explained in Subsection 3.1. Table 4.2 provides additional details on the created metrics.

Table 4.2: Online Items Related Metrics

Online Items				
Number	**Name**	**Formula**	**Description**	**Frequency**
1.1	**Number of online items**	$\left(\sum_{os=1}^{OS} Online\ Items\right)_t$	Number of online items at a certain point in time, t, and its distribution accross the several original seasons, OS.	Weekly
1.2	**Online items distribution accross homologous seasons**	$\frac{\left(\sum_{MS=1}^{MS} Online\ Items\right)_t}{(\Sigma\ Online\ Items)_t}$	Relative distribution of online items at a certain point in time, t, for that year's main seasons, MS.	Weekly
1.3	**Online items variation**			
1.3.1	**Items that went offline**	$\left(\sum Items\ that\ went\ offline\right)_t$	Number of items that went offline at a certain point in time, t, when compared to the previous week.	Weekly
1.3.2	**Items that went online**	$\left(\sum Items\ that\ went\ online\right)_t$	Number of items that went online at a certain point in time, t, when compared to the previous week.	Weekly
1.3.3	**Items that went online for the first time**	$\left(\sum Items\ that\ went\ online\ for\ the\ first\ time\right)_t$	Number of items that went online for the first time at a certain point in time, t, when compared to the previous week.	Weekly
1.4	**Online items YoY % growth**	$\left(\frac{\Sigma Online\ Items_y - \Sigma Online\ Items_{y-1}}{\Sigma Online\ Items_y}\right)_t$	Percentage of online items relative to the same period, t, compared to last year's values.	Weekly
1.5	**Online Items of those that have been created**	$\left(\frac{\Sigma\ Online\ Items}{\Sigma\ Created\ Items}\right)_t$	Percentage of items that have reached the "Online" stage out of those that have been through the "Created" stage, at a certain point in time, t.	Weekly
1.6	**Cumulative online products**	$\left(\sum Items\ that\ have\ been\ online\ at\ least\ once\right)_t$	Number of different items that have been in the "Online" stage at least once, at a certain point in time, t.	Weekly

These metrics will provide insights such as what is driving the growth or decrease of online items at a certain point in time (e.g., the introduction or removal of a season), as well as understand how many products are going online and offline every day. More importantly, this metric group will allow the company to understand if a partner is proving Farfetch with an increasingly bigger product catalogue. This will ultimately be a sales driver, since it dictates the supply that Farfetch is providing its customers to buy.

Stock Availability Related Metrics

As mentioned in subsection 4.2.1.1, one of the mandatory attributes for a product to go online is stock availability. A product may have all the necessary attributes to be available for purchase, but customers will only be able to buy products with associated stock. Consequently, it is also essential to monitor how much stock a certain partner is providing the company to sell. Stock metrics will inform the company if a partner is progressively providing Farfetch with more stock. These KPIs can also be correlated to sales performance, since a product with more stock will ideally sell more times. The created stock related metrics are further scrutinized in Table 4.3.

Table 4.3: Stock Availability Related Metrics

Stock Availability				
Number	**Name**	**Formula**	**Description**	**Frequency**
2.1	**Online Stock Units**	$\left(\sum_{ms=1}^{MS}\left(\sum_{P=1}^{P} Online\ Stock\ Units_P\right)_{ms}\right)_t$	Total online stock units available for a specific partner at a certain point in time, t, and its distribution accross that year's main seasons, MS.	Weekly
2.2	**Online Stock Value**	$\left(\sum_{ms=1}^{MS}\left(\sum_{P=1}^{P} (Online\ Stock\ Units * Unit\ Price)_P\right)_{ms}\right)_t$	Total online stock value available for a specific partner at a certain point in time, t, and its distribution accross that year's main seasons, MS.	Weekly
2.3	**Avg. Stock Depth**	$\left(\frac{Online\ Stock\ Units}{Online\ Items}\right)_t$	Average item stock depth at a certain point in time, t, that a partner provides Farfetch to sell.	Weekly
2.4	**Avg. Online Item Value**	$\left(\frac{Online\ Stock\ Value}{Online\ Items}\right)_t$	Average item value at a certain point in time, t, that a partner provides Farfetch to sell.	Weekly
2.5	**Items w/ stock which are online**	$\left(\frac{Online\ Items}{Items\ with\ Stock}\right)_t$	Percentage of items that the partner has given Farfetch to sell at the marketplace which are not online, a certain point in time, t.	6 months
2.6	**Items that already had stock but never went online**	$\left(\frac{\sum Items\ that\ already\ had\ stock\ but\ have\ never\ been\ online}{\sum Items\ that\ have\ never\ been\ online}\right)_t$	Percentage of items that, at a certain point in time, t, already had stock, but have never been in the "Online" stage.	6 months
2.7	**Items w/o stock that already had stockl**	$\left(\frac{\sum Items\ that\ have\ no\ stock\ but\ already\ had\ stock}{\sum Items\ that\ have\ no\ stock}\right)_t$	Percentage of items that, at a certain point in time, t, have no stock, but already had stock.	6 months

A product with stock may not be online by a multitude of reasons, one of them being partner choice. Metrics such as the percentage of items with stock which are offline will allow the company to understand how many more different items could be available for sale.

Time to Go Online Related Metrics

Hunter et al. (2018) argue that time to market is one of the core priorities for top performers in the fashion industry. This means that closely monitoring the time that a product takes to reach Farfetch's marketplace is a critical metric to closely monitor.

Figure 3.1 describes the route a product makes until it reaches the online status. Despite being useful to monitor the time a product takes from being created until it is available for purchase, Farfetch can only take corrective measures based on insights derived from the analysis of these metrics from the start of the production process onwards, since the item creation and slot allocation processes are entirely managed by the partner. That being said, two different times are going to be monitored - time from the "Item Created" stage to the "Online" stage, and time from the "Ready for Production" stage to the "Online" stage. These metrics are further detailed in table 4.4.

Table 4.4: Time to Go Online Related Metrics

Time to Go Online					
Number	**Name**	**Formula**	**Description**	**Frequency**	**SLA**
3.1	**Average Time from created to online**	$\bar{x}(\text{First Online Date} - Created\ Day)_{SOD}$	Average time that an item that went online at a certain date, SOD, took go from the "Created" stage to the "Online" stage.	Weekly	-
3.2	**Average Time from ready to production to online**	$\bar{x}(\text{First Online Date} - Ready\ for\ Production\ Day)_{SOD}$	Average time that an item that went online at a certain date, SOD, took to go from the "Ready for production" stage to the "Online" stage.	Weekly	6 Days

It is worth of note that, in collaboration with the project stakeholders, an interval of 6 days was set as the SLA for the time between the "Ready for Production" stage and the "Online" stage.

Produced Items Related Metrics

Despite not being a product manufacturer itself, Farfetch still incurs in costs mainly during the production process. Therefore, it is important to analyze how many of the items that have been through the "Produced" stage have actually reached the "Online" stage. Those that have not are: a) products that have been recently produced and will soon be online; b) products that have been produced but will never be available for purchase. Needless to say that Farfetch's main goal is to minimize, or if possible, eliminate the possibility of wasting resources in products that will never be sold.

The main difficulty with monitoring this metric is the lack of historical data. The creation of the several stages a product flows through is one of the results of this project, as described in subsection 4.2.1.1. This means that historical data will only be available since the moment these stages were conceived, which makes the comparison with previous time periods impossible. This metric is further detailed in table 4.5.

Table 4.5: Produced Items Related Metrics

Produced Items				
Number	**Name**	**Formula**	**Description**	**Frequency**
4.1	**Produced items which have been online**	$\left(\frac{\sum Items\ that\ have\ been\ online\ at\ least\ once}{\sum Produced\ Items}\right)_t$	Percentage of items, at a certain point in time, t, that have been produced and reached the "Online" stage.	6 months

Items that Sold Related Metrics

Despite the strong focus on the "Road to Online" that this project has, it is also important to have a grasp on what happens to a product after being available for purchase and provide the company and partners with sales-related metrics. The two created metrics for this case are detailed below in Table 4.6.

Table 4.6: Items that Sold Related Metrics

Items that Sold				
Number	**Name**	**Formula**	**Description**	**Frequency**
5.1	**Items online at least once that sold at least once**	$\left(\frac{\sum Items\ that\ have\ sold\ at\ least\ once}{\sum Items\ that\ have\ been\ online\ at\ least\ once}\right)_t$	Percentage of items, at a certain point in time, t, that have been online at least once and sold at least once.	6 months
5.2	**Items that sold within one month**	$\left(\frac{\sum Items\ that\ sold\ within\ 1\ month\ of\ going\ online}{\sum Items\ that\ have\ been\ online\ at\ least\ once}\right)_t$	Percentage of items, at a certain point in time, t, that sold within one month of going online.	monthly

4.2.5 Routines

The catalogue specialists will perform two weekly routines. The first and most important one is the elaboration of a weekly report on the product catalogue status for each brand. This report will include insights such as the evolution of the number of online products for the partner in scope, as well as an analysis on how the products are distributed across the different stages, leakages and sub-leakages. The report will also include powerful insights on catalogue performance by using some of the advanced metrics developed in subsection 4.2.4. One of the main goals of this report

is to provide both monitored partner and Farfetch with the needed actions to move a product from the offline to the online status.

The developed processes, tools and metrics need to be continuously improved. This will only be possible with feedback from the process owners. To do so, the catalogue specialists will have a second routine solely focused on providing the project team with insights they gather from the continuous utilization of the created tools. This routine is critical for the success of the project, since it will shape the processes and tools into solid and future-proof deliverables.

4.3 Mid-term monitoring vision

As previously mentioned, the work performed in section 4.2 revolved around the analysis of a closed group of B0 partners, motivated by the urgency of delivering the adequate tools for catalogue management monitoring, and by the data magnitude it would be needed to analyze to provide the same monitoring tools for all of Farfetch's partners.

With this said, the biggest pain point with the catalogue monitoring tools described in subsection 4.2.3 is their inability to scale up the number of monitored partners due to the fact that the used software platform - *Tableau* - is not able to process such great amounts of data. Additionally, even if the integration of every Farfetch partner in the developed tool would be possible, catalogue specialists would still be needed to elaborate the partner's weekly report. This process is not scalable at all, since hundreds of people would need to be hired to monitor all of the company's partners.

The achieved developments, however, such as the product availability process, the product workflow and associated stages, leakages and sub-leakages, as well as the key metrics to track, were done so with the baseline of being applicable to any Farfetch partner - boutique or monobrand. This means that the developed concepts and processes supported by the correct product catalogue management tool will allow the company to closely manage and monitor all its catalogue in a much more efficient way.

Furthermore, several pain points related to the product workflow were identified throughout this project. For example, partners with manual product creation find the process very lengthy and overcomplicated. Additionally, as mentioned in Subsection 3.2.3, production via FTP is a very manual and inefficient process, supported by shared folders and spreadsheets scattered across the organization. To tackle the identified pain points, as well as the lack of scalability regarding workforce and data management, a Product Information Management (PIM) tool will be needed.

As mentioned in Section 2.5, a PIM tool can either be developed in-house, or be contracted from a specified vendor. It is important to say that even before the start of this project, a PIM tool was already being developed by Farfetch's engineering team. The knowledge gathered and the developed concepts and processes during the project allowed the project team to identify and convey to the development team additional features that the information management tool should have. This development process, however, is incredibly complex and lengthy, because it involves a complete restructure of the core of the company - the way it manages product data. Based

on communications with the engineering team, it would take two to three years to have a fully in-house developed and integrated tool that meets the company's requirements.

The company needs to make the decision of either waiting for the information management tool to be developed, or, in the mean time, buy a PIM tool from a specified vendor. With this in mind, and in collaboration with Farfetch's engineering team, the project team initiated a deep research work that would, in a first stage, gather all the requirements that Farfetch's PIM tool needs to have, and in a second stage, analyze several PIM vendors in order to determine if the features they offer match with the company's requirements.

In summary, the end goal of this stage of the project is, if the company decides to contract the services of a PIM vendor, it can make such a decision with the information gathered by the project team.

4.3.1 Company Requirements

As mentioned above, the first stage of this section consists in gathering all the requirements that Farfetch's PIM tool needs to have. All these requirements will, ideally, be implemented in the information management tool currently being developed by the company's engineering team. The requirements were gathered based on the knowledge acquired throughout this project and on the already gathered information by the engineering team.

It is also important to differentiate between company and partner requirements. As mentioned above, partners have complained about the complexity of the product creation process. This means that the PIM tool should feature, for example, partner-side access in order to facilitate the manual item creation in Farfetch's database.

From the catalogue management monitoring perspective, the identified requirements are detailed in Table 4.7. The identified requirements for the entire organization are better scrutinized in Appendix B, in tables B.1, B.2 and B.3.

4.3.2 PIM Tool Vendors Evaluation

Having gathered all the requirements the company needs in a product information management tool, it is now needed to select a group of PIM tool vendors and evaluate them depending on their competencies to satisfy.

The contenders were selected based on reports by Forrester Research and G2 Crowd - specialist firms that provide advice on technology purchasing options. Both these firms have specific classification methodologies based on several factors, such as market presence, customer satisfaction and strategy strength, that culminate in the evaluated technologies' relative ranking. Figure 4.3 comprises both Forrester Research and G2 Crowd Product Information Management tool rankings.

Gartner's Magic Quadrant report for Master Data Management (MDM) was also used as a support for this procurement phase, however, MDM is more of an all-encompassing information tool, that not only supports product information management, but also other relevant data for the

Table 4.7: PIM tool requirements from the catalogue management monitoring perspective.

Requirements	
1 - Data Maintenance	
1.1	Tool's search/load/update performance is not affected by information's volume
1.2	Product data can be uploaded in bulk for creation or update of a subset of attributes
1.3	Ability to classify/filter/report an item by stage, leakage or sub-leakage
2 - Product Workflow	
2.1	Workflow ensures that data that is needed at a particular phase of the product life cycle becomes mandatory at that stage, but not before
2.2	Tool allows partner to validate uploaded prices - one-by-one and in bulk
2.3	Workflow can generate notifications and timeout alerts
2.4	It is easy to monitor and report on which items are/were at what stage in the workflow and how long each item spent at each stage of the workflow
2.5	Tool allows users to mark sections of product data as complete vs work-in-progress
3 - Reporting	
3.1	Support filtering and reporting products that have all the mandatory information or products missing details or failing validation checks
3.2	Build reports/graphs (time series) using a mix of metrics of the product workflow
3.3	Build graphs (column bar) with % of total catalogue which went from stage A to stage B
3.4	Build graphs that compare performance in a certain metric (e.g., products online / products produced)
3.5	Build ranking graphs for top / bottom partners for a certain metric (e.g., products online / products produced)
3.6	Create dashboards with the previously created graphs and have the ability to apply localized filters (e.g., filter by season)
4 - Data Quality	
4.1	Support duplicate checking, and can proactively warn users when such records are added

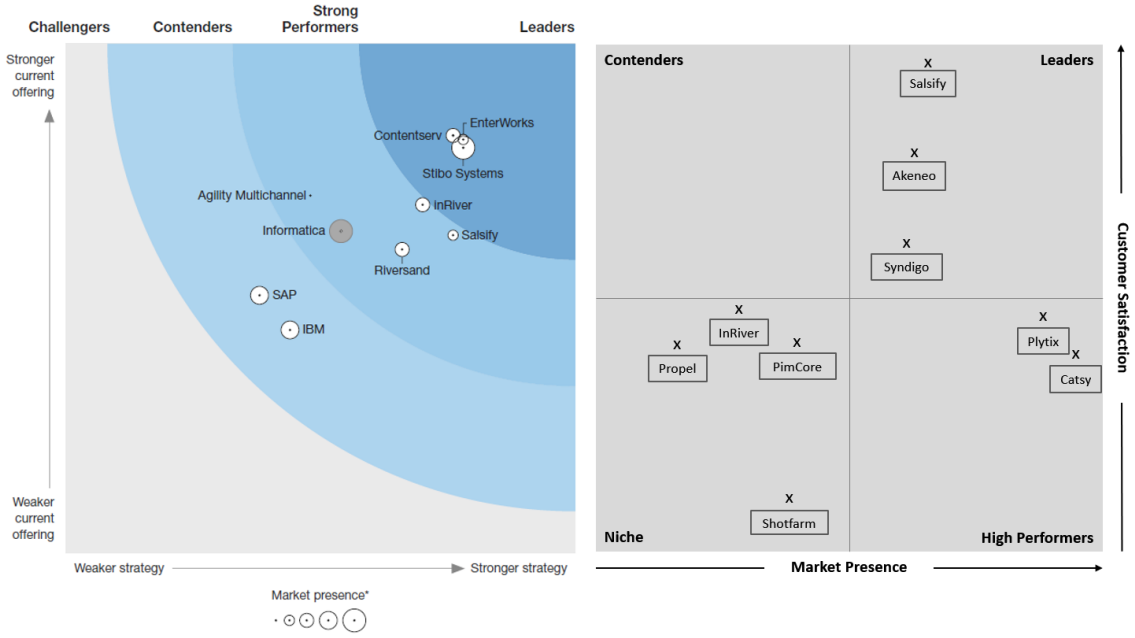

Figure 4.3: PIM Tool Vendors' ranking according to Forrester Research (left) and G2 Crowd (right). Source: Eppinger et al. (2018), G2 Crowd (2019).

business, such as personnel information. With this said, an entire MDM solution will not be relevant for this project, since the company already possesses of similar tools. Figure 4.4 depicts Gartner's ranking on several MDM tools available in the market.

Based on the information provided by Forrester Research and G2 Crowd classification reports,

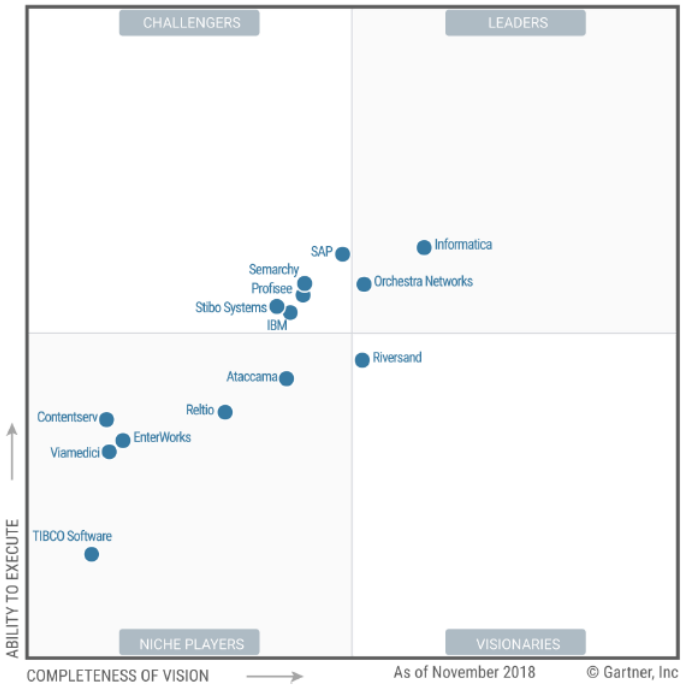

Figure 4.4: MDM Tool Vendors' ranking according to Gartner. Source: Radcliffe and O'Kane (2012).

as well as some information gathered from Radcliffe and O'Kane (2012), three possible contenders for providing Farfetch with the needed information management tool were selected for further investigation.

This investigation process was based on online research by gathering information from customer reviews, success cases from the vendors' clients and information obtained from their own websites. The most efficient way of evaluating if the contenders corresponded to the company's requirements, however, was through demonstration sessions provided by the vendors themselves.

Chapter 5

Results

The results of this project will be divided into the two workstreams presented in Chapter 4 - End-to-End Monitoring and Mid-Term monitoring vision. Firstly, the developed tools for catalogue management monitoring will be presented, as well as an example of data-driven insights that can be drawn from the analysis of the developed advanced metrics for a selected partner. Afterwards, the results of the PIM vendors investigation will be presented. Finally, the impact that this project had in the organization so far, as well as the identified opportunities, will be translated by using the created metrics as a support.

5.1 End-to-End Monitoring

This project produced several deliverables regarding end-to-end catalogue management monitoring. Starting from the product workflow stages, leakages and sub-leakages, followed by the product availability process model, the implemented routines and the developed metrics for advanced monitoring.

To aggregate all this progress, two dashboards were developed using *Tableau*: one with a snapshot status of each item for a selected partner, and another with a historical evolution on the developed advanced metrics. These tools are currently being used to create partner-specific weekly reports that are shared with the partners themselves. It is also important to mention that the developed dashboards satisfy the previously identified user requirements. The tools are accessible via web-browser, allow for quick information extracts to Excel worksheets and the displayed information is refreshed on a daily basis.

5.1.1 Product Availability

As mentioned in Subsection 4.2.3, the modelled Product Availability process needed to be performed much more quickly and efficiently. This is the product availability dashboard's main goal. Figures 5.1, 5.2 and 5.3 are the result of an information extract - one of the identified tool's requirements in Subsection 4.2.3 - performed on the developed dashboard that has been filtered to the specific partner in analysis - Partner A.

It is also relevant to mention that this tool provides the users with a snapshot status of a partner's products, meaning that there is no historical evolution (i.e., currently it is not possible to know for how long a product remained in one stage), due to data storage and performance limitations.

Figure 5.1 represents the relative distribution of a partner's products across the several stages and leakages. The goal of this graphic is to provide the partner with a high-level view on its products' road to online.

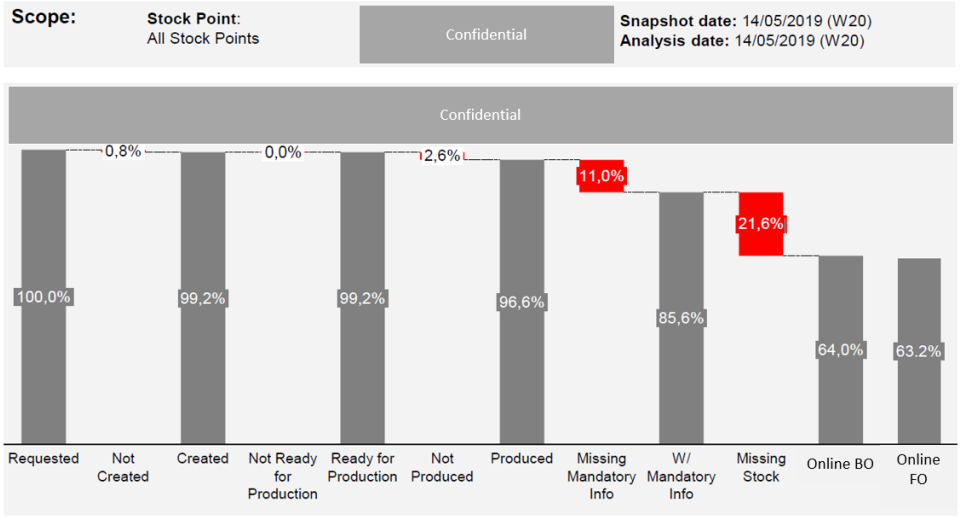

Figure 5.1: Partner A products' distribution across the several stages and leakages for week 20 of 2019.

Figure 5.2 represents how the developed tool provides both partner and Farfetch with a more detailed analysis on the absolute distribution of products across the several leakages and sub-leakages. This will, in turn, possibilitate either the partner or the company, depending on the type of sub-leakage, to take a specific action and force the product to move along the road to online.

Figure 5.3 displays how, for partner A, at the time of analysis, the products are distributed across the several types of actions associated with the different sub-leakages. It is important to notice that products in the "No Action" zone are products that are still within the SLAs and require no special attention.

Finally, in order to allow for maximum actionability, it is also important to mention that the developed tool has the ability to deep dive even further into the several sub-leakages up to an item level of aggregation, which allows catalogue specialists to highlight specific items, if necessary. The layout of the developed dashboard is better represented in Appendix C.

5.1.2 Performance Evaluation

The developed metrics mentioned in Subsection 4.2.4 were aggregated in a *Tableau* dashboard in order to meet the previously identified user requirements. This subsection will serve as an example for the type of insights that the dashboard's users can draw from analyzing the historical evolution

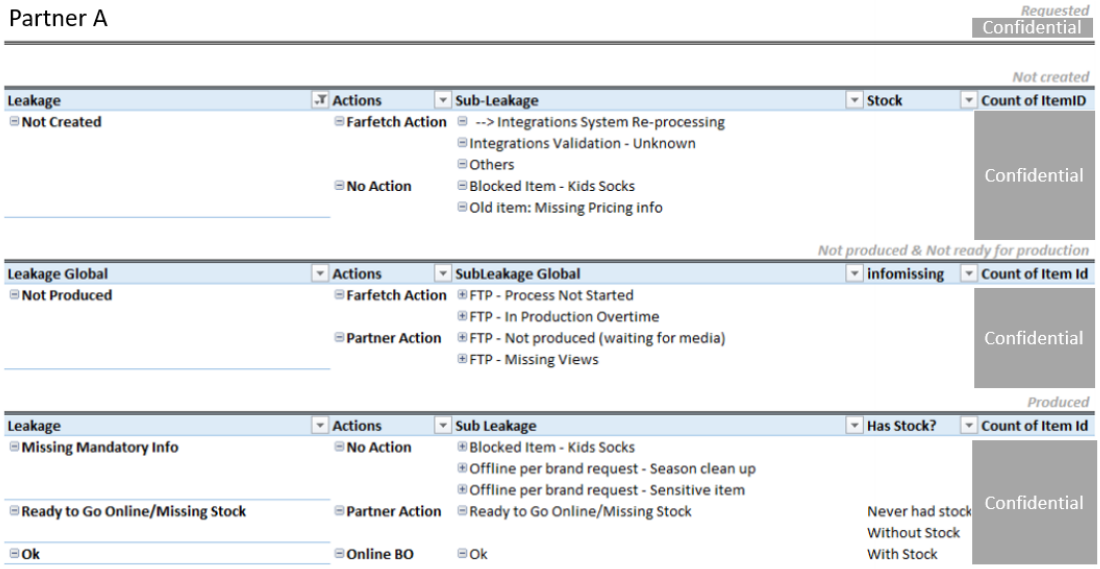

Figure 5.2: Partner A products' distribution across the different sub-leakages.

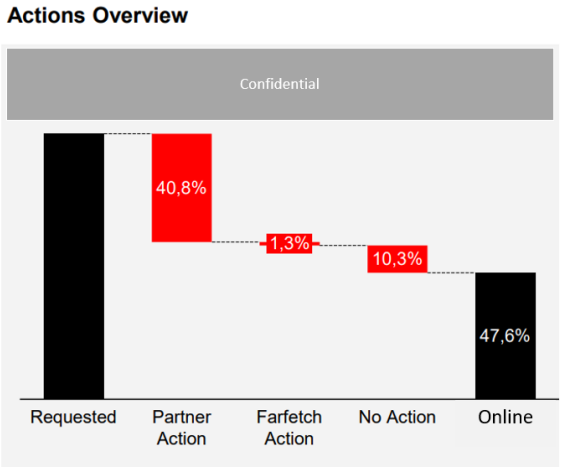

Figure 5.3: Partner A products' distribution across the different types of action (percentage).

of those metrics. Appendix D contains a more detailed view on the layout of this dashboard, with a display of all the developed metrics.

Online Items

Figure 5.4 represents metric 1.2 - online items distribution across homologous seasons. From the analysis of the displayed results we can conclude that for the same week in different years, the ramp up of the AW season has slowed down, and the amount of carryovers represents almost half of the partner's product catalogue. This means that the partner has been providing Farfetch with less new products, when compared to previous years.

Furthermore, 2% of the partner's catalogue in 2019 is composed of seasons marked as "Others". These include seasons older than the main seasons for that year - AW19 and SS19. This

means that 2% of Partner A's online products are seen by the customer as items from older seasons, an undesirable situation that has been reported and corrected.

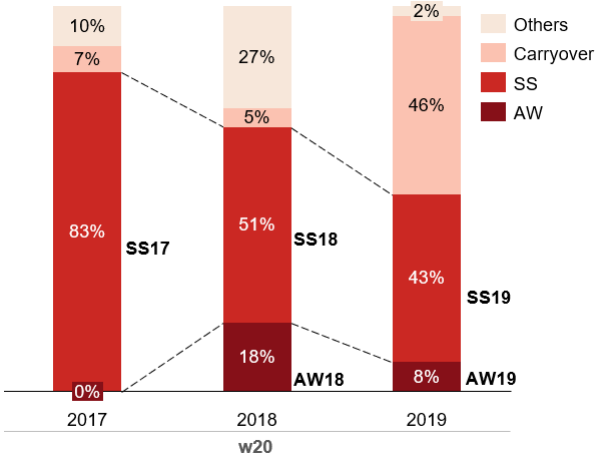

Figure 5.4: Partner A products' distribution across the different types of seasons for week 20 of 2017, 2018 and 2019 (percentage).

When compared to 2018's percentage of "Others", partner A's catalogue seems much healthier, with a 25% reduction in 2019. This was due to a new implemented process to carry the items to a new season.

Stock Availability

Regarding stock availability related metrics, Figure 5.5 represents the evolution of products that partner A has provided Farfetch with stock to sell and are online. As explained in Subsection 4.2.1.1, an item may have available stock to be sold, but still need other additional information to be able to go online in Farfetch's marketplace. This means that the items considered in Figure 5.5 may need additional attributes other than stock to flow to the "Online FO" stage.

Despite this, the company needs visibility on the items that it has available stock to sell, but are still missing mandatory information, which is what the graphic represented in Figure 5.5 provides.

Time to Go Online

Figure 5.6 represents the time it took for the produced products of Partner A to reach the "Online FO" stage since their arrival at the production center, for weeks 17, 18, 19 and 20 of 2019.

From analyzing Figure 5.6 we can conclude that the established SLA of six days may be too ambitious. Despite the increase in the number of products that take less than 15 days to reach the "Online FO" stage, still more than 60% is taking more than 15 days.

The long times for an item to go online may be explained by situations where the partner holds the stock after a product is produced and sharing it with the company only whenever it deems justifiable. As mentioned in Subsection 4.2.4, time to market in the fashion business is one of top players' core priorities. To reduce these times, an incentive system that rewards partners for uploading their stock as soon as possible should be put into place.

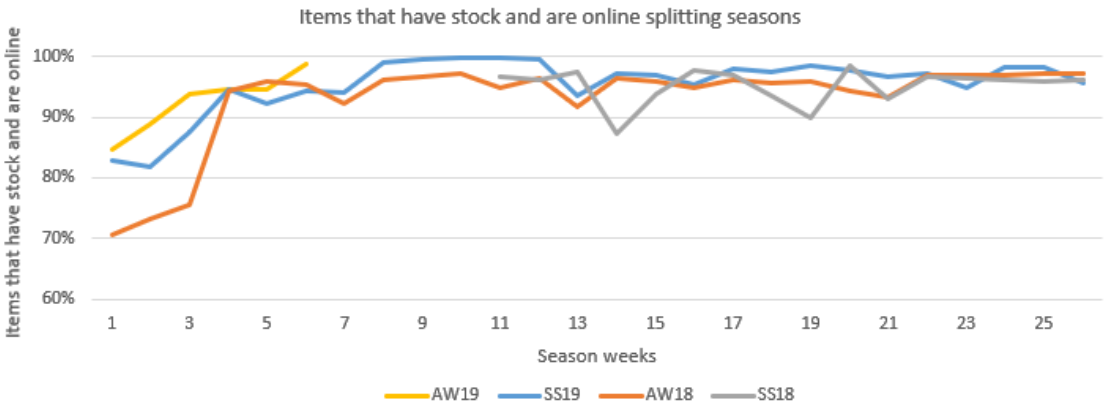

Figure 5.5: Partner A products' with stock which are online evolution throughout the different seasons (percentage).

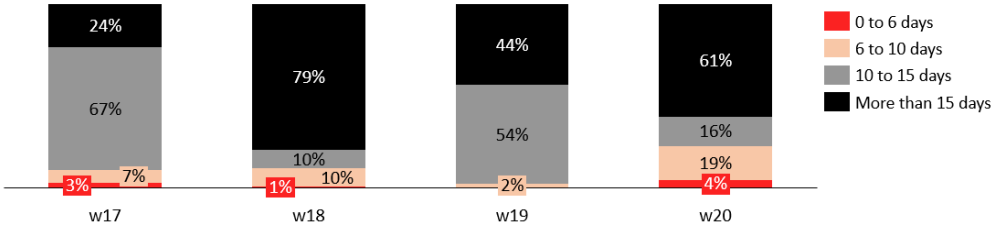

Figure 5.6: Partner A's Time to Go Online distribution across weeks 17, 18, 19 and 20 of 2019 (percentage).

Produced Items

As mentioned in Subsection 4.2.4, it is important to provide the company with visibility on the percentage of items that are being produced but are not going online. Figure 5.7 shows that around 90% of partner A's SS19 products which have been produced have also been online by the end of the season (week 17), meaning that the company has a maximum of 10% improvement opportunity, for partner A only.

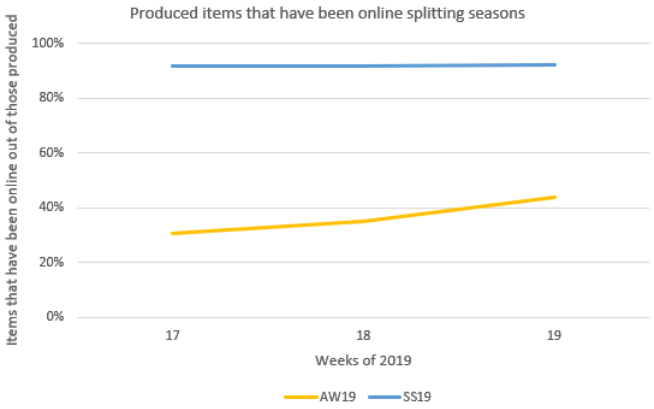

Figure 5.7: Partner A produced products that have been online at least once (percentage).

It is also important to mention that, as explained in Subsection 4.2.4, there is only historical data available for any kind of metric that aims to monitor each of the product workflow's stages since the moment they were conceived, which is why Figure 5.7 starts in week 17 of 2019.

This metric should be analyzed and reported at the end of a season, since throughout a season, a product can go online for the first time at any given time, however, after the end of the season this is much less likely to happen.

Items that Sold

It is also very important to keep track on how well a partner's items are selling, season after season. Figure 5.8 represents the percentage of partner A's items that sold at least once for several seasons.

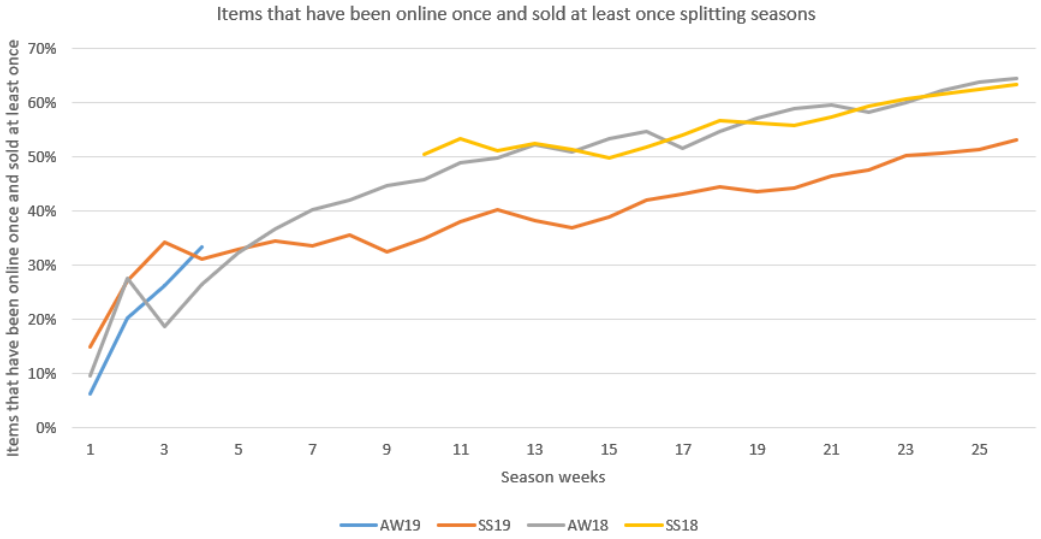

Figure 5.8: Partner A's items that have been online once and sold at least once (percentage).

As with every metric based on items that have been online at least once, this metric should also be monitored at the end of each season, however it is also helpful to monitor how it develops throughout the season.

From Figure 5.8 one can conclude that the percentage of items that sold at least once has decreased for the SS season from 2018 to 2019, however, AW19 is showing signs of improvement in comparison to AW18.

5.2 Mid-term Monitoring Vision

As mentioned in Subsection 4.3.2, the PIM vendors were analyzed based on online research, previous customers case studies and demonstrations provided by the vendors themselves. Figure 5.9 shows the vendors' final evaluation.

By analyzing Figure 5.9 and the more in-depth analysis provided in Appendix E, the team recommends Vendor 2, in case the company decides to buy a PIM solution while the in-house tool is built.

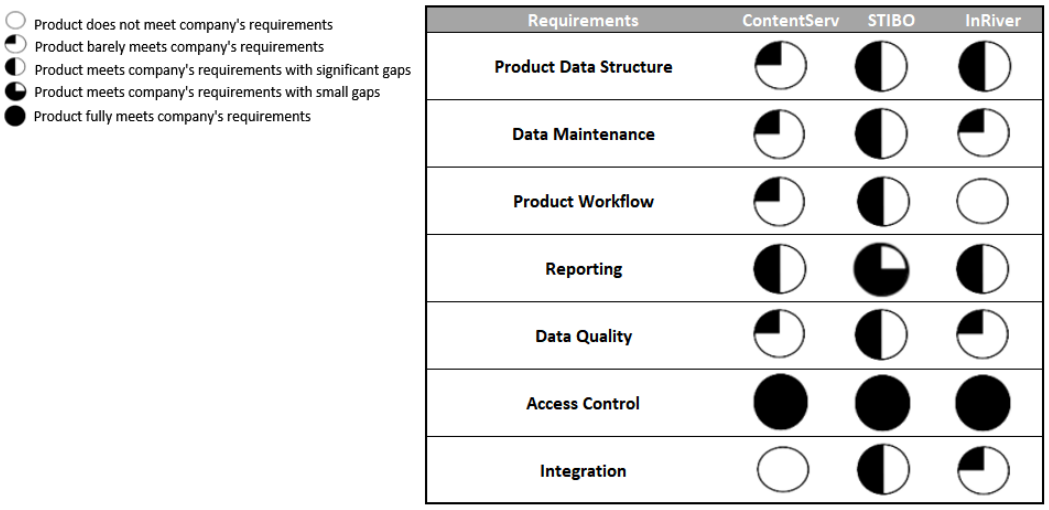

Figure 5.9: PIM Vendors final evaluation results.

It is also important to mention that it would be impossible to find a vendor that would fully meet the company's requirements out of the box. Only the in-house built tool will be able to support this.

5.3 Project Impact and Potential Opportunities

Finally, it is important to clarify the impact that the product catalogue management monitoring project has had so far in the organization as well as the potential opportunities for future improvement that were identified.

The developed concepts, processes and tools have allowed both Farfetch and partner to effectively know which attributes a product was missing to go online, which in turn has helped the company push these items online in a much more efficient and faster manner, effectively cutting in half the monitored partners' average time to go online, from 39 to 19 days, for seasons SS18 and SS19, respectively - Figure 5.10.

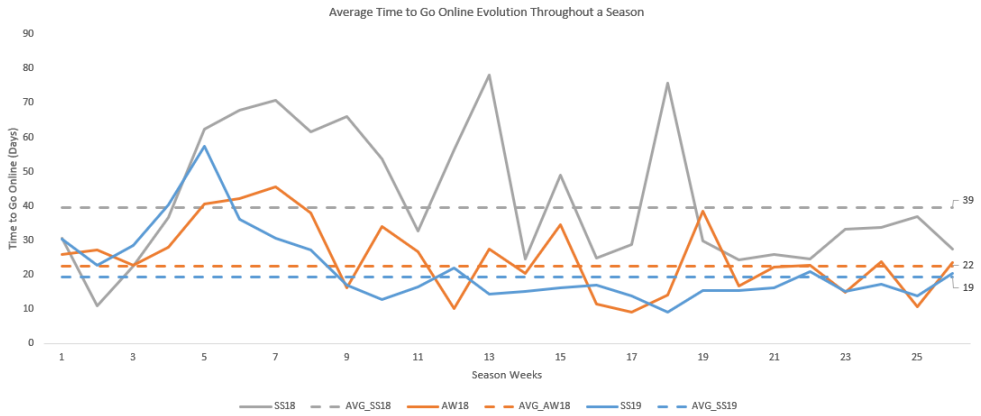

Figure 5.10: Time to go online evolution throughout seasons SS18, AW18 and SS19 for the monitored partners.

Additionally, the provided visibility on the attributes that a product is missing to be sold at the marketplace, has translated itself on a 6% reduction of the overall percentage of products that have had stock at least once but have never been online from season SS18 to SS19 - Figure 5.11.

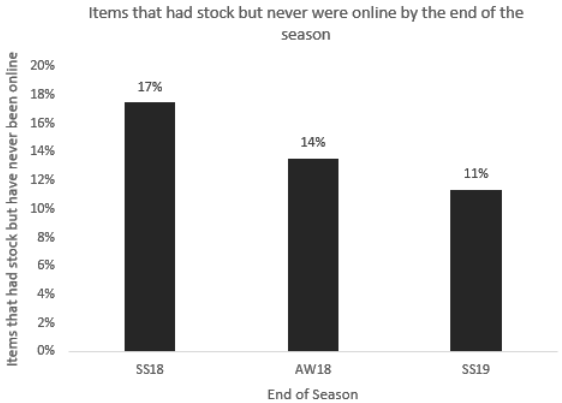

Figure 5.11: Products that had stock at least once but have never been online at the end of a season for all the monitored partners (percentage).

The newly created catalogue health monitoring metrics have also allowed the team to identify several opportunities for improvement Around 10% of the produced products for the monitored partners never made it to the "Online" stage by the end of the SS19 season (highlighted week 21) - Figure 5.12 -, which means the company wasted resources that will never provide any revenue.

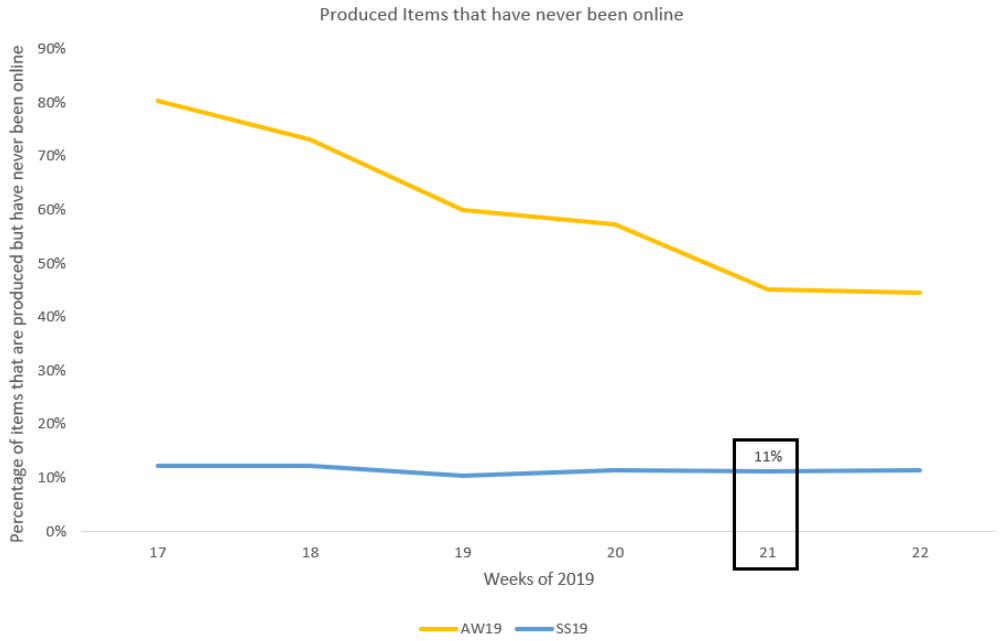

Figure 5.12: Items that have been produced but never went online for the monitored B0 partners (percentage).

Furthermore, a second area of opportunity was identified when the team concluded that an average of 15% of a partner's catalogue had stock but was not online at any point in time - Figure 5.13, which means that a 15% direct sales opportunity was being lost.

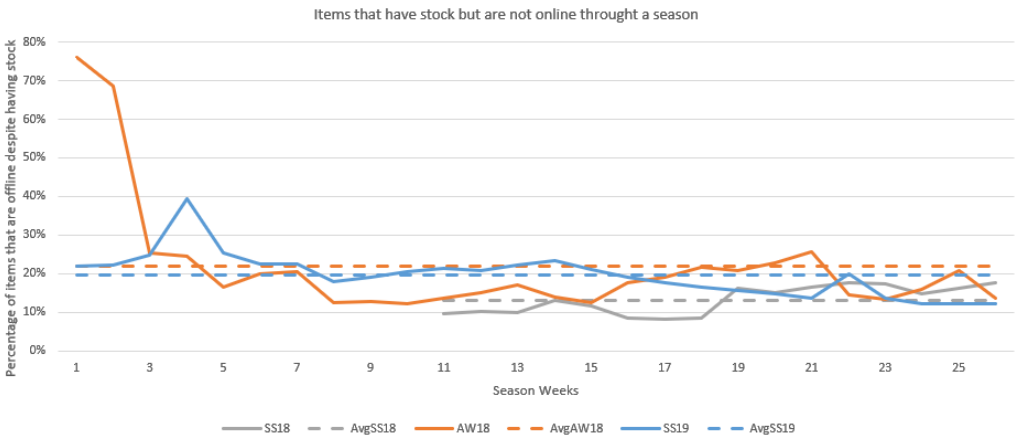

Figure 5.13: Items that are offline despite having stock throughout a season for the monitored B0 partners (percentage).

It is also critical to enhance that the identified opportunities are only achievable through operational excellence, and appropriate catalogue management tools, processes and people that extend to the entire organization and not only a select group of partners.

Finally, as time goes by and product catalogue management monitoring is expanded to additional partners, it is expected that the percentage of items that are produced but have never been online, reaches numbers close to 100% and the time to go online should decrease even further, if possible, to under the established 6 days SLA.

Chapter 6

Discussion

6.1 Project Conclusion

Product catalogue management is at the foundation of a successful online marketplace. From the sellers' perspective, understanding why their items are not selling is essential so that an action may be taken to allow those products to go online. From a customer's perspective, it is of the utmost importance to have a consistent experience while navigating through the items' information pages. These two types of perspective are only achievable with a trustful and accurate product catalogue management.

Before the start of this project, Farfetch did not understand how crucial it was to have visibility on its online and offline catalogue. As a result of that, there were no implemented concepts, processes and tools to provide that so needed visibility to both partner and enterprise. By assembling a team comprised of people from several areas of expertise in the company, this problem could be effectively tackled. The project team delivered the basis of what constitutes a well managed product catalogue management monitoring, by creating the product workflow and providing tools that allow the organization to closely monitor it.

The developed monitoring tools and weekly routines allowed the company to perform a more proactive catalogue monitoring and report back to the partners in order to inform them about which actions needed to be taken for their products to reach the "Online" stage. This resulted in improved relationships with key partners - a great achievement, when compared to the *status quo* before the start of the project, where some of them threatened to leave the company. Additionally, the developed monitoring tools allowed the team to identify a substantial decrease in the average time a product takes to reach the "Online" stage as well as decrease in the percentage of items that never reach it.

Furthermore, it is important to mention that the developed tools do not support the monitoring of every partner Farfetch works with. For this to be possible, a dedicated product information management tool needs to be developed, either in-house, or acquired to a third party vendor. Despite this, the investigation work performed by the project team has also supported the organization considering the buy or make option regarding a tool of this sort.

In summary, we can safely say that the proposed project goals at the start of this dissertation were successfully accomplished. Farfetch is now empowered to conduct a more proactive, effective and efficient product catalogue management monitoring, thanks to the developed concepts, processes and tools that the team dedicated to this project has delivered.

6.2 Future Projects

One of the main drawbacks of this project was the lack of historical data. Since the product workflow was one of the deliverables of the project, there was no data related to how long a product would spend in a certain stage. Despite this, the team has created a database that stores the workflow's stage that each product is at, on a daily basis. This means that in the future, when an appropriate historical baseline is stored in this database, the product workflow can be further optimized. Having historical data will allow the company to understand which stage is the bottleneck of the process, optimize it and further reduce the time to market.

Additionally, the developed product workflow does not differentiate between the first time a product goes online and other instances. For example, if a product has reached the "Online" stage, but its stock and descriptions are removed, the item flows to the "Created and Ready for Production" stage, which reduces the visibility on which products are on the road to the "Online" stage for the first time, and the ones that have already been online but are currently not. One of the workarounds for this situation is to create two different product workflows - one of them solely dedicated to the items going online for the first time.

Finally, one of the most important next steps regarding product catalogue management monitoring is to expand the analysis to the entirety of Farfetch's partners, for which the company will need new tools and a revamped data structure. These requirements can be satisfied with an adequate product information management tool.

Bibliography

Achille, A., Remy, N., and Machessou, S. (2018). Apparel, Fashion & Luxury Group - The age of digital Darwinism. Technical report, McKinsey & Company.

Adams, A. (2013). Luxury Consumers Value Products, Not Buying Experiences. https://www.forbes.com/sites/arieladams/2013/05/23/luxury-consumers-value-products-not-buying-experiences/{#}6a26e4b44fe5. (Accessed on 2019-05-13).

Amed, I., Balchandani, A., Beltrami, M., Berg, A., Hedrich, S., and Rölkens, F. (2018). The state of fashion 2019. Technical report, McKinsey & Company.

Archambault, S. G., Helouvry, J., Strohl, B., and Williams, G. (2015). Data visualization as a communication tool. *Library Hi Tech News*, 32(2):1–9.

Arienti, P. (2019). Global Powers of Luxury Goods 2019. Technical report, Deloitte.

Bhatti, N., Bouch, A., and Kuchinsky, A. (2000). Integrating user-perceived quality into Web server design. *Computer Networks*, 33(1-6):1–16.

Braganza, A. and Lambert, R. (2000). Strategic integration: Developing a process–governance framework. *Knowledge and Process Management*, 7(3):177–186.

Dauriz, L., Remy, N., and Sandri, N. (2014). Luxury shopping in the digital age. Technical report, McKinsey & Company.

del Rıo-Ortega, A., Resinas, M., and Ruiz-Cortés, A. (2009). Towards modelling and tracing key performance indicators in business processes. *II Taller sobre Procesos de Negocio e Ingenierıa de Servicios, PNIS*.

Dell, R. K. (2005). Breaking organizational silos: Removing barriers to exceptional performance. *Journal / American Water Works Association*, 97(6):34–37.

DeLone, W. H. and McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, 19(4):9–30.

Doran, G. T. (1981). There'sa SMART way to write management's goals and objectives. *Management review*, 70(11):35–36.

Eckerson, W. and Hammond, M. (2011). Visual reporting and analysis. *TDWI Best Practices Report. TDWI, Chatsworth*.

Economy, P. (2015). Forget SMART Goals – Try CLEAR Goals Instead | Inc.com. https://www.inc.com/peter-economy/

forget-smart-goals-try-clear-goals-instead.html. (Accessed on 2019-05-13).

Eisenberg, B. (2013). What People Do on Your Site and Why by @TheGrok. https://www.bryaneisenberg.com/what-people-do-on-your-site-and-why/. (Accessed on 2019-05-11).

Enoma, A. and Allen, S. (2007). Developing key performance indicators for airport safety and security. *Facilities*, 25(7/8):296–315.

Eppinger, B., Swedlow, F., Sjoblom, S., and Lynch, D. (2018). The Forrester Wave ™ : Product Information Management Solutions , Q2 2018. Technical report, Forrester Research.

Fry, B. (2008). *Visualizing data: Exploring and explaining data with the processing environment*. O'Reilly Media, Inc.

G2 Crowd (2019). Best Product Information Management (PIM) Software in 2019. Technical report, G2 Crowd.

Grover, V., Jeong, S. R., Kettinger, W. J., and Teng, J. T. C. (1995). The Implementation of Business Process Reengineering. *Journal of Management Information Systems*, 12(1):109–144.

Hand, D. J. (2013). Data MiningBased in part on the article "Data mining" by David Hand, which appeared in the Encyclopedia of Environmetrics . In *Encyclopedia of Environmetrics*. John Wiley & Sons, Ltd, Chichester, UK.

Harris, L. and Dennis, C. (2002). *Marketing the e-Business*, volume 2. Psychology Press.

Hung, R. Y.-Y. (2006). Business process management as competitive advantage: a review and empirical study. *Total Quality Management & Business Excellence*, 17(1):21–40.

Hunter, E., Marchessou, S., and Schmidt, J. (2018). The need for speed: Capturing today's fashion consumer | McKinsey. https://www.mckinsey.com/industries/retail/our-insights/the-need-for-speed-capturing-todays-fashion-consumer. (Accessed on 2019-05-25).

IBM (2016). Infographic: The Four V's of Big Data | IBM Big Data & Analytics Hub. https://www.ibmbigdatahub.com/infographic/four-vs-big-data. (Accessed on 2019-06-29).

Jensen, K. (2012). File:CRISP-DM Process Diagram.png - Wikimedia Commons. https://commons.wikimedia.org/wiki/File:CRISP-DM{_}Process{_}Diagram.png. (Accessed on 2019-05-06).

Jorij, A. (2004). *Product Information Management*, volume 94.

Kaplan, R. S. and Norton, D. P. (1992). The balanced scorecard: measures that drive performance.

Kotter, J. P. (1995). Leading change: Why transformation efforts fail. Technical report, Harvard Business Review.

Kueng, P. (2000). Process performance measurement system: A tool to support process-based organizations. *Total Quality Management*, 11(1):67–85.

Larson, K. D. (1998). The role of service level agreements in IT service delivery. *Information Management & Computer Security*, 6(3):128–132.

Maté, A., Trujillo, J., and Mylopoulos, J. (2017). Specification and derivation of key performance indicators for business analytics: A semantic approach. *Data & Knowledge Engineering*, 108:30–49.

McKinney, V., Yoon, K., and Zahedi, F. 2002). The measurement of web-customer satisfaction: An expectation and disconfirmation approach. *Information systems research*, 13(3):296–315.

Moitra, D. (2002). Managing change for software process improvement initiatives: a practical experience-based approach. *Software Process: Improvement and Practice*, 4(4):199.

Neves, J. (2018). Goldman Sachs | Talks at GS - José Neves: Moving Fashion Forward with Technology. `https://www.goldmansachs.com/insights/talks-at-gs/jose-neves.html`. (Accessed on 2019-05-19).

Okonkwo, U. (2009). Sustaining the luxury brand on the Internet. *Journal of Brand Management*, 16(5):302–310.

Okonkwo, U. (2010). *Luxury Online: Styles, Systems, Strategies*. Palgrave Macmillan.

Parish, R. J. (1997). Service level agreements as a contributor to TQM goals. *Logistics Information Management*, 10(6):284–288.

Radcliffe, J. and O'Kane, B. (2012). Magic Quadrant for Master Data Management of Customer Data Solutions. Technical Report 2012-10-18, Gartner, Inc.

Redman, T. C. (2016). Bad Data Costs the U.S. $3 Trillion Per Year. `https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year`. (Accessed on 2019-05-04).

Rozenfeld, H. and Costa, J. M. H. (2007). Proposal of the BPM method for improving NPD processes. *Product: Management & Development*, 5(1):25–32.

Turetken, O. and Demirors, O. (2008). Process modeling by process owners: A decentralized approach. *Software Process: Improvement and Practice*, 13(1):75–87.

Wirth, R. (1995). CRISP-DM : Towards a Standard Process Model for Data Mining. *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, (24959):29–39.

Zairi, M. (1997). Business process management: a boundaryless approach to modern competitiveness. *Business Process Management Journal*, 3(1):64–80.

Appendix A

Stages, Leakages and Sub-leakages

This appendix aims at providing the reader with a better description of the different stages, leakages and sub-leakages a product can flow through. Table A.1 provides a summary on the different type of stages a product flow through, while tables A.2, A.3, A.4 and A.5 better describe the several sub-leakages associated with the not created, not ready for production, not produced and missing mandatory information leakages.

Table A.1: Product workflow stages

Product Workflow Stages	
Stage	**Definition**
Requested	All the SKUs that were requested by the partner to be created.
Created	Items that were successfully created in Farfetch system.
Created & Ready for Production	**Shooting** - Items that were scanned in at the Production center. **FTP** - Items that have the minimum number of mandatory images in the FTP shared folder to start the Production Process (currently considering 2 images as the minimum).
Produced	Items that have successfully gone through the whole Production process (they have at least one photo uploaded, Description, Data Entry and Category Tree).
Produced & w/ Mandatory Information	Items that have all the availability flags activated, barcodes in at least one size and have no rules blocking all markets.
Online Back-Office	Items for which all conditions to be online (all previous conditions plus stock).
Online Front-Office	Items that are online in the front office (includes having stock).

Table A.2: Product not created leakage and respective sub-leakages

Product not Created Leakage	
Sub-Leakage	**Definition**
Missing Conversion Mapping → Missing Brand ID	Items that were not successfully created because the Partner did not send the item BrandID.
Missing Conversion Mapping → Missing Category ID	Items that were not successfully created because the Partner did not send the item category.
Missing Conversion Mapping → Missing Category Mapping	Items that were not successfully created due to Category mapping reasons.
Missing Conversion Mapping → Missing Size Mapping	Items that were not successfully created due to size mapping reasons.
Missing Mandatory Information	Items that were not successfully created because the Partner did not send at least one of the Product Creation mandatory fields.
Missing Mandatory Information → Pricing	Items that were not successfully created because the Partner did not send the item price.

Table A.3: Product not ready for production leakage and respective sub-leakages

Product not Ready for Production Leakage		
	Sub-Leakage	**Definition**
Shooting	**Not Added to Slot**	Items that are not ready for Production because they were not added to a slot.
	Added & Not Ready to Send	Items that are not ready for Production because they haven't a Ready to Send Date (RTS) defined, despite having already been added to a slot.
	Ready to Send & Not Allocated	Items that are not ready for Production because they do not have an Estimated Time Arrival (ETA) defined, despite having already been added to a slot and had a RTS.
	Slot not Arrived	Items that are not ready for Production because their slot has not arrived yet.
FTP	**Partner did not Send Media**	Items that are not ready for Production because the Partner did not send any images yet.
	Error on Media Integrations Process	Items that are not ready for Production because the Media Integrations Process is with an error, despite the partner has already sent the media.
	Media not in FTP Folder	Items that are not ready for Production because there are not any images in the FTP shared folder, despite the partner has already sent them and there is no error in Media Integrations Process.
	Missing Views in Folder	Items that are not ready for Production because there is only one image in the FTP shared Folder.
	Not Added to Slot	Items that are not ready for Production because they were not added to any slot.

Table A.4: Product not produced leakage and respective sub-leakages

Product not Produced Leakage		
	Sub-Leakage	**Definition**
Shooting	**In Process**	Items that are undergoing the Production process and less than 5 days passed since slot arrival.
	Production Overtime	Items that are undergoing the Production process, more than 5 days passed since slot arrival and are still in the Production Center facilities.
	Left Production Incomplete	Items that are undergoing the Production process, more than 5 days passed since slot arrival and already left the Production Center facilities (already scaned out).
	Item not Arrived	Items that are not produced because they did not arrive in the slot they had been added to.
	Not Produced - Other reasons	Items that are not produced because its Production process was blocked (e.g the Partner sent the wrong size for being photographed, electronical device).
FTP	**Process not Started**	Items that have the minimum number of mandatory images but the quality control process has not started yet.
	In Process	Items that are undergoing the Production process and less than 5 days passed since the start of quality control process.
	In Production Overtime	Items that are undergoing the Production process and more than 5 days passed since the start of quality control process.
	Duplicate	Items that were associated and whose "parent-item", from another partner, is undergoing the Production Process.

Table A.5: Product missing mandatory information leakage and respective sub-leakages

Missing Mandatory Information Leakage	
Sub-Leakage	**Definition**
Portal Flag	For a stockpoint, all items become disabled (e.g. the stockpoint is not live yet, the stockpoint has been asked to be put "on holidays").
Blocking Rules	For a stockpoint, items that are blocked for selling on all markets.
Available Portal Flag	Items that are with the Available Portal flag disabled (the flag is automatically activated when the item is produced but can be deactivated manually anytime).
IsAvailable Flag	Items that are with the IsAvailable flag disabled (the flag is automatically activated when the item is created or associated but can be deactivated manually anytime).
Online Flag	Items that are with the IsOnline flag disabled (the flag is automatically activated when the item is produced but can be deactivated manually anytime - e.g. when the partner requests to put an item offline).
Partner did not send Barcodes	Items that have no barcodes associated because the partner has not sent them.
Error on Integration Barcodes	Items that have no barcodes associated because of an error in the Integrations process, despite the Partner having already sent them.
Barcodes	Items that have no barcodes associated, despite the Partner having already sent them and the Integrations process has been successfully completed (e.g. someone has manually removed the barcodes).
Qualitative	Items that have neither Country of Origin field nor Material field completed.

Appendix B

Product Information Management Tool Requirements

Table B.1: PIM tool requirements from a Product Data Structure and Data Maintenance perspective.

	Requirements
	Product Data Structure
1.1	Product catalogue can be built by Farfetch PIM specialists without need for vendor consultancy or programming
1.2	Different attributes, hierarchies, business rules can be defined at product category level
1.3	Typical but more complex business rules can be defined using minimal scripting
1.4	Tool supports M:M relationships (e.g., a product can belong to multiple categories and/or multiple styles)
1.5	Tool supports global and tenant specific data attributes/hierarchies
1.6	Tool supports multiple versions of attributes
1.7	Data model (meta-data) can be easily extracted/reported to support model management and integration efforts
1.8	Tool supports searchability/findability (e.g., tagging, creating associations, defining synonyms)
1.9	Tool supports definition of blocking rules assigned to products - these rules are time-variant and can be classified and filtered by type (e.g., input by partner vs. input by Farfetch) and nature (e.g., commercial, legal, etc.)
1.10	Tool supports defining description tabs for different types of product details/attributes (e.g. product description, sizing/fit, care instructions, ingredients)
	Data Maintenance
2.1	Tool supports loading/creation of invalid records
2.2	Tool supports definition of prices across time
2.3	Tool supports different prices/discounts to be defined based on size/colour as well as tenant/boutique and geography
2.4	Search/load/update performance is not affected by information volume
2.5	New records can be added via the GUI bulk loads from excel/csv files
2.6	Product data can be uploaded in bulk for creation or update of a subset of attributes
2.7	Product data load templates can be created using the tool's UI
2.8	Prices for different geographies/time periods can be bulk uploaded
2.9	Tool provides graphical hierarchy browser
2.10	Tool provides preview environment for quality control
2.11	Ability to edit images
2.12	Tool supports maintenance of exclusive/VIP brands stock (products inherit VIP or Exclusive flag from their brands marked as exclusive/VIP)
2.13	Ability to classify/filter/report an item by stage, leakage and sub-leakage
2.14	Support for translations
2.15	Brands can control and automate when their products go online/offline (e.g., put items online and offline with one click and schedule for products to go live in specific markets)
2.16	Carryover management: ability to edit current season

Table B.2: PIM tool requirements from a Product Workflow and Reporting perspective.

	Requirements
	Product Workflow
3.1	Workflows can be created/modified by Farfetch PIM specialists without need for vendor consultancy or programming
3.2	Workflow ensures that data that is needed at a particular phase of the product life cycle becomes mandatory at that stage (but not before)
3.3	Workflow can be configured to be dynamic (e.g., routing to different assignees depending on data attributes/business rules)
3.4	Data quality/completeness issues or violation of business rules can trigger alarms for the attention of specific user roles depending on data values or issue type
3.5	External parties can take part in the workflow, create and update data subject to access control rights
3.6	Tool allows partners to validate prices uploaded - one-by-one and in bulk
3.7	Workflow can generate notifications and timeout alerts
3.8	Users can mark tasks as "done" as required
3.9	It is easy to monitor and report on which items are/were at what stage in the workflow and how long each item spent at each stage of the workflow
3.10	Users can search/filter/select and mark products as their work-in-progress to avoid two people trying to edit the same product at the same time
3.11	Allow supervisors to assign products in the workflow to data stewards
3.12	Allow users to mark sections of product data as complete vs work-in-progress (e.g. for photos, descriptions)
3.13	Tool allows reviewer role to approve product data within the workflow, or reject and send back to data steward with rejection message
3.14	Tool supports tracking of performance of the user who solves problems (e.g., how many products he brought from offline to online)
3.15	Tool supports defining targets for individuals and teams and provides dashboards to compare performance of teams that solve problems
	Reporting
4.1	Prebuilt reports/search functionality supports typical reporting needs (e.g., filters, fuzzy search, search for null values, greater/smaller filters, begins with ends with filters, wildcards)
4.2	Data can be searched, and extracted to Excel spreadsheet through the UI
4.3	Tool supports filtering and reporting products that have all the mandatory information or products missing details or failing validation checks
4.4	Build reports/graphs (time series) using a mix of metrics of the product workflow
4.5	Build graphs (column bar) with % of total catalogue which went from stage A to stage B
4.6	Build graphs that compare performance in a certain metric (e.g., products online / products produced)
4.7	Build ranking graphs for top / bottom partners for a certain metric (e.g., products online / products produced)
4.8	Access to graphs and dashboards is controlled by user/user group permissions (e.g., internal, partner) and managed by an admin through the UI
4.9	Create dashboards with the previously created graphs and have the ability to apply localized filters (e.g., filter by season)
4.10	Prices for different geographies/dates can be extracted/downloaded

Table B.3: PIM tool requirements from a Data Quality, Access Control and Integration perspective.

	Requirements
	Data Quality
5.1	Tool supports duplicate checking, and can proactively warn users when such records are added
5.2	Tool helps identify/create goldenn record, and links duplicates to the gold record
5.3	Tool supports probabilistic & fuzzy matching
5.4	Tool can display warning message whenever prices are not meeting certain conditions (e.g., price equal to zero)
5.5	Quality control (e.g., blocking items from going live if info is incorrect)
5.6	Tool supports blocking/inactivating records, with different reason flags.
5.7	Inactivated records are accessible in the UI and in reports but clearly flagged as inactive
	Access Control
6.1	Tool keeps track of audit history at field level (e.g., what changed when by whom)
6.2	Demonstrate capability to support multi-tenancy (e.g., data can be segregated for different partners)
6.3	Data access can be managed at user/user group level and security can be set at record/field level as well as product category level
	Integration
7.1	Tool generates dimension tables (star/snowflake) to support data warehousing needs
7.2	Tool supports creating full extracts of product data to database tables
7.3	Tool supports real time updates through messaging/APIs

Appendix C

Product Availability Dashboard

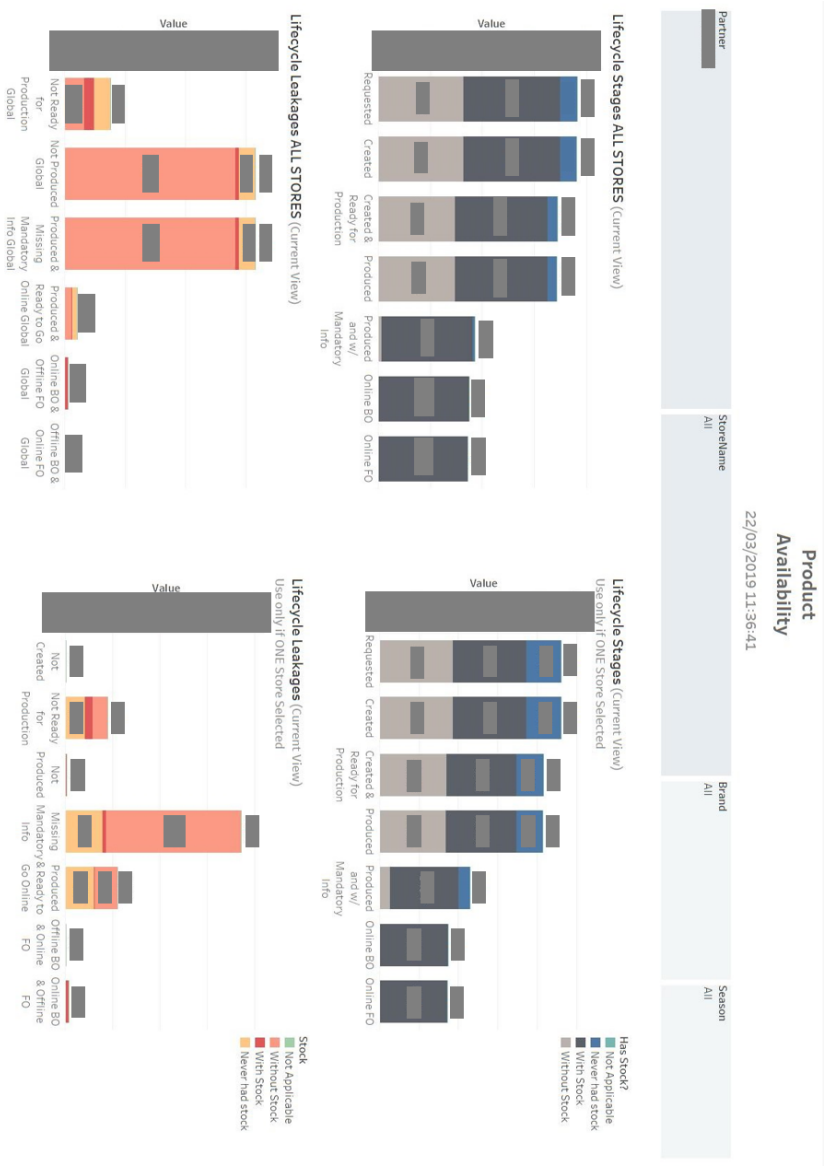

Figure C.1: The Product Availability Dashboard (1)

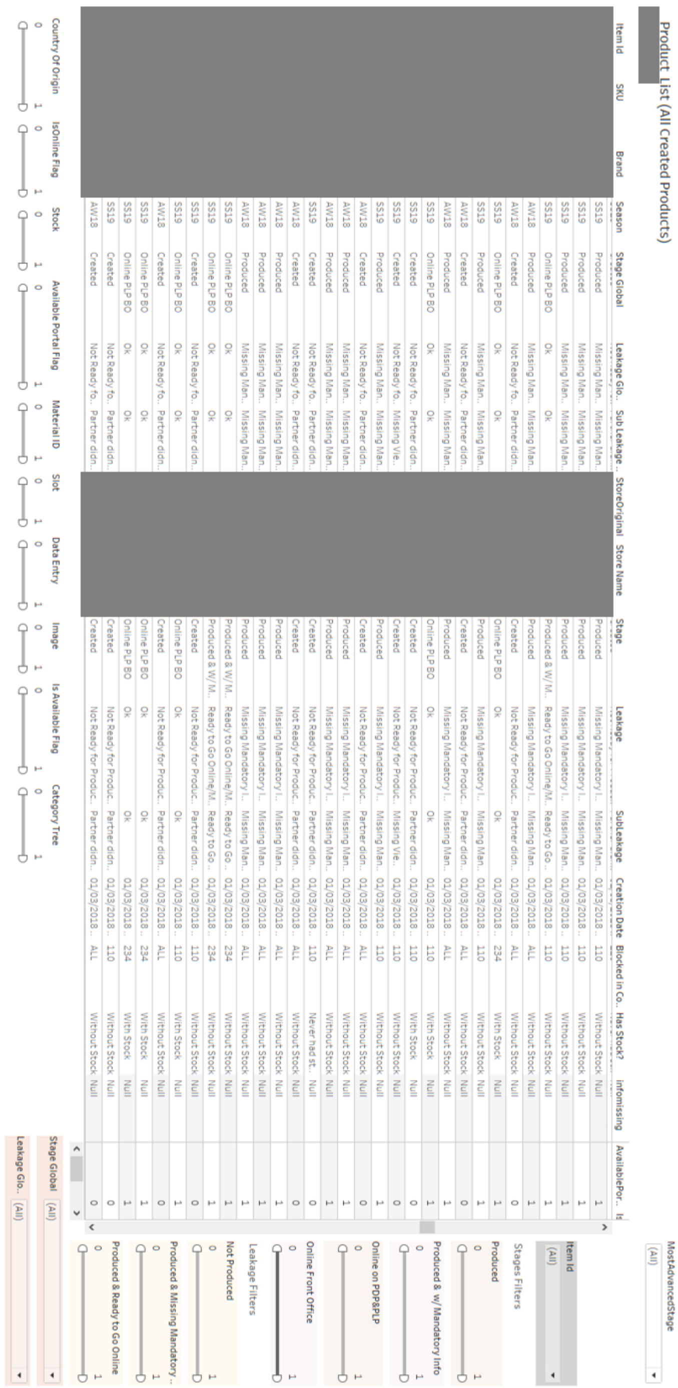

Figure C.2: The Product Availability Dashboard (2)

Appendix D

Performance Evaluation Dashboard

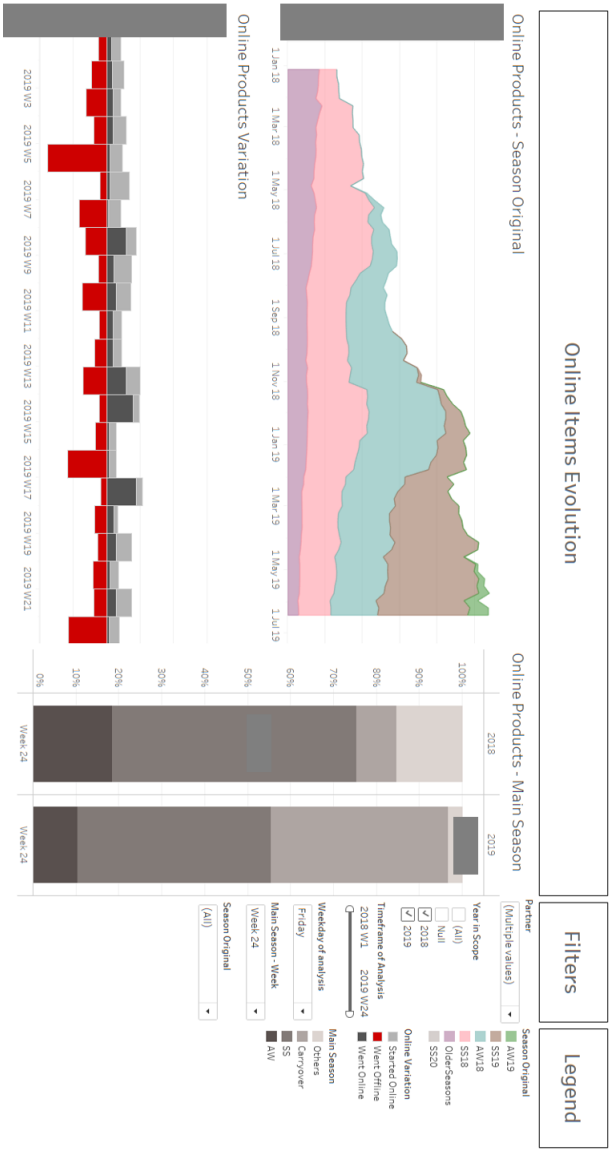

Figure D.1: The Performance Evaluation Dashboard - Online Items Evolution (1)

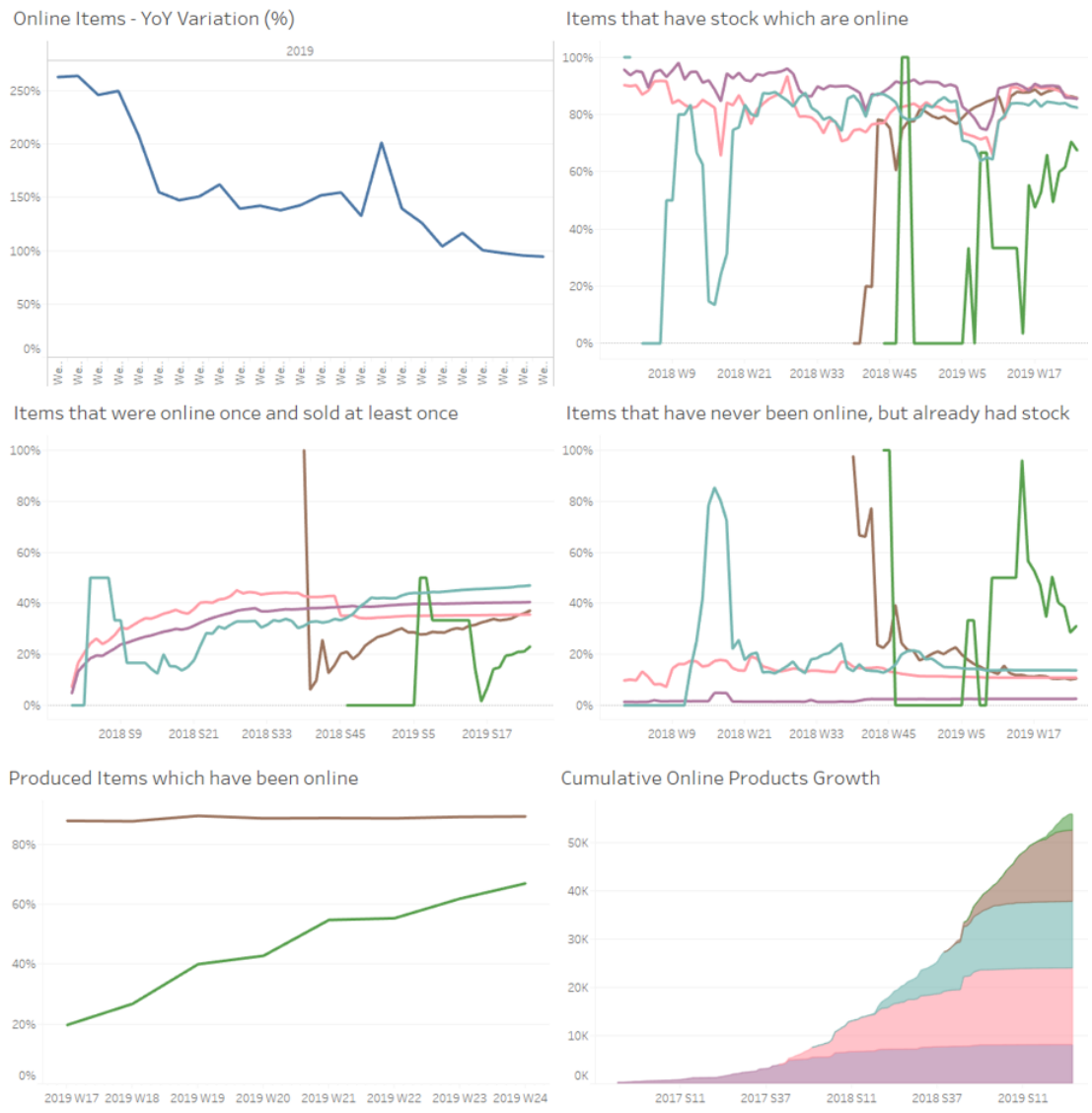

Figure D.2: The Performance Evaluation Dashboard - Online Items Evolution (2)

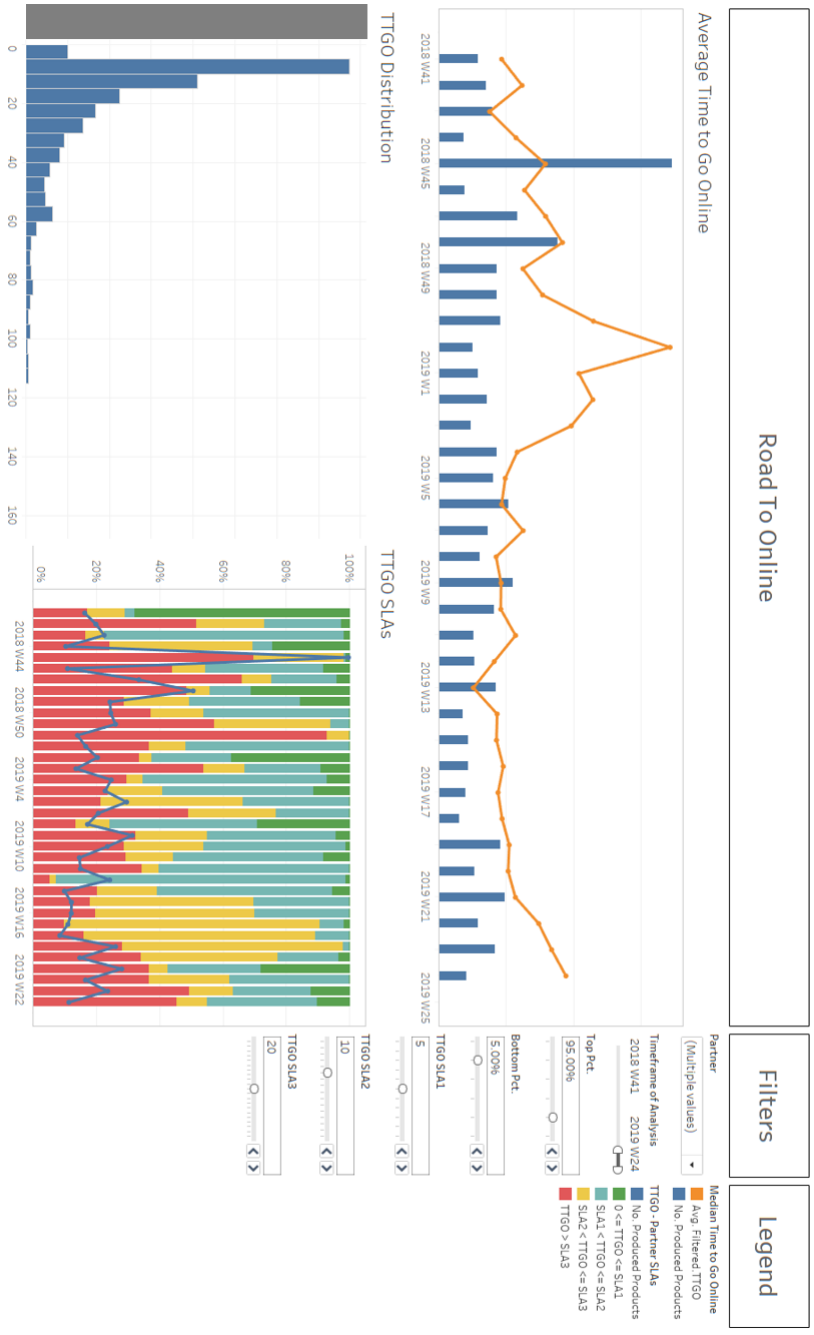

Figure D.3: The Performance Evaluation Dashboard - Road to Online (1)

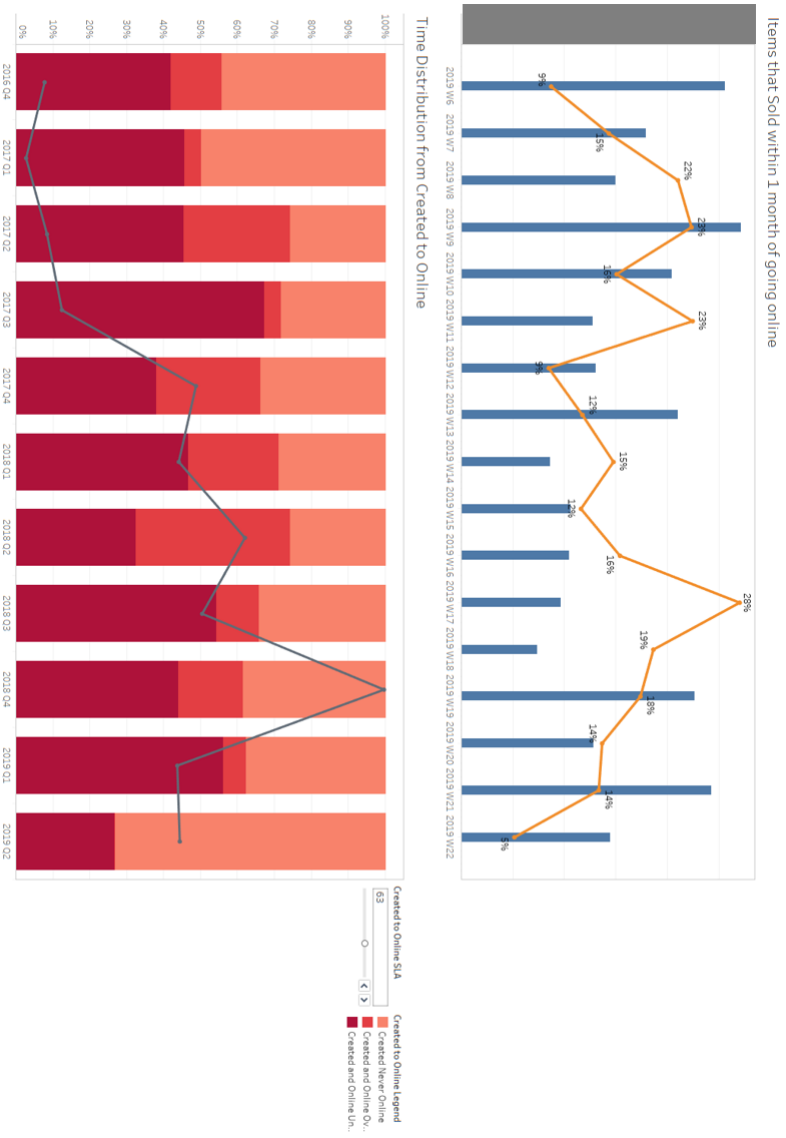

Figure D.4: The Performance Evaluation Dashboard - Road to Online (2)

Appendix E

PIM Vendors Evaluation

Table E.1: PIM tool vendors' evaluation from a Product Data Structure and Data Maintenance perspective.

Requirements		Vendor 1	Vendor 2	Vendor 3
Product Data Structure				
1.1	Product catalogue can be built by Farfetch PIM specialists without need for vendor consultancy or programming		x	
1.2	Different attributes, hierarchies, business rules can be defined at product category level	x	x	x
1.3	Typical but more complex business rules can be defined using minimal scripting		x	
1.4	Tool supports M:M relationships (e.g., a product can belong to multiple categories and/or multiple styles)			x
1.5	Tool supports global and tenant specific data attributes/hierarchies		x	x
1.6	Tool supports multiple versions of attributes	x	x	
1.7	Data model (meta-data) can be easily extracted/reported to support model management and integration efforts	x	x	x
1.8	Tool supports searchability/findability (e.g., tagging, creating associations, defining synonyms)	x	x	x
1.9	Tool supports definition of blocking rules assigned to products - these rules are time-variant and can be classified and filtered by type (e.g., input by partner vs. input by Farfetch) and nature (e.g., commercial, legal, etc.)			
1.10	Tool supports defining description tabs for different types of product details/attributes (e.g. product description, sizing/fit, care instructions, ingredients)			
Data Maintenance		**Vendor 1**	**Vendor 2**	**Vendor 3**
2.1	Tool supports loading/creation of invalid records			
2.2	Tool supports definition of prices across time	x	x	x
2.3	Tool supports different prices/discounts to be defined based on size/colour as well as tenant/boutique and geography		x	
2.4	Search/load/update performance is not affected by information volume		x	
2.5	New records can be added via the GUI bulk loads from excel/csv files	x	x	x
2.6	Product data can be uploaded in bulk for creation or update of a subset of attributes	x		x
2.7	Product data load templates can be created using the tool's UI		x	
2.8	Prices for different geographies/time periods can be bulk uploaded	x	x	
2.9	Tool provides graphical hierarchy browser		x	
2.10	Tool provides preview environment for quality control	x	x	x
2.11	Ability to edit images			x
2.12	Tool supports maintenance of exclusive/VIP brands stock (products inherit VIP or Exclusive flag from their brands marked as exclusive/VIP)			
2.13	Ability to classify/filter/report an item by stage, leakage and sub-leakage	x	x	
2.14	Support for translations	x	x	x
2.15	Brands can control and automate when their products go online/offline (e.g., put items online and offline with one click and schedule for products to go live in specific markets)		x	
2.16	Carryover management: ability to edit current season			

Table E.2: PIM tool vendors' evaluation from a Product Workflow and Reporting perspective.

Requirements		Vendor 1	Vendor 2	Vendor 3
Product Workflow				
3.1	Workflows can be created/modified by Farfetch PIM specialists without need for vendor consultancy or programming		x	
3.2	Workflow ensures that data that is needed at a particular phase of the product life cycle becomes mandatory at that stage (but not before)	x	x	
3.3	Workflow can be configured to be dynamic (e.g., routing to different assignees depending on data attributes/business rules)	x	x	
3.4	Data quality/completeness issues or violation of business rules can trigger alarms for the attention of specific user roles depending on data values or issue type		x	x
3.5	External parties can take part in the workflow, create and update data subject to access control rights			x
3.6	Tool allows partners to validate prices uploaded - one-by-one and in bulk	x	x	
3.7	Workflow can generate notifications and timeout alerts		x	
3.8	Users can mark tasks as "done" as required		x	x
3.9	It is easy to monitor and report on which items are/were at what stage in the workflow and how long each item spent at each stage of the workflow	x	x	
3.10	Users can search/filter/select and mark products as their work-in-progress to avoid two people trying to edit the same product at the same time			
3.11	Allow supervisors to assign products in the workflow to data stewards	x	x	
3.12	Allow users to mark sections of product data as complete vs work-in-progress (e.g. for photos, descriptions)		x	
3.13	Tool allows reviewer role to approve product data within the workflow, or reject and send back to data steward with rejection message	x	x	
3.14	Tool supports tracking of performance of the user who solves problems (e.g., how many products he brought from offline to online)			
3.15	Tool supports defining targets for individuals and teams and provides dashboards to compare performance of teams that solve problems			
Reporting		Vendor 1	Vendor 2	Vendor 3
4.1	Prebuilt reports/search functionality supports typical reporting needs (e.g., filters, fuzzy search, search for null values, greater/smaller filters, begins with ends with filters, wildcards)	x	x	x
4.2	Data can be searched, and extracted to Excel spreadsheet through the UI	x	x	x
4.3	Tool supports filtering and reporting products that have all the mandatory information or products missing details or failing validation checks	x	x	x
4.4	Build reports/graphs (time series) using a mix of metrics of the product workflow	x	x	x
4.5	Build graphs (column bar) with % of total catalogue which went from stage A to stage B		x	
4.6	Build graphs that compare performance in a certain metric (e.g., products online / products produced)		x	
4.7	Build ranking graphs for top / bottom partners for a certain metric (e.g., products online / products produced)			
4.8	Access to graphs and dashboards is controlled by user/user group permissions (e.g., internal, partner) and managed by an admin through the UI	x	x	
4.9	Create dashboards with the previously created graphs and have the ability to apply localized filters (e.g., filter by season)	x	x	x
4.10	Prices for different geographies/dates can be extracted/downloaded	x	x	

Table E.3: PIM tool vendors' evaluation from a Data Quality, Access Control and Integration perspective.

Requirements		Vendor 1	Vendor 2	Vendor 3
Data Quality				
5.1	Tool supports duplicate checking, and can proactively warn users when such records are added		x	
5.2	Tool helps identify/create goldenn record, and links duplicates to the gold record		x	x
5.3	Tool supports probabilistic & fuzzy matching			
5.4	Tool can display warning message whenever prices are not meeting certain conditions (e.g., price equal to zero)	x	x	
5.5	Quality control (e.g., blocking items from going live if info is incorrect)	x	x	x
5.6	Tool supports blocking/inactivating records, with different reason flags.			
5.7	Inactivated records are accessible in the UI and in reports but clearly flagged as inactive			
Access Control		Vendor 1	Vendor 2	Vendor 3
6.1	Tool keeps track of audit history at field level (e.g., what changed when by whom)	x	x	x
6.2	Demonstrate capability to support multi-tenancy (e.g., data can be segregated for different partners)	x	x	x
6.3	Data access can be managed at user/user group level and security can be set at record/field level as well as product category level	x	x	x
Integration		Vendor 1	Vendor 2	Vendor 3
7.1	Tool generates dimension tables (star/snowflake) to support data warehousing needs			
7.2	Tool supports creating full extracts of product data to database tables		x	x
7.3	Tool supports real time updates through messaging/APIs		x	