

Louvain School of Management

How to Develop and Monetize Network Effects In a platform-based company?

A concrete case in the clinical genomic domain with a Next
Generation Sequencing bioinformatics data platform

Master's thesis done by
Sébastien Duchateau

Master 120 credits business engineering, International Business

Thesis Supervisor
Christophe Lejeune
Academic year 2016–2017

Appreciation

I would first like to thank my supervisor, Christophe Lejeune, who agreed to support the realization of this thesis. At the same time, I would like to thank all the faculty of UCL-LSM for teaching me essential lessons for my personal and professional development; my father, for his precious mental support at the end. I would also like to thank Sophia Genetics, the company that allowed me to live an incredible professional opportunity in the US, more precisely my boss, Kevin Puylaert, for the help and the time provided for the training. And finally, I would like to thank all those who participated in the achievement of my studies; special mention to my parents, my girlfriend and my cousin, who always supported me in my choices.

Table of Contents

Appreciation.....	III
Tables	IX
Illustrations	IX
Abbreviations.....	X
1. Introduction	1
1.1 App Store Revolution	1
1.2 Research questions.....	2
1.3 Thesis structure	3
2. Transition to a platform economy	5
2.1 Platform definition	5
2.2 From resources to network and ecosystem assets	6
2.2.1 Network orchestrator	8
2.2.2 VRIO Model.....	10
2.3 Porter's five forces model	11
2.4 The main platform actors.....	13
3. Network effects.....	15
3.1 What are network effects?	15
3.2 Direct network effect	17
3.3 Indirect network effects	19
3.3.1 Definition	19
3.3.2 Platform envelopment	20
3.3.3 Multi-sided platform	20
3.4 Critical mass	22
4. Developing and monetizing network effects	25
4.1 What are the main challenges of monetization?	25
4.2 Finding the value to develop.....	26
4.3 Who to charge?.....	28
4.3.1 Customer Heterogeneity.....	28
4.3.2 Three different models for charging a specific user	29
4.3.3 Chicken or egg	32
a) Strategies for tackling the chicken or egg issue.....	32

4.4 Price structures	34
4.4.1 Membership fee	34
4.4.2 Transaction fee	36
4.4.3 Proportional fee	36
5. Sophia Genetics.....	39
5.1 Research questions.....	39
5.2 Presentation of Sophia Genetics company.....	40
5.3 Qualitative analysis methodology	41
5.3.1 Data Collection	41
5.3.2 Interviews	42
5.3.3 Observations.....	44
6. Analysis and Results.....	45
6.1 Global Data overview	45
6.2 Present Situation	46
6.2.1 How does Sophia Genetics develop network effects?.....	46
a) Data overview	46
b) The Community of Sophia Genetics platform	48
c) Platform utilization	51
d) Integration	53
e) Direct network effects (DNE).....	53
f) Indirect network effects	56
6.2.2 How does Sophia Genetics monetize network effects?	58
a) Data overview	58
b) Transaction fee	59
c) Membership fee	60
d) Monetize network effects	60
6.2.3 Limitations	61
a) Rules Limitation	62
b) Human Challenges	63
6.3 Future perspectives	64
6.3.1 How will Sophia Genetics develop new network effects?	64
a) Community	64
b) Integration	65
c) Direct network effects	65
d) Indirect network effects	66

6.3.2 How will Sophia Genetics monetize network effects?	69
6.3.3 The future limitation of NGS data analysis market	70
7. Discussion	73
7.1 Zig Zag strategy	73
7.2 Direct network effects-Integration and flagging system	73
7.3 Adoption Curve	74
7.4 Bundle and IDS products	76
7.5 Indirect network effect - Future Perspectives	77
7.6 Long-term monetization plan	79
7.7 Limitations	81
8. Conclusion	83
8.1 Theoretical Context	83
8.2 Practical context	84
Bibliography	87
Appendix	95
Appendix 1:	95
Appendix 2:	96
Appendix 3:	104
Appendix 4:	173

Tables

Illustrations

Figure 1: Tab Comparison	5
Figure 2: Economical Revolutions Cycles	7
Figure 3: Traditional Value Chain.....	7
Figure 4: Average Multiplier Comparison.....	9
Figure 5: VRIO Framework	10
Figure 6: Porter 5 Forces Framework	12
Figure 7: Platform Stakeholders	13
Figure 8: Network Effects with Phones.....	15
Figure 9: Uber's Network Effects	16
Figure 10: Types of Network Effects	17
Figure 11: Network Effects Diffusion	18
Figure 12: MSPs vs. other Business Models.....	21
Figure 13: Critical Mass Optimization Path.....	23
Figure 14: Value Creation for LinkedIn	27
Figure 15: Zig Zag Strategy to Reach Critical Mass	33
Figure 16: Ad-Valorem fees according to different categories	37
Figure 17: Data Triangulation	41
Figure 18: Interviewees.....	42
Figure 19: Data Overview.....	46
Figure 20: Development Categories Dynamics.....	47
Figure 21: Development of Network effects Keywords.....	48
Figure 22: Technology Adoption Curve	50
Figure 23: Direct Network Effects at Sophia Genetics.....	54
Figure 24: Leverage effects at Sophia Genetics.....	57
Figure 25: Monetization Keywords.....	58
Figure 26: Monetization Strategy today	58
Figure 27: Monetization of Algorithms.....	61

Figure 28: Two Main Limitations	61
Figure 29: Future Indirect Network Effects at Sophia Genetics	66
Figure 30: Leverage Effects between Oncologists and Pharma Companies	68
Figure 31: Integration and Flagging Leverage Effect	74
Figure 32: Adoption Curve with Chasm	75
Figure 33: Future Indirect Network Effects Leverage	78
Figure 34: Interactions Dynamics	79
Figure 35: Long-Term Horizon Monetization Plan.....	80

Abbreviations

DNE	Direct network effects
INE	Indirect network effects
NGS	Next-Generation Sequencing
IDS	Integrated diagnostic solution
LIMS	Laboratory information management system
ACMG	American college of medical genomics
GDPR	General data protection regulation
TF	Transaction fee
MF	Membership fee
MSP	Multi-sided platform

1. Introduction

The App Store's great idea ten years ago revolutionised traditional business models and opened a new era for generating business entirely based on network effects and platform concept. In this introduction, the App Store story will be briefly analysed. Following this, the key questions behind this thesis will be examined and finally, how this thesis has been structured and why.

1.1 App Store Revolution

“Any businesses where information is a key element
is a candidate for the platform revolution”
(Marshall et al., 2016).

Ten years ago, in 2007, the first iPhone was sold. At this time, keyboards phones from Nokia, Sony Ericson, LG, Motorola, and Blackberry dominated the market. Together, they represented 81.1 % of the market shares (Gartner, 2008). Those multinational institutions had strong distributional logistics, top operational systems, huge turnovers, and a large budget for research and development. How did iPhone succeed in conquering the market within two years? Today, iPhone passes the threshold of 1,200,000,000 units sold and engaged a total of \$738 billion revenue in the last decade.¹ Of course, the product had novelties such as the design and the touchpad. However, something central was different and really creative, something never seen before, something that revolutionised not only the phone industry but many more sectors, something that deeply changed our habits: the App Store. The App Store suddenly created a way to connect developers and consumers on a global basis and potentially over an unlimited range of applications. An active community was born. "If you look at all the things that happened in the world five years ago, the App Store is among the most influential things for humanity that happened in that period,"

¹ Forbes article from McCarthy (2017): Apple has sold 1.2 billion phones over the past ten years

2.

said Phil Libin in 2013, CEO of Evernote platform. Apps developers can directly have access to a community of millions of potential customers and can compete with the biggest companies. On the other hand, consumers have access to a vast marketplace of diversified applications. This phenomenon was totally unexpected and had not been perceived by the competitors. It took them by surprise and clearly this marked the beginning of the Apple supremacy in the mobile phone industry. Moreover, through this concept of community, developers generated a huge amount of revenue, not existing before. In June 2017, Apple announced that its worldwide developers' community has earned over \$70 billion since the beginning of the App Store.²

Apple created the concept of a global platform fuelled by users' interactions. These interactions are the value created within in the community. They are called **"NETWORK EFFECTS"**. Apple invented a new asset based on the mobile phone allowing developers and consumers to directly exchange and trade. Overall, Apple succeeded in combining a tangible asset (iPhone) and an intangible asset (the network) .

1.2 Research questions

In this thesis, I will concentrate on two main questions: how to develop network effects and create an intangible asset based on a platform, and how to monetize such intangible asset. Numerous platforms have appeared in the last 10 years and they are now omnipresent in our society. They are revolutionizing core business models in several industries. They represent today multi-billion dollar firms, the largest market capitalisations in the world, and this in just ten years. The main ones are Apple in the mobile phone industry, Amazon in e-commerce, Google in big data, Airbnb in hotels, Uber in transport, LinkedIn in professional interactions (500 million users for LinkedIn, purchased by Microsoft in 2016 for \$26 billion)³ and many others. A lot of literature has appeared for helping us better understanding the mechanisms behind this phenomenon. It is actually all about enabling a match of interests between categories of people and by initiating numerous active valuable connections between

² Source from Apple's official press release website (2017)

³ Forbes article from Rick Gillis (2017): Microsoft 's acquisition of LinkedIn

them. “You have to make the connection extremely valuable,” says Robert Schmalensee from MIT⁴. The more users get connected to share knowledge, facts, observations, and objects, the more value is created. The key characteristics of this phenomenon are the network effects. These are critical to succeed in building a sustainable platform community. Once this is done, the next main challenge for platform creators is to succeed in monetising network effects. In this thesis, I will analyse the various strategies that lead to make a platform attractive, that lead to the development of networks effects and finally that lead to an auto-fuelled active community of users. I will then analyse the various strategies that lead to monetise these platforms and their interactions.

In summary, I will aim at replying to two questions:

1. How networks effects became the main assets in platform-based businesses and which kind of network effects were created ?
2. Which strategies were thought and developed to monetize those effects?

1.3 Thesis structure

The thesis is divided in two parts: the theoretical part and the practical part:

For the theoretical part, I will first explore the transition from a traditional linear tangible asset builder to a platform-based company. I will develop the context of these platforms and the importance of network effects as a main asset. The concept of Direct Network Effect (DNE) and Indirect Network Effect (INE) will be analysed. Finally, I will examine the main strategies that have driven the monetisation of the networks effects and the theory developed around such strategies.

In the practical part, to illustrate the phenomenon of network effects, I will explore

⁴ Robert Schmalensee is a well-known professor of economics at MIT

4.

the example of an incredible Swiss start up in the bio-informatics health sector, Sophia Genetics. Over the last five years, the health sector also started to be impacted by platform models with network effects. Sophia Genetics is aiming at creating a global platform around the concept of “Data Driven Medicine”. Concretely, an immense amount of data is starting to be generated from DNA analysis in clinical environments over the world. But it is extremely hard to analyse and interpret these data for extracting useful insights. And, paradoxically, the more data the better could be the results and the interpretation. The complexity and the evolution in this domain are such that the collaboration between worldwide medical institutions will be essential to move on and to extract the best of these new disciplines. Network effects phenomenon will be ideal in the context of creating incentives and useful interactions between medical institutions for enhancing the medical progress. Sophia Genetics’ Vision is to create such global platform and such network effects around the concept of “Data Driven Medicine”.

How could a global platform enable a new worldwide collaborative project involving hundreds of medical institutions around the world? How will this have a major impact in medical progress? This is what I will aim explaining through the fascinating story of Sophia Genetics.

2. Transition to a platform economy

The aim of this chapter is to analyze differences in the transition from traditional to platform-based companies to better understand the context and the importance of the network as a main asset. The following tab resumes the key shifts by citing the differences between the two models.

Figure 1: Tab Comparison

Characteristics	Traditional Business Network Approach	New Business Network Approach
Products and services	Relative simple, unbundled, and slowly delivered products and services	Relative complex, bundled, and fast delivered products and services
Value creation	Supply chains with long term connected relationships	Demand networks with quick connect and disconnect relationships
Coordination and control	Hierarchical and central control and decision making	Network orchestration with distributed control and decision making
Information sharing	Information sharing with direct business partners	Information sharing over and with network partners
Infrastructure	Actor platforms with information silos and systems	Network platform with networked business operating system

Source: Heck and Vervest (2017)

I also use the Porter (1979) model and the VRIO framework from Barney (1998) to better characterize a platform-based company's network asset. Finally, I describe the main actors involved in a platform.

2.1 Platform definition

What we observe the most in the literature are the terms 'two-sided' market by Tirole and Rochet (2003) and the 'multi-sided' platforms by Hagiu and Wright (2015). Platform organizations help different groups of users who are potentially attracted to each other. The heart of the business for the two-sided market is to offer a common

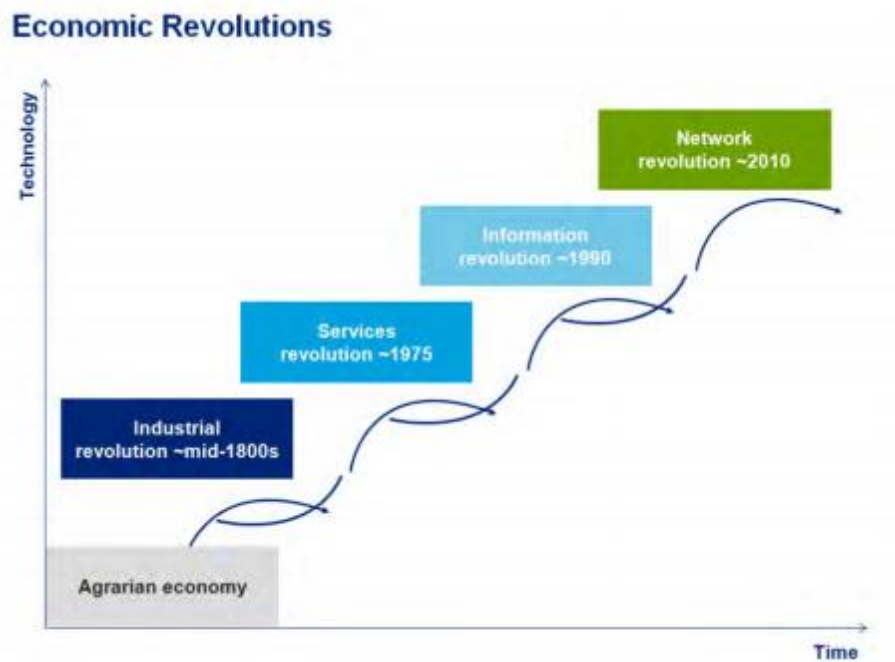
6.

meeting area, virtual or real, to facilitate interactions between members of the two distinct customer groups (Evans & Schmalensee, 2005). On the other hand, multi-sided platforms organizations enable interactions between two or more types of "affiliated" customers (Hagiu & Wright, 2015). I will, in my thesis, use both terms but will focus on virtual platforms. The following definition is the closest meaning to this axis of research. Marshall (2016) defines a platform as an online organization that facilitates the value creation through interactions between what he calls producers and consumers. The platform offers an open, participative online groundwork and sets rules. This can include e-commerce platforms such as eBay that link sellers and buyers, social networking platforms that enable connections between people such as LinkedIn, or commercial search engine platforms that associate advertisers with internet users such as Yahoo (Bakos & Katsamakas, 2008). Those platforms succeeded in creating a huge stock valuation, even though they have limited physical assets. So what is the main asset of platforms-based companies? It is the network asset explained in the following subchapter.

2.2 From resources to network and ecosystem assets

Platform business models have existed since centuries with, for example, local town marketplaces composed of artisans and consumers. However, thanks to vast computer improvements in data processing and cloud computing, platform organizations have been able to evolve online. Consequently, platforms can offer new possibilities of services and unlock innovative markets. "Relational capabilities and new complementarities among a firm's resources and capabilities can be exploited" (Amit & Zott, 2001). Today, platforms have the ability to analyse a massive amount of data and find interpretable insights to increase matches between users. Naturally, this implied changes from the traditional economy called 'pipeline' by Marshall (2016) to the platform economy. This transition led to a new economic revolution—the network revolution.

Figure 2: Economical Revolutions Cycles



Source: Ribaud (2016)

One of the main transition shifts comes from the assets. In a pipeline economy, assets are based on tangible resources with a traditional linear chain. In fact, all the chain is optimized from the raw resources to the customer consumption in order to maximize the value created. It is the internal optimization (Won et al., 2007). "In traditional linear chain, value moves from left to right with on the left the cost and on the right, the revenue. In platform two-sided markets, cost and revenue are both to the left and the right" (Eisemann et al., 2016).

Figure 3: Traditional Value Chain



Source: Reillier (2017)

8.

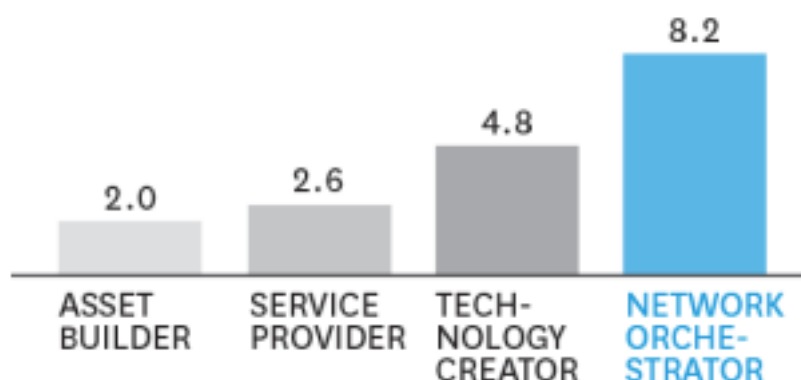
Investors in the last century knew those characteristics and valued firms based on their competencies to generate margins from physical assets (Ribaudo, 2016). Today, investors need to use other KPIs and metrics to value platform companies. In the platform world, the main asset is the network created from the users' interactions. It is the external asset optimization (Marshall et al., 2016). Pipeline companies tend to improve customer experience in an individual, centric way. Platforms develop users' experience in a much more collective approach with constant feedbacks in order to build a sustainable ecosystem (Merchant, 2016). For instance, sometimes a user needs to make a sacrifice by paying a fee to satisfy another one who will be subsidized (Marshall et al., 2016). The ecosystem created by the company through network effects is becoming as much important, if not more, than the product (Cusumano, 2010). You will see in the following subchapter that those network effects can be highly valuable on the market.

2.2.1 Network orchestrator

In 2016, Deloitte consulting firm did an extensive study on four different types of companies according to their core businesses (Ribaudo, W. 2016). The first category is the asset builders; they sell principally physical products. The second type describes service provider companies; they usually charge per hour (e.g., a law firm). The third category is the technology creators; they sell their intellectual property and expertise—like software analytics. The last type defines companies known as network orchestras. They are successful firms that have created a network of interaction between users, releasing an important value. These companies can also sell goods or services. In fact, the combination of their assets with a network infrastructure increases considerably the price-to-revenue ratio according to the study done. Network orchestras and technology creators' business models develop faster, scale faster, and consequently receive higher revenue (Ribaudo, 2016).

Figure 4: Average Multiplier Comparison

AVERAGE MULTIPLIER (PRICE-TO-REVENUE RATIO)
FOR THE S&P 500 COMPANIES IN 2013



Source: Fenley & Libert (2014)

This framework is based on S&P index companies⁵ from 1972 to 2013. The price-to-revenue reached an average multiplier of 8.2 for network orchestra companies mainly due to network effect value creation. In other words, the markets pay 8.2 times the valuation for each 1\$ revenue yielded (Ribaud, 2016). The 8.2 multiplier ratio would be justified by the marginal cost reduction of this type of company (Rifkin, 2014). “Indeed, because of network effects, platforms enjoy increasing returns to scale. Users will pay more for accessing a larger network, so margins improve as user base grow” (Eisemann et al., 2006).

Companies are clearly being valued differently following a new trend of businesses. In 1975, 83 % of the 300 S&P index companies were asset builders. The technique of market capitalization was less complicated than it is today. In fact, the calculation was based on the difference between the total firm's assets and total liabilities. However, in 2015, the situation has changed, given that 84 % of the market is shared between service, technology, and network companies (Ribaud, 2016). The value evaluation comes not only from the horizontal or vertical integration (Teece,

⁵ The S&P index is constituted of the 500 US largest companies. It is recognized as a benchmark for all the US market (InvestSnips)

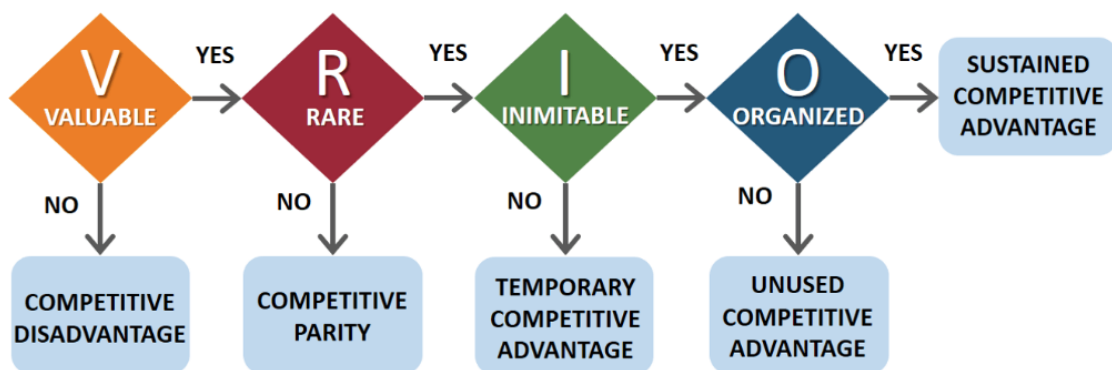
10.

1980) but also from an external integration (Marshall et al., 2016). If the external integration is optimized in the business model, companies will possess a strong competitive advantage according to VRIO framework (Barney, 1998).

2.2.2 VRIO Model

In pipeline firms, competition relates principally to the resource, the production, and the distribution chain. They gain a competitive advantage by cost-reducing all those different categories (Mentzer et al., 2001). In fact, the framework VRIO from Barney (1998) puts in evidence a sustainable competitive advantage by defining a resource valuable, rare, inimitable and organized. In the platform world, the resource applied in the VRIO model is the network.

Figure 5: VRIO Framework



Source: Business 2 you (2016)

As I mentioned before, the network asset is valuable, thanks to the users' interactions. For Heck & Verves (2017), the asset is rare and difficult for competitors to quickly have it in their hands. In a short- and medium-term horizon, a competitive advantage can be sustained for some time because it is difficult to gather a community quickly. However, in a long-term view, any opportunities can be grapped by the competitors in order to enable, as well, the building of a network. But it is very hard to copy due to the complexity of users' interactions and to the large databases progressively built. Indeed, network effects can represent high barriers to entering the market (Amit & Zott, 2001). For the organized criteria in the framework, you will see

in the next chapter that a lot of different strategies have been developed to optimize the network asset exploitation. For Parker (2016), the entire traditional competition is disrupted.

Monopolies are often present in this kind of market. For instance, the battle of social networks between Google+ and Facebook characterize a substantial competitive advantage. Google+ didn't have the chance to be as famous as Facebook because the network effects and the community created by this one was already too strong to compete with. "Unlike other social media offers which have their niches, Google+ was seen as a straight competitor to Facebook, but it was tough to compete against the largest network competitor in this business, which was its biggest hurdle."⁶ Facebook has built a long-term network of users, which is driven by the phenomena called 'lock-in' (Marshall et al., 2016). A strong base of data users is a significant competitive advantage and leads to a winner-takes-it-all spirit. Many 'two-sided network' industries are operated almost exclusively by a single platform (Eisenmann et al., 2006).

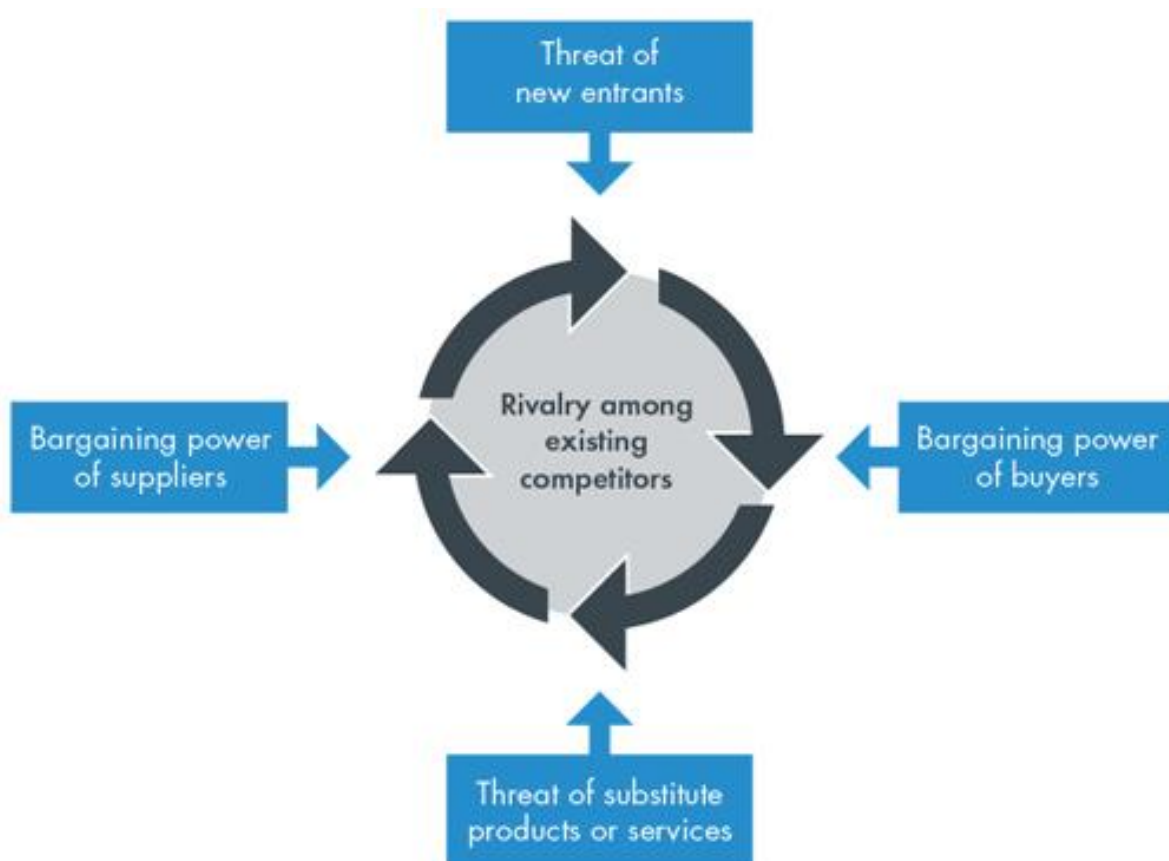
However, in certain cases, platform businesses can grow very fast but can also shut down extremely fast due to negative effects. For this reason, after analysing the competitive advantage, it is crucial to understand what the main differences are in the five external forces behaviour from Porter's (2008) model with platform businesses.

2.3 Porter's five forces model

In the transition to platform economy, the five forces model from Porter (2008) still applies but behaves differently. Grundy (2006) and Amit (2001) claim that the model should be modified in the context of platform markets. In fact, this model doesn't take into account the network effects (Eisenmann et al., 2006). External forces are always considered as negatives for the company (Parker et al., 2005) while they could have a positive impact on network effects in the platform world.

⁶ Statement done by Richard Broughton, research director at Ampere analysis company (2014)

Figure 6: Porter 5 Forces Framework



Source : Chartered Global Management Accountant (2017)

The essential in the platform strategy is to know when exactly the 4 external forces on the figure above are a positive value for the network and when they are a threat (Watson, 2017).

For Van Alstyne (2016), there are two main shifts to take into account when we are using the model:

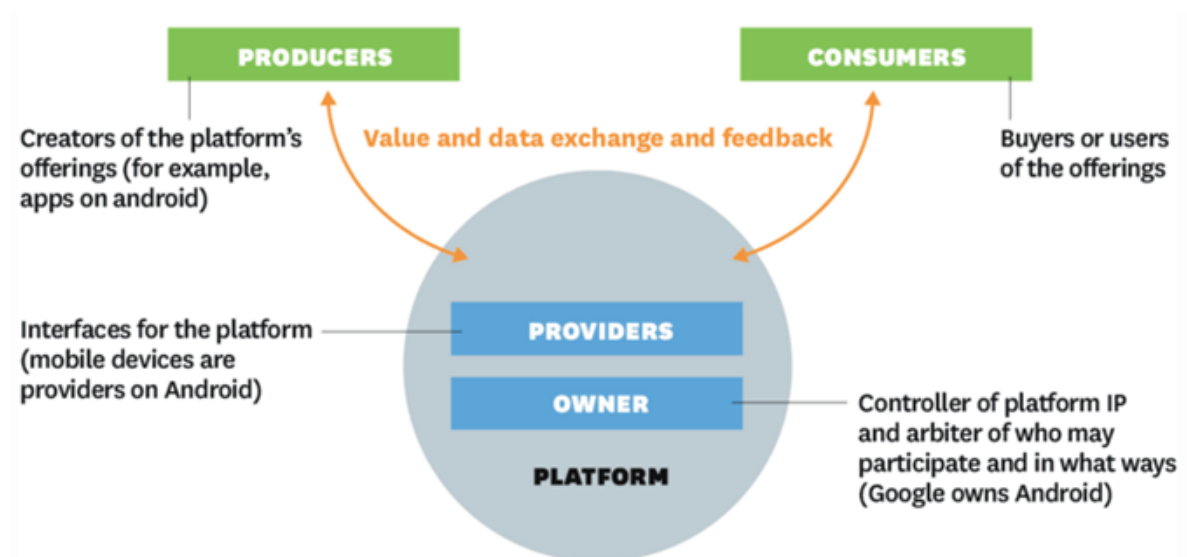
- Industry boundaries are more dynamic in platform businesses. Indeed, companies can grow extremely fast and become multi-sectoral.
- Competitors can sometimes be transformed into a positive external force by becoming complementors for goods or services on the platform. For instance, in some situations, e-commerce platforms can also sell their own products—not to compete with other sellers but rather to bring extra-value by offering complementary goods (Zhu & Liu, 2016).

After analysing a few differences from traditional linear chain to platform companies, we now better understand how networks can become key assets. But who are exactly the participants in those networks?

2.4 The main platform actors

It's important to understand which actors are involved in the ecosystem to capture the dynamics better. The figure below illustrates the relationship between the main actors.

Figure 7: Platform Stakeholders



Source: Choudary (2016), *Harvard Business Review*

First, there is the owner of the platform who has the intellectual property and manages the ecosystem. Second, the provider is the physical interface—such as computers or mobile phones—that hosts the platform. Third, the producers are the ones who generate the offerings. The last key participant is the consumer who uses those offerings (Marshall et al., 2016). Producers and consumers are also called sides (Reillier et al., 2017). Sides are the ones who interact through the platform by exchanging values, data, and information. The last main stakeholders are the partners

14.

that enhanced the value of the network for all the participants by creating strong partnerships with the platform owner. It can be, for example, individuals, communities, and businesses.

Another relevant point is a feedback loop formed by the sides, which is extremely important in the platform world. Feedbacks are a strong asset in the community that reinforces connections between sides. It is the platform manager's duty to take care of the quality of the comments, which is directly related to the network effects (Gawer, 2014).

In this chapter, I first analysed the progressive transition of a company from a tangible asset builder to a network orchestrator. I saw how network became a highly valuable asset in the platform-based company. Then I used the VRIO and the five forces framework from Barney (1998) and Porter (2008) to better understand the competitive advantage and the external forces of a platform. Finally, I described all the different actors involved in a platform and the relations between them to introduce the next chapter.

3. Network effects

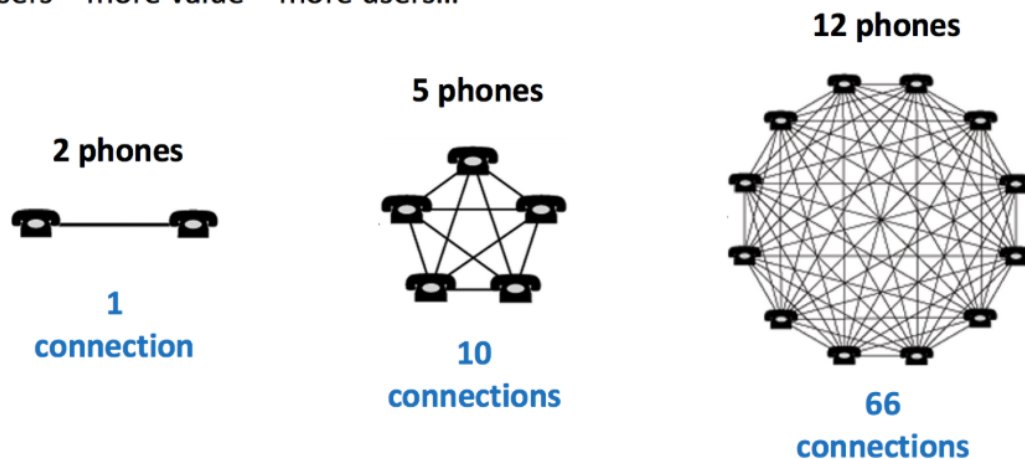
The purpose of this chapter is to study the two main types of network effects, direct and indirect, and their origins. It is the analysis of core concepts that enable network effects to exist. The chapter also describes the relationships between users and the platform. Finally, I define how to reach critical mass by using network effects in a platform context.

3.1 What are network effects?

Network effects are defined as the impact that the number of users of a platform has on the value produced for each user (Marshall et al., 2016). The following figure illustrates network effects with the number of phones. More users mean more value offered because there are more interactions to leverage for the platform's managers. The same phenomenon is observed in platform-based companies.

Figure 8: Network Effects with Phones

More users = more value = more users...



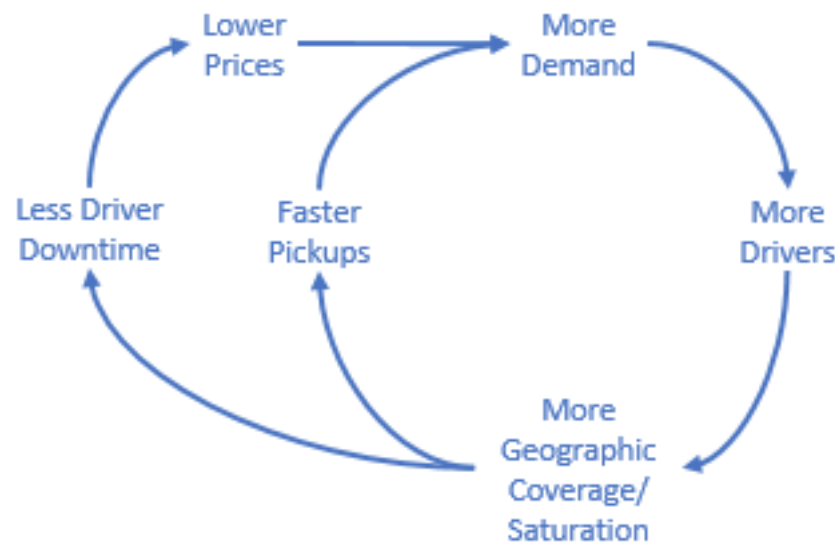
Source: Parker (2016)

If I analyse Uber, for example, the larger the volume of users in the network, the better the match between the offer and demand. The data collection on the platform will also enhance efficient connections between drivers and passengers through the

16.

ranking system. In fact, the more the data becomes rich in information, the more the precision and value increase throughout the community. This effect is typically called “data-driven network effect” (Liu et al., 2016).

Figure 9: Uber’s Network Effects



Source: Zimmermann (2016)

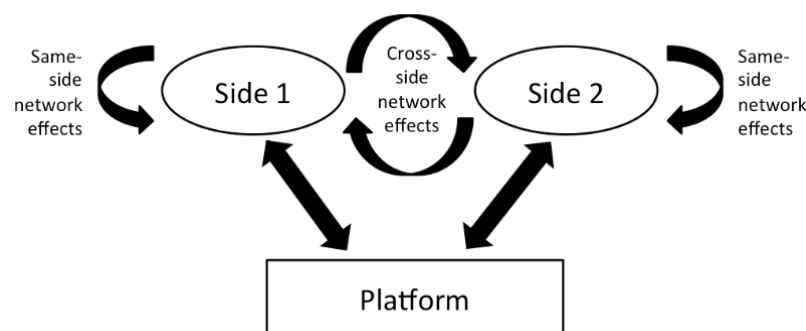
A virtuous circle is created, increasing the number of transaction on the platform. Figure 9 shows the ecosystem’s dynamic through network effects between users. Each line represents a value creation between consumers and producers. If there are more drivers or/and passengers, this will impact directly the service offered to each new user by reducing/increasing pickup times or by changing prices.

Another interesting approach from Belleflamme (2016) states: “Network effects are present if users care about participation and decisions of other users.” There are two dimensions of the participation of other users on the platform: the *volume* of users and the *quality* of each user. In certain platforms, it’s the *volume* which is the principal incentive for a new user to join the platform; on other platforms, it is the *identity* of the participants (Belleflamme et al., 2016). For instance, the Diners club platform company facilitates a transaction between restaurants and customers. In this situation, the volume is important. Indeed, for the restaurant owner, the incentive to join the platform is, to a large extent, dependent on the number of potential customers on the platform, and less dependent on the identity of the participants. On the other hand,

network effects are not just about the volume of the participants in the community but are also about the intensity of interactions between those participants (Reillier et al., 2017). Logically, if users are not ‘active’ enough on the platform, network effects decrease.

The relationship between the size of the network and the value created has been classified into a few different sub-categories to better understand the dynamics. In fact, the literature characterizes two principal types of network effects: the direct network effects and the indirect network effects—illustrated in the next figure. The total network strength is the sum of those two distinct effects.

Figure 10: Types of Network Effects



Source: Torres (2015)

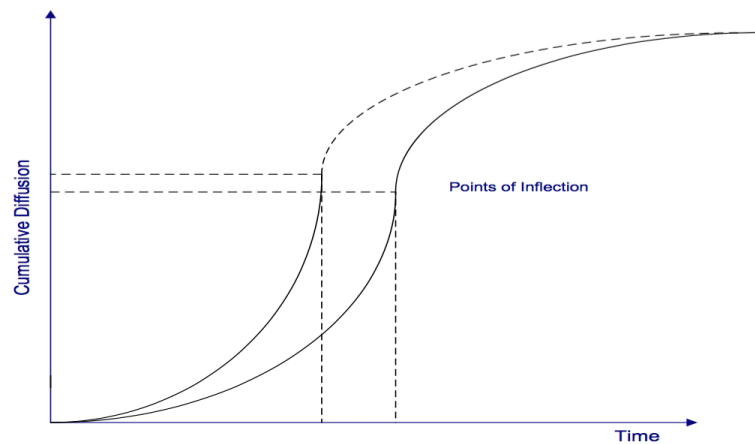
3.2 Direct network effect

Direct network effects, also named same-sided network effect, exist when the value of a technology toward a user depends positively on the volume of other users of this technology (Gawer, 2014). For Marshall (2016), those effects are created by the influence of users from one side of the market on other users from the same side of the market. In other words, it is the impact of consumers on other consumers or producers on other producers. Direct network effects measure the benefits or the costs that users obtain from the presence of other users on the same side, rather than those on the

18.

complementary side (Manchanda et al., 2015). For example, the more the users on a telecommunication network, the more important the ability to connect to other users (Clements et al., 2005). This is also called positive same-sided network.

Figure 11: Network Effects Diffusion



Source: Evans (2009)

Figure 11 illustrates how direct network effects influence the S-curve. According to Evans (2009), DNE act as main actors for a *catalytic reaction* until the point of inflection where the network growth starts to slow down.

Gawer (2014) suggests that specifically, direct network effects form demand sides of the economy of scale. The economy of scale enables a cost per unit improvement with the increase of output units.⁷ The platform approach is concerned with the ability to generate cost optimization when more and more users are added to the platform, thanks to direct network effects. It is essential for a platform business to push economies of scales as first priority by quickly and consistently adding users (Langley & Leyshon, 2016).

However, it is possible also to face negative same-sided effects with a threat on the number of users on one side of the market. The addition of new users on the platform decreases the global value for each user. For instance, the platform *Upwork* connects freelancers and independent workers. If there is a growth in the number of freelancers,

⁷ According to the website Investopedia (2016)

it will attract workers. That being said, if the number of freelancers grows faster than workers, it becomes more complicated for workers and freelancers to find each other. The number of freelancers will decrease, and this will be a negative same-sided network effect. If the effect is too strong, there will not be enough freelancers on the platform, and the number of workers will start to decrease as well. We start to observe that network effects on one side of the platform directly influence the behaviour of the members on the other side. This is called indirect network effects.

3.3 Indirect network effects

3.3.1 Definition

Indirect network effects are essential for a two-sided or multi-sided platform market. They are also called cross-group or cross-sided network effects (Gawer, 2014; Evans, 2008). It is defined by Marshall (2016) as the impact created by users of one side of the market on users on the complimentary side of the market. If a user on side A applies a large positive influence on each user on side B, then side A will be targeted intensively by the platform (Armstrong, 2006). According to Belleflamme (2016), on numerous platforms, we can analyze a specific group of users whose actions influence the value offer of other groups. One example from Reillier (2017) is in operating systems platforms with two distinct groups: the apps developers and the users. The effect results from a virtuous circle where additional developers create more apps; more users will be attracted by the offer, which will again push more developers to join in.

Another interesting definition from Hagiu and Wright (2011) is that indirect network effects are developed if the decision from *side A* members depends on the number of users on *side B* and vice versa. Thus, platforms and network effects reveal a pre-existing fundamental interdependency and complementary relationship between two or more groups of users. The complementary strength is a self-reinforcing loop that gives initial advantages to platform owners (Gawer, 2014).

3.3.2 Platform envelopment

Eisenmann (2010) puts in evidence the notion of 'platform envelopment' based on the complementary relationship between different groups on the platform. In fact, it is a 'multi-bundle' platform approach with a competitive bundled tangible product that leads to an efficient entry in a new market. Platform envelopment can be highly valuable for customers (Eisenmann et al., 2006). A relevant recent example is the online Amazon platform that bought up Whole food for 13.4 billion dollars.⁸ "Without stores and with limited inventory, Amazon possibly has more information on retail goods and their buyers and sellers than many other businesses" (Heck & Vervest, 2007). Amazon will have the possibility to use this kind of information from the customer database to enhance its competitive advantage in the grocery market through Whole Food stores. On the other hand, Amazon will also pull out data from those stores to improve online platform services. It is a win-win situation for Amazon for entering the grocery market that is worth \$750 billion⁹ today. Wal-Mart, a leader in the US market and now the main competitor of Amazon, is already thinking about a new way to innovate as well to respond to this shift.

Indirect network effects lead to economies of scope, which is directly related to the platform envelopment economic perspective. In fact, economies of scope phenomenon occur when the cost of producing the products 1 and 2 jointly is cheaper than producing the products 1 and 2 separately (Teece, 1982). In other words, for an economy of innovation, it is cheaper to jointly innovate products 1 and 2 than to innovate each product separately (Gawer, 2014). This is the core concept of platform-based approach with indirect network effects.

3.3.3 Multi-sided platform

A lot of literatures have analyzed 'multi-sided' groups economic model fuelled by indirect network effects. Later, Armstrong (2006), Rochet (2003), Evans (2009), Parker (2005) and Rysman (2009), Hagiu and Wright (2015) came with another

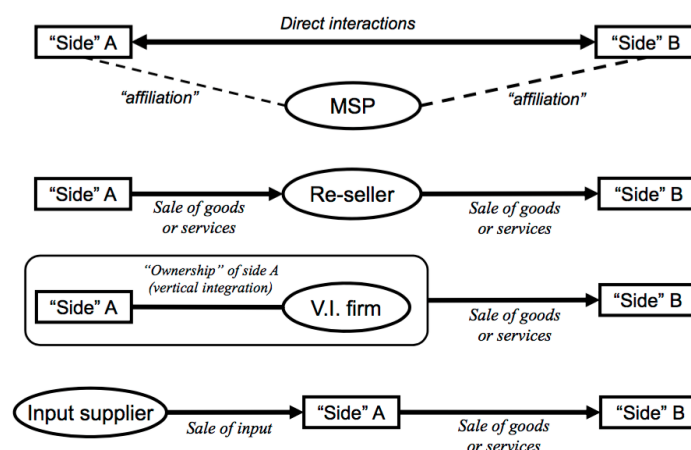
⁸ Source from the New york time article, Wingfield (2017)

⁹ Source from the New york time article, Wingfield (2017)

interesting approach to precisely define what a multi-sided platform or MSP is. The MSP concept provides a general vision with more than two different types of users who are co-dependent and interact through one platform (Henten & Windekilde, 2016). An online supermarket, for example, gives a relevant insight to the notion of MSPs. The supplier—who has the inventory risk on their side—cares about the number of customers that passes every day on the online store. On the other hand, the consumer cares about the products supplied by the store. These are indirect network effects. However, a lot of theoreticians wouldn't approve that this is an INE phenomenon. A more accurate definition of MSPs is given by Hagiu and Wright (2015) to better characterize, in our case, INEs. MSPs have to respect two conditions:

- They have to allow direct interactions and exchanges between two or more distinct groups from different sides.
- Each group from different sides is 'affiliated' with the platform. This is illustrated in the following figure.

Figure 12: MSPs vs. other Business Models



Source: Hagiu & Wright (2015)

In the figure above, we can see that there are four different business models. The one on the top is the multi-sided platform; the last three are the other 'classic' business

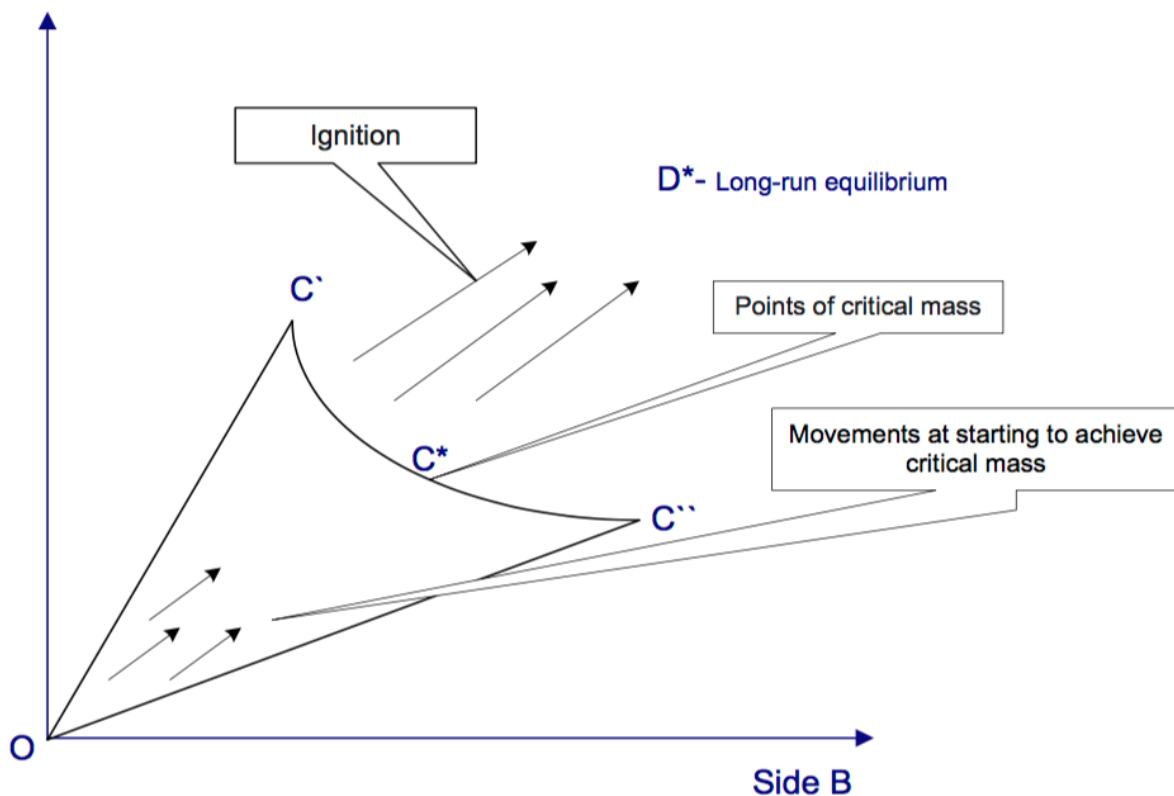
models. The comparison is relevant to analyse what exactly differentiates MSPs from other traditional businesses in order to better understand the dynamics of indirect network effects. By direct interactions, authors mean that users from different sides keep the control on the terms of their exchanges—for example, the price and the quality of the services offered—without passing through any intermediary. By affiliations, authors explain that users from different sides consciously buy or want access to the platform to be able to enjoy the different functionalities. In the case of the online supermarket, users would be to create an account to have access to the store. Thus, we can see that the notion of affiliation is necessary to create indirect network effects (Hagiu & Wright, 2015); however, it is not enough to define MSPs according to the authors. In the case of the supermarket, the affiliation rule is respected but not the direct connections between the supplier and the customer. Indeed, they cannot directly set the rules together—such as the price—without passing through the supermarket intermediary. According to Reillier (2017), letting the users set the price, such as in Airbnb, can enhance network effects. To conclude this paragraph, the indirect network effect is a very broad term that needs a specific context to be better characterized. In fact, the context is different for each platform.

Of course, there are also negative cross-sided effects that happen when the number of users of one group decreases due to the organizational behaviour of the complementary group. Therefore, the question is: What steps must be taken to ensure a sustainable positive network growth? The literature says it is the critical mass.

3.4 Critical mass

As I mentioned in the beginning of this chapter, the total strength of network effects is the sum of direct and indirect network effects. For platform owners, the principal aim of those network effects is to reach the critical mass. According to Reillier (2017), critical mass is reached at a stage where the network growth is **‘self-sustainable.’** Evans (2009) points out that reaching the critical mass is essential; platform managers must ensure critical mass to provoke the evolution of their platforms. If they don’t succeed in reaching this stage quickly, the future of the platform is threatened by a fast decrease of revenue.

Figure 13: Critical Mass Optimization Path



Source: Evans (2009)

The figure above represents the optimized path to reaching the critical mass, which is within the triangle $C'-C''-O$ (Evans, 2008). C' and C'' represent the two groups of customers from different sides on the platform. The optimal balance C^* between the two groups needs to be found by the platform's managers in order to maximize the value of each side. This is again based on indirect network effects and on the 'attraction' that the two sides have for each other. If the platform is going out of the triangle $C'-C''-O$, the balance will not be optimized. In fact, indirect network effects will be too weak to reach critical mass.

Another relevant notion is the 'tipping point' closely related to the critical mass. Malcolm Gladwell's (2000) book analyzed this notion and defined it as "the moment of critical mass, the threshold, the boiling point." According to Reillier (2017), the tipping

24.

point can be applied to the platform strategy. It is described as a specific situation when the company shifts from one state to another one, caused by an increase of network effects. For instance, it could be a shift from a competition state to a monopole state. The tipping point is also an essential notion for platform businesses to consider, as it is an inflection point that generally leads to critical mass and vice versa.

In this chapter I defined what were direct and indirect effects. I also characterized INE through the multi-sided platform approach from Hagiu and Wright (2015). Finally I studied the optimal path to reach the critical mass. But the growth track to critical mass depends principally on pricing (Evans, 2009) as we will see in the next chapter.

4. Developing and monetizing network effects

The aim of this chapter is to analyse different methods to develop and monetize network effects. First, I explore what are the sources of value created by the network effects on a platform. Then I analyse who are the users that specifically generate the value and who are the ones that can be charged or subsidized by the platform. I also study the chicken or egg issue and the Zig-Zag strategy to avoid it. Finally, I develop three different models of monetization.

4.1 What are the main challenges of monetization?

In two-sided or multi-sided platforms, monetization is tricky (Parker et al., 2006). The positive growth dynamic through different sides of the platform paradoxically makes the monetization delicate (Marshall et al., 2016). Platform managers need to think about the optimization of the best price strategy while still attracting both or numerous groups of users. Rysman (2009) analysed the two-sided platform price strategy and said that charging some users on one side of the market depends not only on the costs and the price elasticity of those users but also on the impact and the behaviour of users on the other side.

The interconnections and interdependencies formed through network effects in multi-sided platforms have generated new business models for Skilton (2015) to face challenges. Teece (2010) gives a pertinent definition of a business model which is closely related to our topic: “A business model articulates the logic, the data and other elements that support a value proposition for the customer, and a viable structure of revenue and costs for the enterprise delivering that value”. In our case, the elements to develop are the network effects. The challenge of monetizing, according to (Skilton, 2015), is to be able to produce financial wealth through different approaches to transform know-how, users’ relationships, technology, data, goods, and services into monetary value. The complexity in certain platform businesses is shifting from a non-charging to a charging business model (Teece, 2010). A great example is Google; they

succeeded in creating a business model where consumers don't pay and gain access to multiple services. This phenomenon will be seen more in details in chapter 4.3.2.

Monetizing the access for users could create frictions at the entry and have a negative impact on network growth (Marshall et al., 2016; Reillier et al., 2017). Thus, platform owners need to be extremely cautious and find smart strategies to develop the network effects. Moreover, a lot of very great platforms crashed because they hurried into monetizing the value created before reaching a critical mass of users. This is a dangerous path and would compromise the viability of the platform (Evans, 2008). For instance, there were two main competitors in the social network market in 2005 and 2006—Myspace and Facebook. The owners of MySpace promised the investors that the platform would quickly yield a billion dollars.¹⁰ However, this was absolutely not the case; one tenth of that was generated. To face this huge gap, managers decided to find multiple techniques to monetize network effects quickly by creating services and programs sponsored by other companies that were willing to pay MySpace. However, those services were not interesting for the users and did not add value to the community. If users are not concerned, not diverted, or feel they are being manipulated, they leave the platform (Clemons, 2009). Indeed, MySpace was not attractive enough to reach the critical mass and its users started to move toward Facebook's network. As we can observe, Facebook did the opposite by first reaching a huge critical mass and then performing a late monetization plan. The concept "acquire users first, then monetize" from Reillier (2017) allows less friction. That being said, after many users joined the community, thanks to this concept, the dynamics between network members became more complex and finding sources of value creation became a challenge as well.

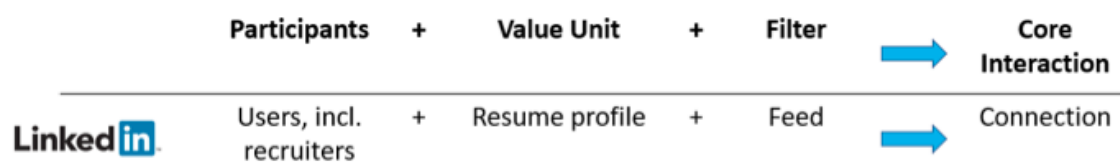
4.2 Finding the value to develop

In this subchapter, I will analyse the different sources of value produced by a platform. According to Parker (2016), there are four different categories of value formed in a platform. The first value created is from the consumers who want to have

¹⁰ Example from the Guardian (2015): MySpace what went wrong?

access to the interface. The second value generated is from the producers who want to be involved in the platform by having access to a large customer community. The third value created is from the consumer and the producers at the same time, enjoying all the facilities furnished by the platform. The last value produced is from the reinforcement of high-quality interactions between users developed by ranking systems on the platform. According to Evans (2008), the main value created comes from the users' access to one side of the platform to other. Another vision would be to involve multiple sides to create even more value (Rochet et al., 2003).

Figure 14: Value Creation for LinkedIn



Source: Creativity and Capitalism (2017), platform revolution a review for start-ups

The figure above shows, first, the main dynamics of a platform that lead to value creation through three main categories, which are the participants, the value unit, and the filter (Marshall et al., 2016). LinkedIn, for example, has created a value between users and recruiters who enjoy all the facilities provided by the platform to connect with each other. All the features—such as the ranking system through skills points—enhance the filter to better match users and recruiters.

Muzellec (2015) states that the value proposition exists mainly in the end-user experience, which is the audience that can be monetized either due to their size or personal characteristics such as demographic, psychology, medical background, interests.

Those values are the pillars for networks effects. They are described as sources of excess value because they would not exist without the platform (Choudary et al., 2016). Platforms interconnections generate vast sources of an excess value that need

to be developed and monetized. If the platform managers don't succeed in yielding sources of an excess value, monetization process might not be conceivable, and the life of the platform would be probably compromised. Moreover, the number of users on the platform is not necessarily enough to determine the value created by network effects. For example, the Zvent company had a large customer base of 14 million visitors each month (Parker et al., 2016). It was a platform providing a listings of events happening each week, connecting event producers doing promotion and consumers—the people interested in those events. Unfortunately, although they had a great community base, the monetizing process was very complicated because producers and consumers did not have any willingness to pay. The company was sold to eBay for an undisclosed price. To face this issue, we will see now that there are different innovative ways to charge specific users on a platform to face the challenge of monetization.

4.3 Who to charge?

4.3.1 Customer Heterogeneity

Numerous platforms have a selective charging model for users to enhance positive network effects or limit the negative network effects. I would like, first, to describe what Evans (2008) called 'customer heterogeneity' as a way of introducing this subchapter. Not all the customers on the platform are equal (Evans, 2008). Some users on one side are more valued by platform managers than those on other sides; Rochet and Tirole (2003) called them 'marquees,' Evans (2008) called them 'influencers,' and Marshall (2016) called them 'super stars' because they have a dominant position on the platform and are able to pull in other potential users to join the community. In this way, platform managers will subsidize those kinds of users to rapidly attract them to the community to enable network effects. This is in line with the thoughts of Van Alstyne (2006) who argued that one of the most important features—when there are pricing decisions to make—is the capacity to develop the cross-network effects and to strongly build loyalty from the side that is attracting the other sides. Armstrong (2006) used a different approach to explain differences within customers: 'multi-homing' users are simultaneously involved in more than one platform for the same service; 'single-homing' users are involved only in one platform. In fact, platforms

manage to increase prices for the multi-homing side and decrease prices for the single-homing side, because platforms compete with each other to appeal to the single-home users (Bakos & Katsamakas, 2008).

Some users value the platform's offer differently. The willingness of each user to pay for the offer is related to the consumer price elasticity. In the traditional one-sided market place, price and cost markup are calculated as a function of demand elasticity and the marginal cost. In the two sided or multi-sided platforms, the pricing on one side will be also based on the demand elasticity of the other sides (Rysman, 2009). According to Weyl (2010), since the other sides are also applying the same elasticity-pricing approach, the pricing on one side depends on the "joint set of demand elasticities and the marginal cost on each side." Indeed, "the low price on one side not only attracts elastic consumers on that side but also leads to higher prices or more involvement on the other side" (Rysman, 2009). In other words, the enlarged value gain from the other side amplifies the effect of having more users on the first side, which can create an even larger price decrease and a quantity increase on this side. In certain situations, price on one side can even be below the marginal cost due to the positive effect created on other sides to attract users. Few models take into account those elasticities to adapt their pricing.

4.3.2 Three different models for charging a specific user

The model from Parker (2006) proposes to subsidize the side where the users are the most price-sensitive and charge the ones that have the strongest increase in demand elasticity on the other sides. A relevant example from Parker (2006) is the PDF format created by Adobe platform. In this platform, there are two groups of users: writers of papers and the readers. The readers are very price-sensitive and would not pay much to have access to the platform. On the other hand, writers have higher price elasticity and want to have access to this huge audience of potential readers; they are willing to pay a certain fee. If the platform managers had decided to charge readers and subsidize writers, it would be much more difficult build a community. Thus,

analysing elasticity of demand and of each side is crucial.

The second model takes into account the user sensitivity to quality. Eisenmann (2006) proposes to charge the users who offer the product's quality rather than to charge the ones who demand it. In this way, the demand side would be first attracted by the high-quality offers from the supply side. Then, the supply would appeal to the large consumer base. For example, game developers on a platform such as AppStore can be pushed to deliver games of high-quality standards, with a royalties charging system on each game sold for platform managers. In fact, paying royalties engaged important fixed costs for game producers, and only the best ones that have achieved big sales volume on the platform will be able to pay for it. This is how certain platforms filter the quality standards. Therefore, the consumers will be attracted to this high-quality game zone. However, to cover the game producers, the platform needs to be sure to reach the critical mass of consumers to guarantee important revenue. This is described as an entry strategy with a pre-commitment investment (Evans & Hagiu, 2006). On the other hand, if the charge was imposed uniquely on consumers, the platform would not have any control on the quality of games, which could rapidly cause negative network effects on consumers.

It is sometimes very complicated to charge the platform users because they absolutely don't have the willingness to pay. Another approach, from Marshall (2016), is to charge third-party actors for access to the platform—these third-party actors do not interact directly with the platform users. An example could be Dribbble company (Parker, 2016). It is a platform-based firm that gathers and connects people from communities of design artists. In this way, artists can expose and show their art to a large community of people who they are specifically interested in. This centralization of interests between users is the core value of the platform. For this reason, platform managers have decided not to charge the users—afraid of negative impacts on network effects. To monetize the platform, managers introduced third-party actors (i.e., companies looking for new talented designers) to pay for platform access. Those companies post-employment listings on the platform. This strategy of monetization forms new interactions that benefit the users who are attracted by exciting jobs. Moreover, users are motivated to post high-quality content so they can be attractive to

potential employers. In this way, the quality value of the platform is automatically improved; the virtuous circle is created. In fact, companies that post employment listings could be considered as advertisers, but it is rather a new source of value creation that benefits all the community. Indeed, charging a third-party agent for access to the platform works if and only if the newly added contents reinforce the sustainable growth of the platform (Marshall et al., 2016). It appears that platform business models consider end-users as 'loss leaders'—because they get the services for free—and business contributors as subsidizers—given that they pay the service providers to reach the community they want (Muzellec et al., 2015). Muzellec (2015) claimed that progressively, in the future, platform network effect monetization plans will be oriented toward B2B (or indirectly B2B&C).

In this B2B charge model orientation, platforms can furnish tools that allow business partners to stand out over the vast number of competitors and appeal to the consumers' attention. In fact, a lot of platforms use this type of charging system through advertising business models. Facebook is an interesting example—where some brands are pushed above other feeds to catch consumers' attention. According to Fish (2009), B2B platforms use consumers' data as a 'private capital.' Indeed, Facebook advertising system targets users based on their personal information and on the geographical position. However, in 2014 and 2015, some polemics were made against Facebook about the limit of research to specific brands except for those who were paying in accordance with the Facebook pricing model. Facebook's issue was this: indirectly performing restrictions to other brands that were not paying. A lot of users were not satisfied with this unequal service. However, thanks to their already-strong community of users, the network went over it. Monetization model is sustainable only when it reinforces network effects (Marshall et al., 2016).

Charging all the users is rare in platform businesses. According to Choudary (2016), platform businesses rarely charge all their users as traditional businesses do. Logically, the reason is that charging all the users dissuade participation and negatively affects the network effects. The famous chicken or egg issue could arise in

32.

those situations because the attraction of each group of users on the platform is difficult.

4.3.3 Chicken or egg

Chicken or egg issue is a central challenge for platform managers. Indeed, in the early stage of the platform, the question is: How can one side of the platform be attracted without having the other side present?

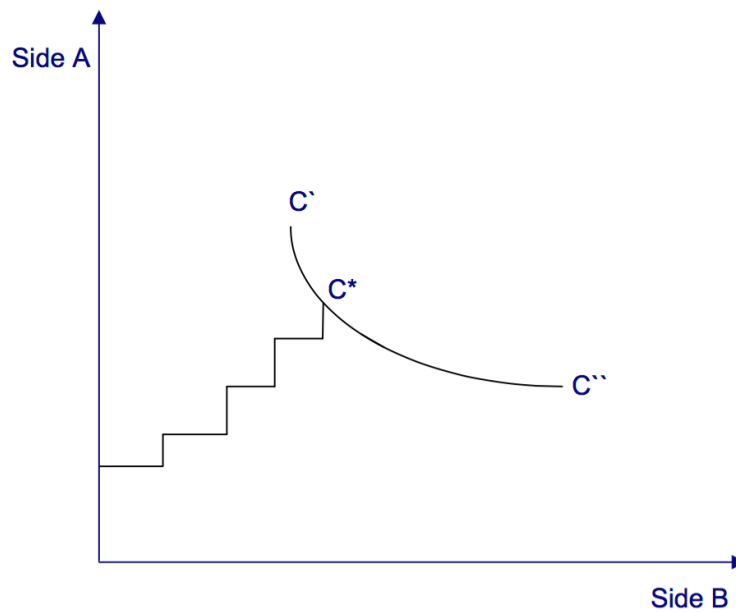
Caillaud (2003), Parker (2005), Rochet (2003), and Rysman (2009), for example, analysed the business model of subsidizing one side and attracting the other side to face the synchronization problem.

a) Strategies for tackling the chicken or egg issue

There are a lot of different strategies developed for tackling this challenge, such as the the 'divide and conquer' strategy from Caillaud and Jullien (2003). The pricing strategy is to subsidize one side (divide) and to catch up the cost with a membership fee on the other side (conquer). In that context, the platform will uniquely make a profit if the indirect network effects between the two sides are strong (Belleflamme & Peitz, 2016). A balance needs to be found in order to set when the willingness of the 'conquer' side to pay will be big enough according to its attraction to the number of the 'divide' side users.

Evans (2008) developed the Zig Zag model to avoid the chicken or egg issue. It involves convincing a small number of people on one side and a small number of people on the other side of the platform—one step at the time. In this way, indirect and direct network effects are built progressively.

Figure 15: Zig Zag Strategy to Reach Critical Mass



Source: Evans (2009)

In the above, C^* represents the critical mass. Step-by-step, a small number of people will attract each other and progressively form bigger communities in order to reach the critical mass. An interesting example is, again, the Diners club platform. The founder convinced only 14 restaurants in Manhattan to join the platform and a few dozens of consumers to accept the Diners club cards. Those numbers were small but already enough to build the first step of the Zig Zag strategy and enable direct and indirect network effects between users of the same local area. This process was referred to as “seed one side at a time to ignite traction” by Reillier (2017). As it worked with small communities by word of mouth, over the next two years, the Diners club platform grew and had over 330 restaurants and 42000 cardholders.

Another interesting main strategy is to focus first on one side of users before opening the access to the other side (Reillier et al., 2017). Evans (2009) calls this a ‘sequential entry’ where different groups of a platform come into the community at different times. Evans (2009) explains that it is exactly the same situation with certain advertising media platforms. The users will first be attracted uniquely by the content

with no advertisements on the platform. Then, progressively, platform managers will decide to add advertising. It comes back to the primarily strong fidelization of one side of a platform. Thus, a sequential entry strategy is elaborated due to the non-positive network effects between the two sides (Clements et al., 2005). Hagiu and Wright (2015) illustrate the sequential strategy with the Amazon case. The company started by offering a vast variety of books online cheaper than the traditional reseller. Indeed, the managers of this platform succeed in implementing the 'negative operating cycle.'¹¹ The fact that customers pay before receiving the product allows the platform to even lower their prices the more. Consequently, a lot of potential buyers were attracted to this platform. Then, Amazon decided to create an online marketplace where individual private sellers were given the possibility of having access to a large buyers' community base and sell their own products. In fact, a large group of individual sellers was immediately attracted to the platform because the vast consumer base was already set (Tiwana, 2013). Thanks to this strategy, the chicken or egg issue was avoided. According to literature, a way to face the chicken or egg problem is to "choose a price structure and not only a price level" (Tirole, 2003).

4.4 Price structures

In this subchapter, I will compare three different main price structures involved in a platform business model, which are the membership fee, the transaction or per usage fee, and the proportional fee. Tirole (2003) and Evans and Schmalensee (2005) stated that membership fee directly impacts the number of customers that will join the platform, and the transaction fee will influence, to a larger extent, the volume of interactions between the users already on the platform.

4.4.1 Membership fee

According to Parker (2016) and Rochet and Tirole (2006), platform managers have three main reasons for using a membership fee:

¹¹ Term used in a the Harvard Business Review article from Andrea Ovans (2015) : What is a business model?

- The platform is having difficulties in charging the interactions. For instance, users can interact once on the platform and then leave without passing through an intermediary for the next interaction. Those direct transactions off the platform threaten the company's capacity to develop the value and the revenue generated. Airbnb is an interesting example; the company struggled with deals off the platform. To avoid this issue, platform managers try to block all the keywords related to contacts' information. In this way, there is a lesser likelihood of users exchanging critical info that could allow them to bypass the platform. Another approach to avoid this issue is to provide an exclusive and highly-valuable service on the platform—i.e., a service which is very hard to get off the platform. For instance, the Upwork company provides services that allow consumers to monitor the freelancer's work remotely on the computer screen (Marshall et al., 2016).
- Membership fee can act as a guarantee to generate cash flow in the early stage of a platform business. It can also catch the value created when the platform is doing large investments by subsidizing one group of users.
- The third main reason is that network effects are also mainly based on the quality of the content on the platform. When the quality becomes a central competitive advantage for a platform, the membership fee pricing structure gets interesting. Indeed, consumers may be disposed to pay for having access to high-quality content. For example, Bsit, a Belgium platform that connects parents and individual babysitters, is related to this topic.¹² Three months ago, managers of this platform decided to propose, as an alternative to a transaction fee, a membership fee because parents wanted to have access to new services only offered via the membership fee. This worked, as the parents had built trust on the platform and were happy to switch to membership fee. The platform provides multiple indicators and rating systems that increase match between users. The success was immediate, and for three months now, usual customers have been progressively shifting from a transaction fee to a membership fee model.

¹² Article du journal le Soir: Alexandre Colleau (2015)

4.4.2 Transaction fee

Charging a transaction fee is a great way of monetizing the value created by the platform without negatively impacting the network effects (Marshall et al., 2016). Indeed, charging only when a transaction happens prevents users from being dissuaded from joining the community except if the transaction fee is too high. In fact, transaction fee influences the user's willingness to interact once on the platform. This type of price structure works for platforms that considerably facilitate transactions between users. It is the case with Fiverr, an online company that instantly connects consumers with freelancers, and the freelancers are paid when each project is fully completed. According to Choudary (2016), this structure is specifically attractive when there is a high intensity of transactions without an important difference in the pricing.

Most analysis on platform pricing is focused on membership fees and transaction fees. However, new researches have been done on the proportional platform fees, which is readily present in the world today (Muthers & Wismer, 2013).

4.4.3 Proportional fee

A famous pricing structure that is commonly used in practice—specifically, in e-market place such as eBay—is the proportional fee or the ad-valorem fees. Platforms that enable transactions between users usually charge sellers ad-valorem fees (Wang & Wright, 2017). They are described as fees that rise proportionally with the transaction price. According to Wang and Wright (2017), this model provides an effective price discrimination process following the Ramsey (1927) pricing theory. In this situation, allowing the platform to discriminate price increases not only its benefit but also the consumer surplus and social welfare because low-demand products are also offered on the platform (Wang & Wright, 2016). Moreover, the introduction of external sellers into platforms—such as freelance programmers in Google Android—has significantly increased the variety of products sold. The majority of those companies use the proportional fee as a pricing structure (Muthers & Wismer, 2013). According to Wang (2017), a small fix fee plus a proportional fee is optimal because when there are a lot of different products involved on the platform, the transaction value is probably relative to the cost value of the product. This is illustrated in the

following figure. As we can observe, ad-valorem fees vary according to the category of the product.

Figure 16: Ad-Valorem fees according to different categories

Amazon				Visa			
DVD	15%	+	\$1.35	Gas Station	0.80%	+	\$0.15
Book	15%	+	\$1.35	Retail Store	0.80%	+	\$0.15
Video Game	15%	+	\$1.35	Restaurant	1.19%	+	\$0.10
Game Console	8%	+	\$1.35	Small Ticket	1.50%	+	\$0.04

Source: Wang and Wright (2017)

External sellers introduced in the platform may know more about the product demand than the platform's managers. In this way, a platform becomes a real place of matchmaking between users and a strong network effects machine (Evans & Schmalensee, 2016). Indeed, platforms act as 'demand coordinators' (Evans, 2003)..

In this chapter, I first enumerated challenges that face the monetization of network effects. I then analysed the sources of value created by the network effects. I also studied the sensitivity of different group of users inside the platform to be subsidized or charged. I also noticed a trend toward charging a third-party user, which is willing to pay to have access to the platform community. Furthermore, the Zig-Zag strategy was explored to develop network effects while avoiding the chicken or egg issue. Finally, I analysed three important monetization strategies involved in platform-based companies, the transaction fee, the membership fee and proportional fee.

5. Sophia Genetics

The purpose of this chapter is to illustrate the theoretical part by referring to a platform-based company in the health sector, called Sophia Genetics. Then, I describe the methods used to perform the qualitative analysis.

5.1 Research questions

As I mentioned in the theoretical part, platform businesses are omnipresent on diverse markets and have revolutionized core business model of specific industries. Over the last 5 years, medicine has also been impacted by this phenomenon—specifically, in the DNA genetic analysis. Genomic data is the central piece of this phenomenon. A huge amount of data is produced every day from next-generation sequencers. More than ever, humans, medical institutions need to collaborate and work together in order to gather the data. However, it seems that technology is moving faster than human knowledge and wisdom. Indeed, the main challenge today is not to improve the technology but to improve human collaboration.

This process of gathering data is crucial to getting useful insights and moving forward in medical precise diagnostics. The platform approach is ideal to face this challenge. I believe that the ability to create communities of people and strong relationships between them through platforms mechanisms could be applied to various main actors in medical institutions. The question now bothers on how to reinforce those connections of users; developing and monetizing network effects are the pillars of the interrogation “How to develop and monetize network effects?” coming from an online SaaS (Service as a Software) platform, taking the concrete case of Sophia Genetics. Moreover, I will use the theoretical concepts that I learned in my literature review to approach the case of Sophia Genetics.

5.2 Presentation of Sophia Genetics company

Sophia Genetics is one of the global leaders of data-driven medicine. In the 1950s, Watson and Crick discovered the DNA structure and function (Watson & Crick, 1953). For many years, this discovery remained at research level due to the difficulty to industrialize the concept and the costs of doing so. In 2007, it was still costing about \$14 million to sequence one human genome; today, it is \$1,000, thanks to the NGS—the Next generation Sequencing.¹³

Sophia Genetics is linked to the industrialization of the DNA sequencing and the NGS discovery. Since 2011, their mission has been to propose solutions for using the NGS DNA sequencing in the real clinical world. The DNA sequencing in the clinical world is done from blood or biopsy and includes 3 phases: the lab preparation with chemical amplifiers, the sequencing itself, and finally, the software part to interpret the huge amount of data coming out of the sequencer. Sophia Genetics has initially focused on this last phase to provide clinicians (oncologists, geneticists, pathologists, biologists, hematologists, cardiologists, etc.) with structured and easy-to-use information on the DNA variants of patients. To do so, Sophia Genetics has developed an advanced clinical genomics software solution that is sold as a service (SaaS) from a cloud-based platform used by over 300 hospitals. The platform (Sophia DDM) analyses genomic data to help hospitals to diagnose patients better and faster across a broad range of illnesses.

Today, Sophia Genetics is involved in the 3 phases of the sequencing process described above. From DNA extraction to data interpretation, Sophia sells 3 different products:

- *DryLab*: Sophia offers the use of the analytical platform
- *Bundle*: Sophia sells the bundle of the chemical sequencing material and the analytical platform
- *IDS*: Sophia offers a full-service solution from biological samples to the sequencing with chemical material and the analytical platform

¹³ Source: US National Institute of Genome Research (2016)

5.3 Qualitative analysis methodology

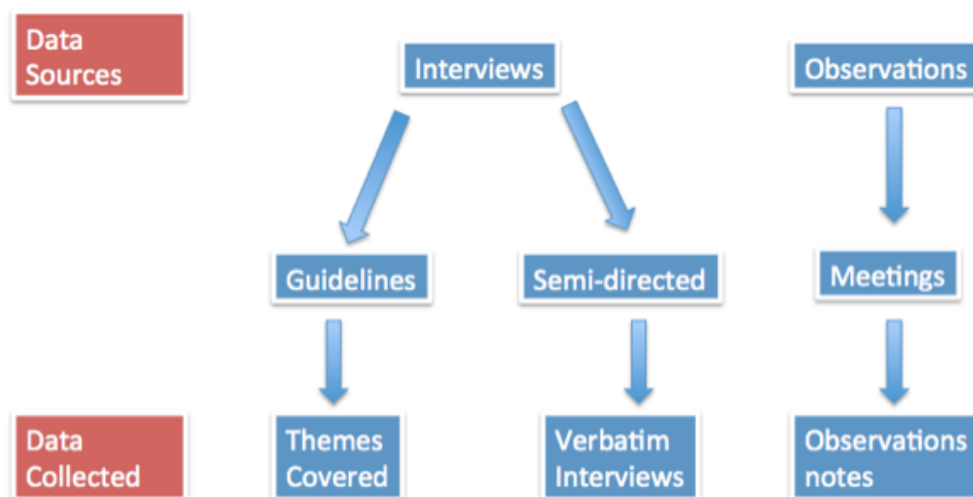
5.3.1 Data Collection

To explore the practical part, I will use a qualitative analysis approach and the method of triangulation through the framework of Koners and Goffin (2007) to collect the data. Triangulation is defined as a combination of methods in the analysis of the same phenomenon (Denzin, 1978). According to Mathison (1988), data triangulation means using numerous sources of data in order to have a specific position about the phenomenon being analyzed. Indeed, 'mixed-methods' or 'multifaceted' approach reinforces the study's comprehension and analysis (Jonsen & Jehn, 2009).

In my case, the data collection will be based on two main sources, passive and active:

- Nine semi-structured interviews (passive)
- Observations during my six-month internship (active, directly involved in the field)

Figure 17: Data Triangulation



Source: Koners and Goffin (2007)

5.3.2 Interviews

I chose the approach of semi-structured interviews (Edwards & Holland, 2013) because, in my case, it was the most interesting way to get insights from my interlocutors. I set the outline of the themes covered through a planned framework. Thus, thanks to more open discussions, new ideas were raised from the interlocutor and, somehow, his answers guided the way the interview was engaged. However, as I mentioned, the interview's main topics were prepared in advance to have the 'interview's skeleton' ready. Furthermore, as I had already done some observations on the field during my internship, I could directly ask relevant questions that led to interesting discussions for the interlocutor and me. The interview guide is in appendix 1.

Figure 18: Interviewees

I₀	Name:	Description:
I₁	Kevin Puylaert	Head of sales outside Europe, Sophia Genetics
I₂	Dr. Shirine Behenda	Product manager at Sophia Genetics
I₃	Dr. Stephane Gerart	Head of the sales in Canada, Sophia Genetics
I₄	Dr. Jurgi Camblong	CEO of Sophia Genetics
I₅	Dr. Jean-Francois Vanbellinchen	Used to be head of the genetic lab at UCL (Catholic University of Louvain), daily user of Sophia Genetics' platform. He is now Subject Matter Expert on the field.
I₆	Dimitri De Boose	CEO of Bsit, online platform directly impacted by network effects
I₇	Kenneth Jones (two interviews)	Daily user of NGS data analytics platforms, used to work at GeneDX, competitor of Sophia Genetics
I₈	Thomas Grivegne	Business analyst director at Bordet Hospital

I chose people with completely different professional backgrounds to enhance the diversity of opinions and to give a better approach of the problem. Unfortunately, I couldn't interview Sophia Genetics' competitors because as I worked there for 6 months, they refused to see me for confidentiality purposes. Therefore, I have decided to analyse Sophia Genetics' platform phenomenon in depth through 3 different axes:

First, I interviewed four people from Sophia Genetics to reach a global approach of the company and to involve a different levels of experiences. Dr. Jurgi Camblong, the CEO-Founder of Sophia Genetics who gave me very interesting thoughts about developing and monetizing network effects. Dr. Shirine Behenda is a product manager who knows, very well, the doctors and geneticists' technical expectations about NGS diagnostic analysis platforms. Kevin Puylaert is a new talent of the company. He traveled all over the world to introduce Sophia Genetics to local markets and has an excellent perception of the NGS needs of medical institutions. Dr. Stephane Gerart has a pertinent commercial and technical approach to NGS genetic analysis and succeeded in convincing first labs in Canada to adopt Sophia's platform.

Secondly, I interviewed two people who use NGS data analytics platforms on a daily basis. Kenneth Jones who has a very good perspective of the platform's value proposition; he used to work at GeneDX, which, indirectly, is a competitor of Sophia Genetics in the US. He has also classified hundreds of variants through the GeneDX's platform. Dr. Jean-Francois Vandellinghen used to be head of the Catholic University of Louvain genetics laboratory, Sophia Genetics' customer. He also worked at IPG (Institut pathologie et génétique), the biggest genetic lab in Belgium, and brought pertinent opinions on the limitations of monetizing the data.

Third, I interviewed two people who were not related to Sophia Genetics but had a more global approach of the platform market. Dimitri De Boose is the CEO of Bsit, described as the 'Airbnb of babysitters.' This platform, which was founded two years ago, is now very successful in Belgium and France. The growth is based on strong direct and indirect network effects. Dimitri gave me interesting insights on how they develop and monetize those effects. I chose this interview because I think it is very interesting to make some parallels between different famous platforms on the market and the Next-Generation sequencing data analytics. Thomas Grivegne is the business analyst director at Bordet medical Institute, which is the biggest cancer centre in Belgium. He is doing the business cost-beneficial analysis about new applications that could be implemented at Bordet. This was relevant to see what he thought about

Sophia Genetics and the main challenges that faced this new technology in Belgium.

To analyse the interview data, my method was to carefully read the transcripts and label all the important pieces called codes. This technique is known as 'primary coding' (Basit, 2003). The aim was to find a conceptualization of patterns by segmenting the data and to create categories (Basit, 2003). Following my research question, I labeled the pieces that were related to 'develop,' 'monetize' and 'limitations' topics. I tried to stay the closest as possible to the transcript's data and gathered all the most important codes to build categories (cf. appendix 2). Afterward, I connected the categories and described how they were related. Then I explained their relevancies and my interpretations. Finally, a comparison was made between the results and the theoretical aspects seen in the literature review.

5.3.3 Observations

The collection of data is from my field notes that I wrote during my internship. I chose to act as participant-observer, in the category of a 'complete participation' (Spradley, 1980), because I was completely integrated into the company and worked as a consultant. I wrote a few pages on specific meetings' main themes and how Sophia Genetics' platform was accepted in hospitals. The observations' data analysis method is similar to that of the interviews, as it involves underlying codes, categories, and concepts (cf. appendix 2). Then, I linked those concepts to the interview data, based on the triangulation approach.

In this chapter, I illustrated the network effects context through a practical case and explained the techniques used to analyse the data. The data were collected through the interviews and my internship. I performed the coding of categories (cf. appendix 2) to aggregate the data in order to get useful insights.

6. Analysis and Results

The aim of this chapter is to analyse all the collected data through the categories built via the coding method. Then, I expose the answers to the questions on how Sophia Genetics develop and monetize network effects.

6.1 Global Data overview

Through observations and interviews coding methods, I analysed three different main chapters: developing network effects, monetizing challenges, and the limitations of the model. Within those chapters, I created categories that gather different codes to better approach the subject (cf. appendix 2). These categories are inter-related and this is explained under the results heading. My interpretation of all this is detailed under the discussion heading. In this chapter 6, I also distinguished two parts, what exists today at Sophia Genetics (The Present) and what could be the future (The Perspectives).

One conclusion came rapidly and unanimously out of the interviews: the future of the company will be around its capacity of generating more and new indirect network effects. This conclusion appears in the development chapter (cf. appendix 2 for full codes description). Indeed, I underlined codes that were oriented toward future visions in green, and in the INE category, almost every code was in green. I will further talk about those outcomes under the discussion heading.

To have an overview of the data coming out of the interviews, I built a tab summarizing the main answers from the five employees of Sophia Genetics. The questions were first related to the development and the monetization of network effects in the present and in the future of the company. Then, I asked questions around the limitations of the model, today and potentially in the future. This tab will be referred to the results of the different chapters. The Figure 19 hereunder represents the overview of data.

Figure 19: Data Overview

	Questions/ Interviewee	I ₁	I ₂	I ₃	I ₄	I ₅
PRESENT	How does Sophia develop network effects?	Platform utilization and standardization	Community	Integration and platform utilization	Community	Community-Integration
	How does Sophia monetize network effects?	Transaction Fee	Monetization of algorithms	Transaction Fee	Transaction Fee and Membership fee	Transaction Fee
	What are the limitations?	Humans	Humans	Humans	Rules & Humans	Rules & Humans
FUTURE	How Sophia will Develop new network effects?	Multi users interactions-treatments optimization match- INE	More interactions-multi users interface with patients involved -INE	Better integration with Labs system	More interactions-marketing tool for pathologists-clinical trials- INE	Better integration-multi users with specialists- INE
	How Sophia will monetize new network effects?	No free model-still charged model	Pharmaceuticals will be charged	No free model-no advertising via pharmaceuticals	Pharmaceuticals will be charged	Specialists charged
	What is the future of DNA platform analysis market ?	Monopoly	Monopoly	Oligopoly	Monopoly	Monopoly

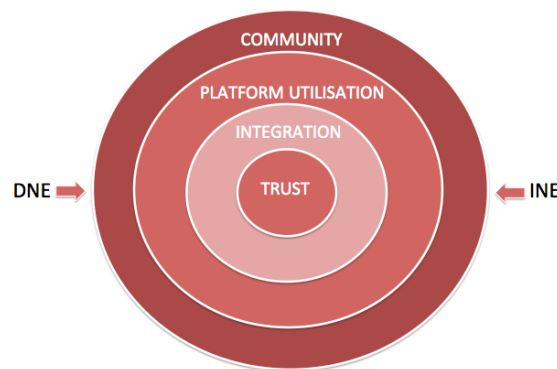
6.2 Present Situation

6.2.1 How does Sophia Genetics develop network effects?

a) Data overview

I have created 3 categories around this topic: Integration, Platform Utilisation, Community. Trust appears centrally in the Figure 20 but it is not a category as such. It's just fundamental to have Trust before creating network effects. Without Trust, there are no platform business model and network effects. Around these 3 categories, I have explored existing direct and indirect network effects and potential new ones in the case of Sophia Genetics.

Figure 20: Development Categories Dynamics



This figure illustrates the categories that are at the origin of DNE (Direct Network Effects) and INE (Indirect Network Effects).

Again, **Trust** between users is critical for a platform to start developing data; without trust, nothing is possible, and the collaborative project cannot be achieved. Trust is not a category as such but it's behind everything.

In terms of categories:

1. **Integration** is probably one the core step for launching network effects. An efficient integration of Sophia Genetics platform in customers' labs is crucial.
2. **Platform Utilisation** means the attractiveness for users for using the platform. If this is achieved, the users will share their experiences and provide feedbacks that will then enable network effects.
3. If the category one and two works, ecosystem of users is progressively created via the platform. It's a **Community**. According to the interviewees, such community was essential in order to create first network effects. Over time, the community becomes itself auto-fuelled by DNE and INE.

DNE and **INE** are the effects created by these categories.

Figure 21: Development of Network effects Keywords

	Community	Platform utilisation	Integration	DNE	INE
Keywords :	Trust	AI-Machine learning	Coordination labs managements	Increase value each new users	Multi-users interaction
	Network orchestrator	Unique Solution	Centralized software	Algorithms improvements	Diversity
	Sharing knowledge	Standardization	Fluid/simple process	Flagging	Leverage effect

Figure 21 outlines keywords for each category and for each network effects. There were parallel statements between interviewees but also different opinions on specific topics such as the flagging incentives, explained in the DNE subchapter.

b) The Community of Sophia Genetics platform

I will start, first, by talking about the community of Sophia because it is the ecosystem that is directly related to the network effects. For few interviewees, it is mainly the community itself that develops network effect. Sophia Genetics gathers an important community of worldwide experts on an online platform. The company is a 'network orchestrator' that enjoys group effects, says Kevin Puylaert. Indeed, the first value of Sophia Genetics is the community, according to few interviewees. This community relies on the economy of contribution where each new user on the platform increases the global benefit of all other users. This phenomenon, in the case of Sophia, is the Direct Network Effect, which will be explained in details later. Dr. Jurgi Camblong claims, "creating a community requires a lot of creativity and determination."

Trust between all the stakeholders involved in this platform is obviously critical.

This is, as said before, a must to develop network effects. The values required for ensuring trust are **transparency, honesty, and ethics**. For Thomas Grivegne, “the most important is too establish trustful relationships”.

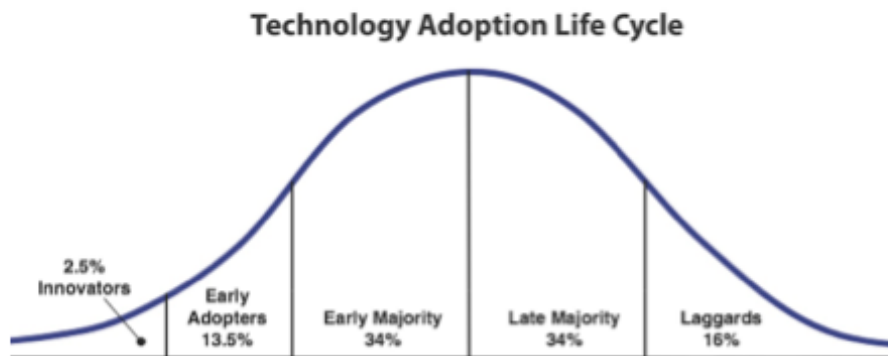
Cedars Sinai, one of our biggest prospects in California, was impressed by the European community of users. The director of the lab, during the meeting, stated that it was a real argument to have a stable and reliable community behind the platform because it guarantees a great quality of services, given that users are not only growing in numbers, but they are also staying on the platform.

Contrary to what could be believed, the community attracts small and big labs but for different reasons. Small labs with less expertise are attracted by the community to learn from the big labs. The Drexel Medical University meeting in April is a relevant example. It is a small university with not a lot of expertise in pathology. They were highly interested in the community aspect because it was a gain of time to have a structured list of variants¹⁴ that responds to their specific patient profile based on the knowledge of other large institutions. In addition, the world of NGS is constantly evolving, and it is very difficult to keep up with the changes, and even more when you are small. For example, the panel of genes being analyzed are becoming larger, reaching 40 to 50 genes in the panel. This is due to the cancer research where interactions and links between genes are being better understood. But the larger the panel becomes, the more difficult it becomes to correctly identify all the mutations. In fact, the community becomes even more essential with those large panels because very few institutions have enough experience to ensure precise diagnostics, and therefore, “they require the expertise of other experts via the platform”, argues Dr. Shirine Behenda.

Interviewees explained the technology adoption curve theory (Rogers, 2010) in order to better understand the community attraction to small and big labs.

¹⁴ Variants are gene mutations that can cause diseases.

Figure 22: Technology Adoption Curve



Source: Rogers (2010)

According to interviewees, its technology is today in the live cycle of **early adopters**. The company claims that the first users (Innovators) were the hardest to convince because there was no community. The second wave of users (the early adopters) were attracted by the services, but also, they started seeing the benefits of the community and wanted to contribute. The third wave of users (the early majority coming now) is considered as, to a large extent, the big labs. These ones don't necessarily look at Sophia Genetics for its solution; they already indeed have important expertise in NGS diagnostic analysis. They only look at the community. They use the platform to see what the others are doing—to ensure they don't miss anything but also to contribute. They all play the game.

An interesting illustration of the above is France. Sophia started gradually entering the market by convincing few very small labs to work on the platform. Progressively, “medium-sized labs were attracted by the services and the community,” argues Dr. Shirine Behenda. Finally, the biggest hospital in France started as well to use the platform because they saw that the community was becoming large, and they just could not afford to be out of it. It is a step-by-step strategy.

Local communities are usually easier to form because users are closer geographically and culturally. Members can interact directly and create links and

partnership. In the case of Sophia Genetics, some consortiums were created locally, which are small communities of institutions that only want to contribute and interact with local partners. Bsit Company is an other relevant example where managers have decided to focus on local geographical areas to progressively create small communities of baby sitters and parents. “The networks we created were first local” says Dimitri De Boose.

On the other hand, international communities are harder to build, given, for example, in the genomic area, that medical institutions have different annotations, standards, and mind-sets. The community can really attract new users or can drive them away. I attended a call at Milwaukee Hospital, where heads of the lab were afraid of the community because they didn’t trust other experts’ expertise on the platform. They claimed they were the best in their specialty, and it was difficult for them to believe that other experts could help them.

Sophia Genetics, to a large extent, solved the issue of the internationalization of its platform by starting with local communities that were ahead of the genomic domain; their first territories of prospection were Switzerland, France, Benelux, UK, Spain, Italy, and Germany. These developed countries were culturally different but used to work together via the European Community concept, promoting lots of interactions in the medical area. “Once this European Community was formed, it was relatively simple to extend to other areas, such as Eastern Europe, Latin America, Canada and Asia Pacific” claims Kevin Puylaert.

c) Platform utilization

For certain interviewees, Network effects are mainly generated due to the platform utilization and the customer’s experience. The platform users enjoy a combination of machine learning and artificial intelligence technologies, which is alimented by a huge amount of genomic data. Each user is uploading their genomic data on the platform to receive the gene mutations analysis of the sample a few hours later. The interface is simple and fluid, which puts the customer experience as a priority. “We want the

platform to be as friendly as possible, to optimize the user experience” argues Dr. Stephane Gerart. The platform uses its own database of 110,000 patients treated so far. The platform proposes as a feature, links to few other public databases on the interface. Sophia Genetics needs also external databases to rely on in order to enhance diagnostics precisions and the diversification of pool of samples. For Kenneth Jones, “using several external databases is crucial”. An important characteristic for Sophia Genetics platform is that it relies uniquely on **clinical** samples with a short turnaround time to get results. This is a strong argument to attract new users because it is a guarantee of quality. The fact of missing variants¹⁵ is extremely serious in the clinical world compared to research because patients’ lives are at stake. However, on the field, some medical institutions such as Virginia Commonwealth University were afraid about the words AI (Artificial Intelligence) and machine learning. Labs directors were concerned about how Artificial Intelligence could treat a huge amount of sensitive genomic data.

The downstream and the upstream analyses are reunited within the same workflow, which is a big advantage for users. In other words, “different processes are gathered on the same workflow to facilitate the user’s experience”, explains Dr. Stephane Gerart. Sophia Genetics constantly needs to think how to optimize the users’ benefit to use the platform. Moreover, there are monthly updates of databases on the platform that guarantees the good management of data. This is also a strong statement for prospects that want to be connected to an updated platform. However, Broad Institute in Boston liked the platform interface but not the black box side of it. The bio-informaticians wanted to have more transparency on how the platform works—specifically, more information on the technology behind it. In fact, they couldn’t trust a platform if they don’t understand its core mechanisms. On the other hand, University of North Carolina was satisfied with the tests we have done with them because the platform didn’t miss any variants. They also loved the interface design. These were strong advantages for them to be furthermore interested in the Sophia Genetics solution.

¹⁵ Variants are gene mutations that can cause diseases that are classify as pathogenic or they can also be classify as benign

The data standardization was also crucial to build a platform of hundreds of hospitals. This standardization keeps evolving, thanks to the multiple users' feedback. Numerous interactions between hospitals and Sophia Genetics are taking place all the time, and these contribute to the optimization of the platform and the services it's providing, claims Kevin Puylaert. In fact, for example, the genomic data generated from Thailand is standardized the same way than the data generated from Chile. The standardization of genomic worldwide data was the key to launch network effects, because it enabled to standardize interactions on gene mutations and made greater the user's experience on the platform.

d) Integration

The integration of an efficient platform in customers' labs is crucial to enhance network effects and convince new customers. For Kenneth Jones " lab efficient integration system is the most important to attract new users ". Indeed, a simple and fluid integration process is necessary to connect labs to each other and to the platform. Currently, Sophia Genetics integration is quite simple, as the software is locally downloaded on labs' computers. According to Dr. Stephane Gerart, this simplicity of integration is critical and pushed new users to buy the product. However, sometimes " a customer needs to pass through excel files to transfer information from the platform to the lab's system and to the patients" claims Dr. Jean-Francois. A step is missing; not all the process is automatized. This issue will be discussed in the future perspectives.

We saw how the community, the platform utilization, and the integration develop network effects. Let's explore now what are exactly those effects in the case of Sophia Genetics.

e) Direct network effects (DNE)

What are the direct network effects in the platform? The more the users on the

54.

platform thanks to trust, integration, community and platform utilisation, the more the global value offered to customers increases. In fact, algorithms are getting better with the number of users, which enhances the gene mutation analysis in order to provide better medical diagnostics for patients.

Figure 23: Direct Network Effects at Sophia Genetics

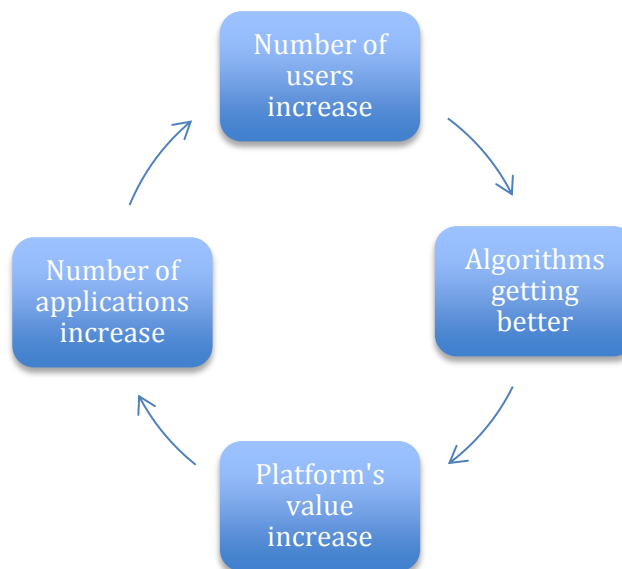


Figure 23 summarizes and illustrates the interviews' data for direct network effect. The more precise the algorithms, the more it benefits the whole community. At Sophia Genetics, the main DNE is the flagging system feature offered by the platform. Flagging is a feature on the interface that allows each user to classify the variant level of pathogenicity¹⁶ according to their own thoughts. Each other user can see, on their interface, the number of institutions that classify the variant pathogenic or benign. However, users cannot see the name of the institutions because it is too sensitive. This flagging system is extremely important to improve the classification of variants. As I mentioned before, diversity of flagging is also crucial; experts need to work in a complementary way in order to flag diverse type of variants and not the same.

How to push users to flag variants? It is, in fact, a relevant question. For now,

¹⁶ Variants can be classify on 4 different levels: Pathogenic when it cause diseases, likely pathogenic category when it will probably cause diseases, likely benign when it is probably non-maligant and finally benign category, where the variant is not dangerous.

according to a few interviewees, users flagged principally for their own interest and not in a community interest. A meeting I had at Cleveland Clinic, one of the best hospitals in the world, according to an international ranking,¹⁷ confirmed this point of view. The lab director was interested in the flagging system but in order to better organize and structure all their data in-house and not to share that information with the community. It is a gain of time because if the user flags the variant once, the next time the variant will appear, it will automatically be classified as the first flag.

The system comforts new users—specifically, small labs with little experience that want to learn and rely on the flags of other institutions. According to Dr. Stephane Gerart, “small labs are happy to have the support of the community through the flagging system”. However, some labs don’t flag because they don’t want to share their research on a certain variant or because they are too scared to misclassify and badly influence others. For Dr. Jean-Francois, “rather to flag a complex variant, customers prefer to not take risks and do not flag” and for Kenneth Jones “It is always the easiest variant that is shared”.

Another direct network effect is the specific service called custom panels¹⁸ that customers request. This effect expands the number of applications offered and benefits other users that could also be interested. For Dr. Jurgi Camblong, “it is a way of leveraging the platform position to validate an application with new users so that this application is then available to any other users on the platform”. The overall satisfaction of the community increases. For example, Sophia Genetics offers hereditary gene mutations¹⁹ analysis and somatic variants analysis,²⁰ so the company covers a lot of different applications in two different domains. The University of North Carolina works with different companies for those applications. The fact that Sophia

¹⁷ Official Hospital ranking from Webometrics

¹⁸ Custom panel: Specific group of genes that does not exist on the market and need to be build specifically for the customer.

¹⁹ Gene mutations that are inherited from your mother, father or both. It can cause rare diseases and cancer.

²⁰ DNA alteration or gene mutation that occurs after conception. It can cause cancer and other types of diseases.

Genetics reunited different applications within the same process was really interesting for them. It also facilitates the product's adoption toward the hospital to use the same centralized software.

f) Indirect network effects

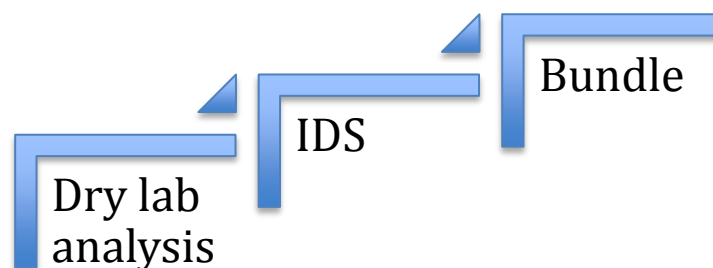
As we saw in the literature review, INE are effects created between different types of users. Interesting community effects similar to indirect network effects according to interviewees happen at Sophia Genetics.

The company was created in 2011, and its first core focus strategy was to provide a high-quality software to analyze NGS genomic data called dry-lab analysis. As we now know, the customer-base grew fast because the service responded well to a global need in hospitals. Sophia Genetics targeted hospitals and labs that had next-generation sequencers to upload the data on the platform. However, two years later, after the launch of the software, the management team decided to attract a new group of potential customers. This new segment of prospects were worldwide labs that don't have sequencers but still want to analyze the gene mutations of their patients. Sophia Genetics created a new community of customer by allowing labs that didn't have a sequencer to send patients' samples (blood, biopsy, etc.) to the closest Sophia Genetics' partner lab that has a sequencer. It is a "win-win situation" says Dr. Jurgi Camblong, where labs having a sequencer that were already a customer of Sophia Genetics for the dry-lab analysis can make more revenue by testing additional tests from other labs. This service called IDS (Integrated diagnostic solution) is working very well and helps labs in Africa, South America, or Asia, for example, that don't have access to next generation sequencers to still be able to have cutting edge gene mutations diagnostic analysis. For Dr. Stephane Gerart, "this offer allows us to have access to a new type of customers". Thus, on the platform, there are institutions having a sequencer and institutions that don't have a sequencer. This is possible because "Sophia Genetics focused first on the dry lab" and progressively built a strong expertise through the community, argues Dr. Jurgi Camblong. Labs that didn't have a sequencer were attracted by this expertise and started to send their samples to Sophia Genetics partner labs—a leverage effect has been created. This would be more complicated if

Sophia Genetics started to target both types of customers because the expertise was not built yet to attract the labs that don't have sequencers. According to interviewees, INE were created because two groups of users on the platform are reunited and attracted to each other.

Sophia Genetics developed a second leverage effect with the bundle solution's approach two years ago. The company is selling the dry lab analysis combined with the chemistry kit²¹ to optimize the analytic solution. In fact, the company realized that they could not just optimize the dry lab analysis but also the chemistry kit that use labs to have the best results. This shift to kits was only possible because they had a dominant position in the dry lab analysis and could use the same customer base for both products. Sophia Genetics became a 'marketplace' says Dr. Jurgi Camblong and succeeded in attracting a new type of customers, thanks to this approach. The Dartmouth Medical Centre meeting is a relevant example because the management team was uniquely interested in our kits' panels of genes with the optimized library preparation²² and not in the dry lab analysis. This allowed us to open a door by first providing a kit, and hopefully, in the future, Sophia Genetics will progressively implement the platform in this institution.

Figure 24: Leverage effects at Sophia Genetics



²¹ Kit: all the chemistry materials required to sequence a specific panel

²² Library preparation: steps required in the lab before sequencing

6.2.2 How does Sophia Genetics monetize network effects?

a) Data overview

Today, Sophia Genetics is monetizing network effects principally through the transaction fee. The transaction fee is the amount that the customer is charged for uploading their samples on the platform. Indirectly, Sophia Genetics has also implemented a membership fee that will be described hereafter. The Figure 25 outlines the keywords for each category built in this chapter.

Figure 25: Monetization Keywords

	Transaction Fee	Membership Fee	Monetize network effects	Other Business models
Key words	Per sample	Validation program	Algorithms	Free model

Figure 26: Monetization Strategy today

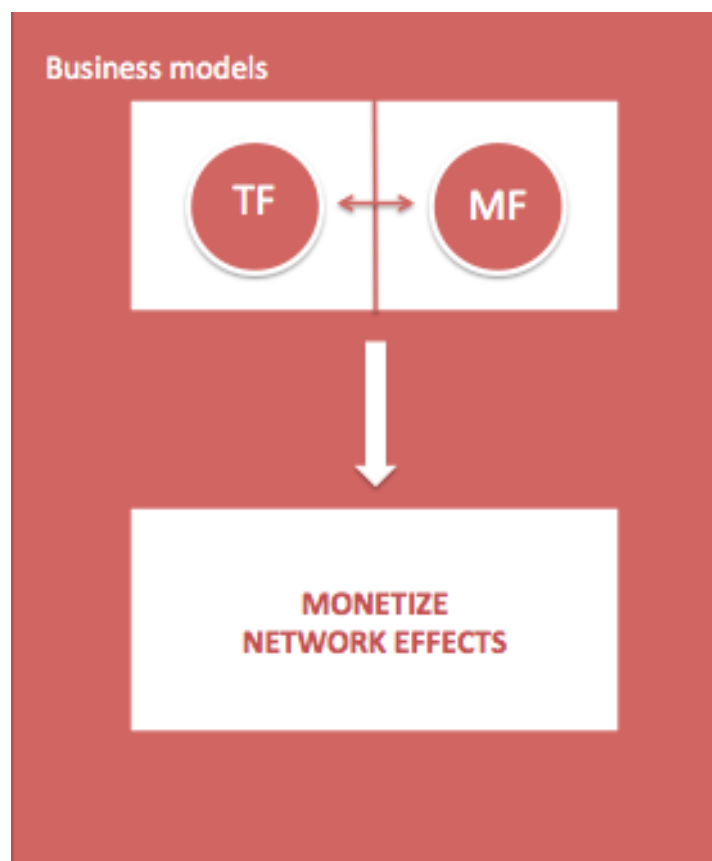


Figure 26 summarizes and illustrates the data from the interviewee. All the codes related to monetization from the interviewees are in appendix 2. TF is the transaction fee model and MF is the membership fee model. In the case of Sophia Genetics, both models are related to the monetization of network effects.

b) Transaction fee

As I mentioned, Sophia Genetics charges customers per sample uploaded on the platform. It is a flexible model because the price is adaptive to the volume of samples uploaded. Thus, “it does not penalize labs that generate small amounts of tests” states Kenneth Jones. Customers enjoy the no-commitment approach of the model and the independency it offers. If one decides to leave the platform because they are not satisfied, they can without any extra-charge. For Kevin Puylaert, “ transaction business fee works well because it proves our high-quality services”. There are no economic barriers, pushing a large number of new users to try and adopt the platform worldwide. Dr. Stephane Gerart claims that it provides also a transparency in the costs structure. Indeed, costs origins are clearer with this type of model. An important reason why Sophia Genetics chose this model fee, according to Dr. Jurgi Camblong, is that global volume the next-generation sequencing tests will increase significantly in the future. That means that the company’s customer-base of 320 institutions will progressively increase the number of tests they generate each year. In this way, the transaction fee will yield more revenue in a long-term perspective. The meeting held at the University of North Carolina Medical Centre was a surprise meeting. It is a very famous hospital in the US, with an important infrastructure, but the lab generates small volumes of next-generation sequencing tests. The transaction fee model attracted them because there were no commitments, and their small volume was not penalized compared to other Sophia Genetics’ competitors that propose an annual license fee such as a membership fee. Therefore, the model works and is appealing to new users, reinforcing network effects.

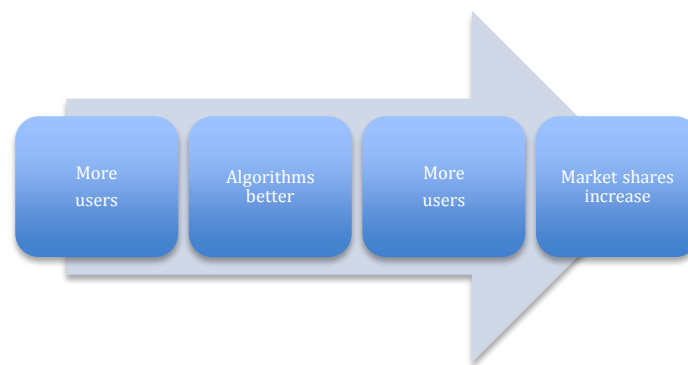
c) Membership fee

Another famous type of fee model in platform businesses is the membership fee. Sophia Genetics does not directly apply a membership fee. However, indirectly, according to Dr. Jurgi Camblong, the validation program provided to each customer is a sort of membership fee. This program is a prerequisite performed by each new customer who wants to join the community. It aims at establishing an analytical performance report as a quality test. The customer needs to pay for this service, but it is not too expensive because Sophia Genetics doesn't want to put economic barriers on the entry.

The membership model, in general, is useful when companies, in their early stage of evolution, need intensive capital. In a short-term vision, membership model will provide important cash flows as customers sign up for monthly or yearly subscription services. Furthermore, the price is usually the same for all the customers. So, in the case of NGS data analysis platform, the membership model fee would be interesting "for big labs with important volumes of tests each year" claims Kenneth Jones. Dimitri De Boose explained that they implemented a membership fee for parents 3 months ago. "It is a big success" states Dimitri and thanks to this, they were able to move to a B2B model by selling membership fees to large companies that want to pay for their employee babysitting services. According to Dimitri De Boose, it was a great strategy to develop and monetize a new network effect. However, for Sophia Genetics, a high membership fee is really not the vision of the company, according to a few interviewees. The vision is rather to attract the most customers as possible worldwide in order to create a powerful community and "cement the company position in hospitals" argues Dr. Jurgi Camblong.

d) Monetize network effects

In fact, revenue generated by the transaction fee and the membership fee is also related to the algorithms improvements. "The monetization is also based on the on the algorithms" claims Dr. Shirine Behenda. The more the users, the better the algorithms become.

Figure 27: Monetization of Algorithms

Thanks to the leverage effect illustrated in Figure 27, using network effects, Sophia Genetics gains important market shares. The company will generate more revenue because it has conquered new market shares. Thus, a monetization has been done on direct network effects. That being said, “Sophia Genetics does not monetize the genomic data that is stock on the platform” says Dr. Jurgi Camblong. International and European laws are very stringent about it. In fact, Dr. Shirine Behenda claims that “Sophia Genetics is not the owner of the data, it is the hospital’s data”.

6.2.3 Limitations

As I already mentioned, genomic data is very sensitive and delicate. Few limitations appeared, in the past few years, on the topic and Sophia Genetics needs to face them. According to interviews, there are legal limitations as well as human mind-set challenges.

Figure 28: Two Main Limitations

a) Rules Limitation

Rules limitations are the legal perspectives that directly impact genomic data. For example, there is a new European rule called GDPR²³ (private policy rules) claims Dr. Jurgi Camblong, which was created last year. It states clearly that genomic data cannot be sold to a third-party under any circumstances in order to protect human privacy. This guideline described two roles: the data controller and the data processor. The data controller, in our case, is the hospital. It has the responsibility to verify that the data is adequately used, respecting specific criteria. In this workflow, Sophia Genetics is the data processor. It means that Sophia can only use the data that respect the rules imposed by the hospital. In general, Sophia Genetics can therefore only use this data in order to continue improving the capacity to diagnostic patients and save lives—but the data could not be exploited further by selling them.

For Dr. Shirine Behenda, “the complexity is that each hospital has their own standards and each country has their own regulations concerning genomic data sharing”. In certain meetings, for example, the hospital said it was forbidden to perform the flagging system because they could not share their work following the institution’s rules. In France, the genomic data cannot go out of the territory. For this reason, “Sophia Genetics implemented the servers in a datacentre in Paris” states Dr. Shirine Behenda.

According to a few employees, this is not simpler than it is on other platforms such as Google. The genomic data is much more complex to manage. Monetizing the platform by charging pharmaceutical companies an access fee will be delicate. Dr. Stephane Gerart says, “ I will find it bad to put forward a type of drug of a certain pharmaceutical brand under the pretext that they have paid”. In any case, all data must be anonymous, and according to law, confidentiality needs to be respected.

In this respect, security access and exchanges were also crucial for Sophia Genetics and the company in this area obtained several patterns. This was one of the key development criteria imposed by the company right from the beginning. Indeed, as we

²³ General data protection regulation

said, trust was a critical factor to build the community, “there was no way that Sophia Genetics could be exposed to hacking or fraud ”states Dr. Jurgi Camblong. The exposure of this in terms of law and reputation was just too high not to attribute a lot of the company’s investment just in this sector.

b) Human Challenges

Data sharing is crucial to medical developments but still remains very sensitive. Medical institutions are in competition with each other—specifically, in the US. “Genomic is considered as a new promising domain, having an important value”, argues Kenneth Jones. Medical institutions are aware that these data are valuable, which compromises data sharing. In some situations, they will keep this value in-house for their own interest.

“Medical institutions tend to fall into competition”, says Thomas Grivegne. However, partnerships are crucial to build collaborative projects. That being said, “working together toward data sharing demands a lot of work for each institution”, claims Dr. Jean-Francois. For instance, flagging specific variants on the platform does not have the same signification for each hospital. They use different criteria to flag variants. Kenneth Jones suggests opting for international guidelines such as ACMG²⁴ to standardize classification of variants. On the other hand, as I already mentioned, some institutions are too scared to misclassify and are therefore not active enough on the platform.

Some ethical issues are also raised on this topic. What can Sophia Genetics do with this data? What image does the company want to portray? For instance, “ if Sophia is ready to pay institutions to flag, it means that the company could be doing a lot of revenue on the work that hospitals achieved”, states Dr. Jean-Francois. For Dr. Shirine Behenda, “institutions are scared that large companies such as insurances will

²⁴ American college of medical genomics

64.

commercially exploit their data". Dr. Shirine Behenda gave an example a few years ago in France where few labs worked together to build a database. The database was managed by a public company owned by the state. However, the company decided to sell this database to a large industrial group at a price they could not refuse. The labs were not aware of this transaction and were very angry, even though they gained royalties. Sophia Genetics' priority one is to always establish trust between users to gather institutions. If trust is broken, "Sophia Genetics loses its community", says Kevin Pulaert.

6.3 Future perspectives

6.3.1 How will Sophia Genetics develop new network effects?

a) Community

One of the last important community to reach is the US market. According to Kevin Puylaert, "the US market is clearly a big challenge". This market in the genomic area is as big as all Europe. In the medical sector, the US is often at the head of the game and they would tend to question European initiatives. Sophia will have to be excellent to gain an US community. Clearly, it's a benefit to demonstrate that there is an European and even global community behind the platform, but it is just not enough. They will have to apply the European strategy, starting with the Innovators, then the small to medium size labs and finally the large labs that will not be able to avoid joining the community. Will the US and European community interact? Possibly but to be seen, answered Thomas Grivegnée.

Sophia Genetics community also needs to work in a complementary way to ensure a wider database with diversified samples. In fact, today, the majority of experts work very much on the most-known gene mutation cases (e.g., breast cancer). In the future, to reinforce Sophia Genetics community, the company will need to select complementary experts that will work more evenly on the 20,000 human genes. This complimentary effect will benefit all the users in an exponential way and will increase the global value of the platform. For Dr. Shirine Behenda, "If all the community gathers this diverse information, each user will have access to the other 19,999 genes"

b) Integration

Labs have systems called LIMS²⁵ that coordinate and gather all the patients' data into one system. For now, Sophia Genetics does not work with this system. Dr. Jean-Francois argues, "it would be a huge plus for Sophia Genetics if the company succeeds to integrate LIMS...some customers want specifically a system like this to synchronize their data with Sophia's platform". Moreover, "if the data is synchronized, this will push oncologists or other actors to fulfil more patients' information on the platform" states Dr. Shirine Behenda. I faced a situation in March with James Cancer Centre Hospital that works with EPIC²⁶ system to centralize all the data. The head of the lab said that Sophia Genetics requires passing at least through EPIC if it wants to be implemented in the institution.

c) Direct network effects

Trust needs to be reinforced in the flagging system to push users to flag and to contribute to the system. According to a daily user of an NGS platform, Kenneth Jones, "more information could be added to the interface on the reasons for other institutions to flag this specific variant pathogenic or benign". This could push the customer to better understand the situation and to create more matches between users. Some interviewees claim that it could be a good idea to reward institutions by offering a discount for next samples uploaded in function of the number of variants flagged. That being said, other interviewees think it is a really bad incentive to pay for flagging. Institutions would be pushed to flag and this could threaten the flagging quality. Flags classification needs to be faultless; it is better to have 1 or 2 correct flags than 50 misclassified flags. Another strategy needs to be formed according to Dr. Jurgi Camblong, to bypass financial incentives and push institutions to flag 'instinctively.' This strategy will be discussed in the next indirect network effect chapter.

²⁵ LIMS: Laboratory information management system

²⁶ EPIC: The most famous laboratory information system in the US

d) Indirect network effects

Today Sophia Genetics has only one type of direct users on the platform, which are the data analysts, but doctors don't use the platform directly. According to Kenneth Jones, "Doctors don't have time to use the platform, they just need information on specific variants to treat their patient". The company wants, in the future, to expose the platform to more types of users such as specialists in oncology or hematology who are in contact with the patient. More types of users on the platforms imply more interactions. Attraction between users will create significant value for the whole community.

Figure 29: Future Indirect Network Effects at Sophia Genetics

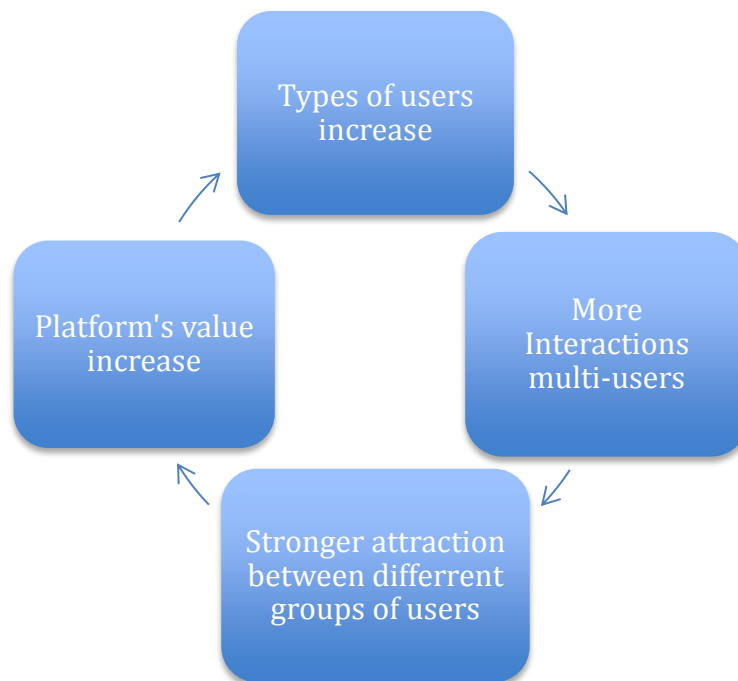


Figure 29 summarizes and illustrates the interview data for indirect network effects. Sophia Genetics wants to attract specialists step-by-step to enhance leverage effects. The platform would become a multi-user place of oncologists, pathologists, biologists, hematologists, and even the patients themselves could have access to the platform. For Dr. Shirine Behenda, "different types of interfaces would need to be implemented on the platform, one interface for each type of user". The aim is to regroup many health actors with different experiences on the same platform to interact in order to share knowledge and to centralize information. How will those

different groups of users be attracted to each other? Why do they want to have access to the platform? Sophia Genetics needs to find a way to gather those people.

The biggest value created through indirect network effects is the gathering of data and information from different groups of users. In the future, clinicians or oncologists who are directly in contact with patients need to be integrated on the platform. Those specialists have information on the treatment outcomes, the phenotypes, and other crucial information. If information of each customer could be progressively regrouped on the platform, it would significantly enhance diagnostics precision.

However, clinicians don't have the time or patience to fulfil the data. Too much paperwork, they say. The platform could be an ideal solution to automatize and make easier the entire data collection process by creating a worldwide portal. Unfortunately, for now, oncologists don't have enough output to upload data on the platform. Sophia Genetics has the duty to create incentives. Dr. Jurgi Camblong gives one example involving indirect network effects to push the oncologists to upload the data. Starting from the fact that pathologists are always seeking for more tests in their labs, Sophia Genetics could encourage them, using the Sophia's platform as a marketing tool to attract oncology tests. The oncologists don't want to do too much paperwork; they want a fast process. The platform could be an effective bridge between the two groups of users. It is a win-win situation, the pathologists are doing more tests and the oncologists are satisfied as well because they can more easily enter information and efficiently order tests. Sophia Genetics is, of course, also very happy because more tests are elaborated on the platform. In this way, oncologists upload the data because it is a real need. There are no financial incentives that could have a negative impact. According to Dr. Jurgi Camblong, it is a 'natural' incentive created by the platform that could generate a huge value for the entire community.

Sophia Genetics could also create a new leverage effect using indirect network effects between oncologist and pharmaceuticals companies by matching treatments. Indeed, today the platform treats around 3000 patients a month worldwide with a

solid tumour. On the other side, Pharma companies are having difficulties in finding patients with a specific molecular profile that match their clinical trials. According to Dr. Shirine Behenda, “ competition between pharmaceutical companies is intense...it takes few months to one year to recruit the rights patients... a patient who has a stage 3 cancer, don’t have time to wait”. Sophia Genetics could implement a partnership with pharma companies to help them in the recruitment. For example, if one pharma company presents a new clinical trial that requires a specific molecular profile, Sophia Genetics could help and say, for example, “based on the platform database, there are five patients in London, two in Paris, and one in Lisbon that match your research” says Dr. Jurgi Camblong. This would be a huge gain of time for the pharmaceuticals company. Sophia Genetics would ask the authorization to hospitals to share the information to pharma companies. Dr. Jurgi Camblong claims again that it is a “win-win situation”, because hospitals would be happy that their patients are registered and identified to have to the latest tests available; this will increase the reputation of the institution. Sophia Genetics would become the bridge between hospitals and pharma companies through the platform.

Figure 30: Leverage Effects between Oncologists and Pharma Companies

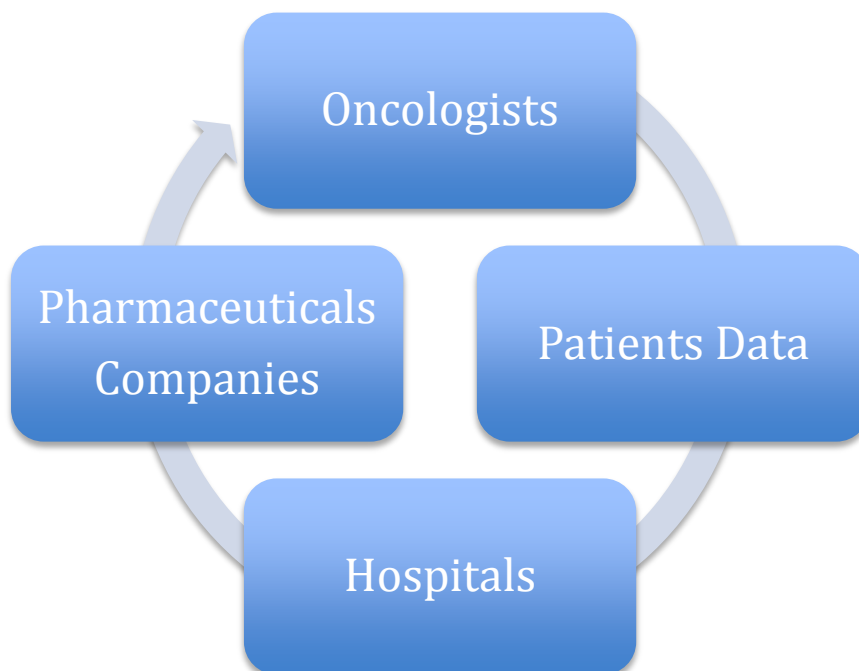


Figure 30 summarizes and illustrates the leverage effects of a potential partnership with pharmaceutical companies. In exchange for this service, to guarantee the match, Sophia Genetics would ask pharmaceutical companies to update Sophia's platform databases with the newest drugs approved on the markets. In this way, the platform will be reinforced with the most complete and up-to-date databases. In fact, large amounts of drugs are released each year, and this makes the task of selecting the right drug harder and harder for oncologists. Oncologists who search for specific drugs that match their patients' molecular profiles would be very satisfied. It will, according to Dr. Stephane Gerart, save a great amount of time and money for the oncologists to have access to an efficient drugs match portal, called Oncoportal". This will also push Oncologists to upload the data from their patients on the platform. Therefore, indirect network effects are created between a multi-user interface including oncologists, hospitals, and pharma companies. This ecosystem formed could be a powerful asset for Sophia Genetics.

6.3.2 How will Sophia Genetics monetize network effects?

In the future, another strategy for Sophia Genetics would be to charge pharmaceutical companies to have access to specific patients. Indeed, as we saw in the last subchapter, Sophia Genetics could match the treatment for clinical trials between patients and pharma companies. In fact, Dr. Jurgi Camblong explains that the Oncoportal's access could be free for the oncologists and charged for pharmaceuticals companies. It would be a valuable service offered to pharma companies and an important source of revenue for Sophia Genetics, according to few interviewees. Furthermore, for Dr. Jean-Francois, "Sophia Genetics could impose a license fee to specific groups of people such as haematologists who want to have access to the platform". The powerful Oncoportal enhanced by oncologists' patients' information could greatly interest hematologists or other specialists. In this way, Sophia Genetics will monetize indirect network effects created by a multi-user interface. According to Dr. Jurgi Camblong, "the most important step in Sophia Genetics' evolution is to create first a solid community and then a monetization of network effects". On the other hand,

if the company would have first monetized and then created a community, it would have definitely been a disaster, claims Dr. Jurgi Camblong.

We have seen a trend in platform companies, over the past few years, toward free business models. Could this phenomenon be efficient in NGS data analysis platforms? Some interviewees claim that it could not be possible. For Dr. Stephane Gerart, “the free model is very complicated”. Free platforms don’t work because the algorithms are too basic and too standardized. In this model, sometimes one algorithm is used for all the different applications on the platform; consequently, the service is not of a good quality. “Charging a price guarantee a certain quality in the data analytics,” argues Dr. Stephane Gerart. On the other hand, some interviewees say that Sophia Genetics could be subsidized by states and in exchange give access free to all the users. Dr. Jean-Francois states, “ people might trust more Sophia Genetics as a public company”. This could enhance the gathering of data and have a significant positive impact in medical diagnostics. Another idea from by Thomas Grivegne would be to charge users who are doing lucrative activities and give in exchange the access free for academics’ users for example. Dr. Jurgi Camblong and Dr. Stephane Gerart are against this idea because it would be too expensive to produce a custom panel²⁷ for academics and the test volumes in this sector are small compared to the clinical institutions. Therefore, the model is not viable for Sophia Genetics.

6.3.3 The future limitation of NGS data analysis market

For the majority of the interviewees, the NGS data analysis platform market will ineluctably go toward a monopole with few main actors. Those companies will have large communities of users and important genomic databases. The barriers to enter the market will be high. In fact, some interviewees think that the same phenomenon of monopole that happen in platform-based companies such as Wikipedia's dominant position in the online encyclopaedia research, could be applied to the NGS data

²⁷ Custom panel: Specific group of genes that does not exist on the market and need to be build specifically for the customer. Commercial panels are group of genes already existing on the market that can be use by different customers. It is cheaper.

analysis market. For Dr. Jean Francois, it is a limitation because maybe those companies will strongly protect their databases and their technologies in order to keep a high competitive advantage in a monopole market.

In this chapter, I analysed the collected data in order to find out how Sophia Genetics developed and monetized network effects. I also studied Sophia Genetics' future perspectives in managing those effects. Finally, I explored few legal limitations on the genomic data sharing between medical institutions.

7. Discussion

The aim of this chapter is to discuss few relevant points seen in the data analysis from chapter 6. I will bring my interpretation and make some parallels with the literature seen in the theory.

7.1 Zig Zag strategy

Sophia Genetics entered the French market and other markets by convincing few labs, for a start, and then progressively bigger labs, similar to the Zig Zag strategy (Evans, 2008) seen in the theory. It is about developing network effects step-by-step to enhance the attraction of new users without necessarily having a strong community at the beginning. This is related to genomic data race and the vision of Sophia Genetics. Indeed, the most powerful NGS platform will be the one with the largest, the most complete and diversified database. This is why Sophia Genetics wants to optimize, as much as possible, the attraction of new worldwide users and create strong direct and indirect network effects. The aim is to build a community fuelled by network effects that will enhance a worldwide genomic database in order to cure diseases.

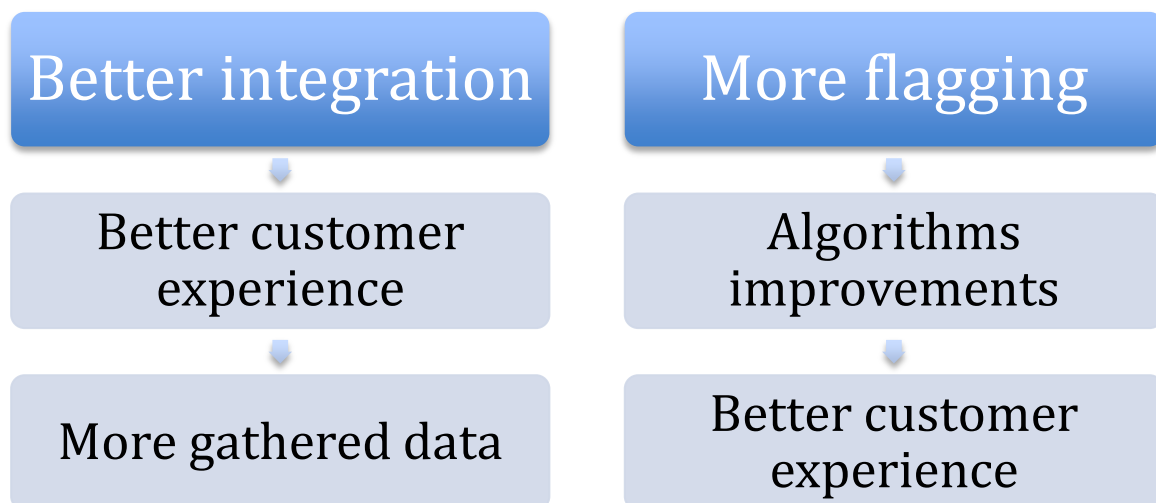
Indeed, the future vision of Sophia Genetics is if, for example, 1000 patients have a lung cancer, 400 survived due to a specific treatment, the next patient on the platform who shares the same genes mutations will receive this treatment. The patient's life will be saved. In fact, one patient in Japan could help a patient in the US to survive by sharing information on the platform.

7.2 Direct network effects-Integration and flagging system

The company efficiently develops the direct network effect by improving the algorithms constantly with the number of users and by implementing the flagging system. Indeed, the attraction of new users on the platform is going fast but could go even faster if Sophia Genetics adopts a better-automatized integration with LIMS local

systems. It seems also that flagging system works a lot in certain institutions and doesn't in others. To face this issue, LIMS system could push the variants classification, as flags will be directly registered in reports through the lab system and not just on the platform. It is also easier and more motivating for oncologists to upload the data because all the information is directly registered in the lab system. Thus, a more efficient integration will enhance direct network effects. For each new user on the platform, the value created become exponential and benefits the whole community. Figure 31 illustrates the leverage effect created by a better integration and flagging system.

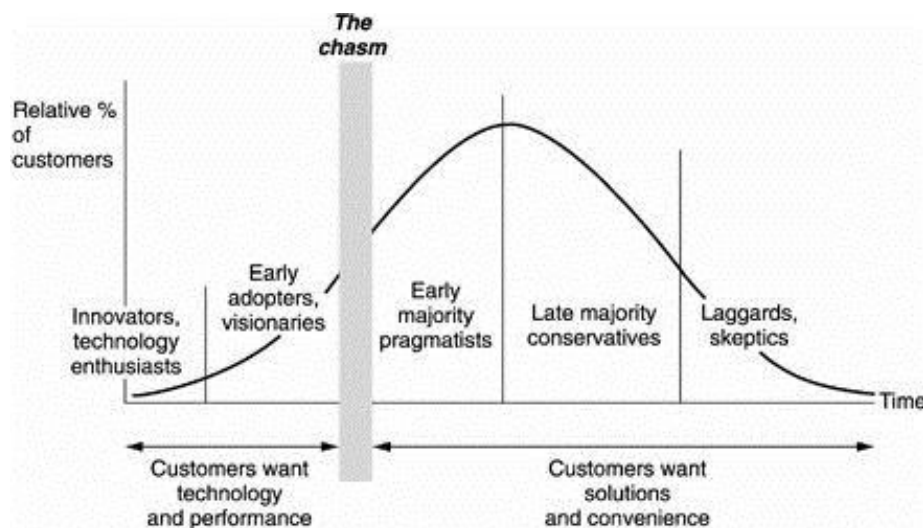
Figure 31: Integration and Flagging Leverage Effect



7.3 Adoption Curve

There are still some difficulties in convincing large labs that already have a strong database and experience—specifically, in the US.

Figure 32: Adoption Curve with Chasm



Source: Eisenmann (2011), Platform and networks blog

In my opinion, if I, once again, use the adoption curve from Rogers (2010), Sophia Genetics still needs to cross the chasm in some markets. The chasm is the point between early adopters and the early majority. In France, the company succeeded in crossing the chasm after 3 years but in other markets, such as the US or Canada, it is still the beginning. With time, hopefully, the early adopters will see the real value of the product, and through network effects, it will tip in the hands of the early majority. The most relevant difference between early adopters and the early majority is that the perceptions of the product change. In fact, early adopters expect a product to be innovative and performing as a new technology gadget. After crossing the chasm, the customers expect a product to bring efficient solutions on daily-basis use, which is convenient, well-integrated until progressively becoming essential in their routines (Eisenmann, 2011). According to the Sophia Genetics customer's interview in Belgium, he was using the platform on a daily basis and it was a tool becoming necessary and indispensable for his next DNA tests analysis. This is why, in my opinion, in certain markets, the platform is now used by the early majority and, progressively, by the late majority.

Interestingly, the critical mass point theory from Evans (2009) is related to the adoption curve. According to Dr. Jurgi Camblong, the community will reach the critical mass when the community becomes essential within each specific application the platform offers. According to Reillier (2017), the critical mass is reached when the platform is self-sustainable. In my opinion, Sophia Genetics platform will be self-sustainable by convincing a large group of early adopters, and like a snowball effect, it will cross the chasm, falling in the hands of the early majority. The threat is that competition appears during this race to critical mass. In the case of Sophia Genetics, there are a few competitors, such as PerianDX, that share the same vision of gathering the genomic data the quickest as possible. However, I agree that it will tend toward a monopole market in the future. Indeed, economic barriers will be very high due to the powerful genomic databases gathered.

7.4 Bundle and IDS products

Sophia Genetics has succeeded in creating new segments of customers through the bundle and the IDS products. For the bundle solution, the company's approach is similar to the 'platform envelopment' (Eisenmann et al., 2010) seen in the literature. Indeed, thanks to 'multi-bundle' strategy, Sophia Genetics platform became more competitive, attracting a new market of users. For the IDS product, the company followed the strategy proposed by Reillier (2017), by focusing on one side of the market first to then attract the other side. It is exactly what happened in the case of Sophia Genetics—by concentrating their expertise first on the software analysis to then attract labs that didn't have a sequencer to join the community. Thanks to this method, the company avoided the famous Chicken or Egg (Tirolle, 2003) issue that often faces platforms in their early stages of life.

However, in my opinion, these two different strategies did not create INE for two reasons. The first reason is that, even if these strategies have unlock a new market, it is the same type of users. Indeed, new users from IDS or Bundle products are still data analysts on the platform and don't bring a new knowledge that could provide others specialists such as oncologists. The second reason will be directly developed in the next subchapter.

7.5 Indirect network effect - Future Perspectives

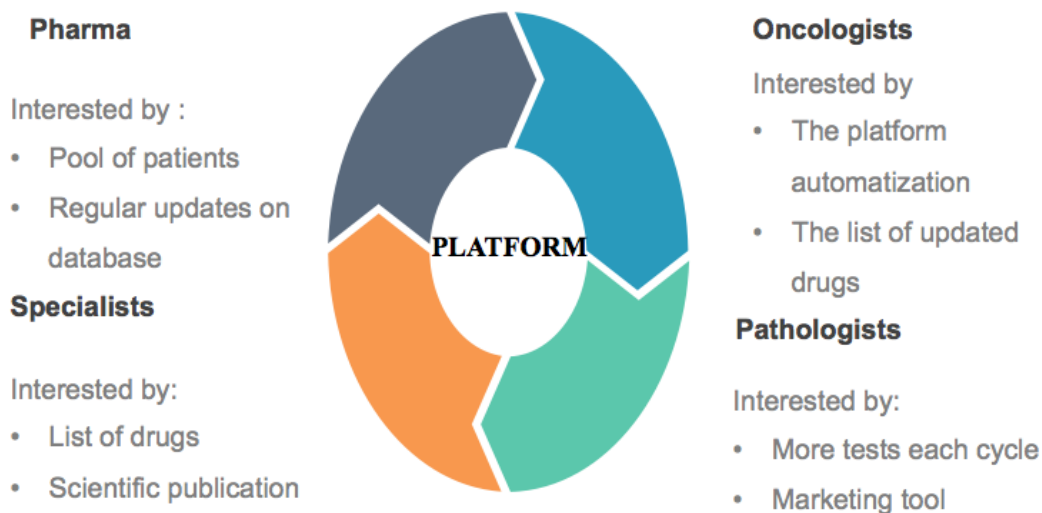
In my opinion, today there are no INE on Sophia Genetics' platform, but the opportunity to create this effect is huge. If we take back the definition from Marshall et al. (2016), "indirect network effects are the impact created by users of one side of the market on users on the **complimentary** side of the market." Users need to be complementary to each other in order to create INE. The IDS community in Sophia Genetics gathers labs that have a sequencer and labs that don't. Those two types of users are not yet complementary because the lab that has a sequencer doesn't necessarily rely on the lab that does not possess a sequencer to be active on the platform. On the other side, the lab that needs access to a sequencer relies entirely on the other lab to be able to have access to the platform. It is a one-sided relationship. This is why I believe that Sophia Genetics is still in between a technology creator and a network orchestrator company (Ribaudou, 2016) as we saw in the literature. To be a network orchestrator platform, DNE and INE need to exist and be managed in the most efficient way in order to maximize the value generated by both effects.

However, according to the interviews, a lot of new ideas are to be developed in the future to go toward real indirect network effects. As you can see in the annex, I put in green all the codes that had future perspectives. In the INE category, 11 codes out of 14 are green, meaning that INEs are in development. Indeed, Sophia Genetics has a lot of possibilities to enable INE. For instance, pathologists and oncologist could be complementary through the platform. Pathologists need oncologists' tests orders in order to be able to use their sequencer and generate precise results through the platform. On the other hand, the oncologist needs the pathologist's lab to perform the test, and the platform to upload patients' information. If pathologists use the interface as a marketing tool, it could enhance the effect. Together, they create a source of excess value (Marshall et al., 2016) that is enabled, thanks to facilities built on the platform. The value in this situation is, in fact, that oncologists upload patients' information that is crucial for the future development of the interface and for the algorithms. In this case, oncologists have enough output from the platform to contribute themselves in the genomic data collection.

Another example is the pharmaceutical companies that need to perform the recruitment of patients for clinical trials based on a specific molecular profile. The two complimentary sides, in this case, are the hospitals and the pharmaceutical companies. Pharma companies indirectly need Sophia Genetics to select the right patients in the hospitals. Hospitals want Pharmaceuticals' new treatments to save their patients' lives through the platform intermediary. The relationship is created and, again, generates a source of excess value. The value is developed by the fact that Sophia Genetics guarantees to match a molecular profile, and in exchange, the pharma companies are requested to upload their latest treatments databases on the platform. In this way, oncologists and doctors can now match their patients' molecular profiles to specific treatments on the powerful up-to-date platform.

I believe that developing indirect network effect is about catching the value created by a complementary relationship of users' interactions through the platform. Furthermore, by increasing the number of complementary sides, the value created could be exponential (Rochet et al., 2003).

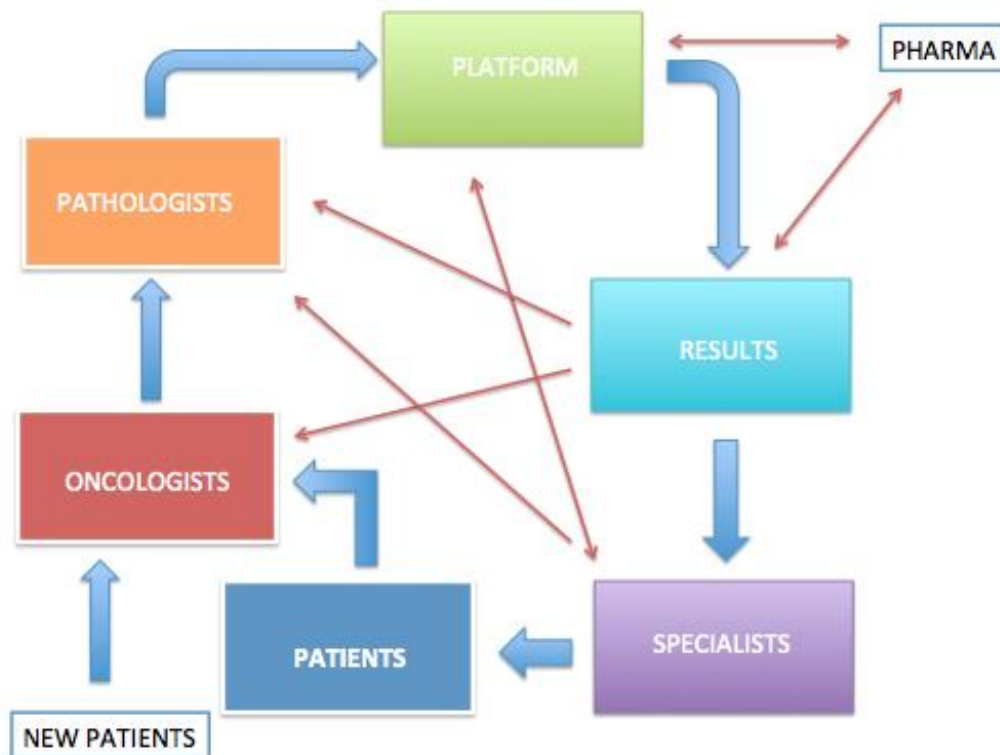
Figure 33: Future Indirect Network Effects Leverage



If Sophia Genetics succeeds in implementing relationships between users in Figure 33, the platform will efficiently exploit INE. Thus, the platform will have a sustainable competitive advantage by respecting the organized criteria from VRIO framework

(Barney, 1998)²⁸. In this way, creating a multi-user interface is the key strategy for Sophia Genetics to adopt. The numbers of interactions will ineluctably increase the platform's value. This is illustrated in the in Figure 34.

Figure 34: Interactions Dynamics



7.6 Long-term monetization plan

It seems that Sophia Genetics followed the theory from (Reillier et al., 2017) of 'users first and then monetize.' Indeed, the board could have decided to implement a much more aggressive charging model by putting adverts or by proposing premium models. According to Evans (2008), it is a dangerous path to take. Dr. Jurgi Camblong thought the same and pushed the opposite strategy; first, build a community, and afterward, progressively implement new ways of monetization. The transaction fee

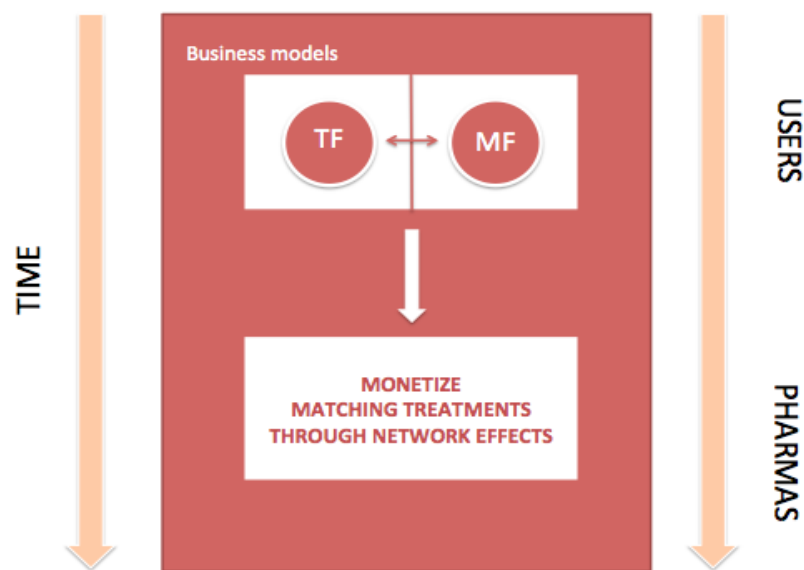
²⁸ I assumed the other criterias, value, rare, inimitability are already respected because the community developed by Sophia Genetics is a strong asset difficult to copy. Chapter 2.2.2 explains the reasons why in general a platform-based company's network asset respects VRIO criterias.

80.

was a great move because it did not put economic barriers at the entry and pushes new users to adopt the platform.

It seems also that with time, the charge could shift to pharma companies or specialists and less toward medical institutions. The price of doing tests could decrease significantly if pharma companies are involved in the process. It is already the case in certain countries where pharma companies sponsor some tests and give all the related medical equipment for free. In the US, competitors' prices are already very low in order to attract new customers. Matching the treatments, thanks to the platform, could be a way to yield a lot of revenue. In exchange, Sophia Genetics could decrease the transaction fees to enhance direct and indirect network effect.

Figure 35: Long-Term Horizon Monetization Plan



We have seen this trend in a few years past in other platform businesses where direct users are less or not charged at all (Marshall et al., 2016). The fact of having access to a lot of different services for a community without paying a cent has revolutionized the business model. It follows Muzellec's (2015) opinion that progressively, monetization plan of platforms will be oriented toward third-party payers through B2B or indirectly B2B&C. I agree and think we are going in the same way toward genomic data platform analysis.

However, as we saw in the limitations, genomic data cannot be sold. It can be used only and only if it is for improving medical diagnostics. The barriers must not be crossed between improving genomic data analysis and the commercial use of those. Sophia Genetics is required to always think about how to ensure trust. If doing partnerships with pharma companies is in order to enhance treatments match and the customer experience on the platform without affecting network effects, then it is possible. The respect of all the privacy and confidentiality of the patient data needs to be well discussed before any partnership agreements.

7.7 Limitations

Today, technology moves faster than the willingness of humans to cooperate. It is interesting to observe that we cannot efficiently interpret yet the genomic data generated by the machines we created. To do so, collaboration between institutions is crucial. That being said, genomic data is a new currency that is exchanged as a commodity. This is why, in my opinion, free models will not be reliable because the evolution will be constant and will require a huge amount of work and investments.

In this chapter, we saw that Sophia Genetics efficiently develop direct network effects. However, they did not yet develop the indirect network effects between different types of users on the platform. We observed that the most important in future would be to involved specialists, doctors, pharmaceuticals company and patients in order to leverage interactions and create a multi-user interface fuelled by INE. Finally, we analysed the future trend of monetization toward a third-party payer such as pharma companies or specialists in exchange of the access to certain pool of patients.

8. Conclusion

8.1 Theoretical Context

We saw how, over a very short period of time, the incredible phenomenon of platform-based companies impacted our society. It has unlocked new markets, new theories, new strategies, and changed the way we think and live. From traditional tangible asset builder to intangible asset network orchestrator, new opportunities have appeared and already evolved significantly.

We studied the concepts of Direct Network Effects (DNE) and Indirect Network Effects (INE). We saw that the multiplication of such effects are key to successfully build a sustainable platform and community-based company. We looked at the difference between the two types of effects, a DNE being the positive (or negative) effect that an additional user on the platform brings to the other users of the same nature (Manchanda et al., 2015), in other words, the more users of the same nature the more value creation brought by the platform (e.g., the more buyers on the App Store, the more applications there will be), a INE being the positive (or negative) effect that a user of the platform creates for an other type of user, in other words it's the creation of a complementary and valuable interaction between two different types of users (e.g., the more drivers they are on UBER, the shorter will be the pick up time for the travellers).

We analysed few strategies to develop network effects such as the Zig Zag theory from Evans (2009) that consists in attracting first step-by-step small group of users and then attracting large groups of users. We described the critical mass milestone for a platform, being the moment when the platform becomes self sustainable (Reillier, 2017).

Monetizing the value created by the platforms was the second key question that we developed and which represents one of the main challenge of platform-based companies. We highlighted the sensitivity of users, either charged or subsidised (e.g., the access to Google is free and it is likely that it would not have worked if the users had to pay for it). We also explored the different monetization strategies, such as the

membership fee, the proportional fee or the transaction fee. The monetisation strategies are evolving so fast that platform managers need to constantly find innovative methods in order to keep the community sustainable.

Finally, we observed the trend toward monetizing the data rather than directly charging the users. In fact, the charge has shifted to third-party payers who are willing to pay an important amount of money to have access to the data (Muzellec et al., 2015). Facebook is a good illustration of this model. This is the 'magic' effect of big data monetization.

8.2 Practical context

Sophia Genetics, the company we analysed in the practical part, is strongly related to the network effects phenomena. This company uses a platform to gather and interpret worldwide genomic data in order to help hospitals to do better diagnostics.

We discovered that Sophia Genetics develops DNE because each new user on the platform increases the benefit of the whole community by constantly improving the algorithms. But the company did not develop INE because there is only one type of user on the platform (data analysts) and no complementary relationships yet created.

Sophia Genetics actual monetization plan to capture network effects is principally a transaction fee, a charge per sample uploaded on the platform. This monetization plan attracts new users because it is a pay per use model with no commitments. The company is likely to introduce other monetisation strategies in the future with the arrival of new users that will be less sensitive to price and that will be interested to have access to data such as pharma companies. In parallel, there might be future pressure for reducing the transaction fee for the existing users; we highlighted the case in the US where there is already pressure on prices.

What is the Vision of the company? Having started as a bio-informatic platform with one type of user, will it focus in evolving towards a multi-users universal platform with several monetization concepts? What image would platform managers want to portray to their users? Will the platform eventually give access to the patients and open a new world of opportunities and objectives? Crucial strategic decisions will need to be taken in the short term to gain leadership and become the platform of reference

in this highly competitive and desired domain. I believe the only way to success will be to achieve an effective collaboration through complementary relationships using INE, between doctors, specialists, institutions, pharmaceuticals companies, perhaps patients, and all main actors in the medical world. Each protagonist needs to have incentives toward collaboration. This will contribute to multiplying the DNE and INE. The platform will then reach the critical mass and become self-sufficient.

Sophia Genetics is a great platform-based model and thanks to such approach, they open the way to fast and efficient data-driven medicine. For sure, in this context, the platform approach combined with network effects will contribute to greatly improve the medicine and save lives.

**“When we have all data online, it will be great for humanity.
It is a prerequisite to solving many problems that humankind face.”**

—Robert Cailliau, Belgium informatic engineer
who created the first browser for Mac

**“You can have data without information,
but you cannot have information without data.”**
—Daniel Keys Moran, American computer programmer

Bibliography

Theoretical part:

1. Reillier, L. C., & Reillier, B. (2017). *Platform Strategy: How to Unlock the Power of Communities and Networks to Grow Your Business*. Taylor & Francis.
2. Parker, G. G., Van Alstyne, M. W., & Choudary, S. P. (2016). *Platform revolution: How networked markets are transforming the economy--and how to make them work for you*. WW Norton & Company.
3. Marshall W. V., Geoffrey G., Sangeet P C. (2017). *Pipelines, Platforms, and the New Rules of Strategy*. Retrieved June 27, 2017, from <https://hbr.org/2016/04/pipelines-platforms-and-the-new-rules-of-strategy>
4. White, A., & Weyl, E. G. (2016). Insulated platform competition.
5. Belleflamme, P., & Peitz, M. (2010). Platform competition and seller investment incentives. *European Economic Review*, 54(8), 1059-1076.
6. Belleflamme, P., & Toulemonde, E. (2009). Negative intra-group externalities in two-sided markets. *International Economic Review*, 50(1), 245-272.
7. Evans, D. S. (2009). How catalysts ignite: the economics of platform-based start-ups. *Platforms, markets and innovation*, 99-128.
8. Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy*, 43(7), 1239-1249.
9. Chu, J., & Manchanda, P. (2015). Quantifying cross and direct network effects in online C2C platforms.
10. Chesbrough, H., & Rosenbloom, R. S. (2002). The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies. *Industrial and corporate change*, 11(3), 529-555.

88.

11. Walton, N. (2017). The Internet as a Technology Based Ecosystem A New Approach to the Analysis of Business, Markets and Industries.
12. Parker, G. G., & Van Alstyne, M. W. (2005). Two-sided network effects: A theory of information product design. *Management science*, 51(10), 1494-1504.
13. Porter, M. E. (1997). Competitive strategy. *Measuring Business Excellence*, 1(2), 12-17.
14. Rifkin, J. (2014). *The zero marginal cost society: The internet of things, the collaborative commons, and the eclipse of capitalism*. St. Martin's Press.
15. Barney, J. B., & Wright, P. M. (1998). On becoming a strategic partner: The role of human resources in gaining competitive advantage. *Human Resource Management* (1986-1998), 37(1), 31.
16. Cusumano, M. (2010). Technology strategy and management: The evolution of platform thinking. *Communications of the ACM*, 53(1), 32-34.
17. Hagiu, A. (2006). Pricing and commitment by two-sided platforms. *The RAND Journal of Economics*, 37(3), 720-737.
18. Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of management*, 17(1), 99-120.
19. Eisenmann, T., Parker, G., & Van Alstyne, M. W. (2006). Strategies for two-sided markets. *Harvard business review*, 84(10), 92.
20. Belleflamme, P., & Peitz, M. (2016). *Platforms and network effects* (No. 16-14). Working Paper Series, Department of Economics, University of Mannheim.
21. Rochet, J. C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the european economic association*, 1(4), 990-1029.
22. Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, 37(3), 668-691.

23. Osterwalder, A., & Pigneur, Y. (2010). *Business model generation: a handbook for visionaries, game changers, and challengers*. John Wiley & Sons.
24. Muhtaroglu, F. C. P., Demir, S., Obali, M., & Girgin, C. (2013, October). Business model canvas perspective on big data applications. *In Big Data, 2013 IEEE International Conference on* (pp. 32-37). IEEE.
25. Muzellec, L., Ronteau, S., & Lambkin, M. (2015). Two-sided Internet platforms: A business model lifecycle perspective. *Industrial Marketing Management*, 45, 139-150.
26. Eisenmann, T. R., Parker, G., & Alstyne, M. V. (2009). Platforms, markets and innovation.
27. Beck, B. (2014). *What Airbnb, Uber, and Alibaba Have in Common*. Retrieved July 05, 2017, from <https://hbr.org/2014/11/what-airbnb-uber-and-alibaba-have-in-common>
28. Amit, R., & Zott, C. (2001). Value creation in e-business. *Strategic management journal*, 22(6-7), 493-520.
29. Rysman, M. (2009). The economics of two-sided markets. *The Journal of Economic Perspectives*, 23(3), 125-143.
30. Teece, D. J. (1980). Economies of scope and the scope of the enterprise. *Journal of economic behavior & organization*, 1(3), 223-247.
31. Clements, M. T., & Ohashi, H. (2005). Indirect network effects and the product cycle: video games in the US, 1994–2002. *The Journal of Industrial Economics*, 53(4), 515-542.
32. Hyun, J. (2016). Competition in Two-sided Platform Markets with Direct Network Effect.

90.

33. Langley, P., & Leyshon, A. (2016). *Platform capitalism: the intermediation and capitalisation of digital economic circulation*. Finance and society.
34. Eisenmann, T. R., Parker, G., & Van Alstyne, M. W. (2010). Platform envelopment.
35. Evans, D. S. (2009). How catalysts ignite: the economics of platform-based start-ups. *Platforms, markets and innovation*, 99-128.
36. Gladwell, M. (2000). *The tipping point: How little things can make a big difference*. Little, Brown.
37. Caillaud, B., & Julien, B. (2003). Competing in network industries: Divide and conquer. *RAND J. Econ*, 34, 309-328.
38. Skilton, M. (2015). Introduction—*The Rise of Technological Ecosystems*. In *Building the Digital Enterprise* (pp. 3-21). Palgrave Macmillan UK.
39. Timmers, P. (1999). *Electronic commerce*. John Wiley & Sons, Inc..
40. Bataineh, A. S., Mizouni, R., El Barachi, M., & Bentahar, J. (2016). Monetizing Personal Data: A Two-Sided Market Approach. *Procedia Computer Science*, 83, 472
41. Kenney, M., & Zysman, J. (2016). The rise of the platform economy. *Issues in Science and Technology*, 32(3), 61.
42. Teece, D. J. (2010). Business models, business strategy and innovation. *Long range planning*, 43(2), 172-194
43. Clemons, E. K. (2009). The complex problem of monetizing virtual electronic social networks. *Decision Support Systems*, 48(1), 46-56.
44. Weyl, E. G. (2010). A price theory of multi-sided platforms. *The American Economic Review*, 100(4), 1642-1672.
45. Rysman, M. (2009). The economics of two-sided markets. *The Journal of Economic Perspectives*, 23(3), 125-143.

46. Hagiu, Andrei, and Julian Wright. "Multi-sided platforms." *International Journal of Industrial Organization* 43 (2015): 162-174.
47. Wang, Z., & Wright, J. (2017). Ad valorem platform fees, indirect taxes, and efficient price discrimination. *The RAND Journal of Economics*, 48(2), 467-484.
48. Henten, A. H., & Windekilde, I. M. (2016). Transaction costs and the sharing economy. *info*, 18(1), 1-15.
49. Sriram, S., Manchanda, P., Bravo, M. E., Chu, J., Ma, L., Song, M., ... & Subramanian, U. (2015). Platforms: a multiplicity of research opportunities. *Marketing Letters*, 26(2), 141-152.
50. Rochet, Jean-Charles, and Jean Tirole. "Two-sided markets: a progress report." *The RAND journal of economics* 37.3 (2006): 645-667.
51. Casadesus-Masanell, Ramon, and Feng Zhu. "Business model innovation and competitive imitation: The case of sponsor-based business models." *Strategic management journal* 34.4 (2013): 464-482.
52. Muthers, J., & Wismer, S. (2013). *Why do platforms charge proportional fees?* Commitment and seller participation.
53. Evans, D. S., & Schmalensee, R. (2016). *Matchmakers: the new economics of multisided platforms*. Harvard Business Review Press.
54. Wang, Z., & Wright, J. (2016). *Should Platforms be Allowed to Charge Ad Valorem Fees?*. Working paper.
55. Evans, D. S. (2003). The antitrust economics of multi-sided platform markets. *Yale J. on Reg.*, 20, 325.
56. Pujol, N. (2010). Freemium: attributes of an emerging business model.

92.

57. Niculescu, M. F., & Wu, D. J. (2011). When should software firms commercialize new products via freemium business models. *Under Review*.
58. Tiwana, A. (2013). *Platform ecosystems: aligning architecture, governance, and strategy*. Newnes.
59. Bakos, Y., & Katsamakas, E. (2008). Design and ownership of two-sided networks: Implications for Internet platforms. *Journal of Management Information Systems*, 25(2), 171-202.
60. Fish, T. (2009). My Digital Footprint *A two-sided digital business model where your privacy will be someone else's business!*futuretext.
61. Evans, D. S., & Schmalensee, R. (2005). *The industrial organization of markets with two-sided platforms* (No. w11603). National Bureau of Economic Research.
62. Hagiu, A., & Wright, J. (2015). Multi-sided platforms. *International Journal of Industrial Organization*, 43, 162-174.
63. Ribaud, M., (2016). Global report on technology and economy
64. Won Lee, C., Kwon, I. W. G., & Severance, D. (2007). Relationship between supply chain performance and degree of linkage among supplier, internal integration, and customer. *Supply chain management: an International journal*, 12(6), 444-452.
65. Merchant, N. (2012). Why Porter's Model No Longer Works. *Harvard Business Review*.
66. Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). Defining supply chain management. *Journal of Business logistics*, 22(2), 1-25.
67. Heck, E. V., & Vervest, P. (2007). Smart business networks: how the network wins. *Communications of the ACM*, 50(6), 28-37.
68. Porter, M. E. (2008). The five competitive forces that shape strategy. *Harvard business review*, 86(1), 25-40.

69. Zhu, F., & Liu, Q. (2016). Competing with complementors: An empirical look at amazon. com.
70. Liu, H., Guo, J., Yu, W., Zhu, L., Liu, Y., Xia, T., ... & Gardner, R. M. (2016). The design and implementation of the enterprise level data platform and big data driven applications and analytics. In *Transmission and Distribution Conference and Exposition (T&D), 2016 IEEE/PES* (pp. 1-5). IEEE.
71. Ramsey, F. P. (1927). A Contribution to the Theory of Taxation. *The Economic Journal*, 37(145), 47-61.
72. Watson, J. D., & Crick, F. H. (1953). Molecular structure of nucleic acids. *Nature*, 171(4356), 737-738.

Practical part:

73. Bryman, A. (2015). *Social research methods*. Oxford university press.
74. Edwards, R. & Holland J. (2013). What is qualitative interviewing? *Bloomsbury Academic*. pp. 2-3
75. Salmen, L. F. (1987). *Listen to the people: Participant-observer evaluation of development projects*. World Bank.
76. Jonsen, K., & Jehn, K. A. (2009). Using triangulation to validate themes in qualitative studies. *Qualitative Research in Organizations and Management: An International Journal*, 4(2), 123-150.
77. Koners, U., & Goffin, K. (2007). Learning from Postproject Reviews: A Cross-Case Analysis. *Journal of Product Innovation Management*, 24(3), 242-258.
78. Mathison, S. (1988). Why triangulate? *Educational researcher*, 17(2), 13-17.

94.

79. Denzin, N. K. (1978). Triangulation: A case for methodological evaluation and combination. *Sociological methods*, 339-357.

80. Basit, T. (2003). Manual or electronic? The role of coding in qualitative data analysis. *Educational research*, 45(2), 143-154.

81. Spradley, J. P. (1980). Participant Observation. Fort Worth: Holt, Rinehart, and Winston.

82. Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.

83. Moore, G. A. (2002). Crossing the chasm

Appendix

Appendix 1:

Interview's guide

Themes	Sub-Themes	Key Words
Your role in the company	What is your activity in the company? Which values do you bring with you? What are the main stakeholders in the company? Which one do you have the most influence? How to better share the value you bring to the company?	Sales Motivation Determination Sharing CCO Communication Attitude
Sophia value proposition	Could you explain Sophia main value proposition? What DDM brings to customer? Tell me more about the community? What are the network effects created on the platform? What is the business model? Do you think to a better one? Is there a lot of interaction on DDM? Do you think about new methods to push customers to comment and share their knowledge ? What about IDS and bundle solutions, influence for network effects? What is the company business, B2B—B2C—C2C? Did this type of shift enhance and better catch network effects?	Direct/Indirect Transaction Sharing knowledge Limits of sharing New business models Iteration Enhance curation Customers satisfaction Bundle C2C ids
Sophia in the future	Is Sophia will be technology creator or a network master? Do you know a way to increase network effects for Sophia? What is your dream about Sophia? Towards one platform on the market?	Gather the data Network orchestrator Dream Match's Monopole

Appendix 2:

Appendix 2.1. Development network effects codes

Platform utilization and direct network effects

Platform Utilisation

Machine learning and AI, crunching data towards precision medicine	Simplicity to use platform/ fluid-friendly user - customer experience first	More external databases to be more precise- to create diversification of pool of samples-HGMD	Unique solution, standardisation of data from different sources(labd-chemistry-sequencers)
Just based on diagnostics, strong argument-Short-turn around of time of results	Data analyst use the platform not the doctors	Data analyst-biologist want upstream analysis, and downstream analysis flagging curation...	Flagging system
Quickly classify variant/ Very important	ACMG guidelines need to be implemented on DDM for US market- Enhance quality curation	Flagging take a lot of time	Sophia DDM could read in the future data base- facilitate work
Clinvar example gathered a lot of data but not always good quality	Other chapter- Good Workflow with Sophia- different processes reunited - don't need other softwarres	Regular update of DDM databases attracts new users	Platform will do all the steps for oncologist automatically, research, variants../ Incentive to fullfill data
Oncologist fullfill data to gain time on oncoportal- own needs	However-Bad frequency on certain variants - pool of samples biased-too mixed up different origins	Platform lock in phenomena- how to opitmise fidelisation	The CEO-Think how we can optimise the utilisation for users to have a real benefit to use it
More test plus AI benefit for all community			
Johns hopkins interested by ACMG guidelines explain why	Darrmouth: want the best anylysis no more missed variants	Broad Institute- scared because DDM black Box - want more transparancy	VCU: Doubts on AI and machine learning interesitng

Direct network effects

Global approach-strong sell team around the world- push new users	Number of users increase- precision increase-value arise	Based on thousands of clinical samples	Algorithms getting better with number of users
Value-High quality NGS diagnostics	Solid pipeline	Flagging incentive- now is gain time and to organize variants principally for each institutions-they do not do it for community reasons but for their own interest	Flagging reconfort new users- more attractive for small lab-related to DNE- Flagging-value inrease for everyone

However-Lot of labs don't flag- too much time to do it- scared for mistakes - when benign they don't do it	Push for trust to be confident to flag-More details on why this variants was flagged to reinforce trust for the next one who will flag	Biologist need to find incentive to curate variants	Discount in function of variant curated-find the customers who flag nicely- Shirine
Pay flagging not good- need to be solidarity spirit- dangerous incentive too pay-bad quality	Flags need to be excellent quality- rather 1,2 good quality to 50 bad quality	Need to choose a strategy to by pass financial incentive for flagging- and do it naturally	
More transparency to attract new joiners	more users-more different offers-more applications- more validate test-more benefit for all the platform		
Cleveland interested in flagging but for their own interest not for community	Moffitt interested by our diversity of our offers and applications can serve for all their interest		

Indirect network effects and Community

Indirect network effects

For now we have those indirect effect- IDS attract new community- accessible platform for all labs- (int manual INE)-Concentrate on one group side of the platform first to attract other groups- avoid chicken or egg issue	B2B Dry lab core business- attract others-leverage effect: Bundle approach thanks to our dominant position in dry lab- DDM become marketplace to other communities(INT man)	Doctors can also exchange but not through platform- we do it for them-interact outside the platform-Why not Forums but doctors don't have time: KP	Now not a lot of automatic INE through platform- More types of users in the future- more attraction-Sophia started just with pathologist and want to integrate other users step by step
Other users such as hematopathologist can be attracted by platform/oncoportal	Hematologist could be interested by scientific report and oncoportal with treatments list-platform want specialist more interactions	Interaction multi-user (biologist, oncologist, patient) different interface for each one- expand diversity-regroup different health actors- Share knowledge from people with different experience-leverage effect	Find attraction between biologist and clinicians
Data analyst need to give value to clinicians and vice versa	Clinicians need to have access and active on platform	Clinicians will be attracted because they have the scientific report and oncoportal	Oncologist don't want to do intense paperwork anymore-need a platform

98.

Oncologists don't have enough output from DDM to fulfill data and to interact- for now we just count on their good willingness to do it	Oncologist need to fulfill data because they need it and it's easy- a real need for example we provide platform and in exchange more interactions with pathologist-marketing tool- attract oncologist-more tests) trust relationship - Those interactions expend diversity of offer and benefit all the community- google example-DDM	Oncologist take the decision of which treatment- need to on platform	Match oncologues desires- with new treatments of pharmas
leverage effect: treat a lot patient per month at Sophia- we could advise pharmas matching molecular profile	Partnership with pharmas to recruit patients for them based on our data base	Partnership with pharmas for clinical trials-hospital will be happy because their patients received new treatments- win win situation-good reputation for hospitals	Propose to pharmas to update their database and Sophia do the bridge with oncologists-win win situation more tests
Collaboration possible with partnerships to have access to platform			
Cedars Sinai interested by IDS	Umiami interest by bundle solutions-test our product-move to our drylab	Drexel small interested by community but specially by oncoportal and specialist opinions-beginning of INE	VCU: Doubts on AI and machine learning interesting

Community

Bring huge community of expert and knowledges to hospital and-Sophia is a network orchestrator- Number one value of Sophia is the community- Contributive economy-if everyone is playing the game-better service for everyone Exemple curie chez Shirine-Create a community requires a lot of and creativity and determination	Standardization of data / allow community- more info- more interaction- push interaction between hospital-leverage effect	Strong marketing- to push the community	Community is attractive for new joiners because they now there is already a lot of feedbacks
Community attractive when large panels of gene because necessity of opinions	Sharing is necessary- towards whole exome- sharing become even more crucial	Big labs with experience less attracted by community than small labs that more advantages-some labs don't see the necessity	Push big lab- it will follow the adoption curve of early adopters, the first do it and the second will do it also because they need it- think Jurgi

Local communities-more interactions because they can directly communicate and trust each other,consortiums-Create a local need to motivate people to join community-example of bsit-global more difficult	International communities harder than loca- more scared for mistakes!	Attract more people step by step	diversity of users benefit all the community- we want them to be involved and active on platform
The value of the platform through the community it is the trust which is crucial in the beggining-Trust is engine for netowrk effects- Establish trust between users on platform and the platform-user profile reinforce trust also- Transpaenrenncy, respect, rules respected, ethical will talk later about this	Diversification in genes curated in comm-self sufficient community if experts for the 20000 genes-we need diversity	Diversification- data base cleaner-complementary in community-better frequency	Critical mass difficult to say for Sophia- maybe one application at the time and Sophia reach CM when it is essential to have it- platform self sufficient and evolve alone
JF vision: Towards data sharing with no competition and no monopolesv	JF vision: Sophia public company, work with states to attract a bigger community and install trust	Research/Academic not a lot of volume- prefer to operate in diagnostics- much more volumes so we can have impact	Jurgi-think about future of a monopole market- example of Wikipedia
Towards a a monopole platform- a lot of barriers to entry	NGS platforms communities market will a monopole in future-KP	Interpretation (build a community as a platform airbnb-same attraction)	
Columbia: Too much advanced on everything tests no community needed-they say they are the best	(Walter reed medical center not interested in community because too secure)	Cedars Sinai attracted by customer base in Europe	Milwaukee didn't trust community of experts

Integration

A good integration in customer services is crucial to enhance community- To facilitate integration, Customer's lab have systems call LIMS (...)- and gather all the patient background data such as sec effects Hospitals want NGS fast to routine-Now de don't have it	This reinforce the Simple and fluid system/process- Integration is necessary to connect labs and DDM-no more excel- automatisation necessary-Customers are rooted with a fluid and efficient -integration system	Moreover-Price get cheaper-NGS analysis in house general movement- we need to integrate them easily	They need a to coordinate Lab management-upstream (reporting) with /downstream-curation - Example of EPIC
---	--	---	---

100.

Reporting generated automatically reporting fast and good quality to gather more data pass through DDM	LIMS integration- push flagging because directly automatize in the reports- Automatisations-More incentive for DATA analyst to see results on the report for clinicians	more facility for integration of publication data via Sophia -automatisation	Easier for Oncologists need to fullfill data/ phenotype/treatment data because automatisations
leverage effect- easier-Integration more Validation program that expend numbers of applications and benefit all users	However need to be careful All the data is mixed, patients with different origins, different ages.. Biased statistics-results less precise on DDM	Big gap between doctors and administration who is making the decision to integrate or not Sophia- both need to be convinced	
Vanderbilt Medical center- attracted by the rapidly of results and get them to routine very quickly thanks to validation program- much more cheaper in house	Darmouth and JH: by pass bioinformaticians difficulties for integration	JAMES cancer hospital- already work with EPIC too much work if Sophia is not in the system	

Appendix 2.2. Monetization codes

Transaction fee and Membership fee

Transaction fee

Main method at Sophia- Charge per samples upload	Small volumes labs are not penalised	Price depends on the panel- adaptive price-No economical barriers	No commitments
More transparency in costs structure-origins	More independant- freedom-flexibility-price personalised	Charge per sample- easier to accept for accounting as typical lab product	Globally the volume of NGS tests in each lab will increase in the future- TF fee model will generate more revenues-long term interesting
In fact If customers pay for transaction fee meanwhile they are giving data- they really need it			
Blood cancer Wisconsin: small liked the aspect	Vanderbilt Medical center liked TF so they could test the product-very small volumes and they signed with us afterwards because they saw the value	UNC small volumes loved TF	Drexel really small volumes

Membership fee

Another type of payment famous in platform as we saw in theory-Sophia does not do membership fee directly	It has advantages-any company in the beginning need money -MF will provide easier-Big volumes of samples-easier for big labs-membership interesting	Interesting in short term	Best example of membership that succeeded because they can sell to B2B easiest contracts
In fact, according to the CEO, Validation program (explain) is a type of membership at Sophia in the beginning-fidelise customer	Against our vision of having more users as possible-puts barriers-We want to "cement" our position in hospitals-economical barriers-Same price for everybody	Difficult for labs to do big budget in advance	
Moffitt prefer annual membership to put in their budget			

Monetize network effects and free model codes

Monetize Genomic data

Huge value is created for sophia users start to flag and curate the database-doctors know this- treat patient each month-community great value also	What we do already: indirect monetisation based on the algorithm-more users- more we gain market shares-more competitive-community better algorithms-leverage effect-virtuous circle	Sophia do not monetize for directly the data because it's very sensitive-We don't monetize data that we stock- not for contracts reason- but for european rules I will describe in limitations subchapter	Charge pharma which do the recruitment for them as explained in Charge because we do the bridge talked in capture chapter
Charge one group of users who want to have access to the platforms such as specialists on the form of license- (maybe subsidised other groups)	(Towards a tendency of charging pharma or other partners and less the patients) Interpretation- US price cheap		

(Others) Business models

First-Need to be extremely competitiveCost/benefit analysis-	In general General platformers: monetization do not negatively impact network effects if we take very small % because there are already a lot of transactions however towards free	Free model is a fundamental model in platform, no barriers but for sophia: Free access to the platform wouldn't work	Free access to a platform didn't work example of Kevin- it's basic algorithms, standardisation no customisation of pipelines- too weak
Access with a charge fee guaranty a quality	Other visions-Charge just institutions with financial interest and not academics- Jurgi says no way-barriers no volume academic	Could let the users charge their own prices- global platform strategy	JF-Charge published results that pass through DDM but difficult because it's not our data
JF- No monetization- Sophia could be plubically subsidied by states			

Appendix 2.3. Limitation codes**Rules limitations**

Legal perspective- European rule GDPR-claims cannot sell genomic data- Example with Jurgi, controller and processor of data	Each hospital have their own standards	Each country has their own rules with data sharing and specific	For hospitals regarding Sophia can only use data to improve diagnostics, nothing else-some time they cannot flag or share because it is noticed in original contracts of hospitals
Not as simple as google or facebook- we cannot push one drug because they paid more	Monetize by pharma publicity ver complex	Medical rules very strict- data confidential	
Blood cancer Wisconsin: Concerned about genomic data securities issues	Walter Reed medical center: Want a data center on campus due to regulations	France need to have data centers within the country	

Human challenges

Data sharing still very sensitive- medical institutions are still very competitive specifically in US- genomics data huge value	Technology moves faster than human willingness to cooperate	We need collaboration not competition-Try to build a collaborative project with states	Lot of work to share- example Flagging not the same notions and relevancies for every lab- different databases
---	---	--	--

Some are Too scared to misclassify	human challenge not tehcnology	Hospitals try to monetize in their own side- and doing in house community	Ethical question- can I give this data like this- can I trust this?- for sophia cn use this data which is not ours
Interpretation-What image we want to give to Sophia? with giving money for flag, not good because it means flagging give us a lot of revenues-Monetize data still need to ensure trust between users!!!	Scared of data exploitation- example with curie from shirine	If loosing trust-catastrophic if labs see that some pharma are favoured	opinion from Thomas- A platform cannot change sociologically human mind -
Idealistic to gather all this data			

Appendix 3:

Interviews

Appendix 3.1. Jurgi Camblong:

Moi d'après mon expérience j'ai vu des clients qui adoraient la communauté et d'autre qui l'appréhendait un peu. Tu dirais qu'on se dirige vers quelle tendance ?

Ecoute, tu ne pourrais pas faire sans la communauté.

La communauté est obligatoire ?

En fait aujourd'hui ceux qui sont prudents ou réservés, ce sont ceux qui estiment qu'ils ont plus de connaissances que les autres et qu'ils ne veulent pas partager leur expertise, comme si moi je ne voulais pas partager les bons spots de surf. Et je me dis si je ne partage pas ces bons spots, il y aura moins de monde et je me réserve les vagues. Sauf que là on ne parle pas de vagues et de spots de surf, on parle de patients qui peuvent être aidés sur la base de l'expertise qui est partagée dans la plateforme. Donc pour prendre une comparaison, c'est un peu ce qui se fait avec d'autres types de plateforme pour l'enseignement avec les MOOCS, les cours en ligne. Ça c'est très important de partager ce savoir parce que c'est le seul moyen de rendre un savoir accessible, et dans notre cas de faire en sorte qu'en faisant levier sur l'expérience de différents utilisateurs et sur leurs connaissances, la plateforme et l'intelligence artificielle puisse en bénéficier et du coup disséminer ce savoir et cette expérience pour que le meilleur niveau de diagnostic soit appliqué partout pour tous les patients.

Et comment motiver ces grands labos qui n'ont pas forcément d'insentive pour lancer cette communauté, parce que ça fait déjà des années qu'ils travaillent sur leurs datas ?

Je pense que le seul moyen c'est, comme avec tout effet de réseau et comme c'est le cas aussi avec Tripadvisor... etc, c'est qu'au début tu as cette courbe de *early adopter*. Donc si tu veux justement dans les domaines où tu as des effets de réseau, c'est encore plus important parce que ceux qui sont moteurs ce sont ceux qui vont être les premiers à partager, qu'ils soient petits ou gros c'est une question de culture, comme ceux qui au début ont enregistré Tripadvisor avec des avis, et bien ensuite toi une fois que tu commences à en bénéficier c'est normal de le faire, donc tu le fais. Et il ne faut pas rêver, pas tout le monde ne le fera, exactement comme dans Tripadvisor pas tout le monde ne le fait. Mais si tu veux on va aller vers de plus en plus de partages. Par contre si on parle de valeur d'entreprise, le challenge là c'est la tentation et c'est pour ça qu'au niveau stratégique il faut bien tenir la barre du bateau, c'est que des fois tu peux aussi te dire que c'est égal et que tu vas accepter que les gens ne veulent pas partager l'information, tu vas accepter qu'ils veulent avoir le *software* en local, mais après si tu

acceptes ça, c'est un *short win*. Au contraire, tu ne te donnes pas la possibilité de tirer la plateforme et via la plateforme de faire le sur les effets de réseau d'un haut niveau du savoir et de la technologie, donc ça veut dire du bénéfice que tu vas apporter aux utilisateurs pour qu'ils soient ancrés à la plateforme et que tu ais de plus en plus d'utilisateurs. Mais au-delà de ça aussi, une fois que tu as la plateforme tu peux commencer à monétiser ça, à monétiser cette position.

Monétiser oui c'est ça. Et ce que j'ai vu dans la théorie c'est qu'il y a les effets de réseau directs, c'est simplement que plus il y a d'utilisateurs sur la plateforme, plus le prochain qui vient sur la plateforme aura une valeur plus grande de la plateforme, ce qui est logique ; et il y a l'effet de réseau indirect.

Le direct c'est typiquement plus de *flags* ont été faits, plus a utilisé la plateforme a des bénéfices ou plus il y a eu de tests que SOPHiA a mis en routine, plus celui qui vient bénéficie de tout le travail qui a été fait ça c'est le premier élément.

Et l'effet de réseau indirect c'est l'attraction entre deux côtés de la plateforme, entre deux types de groupes sur la plateforme. Et là je pense que dans le cas SOPHiA, comme on s'est concentré sur le *core business* et le *drylab* de SOPHiA au début, ce qui est intéressant c'est que vous avez attiré tout une communauté via l'IDS donc les autres labos qui n'avaient pas de séquenceurs eux sont attirés intrinsèquement par la plateforme.

Exactement. Il y a tous ceux-là qui ont été attirés oui, c'est juste. Parce que l'on s'est construit finalement un peu comme Airbnb ou Uber. Non seulement chaque nouvel utilisateur bénéficie de tout ce qu'on avait fait pour les autres mais en plus de ça, c'est devenu un modèle dans lequel on n'a pas investi et sur lequel on peut faire levier, comme tu dis, pour attirer d'autres types d'utilisateurs qui ne sont pas équipés et qui du coup deviennent intéressés par l'utilisation de la plateforme.

Et par rapport au type de monétisation, transaction ou *membership* ? Pourquoi choisir le *transaction fee* dans le cas de SOPHiA ?

Il y avait deux choses. La première c'est que l'on ne voulait pas mettre de barrière économique à l'entrée. Donc si tu étais parti sur un *membership*, ça aurait été à peu près le même prix pour un grand versus un petit. Alors que finalement le bénéfice que l'on apporte à l'utilisateur, c'est celui qu'on lui apporte à chaque fois qu'il diagnostique un patient. Et lui-même est remboursé, ou reçoit une subvention, pour un certain nombre de patients. Donc ça colle mieux déjà à la pratique, c'était pour ne pas pénaliser les petits qui ont des petits volumes et qui sinon auraient dû payer des gros *memberships*. Le deuxième avantage c'est que si tu as confiance en la forte augmentation du volume de consommation, à long terme à la consommation tu vas générer un volume d'affaires beaucoup plus important.

Plus important que le *membership* ?

Exactement. Alors sur le court terme peut-être pas, mais finalement si tu crois en la croissance de l'industrie, au volume tu vas faire beaucoup plus d'argent.

Et un mix des deux ?

Si tu veux quelque part c'est un peu indirectement ce que l'on fait avec le programme de validation. C'est-à-dire que dans cette étape, on leur fait payer pour quelque part mettre quelque chose en production donc c'est presque un *membership* aussi. Mais en même temps le *membership* si on le met comme ça clair et net, il aurait été un peu contre le modèle parce que par exemple nous on préfère qu'il y ait le maximum d'utilisateurs possibles, même au sein d'un hôpital. On se moque de combien il y a d'utilisateurs, à partir du moment où ils utilisent le maximum de volume possible. Et plus tu as d'utilisateurs peut-être, plus tu te dis que tu vas pouvoir générer du volume et plus tu es aussi ancré dans la mesure où si tu as beaucoup d'utilisateurs s'il y en a un qui part, ce n'est pas grave. L'expérience de la plateforme n'est pas perdue. Et plus tu as d'utilisateurs, et à la limite idéalement des utilisateurs qui interviennent à divers niveaux en tirant des données ainsi que de l'interprétation ainsi que des rapports, plus tu arrives à cimenter quelque part ta position dans l'entreprise, dans l'hôpital.

Le but aussi c'est d'intégrer des utilisateurs de différents types, aussi ? Pour encore plus souder la liaison entre les utilisateurs.

Oui exactement. Comme tu as dit, on a attiré tout un groupe de nouveaux clients avec NGS mais ce que l'on veut aussi c'est par couches, aller vers le plus haut. On a commencé par le labo, du labo on est passé au pathologiste, du pathologiste on veut passer à l'oncologue.

Hématologue aussi, qui pourrait être intéressé.

Voilà hématologue. On veut qu'en plus du labo, les spécialistes utilisent notre plateforme. Et un exemple, là c'est un intéressant, c'est que du coup notre plateforme devient même un outil marketing pour certains hôpitaux. Tu vois, on a certains labos de pathologie dans certains gros hôpitaux qui ont besoin d'un outil pour communiquer avec les oncologues, pour faire en sorte que les oncologues leur commandent des tests et les résultats. Et il y a certains oncologues qui commencent à en avoir marre d'utiliser du papier. Donc si tu as déjà une plateforme qui est jolie, qui est utilisée comme on l'utilise dans les labos de pathologie, il suffit de sensiblement adapter quelques écrans de l'interface pour laisser à la limite le labo de pathologie faire sa propre promotion, chasser plus d'oncologues avec le de l'hôpital, et d'enregistrer les demandes directement dans son compte dans la plateforme, et en plus à la fin une fois qu'il a testé le cas il peut même leur donner le résultat du rapport via la plateforme. Et le bénéfice pour SOPHiA GENETICS c'est que le pathologiste est encore plus content donc il est encore plus attaché à la plateforme. Il gère ses interactions via toi alors qu'en plus tu vas augmenter son parce qu'il va pouvoir vendre ses services donc

consommer ta plateforme avec plus d'oncologues. Et puis on commence aussi à voir des oncologues qui renseignent la plateforme avec d'autres sources de données, comme par exemple le stade du cancer, le traitement que le patients à reçu, ce genre de choses.

Les phénotypes aussi ?

Voilà le phénotype, exactement. Qui peut ensuite être du coup *crunché* par l'intelligence artificielle pour apporter encore plus de bénéfice à n'importe quel autre utilisateur.

Donc là ça rentre en plein dans les effets indirects à l'état pur.

Exactement.

Donc c'est ça qui est à développer chez SOPHiA.

Exactemnt, c'est un très gros objectif. Mais en termes de monétisation, quand tu as une plateforme on parle de monétisation de datas. Alors c'est vrai que SOPHiA ne monétise pas les datas, parce que c'est sensible.

Donc dans la transaction ce n'est pas monétisé alors ?

Comment ?

Les transactions par *samples* ce n'est pas de la monétisation ça, pour toi ?

Si si ça c'est monétiser mais ensuite on pourrait presque monétiser les datas qu'on stocke, on pourrait les vendre, on pourrait donner accès à ses datas pour faire de la recherche et ça on ne le fait pas.

Qu'est-ce qu'il y a dans les contrats, quand on signe avec un *customer*, est-ce qu'on peut utiliser leurs datas indirectement, commercialement quelque part ou c'est vraiment strictement interdit ?

Non ça tu ne peux pas, pas à cause des contrats mais à cause de la réglementation. En Europe maintenant depuis un an et demi il y a une réglementation sur la protection des données, la General Data Protection Regulation ou GDPR. Et cette *guideline* fait suite à un étudiant autrichien qui a poursuivi Facebook pour le droit à l'oubli. De ce que j'ai compris, celui qui a fait le lobbying pour qu'il y ait une réglementation européenne qui se mette en place pour respecter l'intimité des données d'individus c'est Microsoft, pour essayer de créer une industrie dans le domaine de la protection des données mais aussi peut-être pour se distinguer de ses concurrents, parce qu'ils pensaient surement qu'ils allaient être plus à même que les autres de respecter ça et à se distinguer, ils ont fait du lobbying pour qu'il y ait cette nouvelle *guideline*. Cette *guideline* définit deux rôles. Elle définit le rôle du contrôleur des données et celui du

processeur des données. Le contrôleur ça doit être l'individu, dans notre cas aussi les patients. Mais dans le cas d'un diagnostic moléculaire à l'hôpital, le patient donne aussi le contrôle des données à l'hôpital. Et ensuite le contrôleur, qui est l'hôpital, doit s'assurer que les données d patient vont être utilisées adéquatement en respectant certains critères. Et dans ce flow, SOPHiA est le processeur des données. Ça veut dire que SOPHiA peut utiliser les données pour la mission que lui a confiée le contrôleur, ça veut dire utiliser les données pour continuer à améliorer notre capacité à diagnostiquer les gens. Mais on ne peut pas exploiter les données par exemple en les vendant.

Mais donc si on met un logo d'AstraZeneca dans l'Oncoportail par exemple ?

Ca ce n'est pas un problème, c'est autre chose. Donc ensuite au niveau de la monétisation, une fois que tu as réussi à créer ta position de plateforme, la question c'est de savoir comment est-ce que SOPHiA GENETICS va monétiser. De 1, le fait qu'on ait plus d'utilisateurs nous a permis quelque part d'améliorer notre technologie ; en faisant ça l'offre est devenue plus importante, on a pu appliquer notre plateforme à plus de besoins d'application clinique et du coup notre portefeuille d'applications a grandi. Donc là c'est aussi un moyen de faire levier sur ta position de plateforme pour valider une application avec de nouveaux utilisateurs pour qu'ensuite cette application soit *available* à tout autre utilisateur sur la plateforme. Ça déjà c'est un des moyens, avec la technologie, de faire levier. Le deuxième moyen avec lequel on fait levier c'est, comme tu l'as décrit, l'IDS. On a pris position de plateforme, mais notre plateforme est devenue un portail d'accès aux capacités des laboratoires. Et un troisième moyen c'est que vu qu'on a cette position de plateforme, comme tu peux l'imaginer dans le futur avec par exemple Uber qui commenceraient aussi à vendre des choses dans les voitures. Nous on a commencé à commencer aussi à vendre, enfin on ne vend pas du Coca-Cola ou des entrées à l'opéra de Paris, mais on vend des de laboratoires. Donc avec cette position de plateforme on peut jouer un peu auEt maintenant le quatrième élément que tu décris, on ne peut pas vendre les données mais vu qu'on *crunch* les données de beaucoup de patients et que parmi eux peut-être 3000 sont en oncologie et matchent potentiellement en fonction de leur profil moléculaire avec des thérapies, et que l'industrie pharmaceutique met sur le marché des nouvelles thérapies mais que les oncologues ne sont pas forcément au courant, et bien nous là on est très intéressants aussi parce que on peut faire la relation entre le traitement et le patient. Et là tu as différents modèles. Ton modèle ça pourrait être d'être rémunéré par exemple sur le recrutement des patients pour les essais cliniques. L'industrie pharmaceutique pourrait nous payer parce que nous on pourrait leur dire « Vous pouvez trouver 10 patients dans tel centre ou tel centre qui pourraient participer à vos essais cliniques ». Mais ça nécessite quand même malgré tout un accord avec les hôpitaux.

Et on pourrait faire ça ?

On peut faire ça bien sûr. Il faut qu'ensuite les hôpitaux soient d'accord parce qu'on utilise leur data. Moi je pourrais dire « AstraZeneca vous avez 10 patients à Paris, 5 à

Bruxelles et 3 à Madrid qui correspondent aux caractéristiques, à des critères d'inclusion pour vos essais cliniques. En échange de l'aide au recrutement, vous nous payez tant ». Et au passage peut-être que d'ailleurs l'hôpital prendrait quelque chose aussi, ou l'hôpital serait content parce que du coup leurs patients ont été identifiés et peuvent rentrer dans les études. Ça ce sont des modèles qu'il faut affiner ensuite. Tant que tu n'as pas validé avec des premiers clients, tu ne sais pas exactement comment est-ce que ça se passe. Mais un autre modèle qui est différent ça pourrait être de ne pas faire payer en plus l'utilisation de l'Oncoportail au client, ce serait toujours le même prix pour l'hôpital, par contre dire à l'industrie pharmaceutique : « Vous savez nous aujourd'hui on voit 3000 cas dans les tumeurs solides par mois. Si vous voulez faire le meilleur *match*, que par exemple les hôpitaux sachent que vous avez des thérapies innovantes ou des médicaments qui ont été approuvés sur le marché, et que l'oncologue voit ce *match*, nous on a deux choix. Soit vous comptez sur ceux qui tiennent des bases de données à jour au risque de rater de l'information, soit on fait un travail de conciergerie pour vous. C'est-à-dire que vous nous payez chaque année pour maintenir cette information sur SOPHiA DDM de sorte que le *match* soit maximisé, vous vous engagez à nous donner l'information pour faire le chaque fois et nous on s'assure qu'il y ai un *match* parfait entre les profils moléculaires des patients qui correspondent à vos critères et vos thérapies ou vos essais cliniques ».

C'est un peu comme Google qui *match* les publicités à certains utilisateurs spécifiques.

Voilà, c'est ça. Tu as raison. Et c'est pour ça que l'important dans toutes les discussions qu'on a au board. C'est vachement important de faire super attention parce qu'à un moment donné, ton père te le diras donc tu peux lui en parler, ces Anglais auraient dit de trouver un modèle de monétisation assez vite. Si on avait fait ça, peut-être qu'on aurait fait plus de revenus, mais on aurait jamais eu le réseau, on aurait jamais eu le potentiel de la plateforme, donc ça aurait été une énorme erreur stratégique.

Oui, la plateforme aurait coulé.

Il n'y n'aurait pas eu de plateforme en fait. On aurait été en concurrence avec les hôpitaux et on aurait. Ça aurait été une erreur. Tu sais, une plateforme ça nécessite beaucoup d'ambition, ça nécessite de très bien connaître l'écosystème et ça nécessite d'avoir des investisseurs qui sont prêts à investir sur le long terme dans une entreprise très ambitieuse qui veut gagner toute cette industrie.

Parce qu'ils disent souvent que dans la plateforme on *reach* le *critical mass*, c'est-à-dire que la plateforme devient auto-suffisante dans le développement. Pour toi quel serait le *critical mass* dans SOPHiA DDM ?

Ca c'est difficile à savoir, il faudrait l'appliquer par application physique. Parce que par

exemple ce *critical mass*, dans notre cas c'est un peu compliqué, tu pourrais réfléchir en termes de *circulating tumor DNA*, quand est-ce que dans le CT DNA tu aurais un flux suffisant pour faire en sorte que la plateforme d'elle-même soit tellement intelligente qu'elle évolue toute seule par le nombre de nouveaux cas que tu rentres qui fassent ce que tu deviennes inévitable. Et ça c'est difficile à mesurer parce que tu as une industrie qui évolue ponctuellement donc on pourrait peut-être faire pour une application donnée.

En nombre d'utilisateurs sur une application ?

Voilà, en nombre d'utilisateurs, en nombre de patients à mon avis plutôt.

Juste si je reviens au *flagging*, tu sais comment on pourrait créer plus d'*insentive* aux gens pour le *flagging* ? Parce qu'au jour d'aujourd'hui, je ne connais pas encore le chiffre du *flagging* mais je me demandais s'il y avait d'autres moyens pour pousser les gens à faire du *flagging*. Ça c'est SOPHiA GENETICS, c'est trouver cette communauté par pour qu'elle travaille pour toi mais pour qu'en ayant des initiatives elle crée quelque chose naturellement, si je puis dire.

Tu as tout à fait raison, mais là il faut faire très attention parce qu'à mon avis, tu as des bons et des mauvais *insentives*. Le mauvais *insentive* c'est dès que tu te demandes de manière un peu perverse si tu ne devrais pas les faire payer, leur donner ça... etc. Ça c'est du mauvais *insentive* parce que ça veut dire que celui qui fait ça n'a pas un réel bénéfice en le faisant. Je vais te prendre un exemple : dans les tumeurs solides, on peut *clusteriser* les patients dans des groupes en fonction du profil moléculaire, du type et du stade de tumeur, du traitement qu'ils ont reçu et de l'efficacité de celui-ci. Ça c'est notre vision pour que dans le futur, quand quelqu'un mettra le profil moléculaire d'un patient dans la plateforme, elle nous dise quelle est le type et le stade tumeur mais également « Vous, Monsieur, votre tumeur fait partie d'un *cluster* de patients dans lequel 10% ont reçu ce traitement et ont survécu au cancer ». Donc là aujourd'hui on compte sur la bonne volonté des gens à mettre le traitement qu'ils ont donné et si ça a fonctionné. Mais eux ont très peu d'intérêt à faire ça puisque la plateforme ne leur donne pas un meilleur *output*. Tu peux te dire « Je vais donner des tests gratuits à ceux qui font ça », mais ça à mon avis c'est une mauvaise approche. Il faut se creuser la tête jusqu'à trouver comment contourner ça. Et par exemple, c'est ce que je t'ai dit avant, ces labos de pathologie ont besoin d'un outil marketing. Les oncologues leur disent « Non, on en a marre de remplir des feuilles sur lesquelles on doit décrire l'état du patient et faire la commande de tests. Pourquoi est-ce que vous ne construisez pas un outil informatique pour faire ça ? ». Du coup pour les labos de pathologies qui eux produisent l'information moléculaire, tu deviens intéressant. Tu leur dis « Mais attendez, c'est ça le besoin des oncologues pour que vous puissiez attirer plus d'oncologues sur la plateforme ? Pas de problèmes, on va vous le faire ». Et en échange, qu'est-ce qu'on gagne ? Non seulement plus d'oncologues pour prescrire plus de tests à ces pathologistes qui travaillent avec nous, mais en plus de ça les oncologues vont remplir cette information pas pour ceux dont nous on a besoin mais juste parce que ça

les arrange.

Oui, effet de réseau indirect.

C'est le réseau indirect. Et là il y a un vrai bénéfice, parce que le gars ne veut pas refaire du *paperwork* et il veut pouvoir gérer ses cas.

Mais c'est remplir les besoins de chacun en ayant un échange.

Exactement. Et donc souvent la difficulté dans les plateformes et dans comment faire levier, c'est de ne pas tomber dans la facilité, de mon expérience. Si tu tombes dans la facilité, tu construis quelque chose d'artificiel.

J'ai fait une dizaine d'interviews là et ils poussaient plutôt à payer.

Mais oui mais voilà, c'est la facilité. Au final, tu ne finis pas par créer quelque chose qui adresse vraiment le besoin d'un utilisateur de manière simple, tout en faisant bénéficier ta plateforme d'autre chose ou un autre utilisateur. Ça ne va pas. C'est là où il faut faire attention parce que quand tu as trop d'argent, tu es tenté de dire « On vient de lever 50 millions, on va faire des choses gratuites pour le truc avec ça ». Quand tu as moins d'argent, tu es obligé d'être créatif.

Ils ont dû bien réfléchir à ça dans les plateformes du marché, finalement c'est ça qui est essentiel. C'est ça le paradoxe aussi dans le sens où j'ai vu plusieurs personnes aux Etats-Unis qui me disait « Vous SOPHiA GENETICS, les gens vous donne des données et en plus ils payent pour ça ». C'est ça un peu le paradoxe de SOPHiA, comme si c'était un vol.

Oui bien sûr mais ça on l'a partout, et encore pire en Europe. Et encore pire au début, sinon on rapporterait très peu de bénéfices. C'est pour ça que ça ne peut marcher que si tu leur donnes un bénéfice. Il faut leur expliquer que non, ils ne nous donnent aucun data, c'est nous qui vous apportons en calculant vos données et en vous donnant un résultat propre.

Au début ça a dû être très dur pour vous, la plateforme ?

Ça a été très difficile, oui. Mais ça c'est vachement important. C'est une question de survie ensuite de te concentrer sur ce que tu dois faire pour que les gens perçoivent dans l'utilisation de la plateforme un bénéfice suffisant pour qu'ils soient prêts à payer. C'est pour ça que c'est culturel. Personne d'autre je crois en *healthcare* n'a réussi à créer un réseau dans lequel tu as une plateforme qui est utilisée par 320 hôpitaux.

Et si tu disais que gratuit c'était dans le sens où on peut faire payer les gens qui ont un intérêt financier derrière les *samples*, et ne pas faire payer pour les académies par exemple, qui n'ont pas d'intérêt financier derrière ?

112.

Ca ne servirait à rien parce que les volumes ne sont pas dans l'académique, ils sont dans la routine. Ca génère plus de volume que n'importe quel académicien. Vas-y j'ai encore deux-trois minutes.

C'est bien merci pour tout, c'était génial, j'ai pris un peu trop de ton temps là. La dernière question c'est une micro-question après j'ai juste autre chose à te demander. On se dirige vers un monopole dans les plateformes de l'analyse NGS, dans le sens où plus on aura une base de données forte, plus il y aura de *main* acteurs sur le marché ? Un peu comme Facebook et les réseaux sociaux.

Je te laisse tirer tes conclusions.

D'accord.

Dans le milieu des plateformes tu en as un dominant, ce n'est pas le seul mais il y en a un dominant. Wikipédia est perçu comme un gentil donc Wikipédia.

C'est bien, je vais faire une belle conclusion sur ça.

D'où l'importance de faire très vite grandir ton réseau et de ne pas uniquement penser à maximiser le revenu, sinon tu es mort.

Je sais, c'est ce qu'ils disent tout le temps. Et vraiment limiter le plus possible les frontières aussi, pour les nouveaux utilisateurs.

Oui exactement.

Et j'ai eu mon promoteur et il m'a dit que ce serait intéressant de comparer avec des concurrents dans la partie pratique mais je me suis fait un peu à OncoDNA, ils m'ont dit que ce n'était pas possible. Je voulais faire une interview mais ils m'ont dit : « C'est quoi ton nom de famille ? »

Tu ne peux pas interviewer les concurrents mais... Bon OncoDNA ils parlent d'un réseau d'oncologues, ce qui est différent.

Oui mais sur OncoSHARE la plateforme était intéressante, je trouvais.

Oui mais c'est très différent quand même. Si tu devais voir nous combien on a d'utilisateurs, eux ils parlent d'oncologues et dans chaque hôpital tu as entre 100 et 1000 oncologues. Mais tu devrais plutôt prendre l'exemple de Pirndx.

C'est ça que je voulais te demander, si tu connaissais quelqu'un que je pouvais contacter dans Pirndx.

Contacté non, c'est sensible là.

Non c'est bien, je vais regarder les informations sur leur site Internet, je vais voir. Parce que j'ai Pirndx écrit en grand sur ma feuille parce que c'est celle-là qui me paraît le plus semblable au phénomène.

Oui exactement, le plus semblable. Après tu aurais aussi Factory Genomics,

l'ancien Omishia mais eux ne parlent pas des données mais ils vont essayer de faire la même chose.

Appendix 3.2. Dr. Jean Francois Vanbellinghen

Dans le mémoire, ce que j'étudie ce sont principalement les effets de réseau qui sont créés sur la plateforme, c'est-à-dire le nombre d'utilisateurs : plus il y a de nombre, plus il y a d'utilisateurs sur DDM et plus la valeur globale de la plateforme augmente logiquement parce que la technologie derrière s'améliore. Donc j'essaie de voir quels sont les effets de réseau sur la plateforme, comment les capturer et comment monétiser cette plateforme. Dans la partie théorique, je prends un peu tous les cas assez connus de type Facebook ou Airbnb et j'essaie un peu de voir ce qu'il se passe dans l'analyse génétique avec le NGS et dans le cas de SOPHiA, et peut-être dans celui d'un concurrent aussi. J'essaie de voir un peu le *business model* que fournit SOPHiA, s'il y a d'autres alternatives... Et toi tu as donc utilisé SOPHiA pas seulement en tant qu'ISIMI mais aussi en tant qu'utilisateur tu disais ?

J'ai indirectement utilisé SOPHiA, c'est-à-dire que j'ai fait utiliser SOPHiA par mon équipe, c'est surtout ça. On a effectivement évalué l'intérêt de SOPHiA chez nous. On a opté aussi pour SOPHiA parce qu'on est parti sur notre solution HCS qui correspondait typiquement à ce moment-là à un besoin de changement au niveau belge et on s'est dit « D'accord, on passe aussi là-dessus ». On avait les CNB, donc on avait quelque chose au point de vue qui était parfait pour nous, et puis après il y avait la plateforme en elle-même. On avait besoin non seulement déjà simplement d'une *dead pipeline* informatique, et ce qui nous posait problème c'est que l'on s'était rendu compte que même si on avait un bio-informaticien, on n'arriverait pas à avoir quelque chose de correct et que justement le fonctionnement idéal de ce genre de *pipeline* c'est si on a beaucoup de *feedbacks* d'un maximum d'échantillons, et forcément on a plus facilement beaucoup de *feedbacks* avec beaucoup d'échantillons si on a beaucoup de clients. Ça a été un argument qui nous a fait choisir SOPHiA GENETICS : on s'est dit « Derrière il y a une équipe qui est beaucoup plus grande, ils ont des données qu'ils reçoivent massivement et dès lors ils peuvent automatiquement améliorer, par un système de *feedbacks*, le système ». Sauf que n'importe quelle institution a du mal et encore plus une petite institution comme on était.

Et si tu devais dire la valeur principale que DDM vous a apporté, le gros plus de SOPHiA DDM ?

Le vrai premier argument c'était le fait que nous n'avions pas de bio-informaticien. Quel que soit ce que SOPHiA DDM allait nous proposer, à partir du moment où l'on nous proposait au minimum un *pipeline* solide et qu'on voyait qu'il pouvait y avoir une mise à jour régulière de celui-ci, pas une mise à jour du *software* ou de la plateforme mais réellement du *pipeline*, nous c'est ce qu'on cherchait. Parce que dans notre système de qualité, on veut avoir un *pipeline* mais on veut aussi que le *pipeline* puisse être revu, amélioré en continu, et SOPHiA GENETICS nous proposait ça. Après effectivement le fait d'avoir une plateforme DDM avec des outils de base très faciles à

utiliser, c'était un plus. D'avoir une visualisation sur ce que la communauté faisait, c'était aussi un plus.

Par rapport à la communauté, est-ce que les labos appréhendent ou est-ce qu'ils sont attirés par celle-ci ? Qu'est-ce que tu en penses ?

Globalement à mon avis on est attiré par la communauté, simplement parce que c'est déjà comme ça que ça fonctionne sans plateforme comme SOPHiA DDM. Quand je devais rapporter un variant sur lequel j'hésitais un peu trop malgré des publications, je partageais cette information avec un ou deux laboratoires belges qui traitaient les mêmes pathologies, ou avec des laboratoires européens ou de n'importe où dans le monde, ça ne me gênait pas. On fait déjà ça.

Sauf que sur la plateforme on ne voit pas le nom des institutions.

Non on ne voit pas le nom mais on a un genre de visualisation qui nous montre qu'on n'est pas les seuls à avoir trouvé quelque chose.

Et ça reconforte ?

Oui c'est typiquement humain, aussi bon soit-on, c'est toujours intéressant de voir que quelqu'un d'autre a trouvé la même chose et a décidé de lui attribuer la même pathogénicité. Hésiter sur ce variant qu'on mettrait bien comme étant totalement bénin et voir qu'un laboratoire l'a mis aussi comme étant bénin, c'est déjà ça, alors quand il y en a 22 à ce moment-là on n'hésite plus du tout. Ça te fait gagner beaucoup de temps et ça te conforte petit à petit dans le fait que tu as un bon jugement puisque ce que tu aurais classé comme 1 se classe comme 1.

Est-ce que c'est plus difficile de convaincre de l'aspect communautaire les grands labos, parce qu'ils ont déjà une grosse base de données, que les petits labos qui évidemment sont plus fervents ?

Oui forcément pour les petits labos c'est plus attrayant, parce que qui dit petits labos dit en général pas d'expérience ou très peu et donc on leur offre une expérience. Pour les gros laboratoires, à mon avis ça a effectivement sans doute moins d'intérêt pour ceux qui travaillent avec des petits panels parce que forcément un laboratoire aura plein d'IPG, je ne sais pas combien ils font de HCS, disons à priori un *run* par semaine ou tous les quinze jours avec cinquante échantillons. Donc forcément sur leur année, ils ont cumulé tellement de données que rien qu'avec celles-là ils peuvent pratiquement déjà résoudre tous leurs cas. Après c'est vrai avec des petits panels, mais quand tu commences à faire des plus grands panels, même si tu es un gros laboratoire il est moins facile d'avoir une expérience sur tous les gènes d'un énorme panel. Si tu partages, c'est d'autant plus facile et ça va sans doute atteindre son apogée avec le *whole exome*.

C'est ça, avec le *whole exome* si on suit ta théorie on sera davantage dans un concept de partage alors ?

Pour moi oui c'est vers ça que l'on doit aller, c'est-à-dire avoir une vraie vision de la totalité de ce qui peut être trouvé dans un génome, et forcément pouvoir après trier par origine ethnique ou avoir des filtres. C'est sans doute un des défauts actuels des statistiques que l'on voit pour l'instant sur SOPHiA GENETICS et sur la communauté, c'est que tout est totalement mélangé. Ça introduit des billets qui sont difficilement appréhendables, tu vois des données qui sont un mélange. Tu sais que ce ne sont que des laboratoires de diagnostic et des laboratoires qui ont validé leur *pipeline* avec SOPHiA GENETICS qui contribuent à ces données-là, mais tu ne sais pas d'où viennent les échantillons, la région géographique, mais aussi quelle est la contribution... J'ai un exemple, je suis tombé dessus par hasard en Turquie : je regardais un variant qui avait été trouvé chez un patient atteint d'une pathologie neurologique. Comme exercice avec le laboratoire turc, on avait fait un panel virtuel avec des gènes liés à ce type de pathologies et on a trouvé un variant qui était a priori pathogène. Donc j'ai dit qu'on avait a priori trouvé le variant. Sauf que ce gène par hasard je le connaissais, il était impliqué dans des fièvres méditerranéennes familiales qui sont forcément, d'où le nom, plus fréquentes sur le pourtour méditerranéen et entre autre en Turquie où il y a une très forte présence. Vu que l'on combine tout sur SOPHiA DDM, par exemple cette fièvre méditerranéenne familiale c'est quelque chose dont on a un kit qui ne contient qu'un seul gène, le gène MEFV, et c'est une analyse qui est faite très souvent. En fait c'est un diagnostic qui est très facile, c'est-à-dire que les gens qui ont une fièvre méditerranéenne familiale, n'importe quel médecin les reconnaissent, le test génétique est simplement utilisé pour confirmer. Donc le nombre d'échantillons qui sont testés positifs est énorme, sur 1000 échantillons tu as beaucoup de chances d'en avoir 1 sur 3 ou 1 sur 2 qui est positif. Toutes ces données-là se retrouvent dans SOPHiA DDM et ont l'air de montrer que dans la population, dans la communauté SOPHiA GENETICS, on a une très haute fréquence de ce variant. On peut se dire alors que ça doit être bénin parce qu'une personne sur dix l'a dans la communauté, mais en réalité ce dont on ne se rend pas compte c'est que cette personne sur dix, c'est simplement parce qu'elle devait l'avoir, parce que la pathologie elle l'avait et que c'est pour ça qu'on l'a testée. Donc ces statistiques de la communauté SOPHiA GENETICS sont à prendre avec des pincettes parce qu'elles sont forcément biaisées car dans la communauté il n'y a que des gens malades.

Donc il faut faire attention.

Oui. Quand j'étais client je ne l'avais pas perçu comme ça, d'ailleurs je ne l'avais tellement pas perçu comme ça que quand l'auditeur qui est venu auditer mon laboratoire m'a posé des questions sur SOPHiA GENETICS et sur cette fréquence-là, en me demandant si j'estimais que la fréquence que j'observais là devait se rapprocher de la fréquence que j'observais moi au sein de mon laboratoire, je lui ai dit qu'on pouvait imaginer que ça tendrait vers ça mais qu'il y avait forcément un effet sans doute

régional qui allait s'y opposer. Mais jamais à ce moment-là je n'avais pensé que j'avais moi dans certains cas des patients pas très bien sélectionnés et donc moi je cherchais une mutation, alors que d'autres ont des patients très bien sélectionnés et ont beaucoup de chances de trouver la mutation. Donc ça donne des statistiques vraiment biaisées.

Dans le réseau, ce que l'on voit dans SOPHiA c'est évidemment qu'il y a deux types d'effets. Tu as en numéro 1 l'effet direct où il y a plusieurs utilisateurs et donc la valeur augmente proportionnellement, et en numéro 2 des effets qui sont moins connus dans SOPHiA DDM, ce sont les effets indirects donc entre utilisateurs qui s'attirent mutuellement grâce au partage de *flagging*, à l'échange de données ou de savoirs par exemple. Mais ça on le voit moins chez SOPHiA pour l'instant, tu es d'accord ?

Je pense aussi qu'il y a encore aussi pas mal de labos qui ne *flaggent* pas.

Exactement.

Il y a deux parties aussi. Il y a la partie des fréquences de la communauté, ce qui dans certains cas peut être utile mais dans d'autres biaisé, et après il y a le *flagging* où tu vois ce que d'autres personnes ont fait. Ça le problème c'est que des fois quand tu vois certains gènes alors que tu te dis que ce sont des gènes relativement étudiés, c'est étonnant de voir aussi peu de *flags*. Ça veut dire que beaucoup de clients sans doute ne *flaggent* pas encore.

Et pourquoi ils ne *flaggent* pas ? Est-ce que tu aurais de solution pour les pousser à *flagger* ?

En fait les gens ne *flaggent* pas parce que dans certain cas c'est énormément de travail de *flagger*, c'est-à-dire qu'ils *flaggent* quand ils ont quelque chose de positif. Ils passent en revue leurs listes et ils disent « Ça c'est pathologique donc je vais le *flagger* », donc ils *flaggent* ça comme un 4 ou un 5. Par contre tous les polymorphismes qu'ils connaissent très bien, ils ne prennent pas le temps de les *flagger* comme étant bénins, parce qu'effectivement c'est soit connus par eux, soit il faudrait qu'ils creusent un peu car l'action de *flagger* ne doit pas être prise à la légère. Ton nom est associé à ça et ça veut dire que tu as suffisamment de connaissances pour considérer que ceci est bel et bien à *flagger* comme un 1, un 2 ou un 5. Mais il y a une vraie responsabilité et je pense que ça c'est aussi un problème. Certains scientifiques se sentent très à l'aise, si ça pour eux c'est un polymorphisme ils *flaggent* ça comme 1, leur nom est associé à ce *flag* dans leur compte et il n'y a aucun soucis. Il y en a d'autres qui ont un peu plus de mal : pour ce qui est très clair ça va, que ce soit très clairement pathologique ou très clairement bénin, mais pour la partie centrale où c'est un peu plus compliqué, plutôt que de *flagger* quelque chose de faux certains ne *flaggent* pas.

S'ils sont sûrs à 80% ils ne vont pas le faire donc, il faudra 100%.

Voilà, et d'un côté je comprends le réflexe. Il s'agit de se dire « Ca je ne suis pas certain, si je le *flagge* ça veut dire que la prochaine fois que je tombe dessus je ne ferais même plus attention étant donné que je lui aurais déjà attribué une valeur de 1 par exemple, or ça nécessite peut-être de faire quand même attention ». Donc ils ne *flaggent* pas ces variants-là en se disant que comme ça la prochaine fois qu'ils retomberont dessus ils seront de nouveau non-*flaggé*, leur attention sera attiré par ça et ils réfléchiront. Je pense donc qu'il y a des variants comme ça qui ne sont jamais *flaggés* tout simplement parce qu'il faudrait peut-être une bonne fois pour toutes prendre une décision, et pour prendre une décision il faudrait passer un jour ou deux, c'est-à-dire faire de la biblio, discuter entre collègues pour savoir ce qu'on fait, si on met le variant en 1, en 2 ou si on le laisse en 3. Tout ce travail-là est quand même important et c'est sans doute aussi un frein au *flaggage*.

Donc en même temps il y a un peu de paresse disons, mais c'est deuxièmement c'est aussi parce que ça demande du travail dans la recherche car il faut être sûr. Oui, il y a une vraie responsabilité.

Parce que ce qui est intéressant avec SOPHiA c'est qu'ils ne travaillent pas pour nous mais en restaurant ce *flagging*, c'est comme si ce travail qui est dur à faire ils le font pour la plateforme, pour la communauté.

Ils le font pour la communauté effectivement. Je pense qu'il y a d'ailleurs des laboratoires en France, il faudrait voir avec quelqu'un qui a travaillé en France moi je ne l'ai jamais fait, qui ne *flaggent* pas parce qu'ils ne veulent pas *flagger*. Ils ont leurs données dans une base de données et ils ne veulent pas donner le contenu de celle-ci en *flaggant*, c'est-à-dire qu'ils ne veulent pas fournir des informations qu'ils ont rassemblées sur des années d'expérience à SOPHiA GENETICS et risquer que SOPHiA, directement ou indirectement, ne les vendent. Là aussi c'est un argument qui se tient parce qu'effectivement une grosse partie de la valeur de la plateforme c'est cette connaissance de *flagging*, et qu'indirectement on la vend, c'est-à-dire qu'on la vante, comme étant un outil intéressant pour le client débutant. Les clients ne savent pas à quel point SOPHiA GENETICS pourrait un jour la vendre carrément en disant « Nous on peut vous donner des statistiques sur la décision de 300 clients pour des millions de variants ».

Il n'y a pas un chiffre pour le taux de *flagging* dans SOPHiA ?

Je ne sais pas comment on peut l'estimer.

Oui c'est ça, parce que je voudrais bien voir sur quelle tendance on se dirige, si on se dirige vraiment sur du *flagging* ou... Je sais qu'en France il existe des clubs avec 3 ou 4 labos partenaires qui partagent entre eux, des micro-consortiums.

Je pense que c'est vraiment une différence d'esprit de certains laboratoires. Certains

jouent le jeu entre guillemets et se fichent réellement qu'un jour ou l'autre SOPHiA GENETICS puisse vendre ces données-là, parce que finalement le bénéfice qu'ils tirent est largement supérieur à leur contribution, d'après eux. Les petits labos ont sans doute l'impression qu'ils ne contribuent pas beaucoup par rapport à ce qu'ils tirent comme avantage, mais forcément les plus gros laboratoires vont sans doute avoir à se dire : « Je donne des milliers d'informations super intéressantes et en plus je paye pour ça ». Ça met donc les laboratoires, en fonction aussi sans doute des pays et de la mentalité, dans des situations totalement différentes. Je sais qu'un laboratoire en France qui fait du HCS et qui normalement ne *flagge* pas parce qu'il est regroupé à je ne sais quel *consortium*, a décidé de commencer à *flagger* parce que finalement une grosse partie des données est vérifiable autre part. Ils se disent donc : « Pourquoi ne ferait-on pas ce *flag* puisque l'on peut vérifier les données autre part ? ». Donc ils vont sans doute commencer à *flagger*. Mais c'est une vraie décision, il y a sans doute une question d'éthique aussi derrière tout ça et les laboratoires français sont beaucoup plus conservateurs et moins tournés vers un partage s'ils ne savent pas exactement ce qu'il en est.

Tu aurais une idée de comment on pourrait pousser ces *flagging* ?

Donner de l'argent. Non, je plaisante.

Shirine me disait que s'ils *flaggent* X nombre de variants, peut-être qu'aux prochaines analyses ils pourraient avoir 20% en moins.

Le problème c'est que c'est aussi le meilleur moyen pour que les potentiels *flaggeurs* non consciencieux fassent n'importe quoi. Le système fonctionne justement sur la bonne volonté, l'intérêt de *flagger*, une connaissance et tout ça rentre dans une équation qui fait que l'on peut difficilement pousser. A mon avis, la seule chose que l'on peut faire pour promouvoir, c'est sensibiliser les clients sur le fait que s'ils tirent profit du *flagging* des autres, c'est dès lors intéressant qu'ils *flaggent* eux-mêmes par solidarité.

Ce n'est pas évident ça.

Non, mais c'est réellement la chose la plus efficace à mon avis. Faire un incitatif financier pour moi c'est quelque chose de dangereux, parce que certains laboratoires, les plus consciencieux, le font bien et le font déjà, et ceux qui se disaient qu'ils ne voulaient pas faire ça vont se dire qu'en prenant une décision un petit peu plus rapidement que ce qu'ils auraient normalement fait, ils se retrouvent avec 20€ par *flag* en plus... Il peut donc vite y avoir une dérive. Certains laboratoires forcément ne dériveront pas du tout, mais d'autres clairement à mon avis se diront : « Je suis intéressé par les gènes BRCA-1 et -2, de toute façon toutes ces données-là je ne m'en sers pas donc je peux *flagger* n'importe quoi ». Ça se gère bien sûr, mais je serais quand

même inquiet de voir un incitant à un *flag*. C'est comme quand on dit que l'on est prêt à payer des reins, il y a des gens qui vont vendre leurs reins alors qu'ils les auraient bien gardés. Si on dit aux gens « Donnez vos reins si vous voulez », tu te rends bien compte qu'il y a beaucoup moins de gens qui veulent donner un rein par rapport à si tu leur donnes de l'argent en échange. A mon avis c'est un peu du même niveau, c'est-à-dire que c'est difficile d'inciter les gens à faire un bon travail aussi anonyme, je ne dis pas que les gens ne travaillent bien que si on les paye, mais là c'est un travail anonyme donc dire qu'on les paye un peu, qu'on leur donne un avantage s'ils *flaggent*, c'est risquer d'avoir un biais. Il vaut mieux qu'il y ait peu de *flags* de super bonne qualité que de nouveau 50% de *flags* avec une qualité moins bonne. Donc à moins avis le seul incitant que l'on peut utiliser, c'est attirer l'attention sur le fait que si les clients vont vers une solution SOPHiA GENETICS pour profiter de la société et de la communauté, un retour est le bienvenu. C'est comme les impôts, on se dit : « Au chômage c'est intéressant d'avoir un salaire, alors ne râtons pas trop quand on doit donner de l'argent ».

C'est pour la communauté donc.

Oui c'est vraiment cet esprit-là, donc pousser financièrement à faire du *flagging* c'est délicat pour nous aussi. Parce qu'automatiquement le client peut se dire « Si SOPHiA GENETICS nous pousse autant à faire du *flagging* et est prêt éventuellement à nous payer, c'est que ça leur rapporte plus ». Et de nouveau ce n'est pas un super beau message envoyé. Mais bon, je suis très sensible à ce genre de message.

Parce que nous aujourd'hui le *business model* c'est de charger par *sample uploaded* en transaction fee, pourquoi a-t-on choisi ça par rapport à un *membership*, une suscription d'un an par exemple ?

Et bien en voyant ce système-là, et c'est là que c'est bien fait, tu as l'impression d'être moins lié parce que tu te dis que finalement si tu souscris pour 10 000€ sur l'année et que tu en fais plus, tu auras gagné, mais si tu en fais moins ou beaucoup moins tu auras vraiment perdu. Donc se dire que c'est adapté à l'utilisation qu'on en fait, pour moi c'était un argument qui me disait que ça me rendait un peu moins dépendant. Après c'est faux forcément parce qu'on se rend bien compte que lorsqu'on signe avec SOPHiA GENETICS un programme de validation, même si SOPHiA nous le donne presque gratuitement au début, c'était le cas pour l'HCS au début, on est en fait d'autant plus lié à eux parce qu'on n'a pas payé le programme de validation, on a eu une réduction de 50% sur le premier kit et toute notre routine, notre accréditation dépendent maintenant de ce *pipeline*. Donc forcément maintenant on va continuer en espérant que les prix n'augmentent pas et tu te retrouves finalement lié. Mais j'ai l'impression qu'à l'utilisation ça a un côté plus attirant pour les petits laboratoires, parce que justement même si les prix sont plus élevés, ça reste quelque chose de très correct. Typiquement quand je vois certains laboratoires qui sans doute avaient planifié de faire 100 ou 150 échantillons sur l'année et qui finalement, parce qu'ils ont trainé dans la validation ou la mise en place, ne vont pas en faire beaucoup, dans un cas où tu as

déjà payé 10 000€, ces échantillons te reviennent super chers. Alors que si c'est payé au coup par coup, tu as pris du temps pour démarrer, c'est de ta faute mais au moins tu n'es pas trop pénalisé.

Et qu'est-ce que tu penses entre ça et un *membership* qu'ils ont appelé le *ad valorem fee*, avec lequel tu as un *fix fee* pendant un an qui est assez faible plus une *fee*, une charge proportionnelle au nombre d'utilisations que tu fais. C'est un peu la même chose ?

C'est un peu la même chose effectivement, mais vu que c'est une solution intermédiaire ça aurait été plus facilement acceptable de payer disons une certaine somme par an supposée ouvrir un nombre minimum d'échantillon et des maintenances de la plateforme, et puis pour tous les échantillons supplémentaires avoir un prix dégressif. Ça aurait été pour moi plus séduisant que d'avoir une grosse somme, mais de nouveau ça dépend. Moi j'étais dans un laboratoire pas très grand et j'avais beaucoup d'espoir qu'on se développe mais je n'aurais pas voulu prendre énormément de risques. Il faut bien se dire aussi que derrière ce genre d'investissement, il y a forcément une défense de budget dans le labo, au niveau de l'administratif. Donc de nouveau quand ce sont de gros laboratoires avec une très bonne connaissance qui savent qu'ils n'augmenteront pas forcément leur débit d'une année à l'autre, c'est très facile de budgétiser quelque chose d'une grosse somme parce qu'ils ont la certitude qu'ils vont revenir dans leur prêt. Par contre quand ce sont des laboratoires qui débutent, on en a énormément en NGS, si on leur dit qu'ils vont devoir sortir 5000€ de base et que s'ils font plus de 100 échantillons ce sera facturé 30€ l'échantillon, c'est facile s'ils sont sûrs d'arriver à ce petit nombre mais s'ils n'y arrivent pas ils auront payé beaucoup plus. Et de nouveau, beaucoup de laboratoires ont des soucis de budget fixe. Moi pour SOPHiA GENETICS j'espérais avoir par exemple au final 500 échantillons qui me coûteraient 200€ donc ça doit faire 100 000€, or budgétiser une dépense de 100 000€ d'office comme ça sans garantie que j'aurais les échantillons pour être rentabilisé, c'est compliqué, et ça peut être aussi facilement refusé par certaines institutions. Par contre un genre de facturation à l'usage, ça ressemble plus à un réactif et c'est dès lors souvent des budgets différents : ça ne paraît pas être un investissement de *software* mais ça paraît être un produit de laboratoire où plus on aura d'analyses et plus on achètera un produit, ce qui, je suppose que tu le sais sans doute mieux que moi, pour la comptabilité n'est pas la même chose.

Exactement oui, Stéphane l'a dit aussi. Et est-ce que tu verrais un modèle où SOPHiA pourrait être gratuit pour certains utilisateurs comme l'académie, et payant pour les institutions qui ont un intérêt financier derrière ?

Je réfléchis parce que je n'y ai jamais pensé. Tu imaginerais que l'académie pourrait accéder gratuitement à... à quoi ?

Elle pourrait *upolader* ses *samples* gratuitement.

D'accord mais donc pour des kits biens définis pour lesquels on a des *pipelines* développés. Pas pour des *pipelines* en développement.

Exactement oui.

En général j'imagine que la majorité de la partie universitaire focalise plus son attention justement sur des choses en recherche et donc un peu à côté de la routine et sans kit déjà disponible. Donc à mon avis, on se retrouverait à ne pas avoir énormément de clients car il ne doit pas y avoir beaucoup de laboratoires de recherche qui utilisent une solution déjà disponible dans le commerce. Ça doit sans doute arriver, mais de mon expérience tout ce que l'on faisait en recherche en parallèle dans les différents labos par lesquels je suis passé, on était bien loin de respecter des kits parce qu'on essayait de faire au moins cher ou on essayait de trouver de nouveaux gènes et dès lors on avait qui des fois étaient mêmes secrètes et qu'on aurait pas aimé partager avec un client qui serait forcément un peu commercial. Donc à mon avis c'est compliqué, même si l'idée d'amener des universitaires dans le jeu serait intéressante. Le problème aussi c'est que dans ces cas-là il faudrait faire attention avec le *flagging* parce qu'en général ils n'ont pas du tout la même notion que des laboratoires de diagnostic donc on devrait faire carrément des bases de données différentes, de *flagging* en tout cas.

Parce qu'il y a des plateformes gratuites et elles ne marchent pas très bien non plus. Donc le prix n'est finalement pas une barrière pour les utilisateurs.

Le problème c'est que les plateformes gratuites sont justement des plateformes sur lesquelles tournent des algorithmes de base, sans aucune customisation.

Oui ils n'utilisent qu'un *pipeline*.

C'est ça, donc tu as des possibilités de régler j'imagine mais tu dois le faire toi-même. Si SOPHiA DDM réussissait à ouvrir un accès à des personnes qui font tourner des *pipelines* qu'elles doivent régler, parce qu'on aurait quand même des réglages accessibles aux clients... En même temps c'est aussi quelque chose dont je ne connais pas très bien la vision que l'on a concernant SOPHiA GENETICS, mais l'une des choses qui a été utilisée comme argument de vente pendant très longtemps et qui est encore utilisée pour l'instant, c'est que la recherche ce n'est pas du diagnostic et que SOPHiA GENETICS fait vraiment des choses au niveau diagnostic. S'il tu mélanges avec la recherche, le message que tu vas envoyer ne sera peut-être plus aussi clair, même si effectivement tu rattrapes beaucoup de clients de recherche, est-ce que tu ne dilues pas un peu cet aspect qualité que tu montres d'habitude aux clients diagnostic... ? Et de nouveau je pense que la spécificité de SOPHiA GENETICS qui moi m'avait plu, c'est que ça faisait vraiment plateforme dédiée au diagnostic, par rapport à des plateformes qui effectivement n'incluaient pas le travail de validation d'un *pipeline* et qui demandaient moins, mais où il y avait un *free access*, enfin tu payais ton *software* mais après tu

faisais tourner autant que tu voulais, et où tu devais tout construire toi-même et où ce n'était pas facile de voir ce qu'il allait y avoir à régler. Dans ma vision de SOPHiA DDM et de SOPHiA GENETICS, je trouverais ça bizarre que l'on ouvre SOPHiA DDM à de la recherche. Que SOPHiA GENETICS s'ouvre à la recherche via autre chose, c'est peut-être très intéressant, mais que la plateforme elle-même le fasse, non. Il faudrait vraiment deux choses différentes et il ne faudrait à mon avis pas mélanger les informations qui viennent de la recherche de celles qui viennent de la routine. Donc finalement qu'est-ce qui serait en commun... ?

Ils disent que c'est l'intégration du système qui est hyper importante.

L'intégration du système ?

Dans le labo. Tu vois le modèle LIMS ? Tu avais ça toi par exemple ?

Plus ou moins tout le monde a des systèmes LIMS dans les grosses institutions mais effectivement c'est une des choses sur lesquelles SOPHiA GENETICS doit travailler, c'est-à-dire que les liaisons au système LIMS institutionnel doivent exister. Parce qu'encore une fois, on vend SOPHiA GENETICS et SOPHiA DDM comme étant un gage de qualité parce que sont envoyés et qu'on a un outil qui nous permet de traiter les variants sans fichier Excel, qu'on a des bases de données qui se constituent petit à petit avec notre *flagging* de nouveau sans fichier Excel, sans *listing*, sans rien. Mais au moment où tu veux transférer les données que tu as interprétées dans le rapport final que tu vas envoyer au client, tu te retrouves à devoir à ce moment-là repasser à du total manuel ou à des exportations Excel et du copier-coller, et là tu te dis qu'il y a une étape qui manque. Effectivement une intégration à des systèmes LIMS, quels qu'ils soient, est un super argument pour beaucoup de clients, à mon avis. Certains sont venus chez SOPHiA GENETICS parce qu'on leur a dit que ça faisait partie des choses qui allaient se développer, mais d'autres attendent sans doute toujours et n'ont encore rien signé. Stéphane, un client à Montréal, nous dis que ça serait bien d'intégrer les données et d'avoir la possibilité de structurer pour le client une espèce de petit fichier exportable et lisible par leur système LIMS, histoire d'avoir une traçabilité complète. Ce serait un énorme plus pour SOPHiA GENETICS.

Ca pourrait augmenter considérablement l'effet de réseau, l'intégration, aussi.

Oui je pense. Après là c'est juste l'étape qui te manque entre les données que tu as générées et interprétées et le rapport final, tu risques d'oublier de rapporter quelque chose ou de faire un mauvais copier-coller des données du patient à côté. Donc là c'est plus sur la qualité que sur la partie réseau parce que ça ne permettrait pas plus de données que ça, à mon avis. La seule chose c'est que les clients seraient peut-être plus enclins à *flagger* parce qu'ils se diraient : « Je sais maintenant que mon *flag* est

transféré aussi dans mon système LIMS, donc le rapport du clinicien est très clair puisqu'il voit que pour tous les variants que j'ai bel et bien attribué une valeur ». Ca pousserait peut-être un petit peu au *flag* et ce serait bon pour ce côté réseau, mais sinon il n'y aurait pas plus de valeur ajoutée au système .

Et est-ce que tu vois d'autres façons de monétiser l'effet de réseau ? On pourrait mettre des firmes pharmaceutiques dans la plateforme et dans l'OncoPortal par exemple, on pourrait mettre leur logo pour faire un peu de publicité.

Je dois bien avouer que moi je suis responsable de labos de diagnostic et je vois que SOPHiA GENETICS commence à se faire sponsoriser par des firmes pharmaceutiques, je leur dis tout de suite que j'arrête. A mon avis, c'est des fois peut-être un peu cliché mais on a toujours l'impression que comme il y a toujours beaucoup d'argent, beaucoup de choses sont possibles et on peut acheter beaucoup de choses. Donc on associerait le logo et on se dirait : « J'ai vu que vous aviez apparemment un genre de publicitaire, pouvez-vous me certifier que vous n'avez pas aussi vendu avec ça les variants que je *flagge* ? ». Pour moi en tout cas ce serait vraiment horrible, je verrais ça comme une vraie perte de l'esprit de départ qui est que l'on travaille pour du diagnostic. J'avais déjà du mal quand j'avais travaillé un peu avec des boîtes qui prenaient contact avec des firmes pharmaceutiques pour essayer d'avoir des projets. Je sais bien que celles-ci demandaient à quel point nos bases de données étaient disponibles pour leurs études, et je m'étais déjà dit que c'était chaud parce que toutes ces données-là il faut forcément qu'elles soient couvertes par un formulaire de consentement que le client a signé et je n'étais pas toujours très à l'aise avec ça, avec les questions qui étaient posées. Je serais encore client et je verrais des firmes pharmaceutiques sponsoriser SOPHiA DDM, je ne sais pas du tout si la majorité des gens que je connais réagirait comme ça mais...

Et tu ne penses pas qu'il y aurait d'autres manières de monétiser ? On ne pourrait pas intégrer d'autres types de personnes dans la plateforme ? C'est-à-dire qu'il y aurait les docteurs, les généticiens... etc.

Voilà, ça je pense que c'est quelque chose qui est sans doute intéressant, la porte d'entrée des pathologistes est sans doute la première à voir parce que dans la structure des généticiens que j'ai déjà fréquenté, en Belgique mais aussi autre part, en général le médecin généticien n'est pas la personne qui va travailler sur les données. En général, il y a un technicien ou deux qui travaillent sur les données et il y a un scientifique qui va valider, faire la recherche bibliographique, créer le rapport, et ce sera ensuite contresigné par un généticien. Par exemple ici à l'UCL, et également à l'IPG puisque je connais le mode de fonctionnement, il n'y a pas un seul médecin qui a accès à SOPHiA DDM, qui veut y avoir accès même, eux veulent juste avoir le rapport déjà pré-signé par le scientifique. Ça c'est pour la génétique. Pour la pathologie c'est vraiment très différent parce qu'eux justement ont toujours plus eu envie d'être intégrés dans la partie labo. Des fois c'est un peu drôle parce qu'ils n'ont pas vraiment les connaissances et ils forcent un peu le passage, mais à mon avis avec l'OncoPortal,

on se rend compte que de plus en plus de données sont à destination du clinicien et pas vraiment pour...

Oui pour les phénotypes.

Pour la partie phénotypes oui mais aussi simplement pour les drogues qui sont disponibles. Il y a des scientifiques qui sont dédiés à la partie pathologie et qui petit à petit commencent à connaître les drogues, qui peuvent proposer. Mais à mon avis il y aurait là un vrai intérêt, c'est que la plateforme soit accessible également par des cliniciens. Je sais bien qu'à Liège ça a été une demande qui nous a été faite spécifiquement par nos clients. Ils nous ont demandé : « Est-ce qu'on pourrait imaginer que les pathologistes ou que les hématologues (parce que là c'est un laboratoire qui travaille pour la partie pathologie mais aussi pour la partie hématologie), avec qui on traite souvent pourraient avoir un accès ? ». Je leur ai dit qu'il n'y avait pas de soucis, la seule chose c'est qu'on ne peut pas restreindre énormément leur accès car on ne dispose pas actuellement de droits de restriction énorme. Donc ça veut dire qu'il y avait un vrai intérêt à ce qu'une partie des données de SOPHiA DDM puisse être visualisable, parce qu'eux à ce moment-là voyaient surtout la partie rapport OncoPortal.

Donc il commence à y avoir une interaction entre les utilisateurs ?

En tout cas le responsable du laboratoire qui fait les analyses avait l'impression que les hématologues, entre guillemets les plus fréquemment contactés par lui, auraient un intérêt à avoir un accès à ça. C'est là où le travail serait plus intéressant s'il est fait par un hématologue ou éventuellement un jour l'oncologue, parce qu'il aurait le rapport du scientifique mais aussi un accès à toutes les drogues possibles et imaginables qui sont en combinaison avec cette tumeur. Et ça c'est difficile.

Donc les hématologues seraient tentés d'aller sur la plateforme. Et est-ce que les docteurs, les généticiens, pourraient aussi être attirés par l'approche des hématologues ?

A mon avis moins parce qu'il faut bien se dire qu'avec les généticiens il n'y a pas de traitement derrière, les traitements sont donnés par d'autres. Le généticien ce qui l'intéresse c'est d'avoir la mutation qui peut expliquer la pathologie parce qu'après il va essayer de la chercher dans le restant de la famille. Si le travail du scientifique est bien fait et que le rapport dit « Nous n'avons pas trouvé de mutation » ou « Nous avons trouvé cette mutation qui est causale », ça leur suffit. Ils n'ont pas besoin d'interagir avec la plateforme, en fait. Ce serait plutôt les pathologistes, les oncologues ou les hématologues qui auraient un intérêt parce qu'ils pourraient aller voir le rapport OncoPortal, c'est-à-dire le rapport qui explique toutes les associations entre certains variants identifiés chez le patient et des drogues que eux maîtrisent et manipulent

régulièrement. Donc là j'ai l'impression qu'il y aurait un vrai intérêt, aussi pour le patient lui-même, tout simplement parce qu'il aurait un médecin qui pourrait avoir accès à la totalité de la liste. Il faudrait faire ça de manière un peu sympa parce que générer un énorme rapport OncoPortal, leur envoyer et espérer que l'hématologue va commencer à jouer avec les pages, ça ne marche pas parce qu'ils ont sans doute dix minutes par patient. Il faudrait avoir un système où ils auraient accès à ça et un outil qui leur permettent de très rapidement extraire les données.

Mais donc il pourrait y avoir des interactions avec ces gens qui seraient intéressés.

Et dès lors ça veut dire qu'il pourrait y avoir une licence d'accès qui pourrait être monnayée parce que pour l'instant le compte a toujours eu accès gratuitement avec un nombre illimité d'utilisateurs, mais on pourrait imaginer que quand ce sont des clients extérieurs à l'institution, car ce sont des fois des hématologues qui ne sont pas forcément dans l'institution mais qui sont dans des hôpitaux avec qui elle travaille, il y ait une licence. Maintenant ce sont encore des groupes, des domaines dans lesquels ils ont des réunions multidisciplinaires pour prendre la décision pour le patient donc ils combinent toute une série d'informations, dont des fois l'information de cryptage génétique. Il faut donc voir aussi si l'intérêt ne serait pas plus simplement qu'il y ait une licence associée à ce groupe, je ne pas comment ça peut fonctionner.

Ce serait intéressant, ce serait une nouvelle façon de monétiser et de créer l'interaction dans les réseaux indirects.

En plus à mon avis c'est aussi parfois intéressant de lancer des ponts vers des médecins, parce que dans certains cas les laboratoires sont très conservateurs car ils pensent que SOPHiA GENETICS pourrait leur voler leur travail et ça peut se sentir chez certains bio-informaticiens. Ça veut dire que des fois quand on essaye d'accéder à des laboratoires, l'accès est très difficile parce que soit le responsable du laboratoire soit le bio-informaticien dit « Oui ça à l'air intéressant mais on fait largement aussi bien chez nous ». Alors que pouvoir accéder à des oncologues en leur disant « Voilà, on a un système qui peut être accessible pour vous, même si forcément il faut que vous travailliez avec un laboratoire qui également l'utilise », ça peut peut-être nous créer des points d'entrée vers le laboratoire.

Et une dernière question, dans le futur tu dirais que ce serait quoi le marché de l'analyse NGS, tu verrais une plateforme globale qui aurait un certain monopole dû à sa base de données ou plutôt comme maintenant une compétition entre plusieurs plateformes ?

Dans mon rêve, j'imaginerais que ce ne serait ni un monopole, ni une compétition, mais un partage entre différents instituts de recherche qui développent des solutions et qui ont certainement pleins de données à partager, et nous qui rassemblons des données d'une autre valeur parce qu'elles sont pour l'instant fort liées au diagnostic.

Mais le seul moyen de rassembler ces données ce serait d'arriver dans une optique qui n'est pas la monétisation mais la coopération donc.

Je suis quand même issu du monde de la recherche avec du diagnostic mais c'est la première fois que je travaille dans une firme privée, donc j'ai quand même toujours cet idéal qui est d'imaginer que monnayer une fréquence d'un variant dans la population c'est quand même fort monnayer quelque chose sur lequel on porte un brevet sur un gène. Après j'espérerais plutôt qu'on arrive à un système où au niveau européen, SOPHiA GENETICS puisse avoir des subventions tout simplement parce que SOPHiA gère également des bases de données publiques européennes.

SOPHiA devrait passer dans le public pour que les gens lui fassent plus confiance ?

Je suis certain qu'il y a sans doute les deux écoles : il y a des personnes qui auraient plus confiance en un SOPHiA GENETICS public qu'en un SOPHiA GENETICS privé et inversement. Le côté privé on se dit que la qualité est meilleure, qu'ils ont plus de moyens. On avait eu une conversation avec quelqu'un d'Autriche avec Connie, où une des questions de ce gars qui n'était pas du tout sympa et un peu agressif était justement de savoir si SOPHiA GENETICS pouvait collecter les données pour après les monnayer. Il trouvait que le système consistant à laisser aux clients la possibilité de publier les données qu'il a collectées via l'outil SOPHiA DDM, et que SOPHiA GENETICS puisse créer un outil pour publier facilement les données qu'il a lui-même générées dans son laboratoire, était quelque chose d'important. Finalement il n'avait pas tout à fait confiance parce qu'il voulait vraiment savoir si le *business plan* de SOPHiA GENETICS c'était un jour de vendre ou non ces bases de données. J'avoue que ni Connie ni moi ne pouvions répondre, on ne savait vraiment pas du tout ce que serait l'avenir, mais c'est vrai que les clients chez nous payent une analyse, c'est-à-dire un service, trouvent des variants et aimeraient bien les publier. Chez Alamut par exemple ils peuvent publier, envoyer un formulaire automatiquement. Il y a un type d'ouverture que SOPHiA GENETICS pourrait montrer, c'est ce genre de choses, mais est-ce que c'est dans le *business plan* de SOPHiA, je n'en sais toujours rien. Je ne sais pas du tout s'il vaut mieux justement pouvoir monnayer réellement ces informations-là ou si on se développera en se disant...

Elle pourrait monnayer peut-être les publications, sorties de SOPHiA.

Oui mais de nouveau on arrive à ce moment-là à un système où ce qui a été trouvé n'est pas la propriété de SOPHiA GENETICS puisque ça l'a été par un client qui a utilisé des patients et un outil, qui est SOPHiA DDM. Donc que SOPHiA GENETICS monnaye quelque chose qui a été trouvé par quelqu'un qui a déjà payé SOPHiA pour le trouver, c'est quelque chose qui n'est pour moi pas le bon modèle. Je suis peut-être des fois fort utopiste mais j'aime bien cette notion selon laquelle le patient et le client

n'appartiennent à personne. Le client a payé un outil pour mettre en évidence des mutations ou des variants chez un patient, donc ces variants auraient intérêt à revenir gratuitement à la communauté. De toute façon SOPHiA GENETICS en tire déjà avantage puisque pour l'avoir, le client a payé, et en plus il l'utilise parce que la communauté grandit et qu'il peut avoir de nouveaux clients en leur disant : « Nous avons des millions d'informations auxquelles vous pourrez avoir accès gratuitement si vous travaillez avec nous ». Si en plus on veut monnayer un export vers des bases publiques, ça me paraîtrait bizarre. Mais à un moment quand on en sera sans doute à séquencer systématiquement le génome parce que ce sera moins cher, il faudra bien avoir des bases de données les plus larges possibles parce qu'on a cet aspect régional dont il faut tenir compte. En même temps j'imagine que quand on aura suffisamment de données, on pourra également déterminer l'origine ethnique du patient avec les données contenues dans le *whole exome*, donc il y aura sans doute des corrections qui seront déjà faites, mais là je sais pas du tout si on va vers une guerre entre gros fournisseurs. J'imagine que des Google ou autre auront intérêt à essayer de monétiser ce genre de données massives, est-ce que nous devons nous préparer en partageant avec des institutions publiques pour qu'il n'y ait pas un monopole ni une bagarre entre deux ou trois géants dont on pourrait faire partie, je n'en sais rien.

Le futur le dira.

Oui. Mon rêve serait forcément d'aller vers des choses publiques mais bon, c'est un rêve.

Je ne sais pas si Yourgi serait content.

Je ne sais pas du tout. Des fois travailler avec des choses publiques ça peut ramener aussi de l'argent parce qu'à mon avis on peut peut-être plus facilement convaincre un pays entier de travailler avec SOPHiA GENETICS en disant « SOPHiA GENETICS rassemble des données, on les publie dans des bases de données publiques donc on n'a pas de secrets. Si vous voulez travailler avec nous et que votre Etat fédéral le veut, et bien on est partants ». A mon avis ça a le même intérêt que de tout cadenasser et de vendre.

Oui ce sont deux approches différentes.

... à Yourgi pour choisir celle qui fera perdurer SOPHiA GENETICS et qui aidera les patients.

J'espère, en tout cas merci c'est top, j'ai toutes les infos, tu donnes une bonne perspective des choses en tant qu'utilisateur.

Appendix 3.3. Dr. Shirine Behenda

Alors je suis product manager à SOPHiA GENETICS et ça fait plus de deux ans que je suis chez SOPHiA, plus précisément depuis mars 2015.

Et tu gères une équipe, par rapport aux autres la communication est facile comme c'est une petite boîte ou est-ce que tu dois passer par d'autres étapes pour parler au boss par exemple ?

Alors ce qui est bien chez SOPHiA c'est que la communication est directe, en tout cas jusqu'à maintenant. Après c'est vrai qu'il se peut qu'il y ait des choses qui changent parce qu'on va grandir, donc on sera obligé de passer par d'autres procédures qui feront qu'on sera forcé de passer par un intermédiaire avant de pouvoir discuter avec le n+1, +2 ou n+3. Mais jusqu'à maintenant, j'ai toujours pu interagir directement avec tout le monde. C'est assez agréable et ça permet d'aller très vite dans le travail et donc de gagner en efficacité. Parce que si tu passes par des intermédiaires, le temps d'avoir une réponse ou même que l'information monte d'un cran et d'un autre cran, tu perds des semaines.

C'est ça, donc ça c'est assez direct. Et au niveau de la valeur que toi tu apportes, tu dirais une valeur qui te tient à cœur, que tu mets dans l'entreprise, pour toi au niveau professionnel ? Détermination ? Quelque chose qui t'intéresse, une valeur que tu voudrais apporter à la boîte ? Kevin lui a adoré que tout aille vite, que ce soit vraiment rapide, le mécanisme. Est-ce que toi tu as une valeur que tu aimes bien, que tu apportes dans l'entreprise tu penses ?

Alors en ce qui concerne l'entreprise moi ce que j'aime c'est l'efficacité des équipes par rapport au domaine dans lequel on est. On est donc dans un domaine médical et ce qui est génial c'est qu'on a des équipes qui bossent à fond dans ce qu'elles font, elles le font bien, sans tricher ni rien, il y a une totale transparence par rapport à la partie technique. Et c'est pour ça que j'ai rejoint SOPHiA, c'est parce que j'ai vu comment les équipes travaillaient et j'ai discuté avec certains de la bio-info, le travail qu'ils fournissaient et la rigueur. Donc j'ai aimé la rigueur, cette rigueur qu'on retrouve en académie aussi.

Donc c'est la rigueur qui te tient à cœur.

Oui.

Génial c'est intéressant, Kevin avait dit un peu la même chose aussi. Moi je m'interrogeais, enfin sur mon mémoire j'essaie de voir comment les réseaux d'effet sont créés à travers les utilisateurs sur des plateformes. Donc dans la partie théorique, je regarde un peu tout ce qui est Uber, Airbnb, quel est l'effet de

réseau créé par les utilisateurs et comment le capturer et le monétiser. Donc je prends le cas de SOPHiA et d'autres boîtes et je m'interroge vraiment sur cet effet de réseau créé par les utilisateurs. J'ai une question par rapport à quand tu dis SOPHiA, quelle est la valeur SOPHiA, la plateforme DDM qu'est-ce qu'elle représente pour les consommateurs ? Si tu dois la vendre, quelle est la *value proposition* ? Qu'est-ce qui définit vraiment SOPHiA DDM, si tu parles à un client, qu'est-ce qui est la principale aide de la plateforme ?

Alors franchement je dirais que la valeur de SOPHiA GENETICS c'est la communauté, le fait de pouvoir finalement regrouper tous ces acteurs de la santé autour d'une même plateforme, on va dire une plateforme rotative où en fait tout est *get together*. Tout est rassemblé dans une sorte de grande base de données et c'est ça la valeur de SOPHiA, c'est vraiment la communauté, le fait que chacun de ces experts joue le jeu de *flagger* les variants, de donner des commentaires... etc.

Et tu n'as pas peur qu'il y ait des hôpitaux un peu réticents au niveau de la communauté ou est-ce que les hôpitaux sont plutôt enthousiastes quand on approche avec cette communauté ?

Alors il y a les deux types en fait. Tu as une partie qui est très enthousiaste, qui voit vraiment le côté bénéfique de la communauté, donc le partage de données, le partage d'informations ; et l'autre côté qui est plutôt réfractaire. Pourquoi, tout simplement parce qu'ils se disent qu'il se peut que tout le travail qu'ils vont fournir eux gracieusement va être monétisé par l'industriel. Et ça on a vu le cas en effet, par exemple Curie, qui est un institut qui ne veut pas travailler avec nous parce qu'ils ont déjà eu une première expérience... Enfin il y a deux raisons, d'une part ils ont déjà leur plateforme de bio-info, ils ont plus de 100 bio-informaticiens, mais surtout parce qu'on est un industriel et ils nous voient plutôt comme un requin qui veut récupérer de l'argent de leurs travaux fournis gracieusement.

Et est-ce que vous avez poussé un peu ou c'est impossible à Curie ?

Impossible à Curie, à l'heure actuelle. C'est vu, par Yourgi, moi-même, par présentation...

Donc c'est appréhender le fait qu'on pourrait monétiser ces données ?

Parce qu'en fait ce qu'il s'est passé, c'est qu'ils ont fait quelque chose de similaire pour les jeunes BRCA. Ils étaient plusieurs à l'échelle nationale à remplir une base de données pour BRCA-1 et -2, c'était déjà il y a un petit bout de temps, et ils ont répertorié tous les variants identifiés chez les patients, certains étaient pathogéniques, certains bénins. Ils ont travaillé dessus, ils ont fait une *curation of the data*, et ensuite c'est Linxiarm transfer qui a récupéré cette base de données, donc qui fait partie de l'Etat, et c'est eux qui devaient gérer cette base de données. Qu'est-ce qui s'est passé, un industriel est venu voir et a dit « voilà on vous donne tant d'argent, tout ce que vous voulez et on récupère la base ». Et Linxiarm a vendu cette base de données au désarroi de tous les petits qui ont travaillé là-dessus. Même si certains, de ce que j'ai compris,

perçoivent des royalties mais la pilule n'est vraiment pas passé parce que tout le monde n'avait pas été concerté a priori pour la vente de cette base de données qui était nationale et publique, quand je dis public c'est-à-dire que ce sont des instituts publics financés par l'Etat, et maintenant c'est passé entre les mains d'un industriel. En plus du coup tous ceux qui travaillent avec nous, sur BRCA notamment, n'ont pas le droit de *flagger* les variants en question, parce que tous les *flags*, tous ce qu'ils doivent *flagger* et annoter, ils doivent l'annoter dans la base de données qui a été rachetée par Quest.

Ça ils n'ont pas le droit de *flagger* non plus ?

Voilà, donc tous ceux qui travaillent avec nous normalement ne *flaggent* pas les variants BRCA-1 et -2.

Donc là il y a un manque d'interaction alors.

Voilà, après c'est à l'échelle nationale en France, c'est l'exemple que j'ai où vraiment ils étaient assez réfractaires pour faire un partenariat de ce type à l'échelle internationale communautaire, parce qu'ils se disent « une fois que vous aurez tout ça, on ne sait pas ce que vous allez en faire, vous pourrez très bien revendre. C'est nous qui avons payé, tout ça c'est de l'argent public, même nous on paye ces analyses, vous allez du coup monétiser ça ».

Mais on le fait déjà en fait, on monétise déjà avec les transactions, enfin avec le *business model* que SOPHiA a, c'est déjà une sorte de monétisation dans le sens où sur chaque *upload* uploadé il y a quand même une monétisation, c'est comme ça qu'on monétise un peu le réseau quand même.

Alors oui et non. Je ne suis pas tout à fait d'accord avec toi. Si tu veux il y a deux types d'analyse. Il y a une analyse algorithmique où on apporte l'analyse algorithmique, donc là en effet tu payes un service puisqu'on va fournir un algorithme fait sur mesure, donc tu peux très bien acheter un *software* et analyser tes données avec ce *software* et si tu veux c'est pareil. Là, la différence c'est que tu le fais par analyse et non pas en achetant un *software*, et la deuxième différence c'est que ce *software*, cet algorithme que l'on a mis en place, évolue. Il évolue grâce en effet à la communauté, c'est-à-dire qu'une fois qu'on améliore les algorithmes, tout le monde en bénéficie. Donc étant donné qu'ils payent de toute façon un forfait par analyse, et bien ils sont assurés de systématiquement voir leur algorithme s'améliorer, ce n'est pas quelque chose qui est figé donc c'est quand même un argument très positif par rapport à quelqu'un qui achète un *software*, qui doit se débrouiller, qui doit faire après toutes les revalidations s'il y a une mise à jour de *software*. Donc la monétisation est vraiment basée sur l'algorithme, et en effet on utilise la communauté pour améliorer encore plus nos algorithmes.

Ce sont les réseaux d'effets direct en fait, puisqu'il y a deux types d'effets de réseau. Le premier est direct, c'est-à-dire que plus il y a d'utilisateurs sur DDM, plus la valeur globale de la plateforme augmente parce que l'algorithme devient de plus en plus performant. Le deuxième réseau d'effets ils appellent ça le *indirect network effect*, c'est entre les utilisateurs en fait, comment créer un réseau où plus d'interactions se font directement entre les utilisateurs. Et ça ce serait plutôt via le *flagging*.

Alors tu as raison, le *flagging* par contre tu ne peux pas savoir qui a *flaggé*. Mais ce n'est pas grave dans le sens où cette information nous on l'a et ils peuvent nous solliciter pour nous demander : « Voilà j'ai vu que ce variant a été *flaggé* par deux de mes confrères internationaux, et je suis assez perplexe là-dessus, je voudrais vraiment discuter avec eux ». Et grâce à ce réseau, cette communauté, on a la possibilité de demander à ces deux experts qui ont *flaggé* par exemple un variant particulier : « Il y a un expert qui est perplexe et il souhaiterait discuter avec vous. Souhaitez-vous que l'on vous mette en communication pour en discuter plus amplement ? ». Et ça c'est quelque chose de très fort, parce qu'actuellement mise à part les publications que tu peux réaliser quand tu es dans un groupe de travail, donc tu choisis déjà tes partenaires pour rédiger un article, là on est dans un système qui est complètement dématérialisé, donc c'est vraiment : « Je suis surpris, je vois un variant qui est *flaggé* par deux de mes confrères internationaux que je ne connais pas, et je voudrais discuter avec ces gens-là ». Et là on donne aussi une possibilité de les mettre en communication pour qu'ils puissent peut-être établir une nouvelle collaboration, discuter plus amplement de pourquoi ces variants ont été *flaggés*, pathogéniques ou non-pathogéniques ou bénins.

Et ce n'est pas possible au jour d'aujourd'hui de mettre un forum ou une espèce d'outil sur la plateforme pour que les gens interagissent directement entre eux ?

Alors ça ce serait formidable c'est vrai, c'est une très bonne remarque. Maintenant il faudrait savoir dans quel contexte on pourrait mettre ça en place. Il faudrait déjà leur demander leur avis, est-ce qu'ils souhaiteraient avoir la possibilité par exemple de voir les noms de ceux qui ont *flaggés* ? Je sais qu'il y en a certains que ça ne dérangerait pas, d'autres seraient plus embêtés parce que certains vont *flagger* des variants pour leur propre compte, c'est-à-dire qu'ils ne veulent pas être responsables de l'étiquetage d'un variant ; ils sont protégés par un écran et se disent « Ca je le fais pour mes analyses, pour mes patients à moi » mais ils ne vont pas forcément publier l'étiquetage sur un papier pour un journal officiel, parce qu'ils n'en sont pas sûr, parce qu'ils n'ont pas fait d'étude approfondie assez importantes pour dire « Je suis sûr que tel variant est pathogénique parce que j'ai vu que dans une cohorte de mille patients, tous ceux qui avaient cette mutation génétique répondaient très mal à un traitement ». C'est pour ça qu'il y en a beaucoup qui sont un peu gênés par le *flagging* international, c'est-à-dire visible par tout le monde, et ils préfèrent les clubs, des réseaux très petits, très restreints de quatre ou cinq hôpitaux qui se connaissent très bien, un *consortium* où ils vont pouvoir partager le *flagging* parce qu'ils vont pouvoir se concerter, discuter de pourquoi untel a *flaggé* tel variant de cette manière.

Et ça existe déjà ?

Les clubs ? Oui tout à fait.

Et donc c'est juste les cinq-six hôpitaux et là on voit le nom des gens qui *flaggent* dans ces *consortium* ?

Tout à fait, en fait quand tu es dans un club, ce sont des clubs où ils disent « Moi je travaille avec trois autres partenaires, les hôpitaux X, Y et Z » et là ils partagent leur *flagging*. Donc là en effet tu vois qui est-ce qui a *flaggé*, quel hôpital a *flaggé* untel, tel variant... etc. Mais ça reste à l'échelle du club, locale.

C'est intéressant, et ce serait donc vraiment pour l'instant au niveau local et ça pourrait grandir au niveau global si c'est des petites entités locales partout ?

Oui mais l'entité globale c'est ce qu'on a actuellement, c'est le *flagging* international.

Oui mais si on arrive à porter l'interaction du niveau local à global, tu vois l'interaction au niveau des petits *consortium* qui est très intéressante ?

Alors par contre ceux qui ont un accès aux clubs ne peuvent pas voir le *flagging* international, ce qui est normal parce qu'ils ne jouent pas le jeu. Notre but c'est vraiment de rendre visible l'étiquetage de tous les variants à l'échelle internationale. Si certains veulent se restreindre à une population plus petite, c'est-à-dire un club plus restreint, et bien ils jouent le jeu en disant « On l'a accepté, on l'a fait malgré le fait que ce soit contre nos valeurs », mais à ce moment-là on leur dit « Du coup vous n'avez pas accès au *flagging* international ».

Oui donc ça fait partie du jeu aussi. Et si on revient à la monétisation, toi tu me dis que c'est vraiment l'algorithme la monétisation en fait ?

Oui.

Si je te demande comment SOPHiA monétise l'effet de réseau, tu me réponds ?

Alors ça va dépendre de ce que tu entends par « l'effet de réseau ». Tout ce qui est *flagging* ?

L'effet de réseau c'est vraiment l'effet, la valeur qui est créé(e) par les utilisateurs. Si un utilisateur arrive sur la plateforme, le deuxième sera plus tenté de venir sur la plateforme parce qu'il y a déjà le premier, et ainsi de suite... Plus il y a d'utilisateurs, plus la valeur de la plateforme augmente pour le prochain utilisateur qui va venir sur celle-ci, parce que les algorithmes deviennent plus puissants.

Oui ça c'est l'effet direct ?

Exactement oui, et donc ma question c'est est-ce que SOPHiA monétise cet effet ? Comment arrive-t-on à lancer ce mouvement d'utilisateurs tout en ayant un *business model* qui marche ?

Là le *business model* c'est la monétisation de l'effet direct, mais il n'y a pas de monétisation de l'effet indirect. Donc ta question elle est posée par rapport à l'effet indirect ou globalement ?

Direct on est d'accord que c'est la transaction par *sample* ?

Oui.

Et indirect on n'en a pas encore ?

Non.

Et qu'est-ce que tu penses de ça ?

Moi je serais contre parce que les données appartiennent aux hôpitaux, ils ont payé un service qui est celui de l'algorithme et on leur a vendu le fait qu'ils appartiendraient à une communauté SOPHiA où ils pourraient étiqueter des variants et voir ce que d'autres experts internationaux ont étiqueté, donc ça irait à l'encontre de nos valeurs précédemment établies et en contradiction par rapport à ce qu'on met en place, dans les valeurs, dans les *call values* de SOPHiA GENETICS.

On pourrait dire aussi que l'on pourrait faire différents *pricing* par rapport à la plateforme, gratuite pour une certaine population et payante pour le monde académique par exemple.

Alors je dirais plutôt gratuite pour le monde académique et payante pour les industriels, oui.

Oui avec ceux qui ont un intérêt financier derrière.

Par exemple les boîtes pharmaceutiques, notamment.

Exactement.

Mais là encore il faut faire très attention parce que l'on rentre dans un domaine assez délicat car nous ne sommes pas propriétaires des données, ce qui fait que même si c'est un effet indirect, tu vends le travail de ceux qui ont *flaggé*. Donc je ne suis pas sûre que ce soit une image pour SOPHiA.

Pourtant quand on voit les autres plateformes comme Facebook ou Google, l'accès est gratuit mais après ils prennent quand même nos données, comme si la donnée devenait une *currency*, une monnaie pour eux, c'est ce qu'ils vendent.

Mais par contre ils n'ont jamais dit « Vos données ne seront jamais utilisées dans la vie ». Ce n'est pas le cas.

Oui c'est vrai.

Alors que SOPHiA l'idée c'est « Nous ne sommes pas propriétaires de vos données ». Maintenant on peut peut-être jouer sur les mots et revoir la loi entre guillemets et ce que nous avons mis comme information dans le règlement et les contrats avec les différents partenaires : est-ce que réellement on parle de données génomiques, donc de choses matérielles ou de ce qui est plutôt dématérialisé c'est-à-dire le travail de *curation* que les biologistes auront fait, qu'en est-il réellement ? Est-ce qu'on peut parler de données de patients ou pas, et dans ce cas est-ce que l'on peut exploiter cette information-là ? Et si on l'exploite, à quel prix, dans le sens ou par rapport à notre image est-ce que nous avons beaucoup à y gagner, ou un peu à y gagner et un peu à y perdre, ou juste à y perdre ? Ça ce sont des questions auxquelles il faut répondre et voir ce qu'il en est en termes d'analyse de risques.

Parce que par exemple une pharmacie que l'on charge et qui mettrait ses conseils par rapport aux médicaments et aux traitements précis dans la plateforme, peut-être que l'on pourrait donner un accès gratuit à certains utilisateurs parce qu'on charge de l'autre côté. Le pharmacien mettrait gratuitement ?

Non l'industriel mettrait ses propres *drugs* par rapport au traitement, donc ça ferait la publicité pour pousser les hôpitaux à prendre une certaine *drug*, un certain médicament. La pharmacie serait intégrée dans la plateforme et mettrait ses propres médicaments. Il y a deux façon de voir les choses pour les boîtes pharmaceutiques, c'est un peu complexe parce que quand elles développent une *drug* en général, pour faire la promotion elles vont faire du recrutement de patients et elles passent par différentes phases. Phase 1, elles recrutent des patients en bonne santé en général pour voir si c'est toxique ou pas. Si ce n'est pas toxique, elles passent en phase 2 sur une petite cohorte de trente patients malades parce que ça coûte des millions. Et si ça marche bien, elles vont prendre une cohorte plus grande, c'est ce qu'on appelle la phase 3. Le problème c'est qu'en général il y a une concurrence assez rude entre les boîtes pharmaceutiques et elles ne veulent pas forcément que leurs résultats soient divulgués. Donc il est vrai que tu peux, ça avait déjà été discuté en interne, dire que l'on peut discuter avec les boîtes pharmaceutiques, si elles cherchent un profil de patient bien précis avec par exemple, elles vont passer du temps à chercher et à recruter leurs patients. J'avais demandé à l'époque, quand j'avais rencontré quelques personnes à Lascaux, combien de temps ils mettaient pour recruter les gens et parfois ils me disaient « entre quelque mois et un an ». Or quelqu'un qui a un cancer somatique n'a pas un an pour vivre, en général il peut mourir même très vite en moins d'un an. Donc c'est vrai qu'il y a l'idée de pouvoir aider parce qu'on a cette communauté et qu'on a des informations sur le profil génétique des patients, on pourrait très bien guider les

boîtes pharmaceutiques en disant « Il y a dix patients français à Paris dans tel hôpital qui ont exactement le profil génétique que vous recherchez, vingt autres qui sont en Espagne, cent autres qui sont aux Etats-Unis ». Du coup, elles iraient contacter directement les hôpitaux pour leur dire « Vous avez des patients qui pourraient nous intéresser ». Donc on leur faciliterait cette recherche-là, le recrutement de patients pour leur *drugs* qui sont en développement. Ça peut être un moyen de dire « on vous aide, mais du coup vous payer », ce qui monétise en fait le recrutement de patients puisque tu facilites le travail de recrutement et tu le raccourcis.

Ca ça pourrait être possible.

Ça peut être possible, après ça paraît assez simple mais dans la réalité elles ne vont pas demander que le profil génétique mais aussi la partie clinique. Elles vont demander par exemple que ce soit des femmes ou des hommes entre 20 et 40 ans, ou qui n'ont jamais subi de traitements chimiothérapiques dans leur vie. Ces informations si on ne les a pas, on ne peut pas les fournir, on ne peut fournir à l'heure actuelle que la partie génomique, mais ni clinique ni phénotypique parce qu'on en est dépourvu, SOPHiA DDM en est dépourvu.

Ça c'est intéressant aussi, je ne savais pas. Si on revient au *flagging*, comment tu pousserais *l'insentives* de ceux qui utilisent la plateforme à *flagger* ? Est-ce que les gens interagissent beaucoup par rapport au *flagging* ? Est-ce qu'il y aurait des moyens de pousser davantage le *flagging* ? Qu'est-ce qui se passe par rapport à ça ?

Je ne pourrais pas te dire en termes de pourcentage, je sais qu'en France certains utilisent de manière massive le *flagging* parce qu'ils y trouvent une utilité, un gain de temps surtout pour le moment et ils y retrouvent leur variants très rapidement quand c'est *flaggé*.

Donc ils ne le font pas de manière altruiste ? C'est vraiment pour eux, pour leur système ?

Totalement oui, ce n'est vraiment pas de manière altruiste du tout. Ils y trouvent un intérêt qui est un gain de temps.

Donc ce n'est pas tellement un réseau...

Non. Par contre par rapport au deuxième point que je vais aborder, moi je travaille là-dessus donc c'est la nouvelle plateforme d'oncologie où l'on va mettre de nouvelles informations, de nouveaux champs dans SOPHiA DDM qui va permettre de capturer ces informations concernant les traitements précédents, ou quel est le traitement qui a été utilisé pour ce patient, ou quelle était la charge tumorale de cette tumeur... etc. Cela va permettre d'enrichir notre base de données en termes phénotypiques et cliniques et qui pourra répondre à la première question par rapport aux boîtes pharmaceutiques. On pourra le faire plus facilement. Le second point, par rapport maintenant aux hôpitaux, quand tu dis « effet de levier » et permettre un plus grand nombre de

flagging, je dirais que cela sera possible quand ils verront un plus grand avantage. Qu'est-ce que je veux dire par là, et bien s'ils voient, surtout dans le somatique je dirais mais aussi dans le mais il faut que l'on pousse un peu plus le travail actuel, c'est que là ils n'ont que le *flagging*, ce qui est très bien pour eux car ils vont tout de suite voir leur variants (« mon variant est là, il est muté, je sais ce qu'il a comme pathologie et je sais ce que je vais pouvoir lui administrer »). On va pouvoir améliorer ça quand on aura intégré les informations phénotypiques et cliniques notamment : quand ils verront qu'ils auront une meilleure estimation du traitement à poser, ils auront quelque chose de beaucoup plus clair, et là ils vont faire ce travail de *flagging* de manière beaucoup plus sérieuse. Et ils vont aussi notamment rentrer les données phénotypiques et cliniques. Ça par exemple on ne l'a pas encore mis en place mais une fois que ce sera fait, je sais pertinemment que très peu de biologistes le feront parce qu'ils n'ont pas le temps, que ça ne va pas leur servir et qu'ils n'ont rien à y gagner. Donc pour pouvoir les inciter, il faut que derrière ils y trouvent un bénéfice. Le premier bénéfice, mais ce sera on va dire au moyen terme, c'est qu'une fois qu'on aura une base de données cliniques et phénotypiques assez importante, ils verront qu'ils auront une aide à l'interprétation qui sera plus intéressante qu'à l'heure actuelle. Pour les cas où il n'y a aucun protocole d'établi, par exemple un patient qui n'a aucunes mutations pour lesquelles il y a des traitements disponibles, ils n'ont rien à lui proposer et en gros c'est chimio, et bien si on arrive à avoir toutes ces informations phénotypiques et cliniques et que dans d'autres pays certains ont traité des cohortes de manière non conventionnelle et que ça marche, et bien ils auront cette information.

Et ça c'est vous qui la fournissez ?

Non, on va fournir des informations mais ce sera comme le *flagging*, ce sera une information fournie par la communauté. Mais pour le moment il va falloir que l'on trouve quelque chose qui existe déjà pour se baser dessus et ensuite inciter les utilisateurs à rajouter des informations supplémentaires pour que l'on puisse enrichir et rendre cette information plus pertinente et donc faire un effet de levier encore plus important, qui sera exponentiel.

Mais c'est juste par rapport à l'insentives que ces gens remplissent ces données ? Quel est le bénéfice pour eux ? Pourquoi ils le feraient ?

Exactement, pourquoi ils le feraient ? A l'heure actuelle si l'on oublie l'oncologie, si l'on oublie les champs cliniques et phénotypiques qui pourraient être un bénéfice sur le moyen terme, si tu veux vraiment inciter les gens, je dirais qu'une des façons est de dire que tous ceux qui *flaggent* les variants auront 20% de réduction pour toutes leurs analyses. Systématiquement. Je sais très bien qu'un certain nombre de clients vont en effet *flagger*, étiqueter, pour avoir ces 20%. Le problème c'est que c'est à double tranchant parce que tu peux très bien avoir un phénomène inverse, c'est-à-dire qu'au

lieu d'avoir un étiquetage de variants correct, il va être complètement faussé. On pourrait entrer dans une phase où on sera obligé de faire du *curation* parce qu'il y aura des choses qui seront fausses. Donc finalement on ne sera peut-être pas gagnant.

C'est à double tranchant donc, c'est sur les deux côtés... Comment on pourrait avoir une standardisation de la qualité pour qu'on ait quelque chose de très bon ?

Tous ceux qui ont déjà *flaggé*, qui le font de manière assidue et juste, on peut par exemple avoir une équipe en interne qui fait du *curation*, et ceux qui *flaggent* bien on peut leur dire « Bravo, ce mois-ci vous avez le droit à cinquante analyses gratuites parce qu'on vous remercie pour votre *flagging* qui va aider la communauté à mieux interpréter les données ». Donc là c'est vraiment un remerciement, c'est une forme qui n'est pas généralisée mais localisée et spécifique à certains clients qui le font de manière systématique. On peut voir ça par exemple sur six mois, quels sont les clients qui ont *flaggé* et étiqueté de manière homogène, qui ont travaillé sur SOPHiA DDM de manière assez quotidienne.

Oui, c'est un moyen d'assurer une certaine qualité.

Exactement. Après il faut quand même derrière que l'on ait une équipe qui vérifie juste pour s'assurer que c'est logique et que ce qui a été *flaggé* est correct, voire même peut-être en appeler certains si on voit par exemple qu'un même variant a été *flaggé* par dix experts différents dans deux catégories opposées, un pathogénique et un autre bénin. Peut-être que ce serait bien de prendre l'initiative de contacter et de comprendre pourquoi ce *flagging* a été fait ainsi, dans quel contexte, est-ce que c'est par rapport à une ethnicité, un fond génétique particulier qui expliquerait pourquoi dans cette population tel variant serait plus pathogénique ou bénin dans l'autre cas...

On définit le *critical mass* comme le moment où la plateforme atteint un niveau où elle est indépendante, elle est *self-sufficient*, c'est-à-dire elle marche toute seule, donc quand elle a atteint un niveau de masse d'utilisateurs où c'est automatique. Et ce que tu penses que si dans cinq ans l'on arrive à ce *critical mass*, il faudra toujours pousser les gens à faire ce *flagging* ? Est-ce qu'il n'y a pas un moyen d'attraction naturelle entre les utilisateurs pour qu'ils le fassent naturellement ? Qu'ils le feraient vraiment parce qu'il y a quelque chose d'attrayant à le faire, pas qu'on le pousse mais que le nombre d'utilisateurs ou la communauté font qu'on participe vraiment à ces partages de savoir.

Si je comprends bien, pour moi *self-sufficient* c'est qu'on a une base de données assez conséquente, avec pas mal de variants, avec des fréquences qui donnent aussi une information, les fréquences de variants dans le compte ou dans la communauté, dans le *run*, qui aident aussi à l'interprétation. Je dirais qu'avec le temps, on aura amassé assez d'informations, après ça dépend si l'on veut faire ça sur un an ou sur cinq ans. Sur cinq ans ça peut se faire de manière passive avec des gens qui *flaggent*, après il faut trouver des gens qui *flaggent* plusieurs gènes à la fois, c'est-à-dire des gens qui soient experts

sur les 20 000 gènes que notre génome contient.

Il ne faut pas juste qu'ils *flaggent* BRCA-1 et BRCA-2 donc.

Non voilà, et il faudrait du coup avoir une communauté assez diversifiée et des experts pour nos vingt-mille gènes. Si c'est le cas, on a nos X experts pour nos 20 000 gènes, disons 20 000 experts avec chacun un gène de prédilection, chacun va *flaggé* parce que ça lui fait un gain de temps. Du coup comme la communauté réunie toutes ces informations, tout le monde aura accès à ces 19 999 autres gènes. C'est là que je dirais *self-sufficient* parce que tous les gènes sont couverts par des experts, par au moins un expert dans tous les cas qui a l'expertise de ce gène.

Là ils régleraient l'attraction d'être sur la plateforme.

Oui parce que du coup tous les gènes sont couverts. Avec en plus une couverture dans le sens où les algorithmes utilisés sont très bons, avec peu de faux positifs, donc tu as vraiment une base de données propre et des fréquences plus justes.

Mais ça c'est plutôt sur le long terme.

Oui là c'est plutôt sur le long terme. On n'a pas encore atteint les 1000 experts, et parmi ces 1000 experts je suis pratiquement sûre qu'il y en a beaucoup qui partagent exactement les mêmes gènes de prédilection.

C'est ça, il n'y a pas de diversité.

Non, après il y a certains gènes qui sont à l'heure actuelle pas du tout étudiés, ou moins connus. Donc ceux-là resteront encore un bon petit moment tous seuls, inconnus, orphelins.

Comme dans cinq, six, sept ans on sera que sur du exome, on pourrait se concentrer que sur ceux-là.

Oui mais alors l'exome c'est pareil c'est un outil de travail, mais encore une fois derrière il faut avoir les connaissances pour pouvoir exploiter un exome. Donc si tu fais un exome, tu vas te retrouver avec 30 000 variants, qu'est-ce que tu fais de ces 30 00 variants ? Comment tu vas classer ces 30 000 variants ? Comment tu vas les prioriser ? Pourquoi tel variant va prendre le dessus par rapport à un autre ? Qu'est-ce que tu vas regarder en premier, ce sont les variants qui sont connus pour être pathogéniques déjà. Après le reste qui existe, ils sont simplement inconnus à la communauté médicale et scientifique parce que pour le moment personne n'a travaillé dessus, parce qu'il n'existe pas de littérature à ce propos.

J'ai une dernière question, pourquoi on a fait un modèle de transaction et pas une licence par année pour les utilisateurs ? Tu sais ?

Bonne question. Je dirais parce que tout simplement on gagne plus d'argent en faisant par transaction que si on avait un *software* annuel. Si on disait « Vous payez 100 000€ par an », déjà 100 000€ il faut pouvoir les déboursier, après on peut toujours dire que payer en quatre fois ou tous les mois n'est pas un souci, mais même pour le client c'est plus intéressant parce que s'il y a une année pendant laquelle il n'a pas autant d'analyses, il ne va pas payer 100 000€ pour dix ou mille analyses, ce n'est pas intéressant. Pour lui c'est beaucoup plus avantageux de payer par analyses que par an. Déjà il y a une plus grande flexibilité pour le client, s'il ne veut même pas analyser avec notre *software* et bien il peut faire autre chose, ce n'est pas un problème. S'il paye par an, il sera contraint de payer de manière annuelle pour un nombre d'analyses qui aura peut-être chuté ou augmenté. Après c'est vrai que pour ces grands laboratoires qui dépensent 600 000 ou 1 000 000€ par an, en France par exemple le CHU de Nantes qui a signé avec nous un marché négocié, qui utilise notre HCS, fait tout le drylab avec nous et dépense à peu près 500 000 à 600 000€ par an, du côté commercial on leur a dit que s'ils dépassaient un certain montant, ils auraient une réduction. On est flexible, donc eux ont tout à gagner.

Comme ça le prix est plus adapté au client spécifique.

Exactement, et par rapport notamment à sa demande à un temps *t*. Il ne va pas payer 100 000 ou 20 000€ d'emblée alors qu'il ne fait pas d'analyses. S'il y a un mois durant lequel il ne fait pas d'analyses, ça peut arriver même si de nos jours c'est de plus en plus rare, au moins il ne payera pas un sou.

Génial, parfait. Si tu vois SOPHiA dans le futur, tu crois que l'on se dirige vers un marché avec une plateforme globale dans l'analyse du MJS ou bien on se dirige plus sur plusieurs plateformes. Est-ce que tu penses qu'on se dirige vers un marché de monopole ou qui restera compétitif dans le futur ?

Dans le futur ?

Oui. Si on continue sur cette lancée de pomper un peu les données de tous les hôpitaux, est-ce que tu penses qu'une plateforme va se démarquer par rapport aux autres parce que la base de données va être beaucoup plus imposante que ses concurrents ?

Alors je pense que oui, c'est vraiment notre force. C'est pour ça que je t'ai dit que SOPHiA GENETICS c'était avant tout la communauté pour moi. C'est vraiment sa valeur, contrairement à tous les autres. On a une communauté et on travaille avec des fichiers de dépôts, les fichiers FASTQs donc les *row data*, et on s'occupe de tout le *process*. La communauté c'est vraiment le gros point. Et je dirais qu'en effet si on met tout ça en place, il y aura une très forte plus-value pour SOPHiA.

Exact, et là c'est encore le travail à faire donc.

Et il y a beaucoup de travail.

Mais l'impact serait majeur.

L'impact serait majeur oui, ce serait bien pour nous.

Parce que là c'est l'effet de réseau pile dedans, s'il y a vraiment un échange par rapport à des données.

Tout à fait. Après le problème c'est qu'il y a des hôpitaux qui n'informatisent pas, qui ne dématérialisent pas du tout et qui sont encore à la feuille de papier. Et il faudrait que l'on touche les cliniciens aussi parce que ce sont eux qui donnent les traitements. Donc il faudra faire plusieurs types d'interfaces, une interface pour les biologistes pour qu'ils aient accès à tous les variants, une interface pour les cliniciens car eux ne veulent avoir que les réponses à leurs questions et vont dire par exemple : « Je veux savoir si mon patient a une mutation BRCA-1 ou -2 et PALB2 », point à la ligne. Le biologiste lui va utiliser le HCS, mais bien sûr il y a plus de gènes que BRCA-1 et -2 et PALB2, donc il va répondre à sa question et l'idéal ce serait d'avoir une autre interface où le biologiste va pouvoir cliquer et envoyer au clinicien. Et le clinicien aura une interface avec les données envoyées par le biologiste. Et là, il va pouvoir dire « D'accord, il a telle et telle mutation BRCA, donc parfait je vais pouvoir traiter mon patient avec telle chimio ». Et si nous on met ça en place aussi, on est loin de tout ça mais ce serait génial, l'idée ce serait de mettre un choix multiple ou faire du *free text* et ensuite SOPHiA, l'utilisateur officiel de SOPHiA va intégrer ces informations en *free text* pour les exploiter et créer elle-même un choix multiple. Et ce sera basé sur la connaissance non plus que des biologistes mais aussi des cliniciens. Et c'est là que l'on va pouvoir monter toutes nos données phénotypiques et cliniques.

Et l'intégration comment on appelle ça ?

Le LIMS. Ce sont deux aspects en fait. Je pense qu'il faut vraiment que l'on puisse donner accès aux cliniciens. J'en avais déjà parlé l'année dernière, je me rappelle quand Yourgi avait présenté pour la toute première fois ses diapos avec la vision et la dernière diapo avec le patient où il disait qu'on allait pouvoir aider à mieux interpréter les datas et mieux aider les oncologues à prendre la bonne décision. J'avais dit que si on voulait vraiment ça, il fallait que l'on donne accès à l'interface aux cliniciens puisque ce sont eux qui détiennent ces informations de traitement, ce sont eux qui prennent la décision du traitement et non pas les biologistes. Le biologiste il fait le travail, il répond à une question mais c'est l'oncologue qui décide du traitement.

Donc on pourrait encore intégrer un autre type d'utilisateur sur la plateforme ?

Exactement, et c'est ça qui serait bien. Mais le problème, c'est qu'il faudrait qu'on ait plus de monde et c'est compliqué. Mais ce serait super.

Parce que là ce serait un effet de réseau encore plus grand en fait ?

Oui beaucoup plus grand puisque tu as pratiquement tous les acteurs de santé. Et j'irais même encore plus loin, ce serait de permettre de donner accès à l'interface au patient lui-même pour qu'il puisse voir lui-même son compte-rendu sur Internet, donc tout dématérialisé.

Donc là on serait sur le B2C.

Et pouvoir même discuter via l'interface. Le patient pose une question, il envoie un mail via l'interface, par exemple concernant les effets secondaires qu'il a de la chimio en disant « Voilà, j'ai eu ça et ça ». Ça c'est aussi important comme information, de connaître les effets secondaires du patient. Je vais encore plus loin. Ça peut être hyper intéressant puisque du coup on pourra dire qu'il y a tel traitement qui fonctionne, que les dix-mille patients ont survécu un an de plus, mais ce serait encore mieux de savoir que par contre il y a des effets secondaires, par exemple qu'on a vu que pour 5000 patients il y a eu vomissements, fatigue, perte de cheveux et que pour les 5000 autres c'était aussi vomissements mais sans perte de cheveux ou sans fatigue. Ça permet d'avoir une idée globale des effets du traitement.

Est-ce que si moi en tant que clinicien j'arrive sur la plateforme, est-ce qu'il n'y a pas quelque chose sur la plateforme qui va naturellement me pousser à mettre les phénotypes ? Sans qu'on me pousse ?

Ça pour le clinicien clairement ça sera si lui il a un intérêt personnel, si ça lui permet de moins chercher l'information donc de gagner du temps. Il n'aura pas besoin de perdre du temps à chercher dans les bases de données les essais cliniques, les traitements disponibles, ce qui a été fait dans la littérature mais qui n'est pas répertorié comme *FDA* ou *EMA approved* parce ça n'a pas encore été approuvé car il n'y a pas eu encore une très grande cohorte de patients et que ça n'a pas été prouvé avec une évidence de niveau 10. Si on lui apporte cette information en encoportal mais améliorée, c'est-à-dire dans tout ce qui en gros n'est pas connu... C'est ça qui est le plus intéressant parce que sincèrement, l'encoportal à l'heure actuelle il est vraiment à améliorer, les indells qui n'y sont pas et c'est une catastrophe, et notamment toute la partie recherche parce que les cliniciens actuellement ils vont chercher dans la littérature.

Il y a un client qui aime bien Encoportal aux Etats-Unis ?

Oui c'est vrai, mais s'ils l'utilisent au quotidien ils vont peut-être déchanter après.

Il est tombé amoureux apparemment.

Oui et bien j'aimerais bien travailler avec lui tu vois dans le futur, ce serait intéressant. Mais nous on apporte au clinicien tout ce côté qu'il doit faire, c'est-à-dire lire la littérature. Par exemple il a un patient avec une mutation, mais cette mutation est inconnue au bataillon, lui ne la connaît pas et il n'y a pas de *drug* disponible, il n'y a rien. Et bien ce qu'il peut faire, c'est aller chercher la mutation sur Internet. Il va tomber sur des papiers, de la littérature, des articles ; il y a peut-être dans le monde dix

laboratoires, dix cliniciens qui ont eu des patients avec la même mutation, et bien là il va le voir et se dire « Super, il y a dix patients qui ont exactement la même mutation, qu'est-ce que leur clinicien a décidé pour eux ? ». Parce que lui c'est la première fois qu'il a un cas clinique avec cette mutation qu'il ne connaît pas du tout, et il va passer du temps et du temps à chercher. Et si nous on peut faire ce travail, et bien on lui fait gagner du temps, et c'est là qu'il va se dire qu'il y a de la valeur. Je vais en effet fournir mes informations phénotypiques et cliniques parce que je sais que derrière SOPHiA va mouliner pour me donner les trois propositions thérapeutiques les plus pertinentes, et après les trois de niveau on va dire intermédiaire parce que deux articles parlent de ça, sachant que dans les trois premier cas il y en a peut-être dix qui ont des valeurs à peu près similaires, c'est-à-dire que les patients ont survécu. Et là il va gagner du temps et il aura une information mâchée...

Mais il ne serait pas attiré aussi par ce que pense l'hôpital, le CHU de Nantes par exemple si je suis un docteur des Etats-Unis ?

Là encore avec le *flagging* personne ne sais qui *flagge* quoi.

Je sais mais dans le futur, est-ce que ce ne serait pas aussi des choses comme ça qui pourrait attirer plus d'utilisateurs ?

Le problème c'est que la législation n'est pas la même d'un pays à l'autre, aux Etats-Unis c'est la FDA, en France c'est EMA, en Angleterre c'est NICE, donc c'est très diversifié. Je pense qu'à l'échelle du pays oui, donc là en plus il va falloir prendre en compte la législation de chaque pays, parce que certaines *drugs* seront peut-être *approved* aux Etats-unis mais pas du tout *approved* en Europe, donc il faudra déclasser ce type de traitement en disant « Là ce n'est pas *approved* donc ce n'est pas de niveau 1, peut-être de niveau 2 voire 3 ». C'est un gros travail et sincèrement, soit il faut travailler avec des partenaires qui sont rigoureux et sérieux, mais si on veut vraiment aller vers là il faudra faire ce travail un peu par nous-mêmes.

Génial, merci. Au départ je voulais vraiment me concentrer sur SOPHiA mais mon promoteur m'a dit que ça pourrait être intéressant aussi de comparer avec des concurrents. Ce qu'ils utilisent comme *business model*.

Les concurrents dans le domaine de la santé ?

Non dans l'analyse NGS.

Oui c'est ça dans notre domaine, parce que tu me parlais d'Uber...

Ha oui non, en théorie j'ai pris Uber, Airbnb enfin les classiques, et après en pratique je prends le cas de l'analyse génétique avec le NGS. J'ai envie de voir ce que les autres utilisent, voir un peu ce qui marche bien. Mais je pense que

144.

SOPHiA a été dans l'article de MIT, pour avoir un *business model* innovant.

Tout à fait. 30ème sur 50 je crois, si mes souvenirs sont bons.

J'essaie de contacter la journaliste mais elle ne me répond pas.

Appendix 3.4.Kenneth Jones n°1

What do you think about ACMG guideline?

Yes it is essential for US medical institutions to follow those standards. Sophia needs to adapt to those requests if the company wants to be dynamic on the US market. At GeneDX we use those guidelines to be present and respond to customer demand.

How was the platform at GeneDX that you use to play with?

Yes it was simple to use. It was an in house developed platform. Using databases from GeneDx and other public databases as Cosmic, Clinvar to be the most updated as possible. Using external databases is also crucial to be aware of what is happening on the market.

Did you see the difference between Sophia Platform and GeneDX?

I think, Sophia needs to add more databases! At least 20 databases for cancer are needed. At GeneDX with have a lot of different databases to have more resources as possible to give the best curation.

What do you think about the community aspect, hospital do want to share their data?

I Know some hospitals don't want to share their data. However, if hospitals want to share, Clinvar is the main hub, because it's centralize information They don't want to share, because there are scared to classify a variant pathogenic for example, and all the other ones benign. They are scared to do a mistake in the classification. This is creating liability issues and could threat the hospital's reputation.

It is their own interest to not share the data?

Yes the classification data. Or some institutions don't have the resources to share. It demands a lot of work, one employee full time to share the data. They do need to be aware about juridical and liabilities issues, if they something is wrong about the genomic data, they could be sued.

How to increase network effects ?

It is the integration system, which is the most important. The lims system is crucial. Lims system is essential also because the integration of NGS data analytics platform needs to be more fluid and gathering the all the data (phenotypes, treatments...) through labs could be huge. A lot of prospects, institutions want this before the collection of data. It's enhanced simplicity of use! Other competitors go also in this integration direction, doctors don't want just to start from the FASTQs, they want all the upstream analysis. For instance, how they are running DNA extraction; how they go into the library preparation

What do you look the most on NGS platforms ?

For variant curation, the most convenient is how quickly classify variants. What is the fastest way to classify variants? A variant that is not yet classify, a new variant. What do you do? Search on pubmed, google just to extract informations from public

databases. This the longest work to do, read publications to better classify variants. If the Software could read publications and do this work for you, it could be amazing.

Do you like to see the community interaction on the platform?

It is useful, to see how many other institutions pre-classify. But is always the easiest one that are shared to take less risks. If the machine classify pathogenic, you need to rely on the machine, still a big step for human mind. Platforms need to find mechanisms to push people to share data. I heard Sophia users are just sharing through flagging variant classification. It is just the numbers of hospital that is given to users that flagged this variant the pathogenicity level. The platforms is helping to share the data.

What do you think are differences between Europe and USA?

It is still the ACMG guidelines. There are a lot of differences in the standards. Sophia needs to adapt to the US standards, such as HGMD data base (bought by Qiagen). They need this kind of database.

What do you think about Sophia's future vision, about gathering all the genomic data bases?

All the commercial lab use Clinvar as a worldwide comparison already. They share those data on this platform and other database. Labs learn from their inside database, from their own data and external databases such as Clinvar. The thing is, you cannot trust Clinvar because there are too many mistakes, not clean, contradictions between labs. GeneDX have specific guidelines to filter and trust good labs that share the data on Clinvar so GeneDX can make good decisions.

What do you think about the business model NGS platforms?

It depends on the hospital, how much money they can make of it. Every analysis is getting cheaper and cheaper with technology improvements, market is competitive. The business model needs to be competitive and offer competitive price, cost benefits products. Its depends on how many samples; If big volume, membership fee is interesting, if small volume transaction fee is rather more interesting. Everything is cost/benefit. On the market, main actors are doing both models of fee. Some are doing annual membership, others are doing transaction and for now, both works.

Do you think we could add a pharma on the platform and do some advertising?

Oh yes. Pharmacogenomics is getting bigger and bigger, pharmas need more informations about genes mutations. Do some partnerships with Pharmas to get even more customers, it is about data driven medicine. Pharmas are looking for partners and are attracted by NGS platforms to collect data. We need to see the good equilibrium of partnership we could do with them.

How many samples per year you developed at GeneDX through NGS analysis?

Around 100'000 thousands, we have developed a team of BioIT and IT team to do the analysis. We also offer Micro array not only NGS.

How does it work at GeneDX for the pricing?

Price depends on the panel. Usually it is 3000 \$. Everyhting is about integration and the service, if labs are happy and financially happy with a simple process to do the

analysis through ACMG guidelines, they will be satisfy. 3000\$ could look expensive but if the process is working, labs will use the service because it makes the doctor life easier. It is also important to target small labs that really need the solution.

Do you think we are going towards a monopole market?

In 5 years, the NGS data analytics platform market will be still in competition and data sharing still a big issue in the medical sector. However, we can see network effects happening, let's see what is the future reserve.

Who use the platform?

Doctors don't use directly the platform, we use it, do the curation, and send the reports to doctors who just signed it. Curation is very important, useful and huge value information for the institution. GeneDX clinical lab was the biggest curation lab in the US.

How to push them to curate the data?

Don't give the pre-classification, they would be forced to do it and they need to follow ACMG. ACMG is not doing a good job though but we use it for its standardisation

Appendix 3.5.Kenneth Jones n°2

And so they have like a system integrated in the labs also ?

Yes they have a full pipeline including the accession and lab management system improvements needs to be done. Like the tracking, the accession and the logistics and the lab tracking yet. lab tracking is very important.

I don't think they're using machine learning right?

No I don't think so. But everyone is going to that direction.

And other competitors such as Watson IBM they take the VCF files and these guys already have some issues with the alignments; so we prefer to take the FASTQs. So I know this, what of about GeneDX, what was the charging model, what was the business model about? Did you charge by samples also?

So Gene DX is a commercial laboratory. And so we get samples from hospitals and from doctors, different doctors. So they send us the samples and then we extract the DNA. We do everything and then we send them back the report. We have a lot of directors; we have a lot of doctors. The people I'm working with everyone has a PhD or a doctorate. We do all the process, all the testing and then they just have to send us a report.

Ok did you charge per report?

Yeah we charge per case.

So if you want to do a test it would be like 3 to 4 thousand dollars.

It depends on the panel, if it is a whole exome, 10,000 per work.

So what do you think at Gene DX, if you think like per sample what's the best solutions to charge people or do you think maybe like a membership fee would be better?

It depends. The membership fee, so is there a limited number of samples?

It's limited but it's a membership you know, so they don't need to pay for each sample they send.

I don't know if they would be interested. At least in the beginning, because they don't want to pay a large amount if they only have a few samples. So, so maybe like, maybe you can have that kind of program later, later you know if they reach 100,000 samples per year. Then you can have some kind of membership.

You're not too expensive compared to your competitors?

We are about the same competitor lab core.

What about the history of Gene DX, did you see like an evolution at like 5 years ago what happened? How was the growth of GeneDX?

Ok so, so let me tell you like usually there are 3 criteria for the client to pick us: first is how easy, how easy it is to order using the patient portal. So number 1 is how easy they can order compared to other companies, their request or.... They want to order a test, they want to do it fast and easy right. So the patient portal has to be user friendly. Number 2 is pricing, how expensive it is. Number 3 is: what's the turnaround time? I

think now the standard is about 2 to 3 weeks for most labs. Some, some labs their turnaround time is like 3 or 4 months.

And customers were not saying they don't need the results like immediately?

Everyone wants their results immediately. So we have more than 300 panels at GeneDX. People, people can pick from, so the more selections the better.

You think because a hospital is going progressively to choose to not send it out and do it in-house?

I think most hospitals they want, eventually maybe not now; eventually maybe 5 to 10 years later they want to do this in-house. I think it's more expensive to send it to someone else. You know the next generation sequencing is getting cheaper right. So if it's getting cheaper, then they can do it in-house and they just need the infrastructure.

Do you think it's not a paradox to charge them and in exchange we use also their data to improve our technology?

Yes, take this example called You Gene, they have a new platform, so basically, people can create and then they sell the information.

They can sell the information to whoever wants to buy it. So you can buy a whole bunch of variance that is curated. So you can curate and if someone buys it, you get paid.

- Oh it's because right now people like because we have this flagging system on DDM and basically when they do this, they work for us because they improve you know the data created on the platform.

- **Yeah but maybe another model would be, more they curated it manually, the more discount they get for curation. I don't know, one variant including one free sample or half price for the next sample. Maybe we're thinking about how to push people to a flag. The more curations they do, the more discount they get or they get 50% off for each sample of each variant.**

- ME: I talked to one of the biggest lab medical directors in Belgium; he told me if he sees like a logo of a pharma on the platform he just quit the platform because he's seeing like we use his data commercially and sell it. So he doesn't want that we use his data for commercial purposes, so it's, it's very difficult to monetize this.

Did he say why? Why he doesn't want?

Because he thinks that it's a scam because it's, it's his data and he doesn't want his data to be used for commercial purposes. I don't know from the legal perspective; I don't know if it's illegal to use it but to sell it to another, like sell it to a pharmaceutical company.

What if you use it? You know it's ambiguous because you could create a partnership with a pharma you know.

Yeah it's those data is confidential. It's just might be quite so tricky because genetic data is so sensitive. Maybe you guys can do it for free and then every time, let's say

they use it right; and then this is completely out of the box. So let's say you guys can install the software, Install the DDM for free. But every time when they have to create it, let's say that they curated it manually, Let's say they just use the classification from your machine learning then they don't have to pay for anything.

I've thought about this but you think like people are not going to just curated the data badly? Maybe they are not going to you know like to put 100% commitment about this.

Well that's, that's why you need to build a platform for the ACMG guidelines so they have to use the data to curated it. Following the software, like when you build your software for the ACMG guidelines. So once they create it using the ACMG guideline, then they can submit it. Once they submit it, they get charged for that. They won't get charged for the whole samples. They only charge for like per variant; I'm sure most of the sample have variants. They can create a system charged by curated variant. Then if there is no curation done, then its negative priced ect...

But ACMG could like guarantee that it's a good, a good creation, like a good creation?

Well with guidelines, so i think if you follow the ACMG like if you build that platform for them to use, then they have to follow it; right? They have to check all the publication, they have to check different data bases that you show them, they have to check everything, they have to check the frequency. So you have a list of seven or eight things which I know in detail like one of the things you have to check then if you have the viewing platform then of they submit it through the platform, then it should be good curation. If you don't have a good platform for them to use to curate it, then they do it randomly. Then it may not be good. it's interesting to see how we can improve this.

Yeah what do you think about; I know you don't know well Sophia but just the differences between Sophia genetics per sample model charge and Gene DX, you think it's quite the same? We don't do the reporting but like i think it's quite the same business model we just send a case, then we charge for it?

You guys have three different products, which one is I think the three of them it's per cases; I think it's the same as Gene DX. You do IDS so basically I think that Gene DX and Sophia similarities are on the IDS such as Labs sending the samples to us when they don't have a sequencer. Like their doctors send it to us and then we send them the report. Definitely I think like the report is important, is very important because no one wants to do the reporting but instead they spend a lot of the time writing the paragraph. Because competitors are doing it also. Most of the competitors they already have a report and they generate the report after the curation. The price of IDS is going down also so we need think about his also. I think you guys can be more flexible because you are still developing new products all the time compare to GeneDX. Sophia Genetics has 3 different business models through 3 different approaches, which is interesting. The thing is to improve the software is to get the upstream such as the accession and the logistics and the lab management. And the downstream will be the ACMG creation for the variant.

Appendix 3.6. Kevin Puylaert

Hi Kevin, thanks for your time. Can you first explain me what is your activity at Sophia ?

My job as head APAC at Sophia Genetics is to manage sales outside Europe. Concretely, our job is to make first contacts with prospects and to establish a professional relationship through a long-term view to finally sign contracts. We have a global approach with each country, trying to show our international strength. We are today in 50 countries and work with 350 institutions, each one seeking for new opportunities in next-generation sequencing data analysis.

As a manager in your company which values do you fight for?

The most important value for me is the attitude in front of a customer. You need always to be nice and clear, but the most relevant attitude for me is the optimism. I like people who are optimistic and rigorous. Moreover, I like people who strongly believe in the product we sell. Customers need to feel our passion and determinism about Sophia.

What are the main stakeholders in the company?

Of course, the customers, the direction committee, the expert committee and the sales team. As a relatively small and dynamic company, communication between employees is efficient. As I mentioned, communication is very efficient in the company. The common core values are well shared in the company. I would like to emphasise on the recruitment, as head of the sales teams outside Europe, I am constantly doing interviews (company growing so fast) to potentially recruit new talents and what I am looking for the most is this new talent could or not share our core values at Sophia Genetics. This is super important. As we are in a multi-stakeholder company a lot of diverse opinions are created and for me, we need to reinforce common values to be the most as possible reunited and strong.

Could you explain Sophia main value proposition to customers?

For me, the most important value proposition to customers is that we bring to them a unique solution, which is the access to a platform that proposes super precise NGS diagnostic analysis based on thousands of different **clinical** patients around the world. I would like to emphasise on clinical patients, which is the most relevant source of information today and on the **standardization of the data** that we do in-house to deliver the most useful insights from big data. We bring to the labs and hospitals via the platform a huge community of experts and knowledge that allow labs/hospitals to strongly enhance their NGS diagnostic analysis in the clinical world.

Tell me more about the community...

The community regroups today 350 institutions around the world with customers based on 110,000 patients already analysed throughout the world. Thanks to this community we try to gather the most data as possible to improve our algorithms through machine learning and artificial intelligence. Today, institutions see how uniquely many other institutions flagged the variant pathogenicity. We hope to share

more informations in the future, but data sharing is still a ambiguous subject, some intitutions move towards more data sharing others prefer to back off, it depends... However from my experience, insitutions are very interested by the community aspect and do not apprehend this charachterist. Doctors know that this solution is the future and want to be part of it.

What the network effects created on the platform?

Mostly direct network effects. What I've seen is that the 110 000 patients and 350 institutions is a key argument for other insttiutions to join. When I m going to visit a new prospect, the doctore is is impressionate about the number and want to join us because the value created on the platform become bigger and bigger as the diagnostics is moore precise due to the increase numbers of patients analysed. This is a huge effect on the platform, it is a virtuous circle of users. For instance, we conquer south America very rapidly because they saw our hug value brought through the community. Same with autralia...They saw the value of genomic data and how we face new challenges This is the main network effect created. There is still not a lot of interactions between users. We could create for instance a forum on the platform, but I think that people don't have time to add comments etc... We already starting to have more exchanges, but it's the beginning. Still achallenge to motivate doctors to be interactive on the platform.

Today, what we can see is that some oncologists in europe start to add some comments, some data to the platform. Different institutions shared some thoughts about specific cases but still nor through the platform. We wil see how sharing data evolves in the future.

What is the business model ?

We charge per sample uploaded on the platform, transaction fee business model which currently works well because this prove our quality services. If insitutions are not happy with our services, they can quit whenever they want,there is no annual memmbership. The two main actors on the platform are the institutions that upload the data, and the people who analysing the data. Maybe in the future we will open the platform to other actors as pharma that want to advertise their drugs, but I don't think so because we are losing our objectivity and agnostic approach. Maybe there is other ways to charge and monetize the networks effects by adding other agents on the platform but transaction fee is currently working well, network is growing aggressively.

What about your bundle solutions and IDS, does those product influence network effects?

Yes our bundle solution with the chemistry bring a great competitive advantage on the platform. Of course, our core business is the dry lab analysis, but bringing the bundle solutions, means the chelistry, convinve new intitutions to join the platform and the community. Our pricing is competitive on the market, with a top quality. People like to buy in bundle solutions. Another smart move, is the IDS product, we have created new network effects by allowing another type of customer who didn thave the instruments and the sequencers to send dna analysis to partners lab that already work with us. They were attracted by this service because we had already a strong expertise with labs and hospitals that already had a sequencer. This side attract the other one. In this way we face the chicken or egg issue, by attracting one siede of the marjet withoou dry

lab and the other type of customers with this side. However, I think we are still in a B2B business and not C2C.

Is Sophia a technology creator or a network master?

Definetly a network master. We have strong network focus with a core value of data-driven medecine, meanwhile we are a tech company! We still to find new ways to push oncologists to share their thoughts on treatments on the platform, push interactions and enhance the network effects but it already happening progressively, we are going towards it with the oncoportal. In the future, Sophia will become more and more precise and will have a huge data base that will bring a great competitive advantage on the market. Our dream is to match cancer variant mutation in all other world and furnished an optimised precise treatments.

Towards one platform on the markert?

Yes definetly, I think one NGS platform will dominate the market due to stong netwek effects and user base created. This will leads to a monopole market because barriers at the entry will be high.

Appendix 3.7. Dimitri De Boose

Moi je suis en train de me concentrer sur les réseaux d'effet sur les plateformes, la façon dont les utilisateurs s'influencent entre eux, ce qui se passe sur les plateformes au niveau des réseaux créés puisqu'il y a des réseaux directs et indirects. J'essaie de voir comment les capturer et trouver les différentes façons de les monétiser, dans quelle direction on se dirige en ce qui concerne la monétisation des réseaux d'effet. Et donc j'applique ça dans le cas de SOPHiA GENETICS qui utilise une plateforme pour rassembler toutes les données médicales et qui utilise aussi l'engouement et l'attraction que peut avoir une plateforme dans le domaine médical. L'idée ce serait d'appliquer le même réseau d'effet qu'Uber par exemple aux hôpitaux, c'est ce que j'étudie. Je pense que toi tu es aussi pile dedans avec Bsit.

Comment vous faites pour attirer des nouveaux utilisateurs ? Quelles sont les attractions sur la plateforme ? Quelles sont vos *metrix* ? Quel est le *business model* par rapport à ce que vous utilisez sur la plateforme ? Comment vous chargez ? Comment on charge ? Quelle est ta position dans l'entreprise ?

Je dirais tout ce qui est le développement du produit, tout ce qui est levée de fonds et puis après la coordination de l'équipe, mais on n'est que vingt. L'effet de réseau c'est quelque chose d'hyper important sur lequel on essaye de travailler beaucoup. Mais je peux commencer par le *business model* peut-être si tu veux, comment est-ce qu'on monétise. Alors au début on ne faisait que de la monétisation par transaction, que du baby-sitting, maintenant on fait de plus en plus de et là on prenait l'équivalent de 15 minutes de baby-sitting. Donc c'est le parent qui décide combien il paye à l'heure, en Belgique c'est 7€ en général, et nous on prenait l'équivalent de 15 minutes de ça, donc un quart de 7€, et ça par transaction. Donc ce sont des *fees* assez bas.

Et est-ce que vous donnez des conseils par rapport aux prix ou ce sont les parents qui décident ?

Non c'est vraiment le parent qui décide et une des choses que l'on doit faire évoluer c'est de lui dire « Dans ta zone en moyenne, c'est autant ». Mais en général les parents connaissent assez bien le prix qu'ils payent d'habitude pour leur baby-sitter.

Le fait de leur laisser de la liberté par rapport au prix, tu penses que ça crée des réseaux en plus ? Le fait de fixer leur propre prix.

Le baby-sitting existe depuis toujours et en général c'est plus le parent qui fixe son prix que l'inverse donc on a juste un peu voulu répliquer ce qui se faisait déjà dans la pratique. Maintenant on évolue petit à petit. Les baby-sitters ou les *sitters* d'une manière générale vont être labélisés avec des *skills* qui sont validés. Et une fois que tu

as ton *skill* spécifique validé, là le prix pourra être choisi par les baby-sitters. Par exemple si elle a fait puériculture elle peut dire « Moi je suis puéricultrice donc c'est 10€ de l'heure », et pour les parents du coup il y a une vraie valeur en plus qui est ajoutée. C'est dans ce cadre-là que les gens pourraient fixer leur prix. Chez nous dans le baby-sitting la chose la plus importante, en tout cas ce que tu acquiers, c'est la confiance, tout est autour de ça. Et je dirais que notre *value unit* c'est la confiance, le *network effect* chez nous c'est vraiment autour ça. Si quelqu'un fait une action complémentaire, donc fait un baby-sitting, remplit bien son profil, il y a chaque fois un peu plus de confiance sur la plateforme. Pour moi c'est ça la valeur de notre plateforme, c'est la confiance qui est créée.

Ca c'est le numéro 1 donc?

Oui c'est le numéro 1. Donc tout ce qu'on fait maintenant c'est de voir comment créer plus de confiance, ça part du profil de la *sitter* qu'on est en train d'améliorer, avec des *skills* en plus. Forcément si la baby-sitter a été chez plusieurs parents avec des bonnes notes, son profil augmente donc un nouveau parent va avoir beaucoup plus tendance à l'utiliser. Chez nous l'effet de réseau se fait vraiment autour de la confiance. Et si on sort dans des nouveaux endroits, dans des nouvelles villes, des nouveaux pays, au début il n'y a rien du tout, il n'y a pas de *rating*, pas de notes, et bien c'est clairement plus compliqué. Tandis qu'en Belgique maintenant des baby-sitters sont utilisés partout et donc ça s'accélère vraiment. On a compris ça et on bosse sur la confiance.

Et la confiance vient principalement des *ratings* qui sont accordés aux baby-sitters ?

Alors chez nous on exprime la confiance sur quatre niveaux, les baby-sitters sont triés. Tu as d'abord tes baby-sitters à toi, en général c'est ceux en qui tu as le plus confiance. Ensuite tu as les baby-sitters de tes amis, ça c'est vraiment très puissant, tu vois chez qui la baby-sitter a déjà été par ailleurs ; là on fait le partage des agendas téléphoniques. Après il y a les baby-sitters de tes cercles de confiance ; là on bosse avec des associations d'expatriés, des choses comme ça, qui eux regroupent les parents et les baby-sitters chez qui ils ont déjà été. Et enfin tu as les baby-sitters qui sont près de chez toi ; donc la proximité est également un facteur de confiance important. Chez nous en fait il y a tellement de baby-sitters que tu vas avoir une baby-sitter à 100 ou 200 mètres de chez toi donc tu pourrais la rencontrer, tu sais qu'elle habite dans ta rue. Le fait d'habiter dans le même quartier rassure déjà les gens. Ça c'est ce qu'on avait au début, là-dessus on rajoute maintenant en plus, c'est ce qu'on est en train de développer, tout ce qui sont les *skills* propres à des baby-sitters et des labels qui mettent un cachet dessus en disant « Identité vérifiée », « Diplôme réussi »... etc. C'est

vraiment là-dessus qu'on bosse pour le moment.

Donc il existe des réseaux d'effet direct, dans le sens où plus il y a d'utilisateurs sur la plateforme, plus ça va attirer d'autres utilisateurs, et il y a aussi les effets de réseau indirects entre les baby-sitters et les parents, qui s'attirent l'un et l'autre. Quelles sont les attractions sur la plateforme ? Ils peuvent parler entre eux ? Il y a des forums ?

Il n'y a pas de forum, non. Pour les baby-sitters, les ambassadeurs, on a créé un groupe Facebook pour qu'ils puissent échanger entre eux. Après ce qu'il se passe c'est qu'assez facilement les baby-sitters et les *sitters* se mettent sur la plateforme et s'inscrivent car ça leur fait des jobs en plus sans rien faire, c'est assez simple pour eux. Ensuite nous on ouvre des zones qu'à partir du moment où il y a un nombre suffisant de baby-sitters dans une zone pour que quand le parent fasse sa recherche, il se dise « Whaou, j'ai tout ça près de chez moi ».

Ca c'est un peu le challenge « *chicken or egg* » ?

Oui. Chez nous ça reste un challenge mais c'est assez simple de le résoudre. Il faut vraiment commencer par faire en sorte qu'il y ait assez de baby-sitters dans une zone donnée : on peut leur dire qu'on va bientôt ouvrir... etc. Donc ça on arrive à assez bien gérer, ils en parlent à leurs amis et au final ça ne nous coûte pas cher du tout et on a beaucoup de *sitters* qui viennent facilement. Après une fois que l'on a ça, il faut communiquer aux parents, et c'est là qu'on bosse pour faire ça encore mieux. On a un système où les baby-sitters peuvent donner elles-mêmes un code et donner accès aux parents à une petite promotion à la première utilisation Bsit, on emploie la presse, pas mal de Facebook.

Et comment vous attirez les baby-sitters dans un premier temps s'il n'y a pas les parents et vice-versa ? Parce qu'au début c'est dur d'attirer un côté sans qu'il y ait l'autre.

Au début c'est assez simple d'attirer les baby-sitters parce qu'ils sont assez connectés donc en général on envoie des choses sur les murs de leurs écoles, de l'univ des choses comme ça, en disant « du baby-sitting, inscrivez-vous », et ça va assez vite. S'il faut on fait une petite campagne Facebook très ciblée et ça accélère encore plus la partie. Franchement ça se fait assez naturellement. Et par rapport au modèle, parce que je n'avais pas fini là-dessus, il y a une partie sur la transaction et ce qu'on a lancé depuis avril c'est un système d'abonnement, une souscription. En fait les parents payent un abonnement mensuel et ils ont un accès illimité à la plateforme ; et ça marche assez bien, on est bien au-dessus des objectifs avec ça. En souscrivant à l'abonnement, ils ont accès à des services en plus, pas juste au baby-sitting mais à de la garde après l'école, si tu veux avoir quelqu'un qui t'accompagne pendant tes vacances, des choses auxquelles

tu n'as pas accès juste avec le système transactionnel. Ce sont des services *premium* donc. Oui exactement, des services *premium*. Et pour y avoir accès, il faut prendre un abonnement.

Et donc le modèle est beaucoup plus attractif que le modèle transactionnel pour les clients ?

En tout cas pour nous ça nous permet d'avoir du cash, c'est plus stable. Après ça ne fait même pas trois mois qu'on l'a lancé, ça marche très bien mais on va voir après ce que ça donne dans le temps. Parce que le plus vrai gros challenge c'est le *platforme lickage*. Dans tout ce qui est service c'est vraiment un challenge parce que la baby-sitter qui habite à 100 mètres de chez toi ce n'est pas évident de faire en sorte qu'elle recommence à utiliser la plateforme.

Et comment vous faites ça ?

Il y a plein de choses. En gros il faut essayer d'offrir une expérience dix fois meilleure via la plateforme que si tu dois envoyer un SMS. Donc la manière de faire déjà tu peux inviter hyper rapidement la baby-sitter, si elle n'est pas disponible tu le vois donc tu ne perds pas de temps, tu peux inviter plusieurs baby-sitters à la fois donc tu ne perds pas de temps, tu peux payer la baby-sitter via l'application donc fini de s'embêter avec du cash et des choses comme ça, tu peux prendre une assurance complémentaire pour ton baby-sitting, ce que tu n'as pas si tu passes hors de là.

L'assurance ça rassure les gens ? Les gens prennent ?

Oui, 20% des gens prennent. Donc ce ne sont que des choses qui rendent l'utilisation de la plateforme vraiment attractive par rapport au simple service. C'est rendre l'expérience la meilleure possible. Exactement. Mais bon, on sait qu'il y a beaucoup de gens qui *bypass*, on continue à travailler là-dessus tout le temps de plus en plus, on a tout un programme de *loyalty* aussi. Chaque fois que tu utilises Bsit, tu gagnes des *sitties* chez les baby-sitters et des *boonies* chez les parents, et nous on fait des partenariats avec des marques, ce qui permet d'échanger ces points contre des cadeaux. Donc ça c'est aussi quelque chose en plus.

Et est-ce que tu as déjà pensé à rendre la plateforme gratuite ou bien à ajouter un service, pas de publicité mais par exemple des industries dans le monde des bébés qui pourraient conseiller d'acheter des Pampers ou des choses comme ça sur la plateforme. Je ne pense pas que ça changerait quelque chose. De nouveau le prix est vraiment très faible, c'était l'équivalent de 15 minutes de baby-sitting.

Donc ça ne bloque pas, ce n'est pas quelque chose qui effraie le consommateur, la plateforme ?

Non, c'est juste que les gens ont des anciennes habitudes en fait. Ils ont l'habitude d'appeler, d'envoyer des SMS. Il faut juste arriver à faire évoluer ces anciennes habitudes, c'est vraiment ça le gros challenge. On sait qu'en France il y a une application de baby-sitting qui s'appelle Baby Sittor, qui était un groupe Facebook à la base, et ils me disent qu'ils ont énormément de *lickage* alors que c'est complètement gratuit donc ce n'est pas lié à ça. Pour certains parents oui, mais ce n'est pas la majorité.

Et est-ce que vous vous voyez quand même mettre un service de publicité sur la plateforme ?

En fait on l'a d'une certaine manière puisqu'on fait des partenariats avec des marques et ça se retrouve dans le programme de *loyalty*, c'est comme ça qu'on le présente.

Et les marques vous donnent de l'argent ?

Les marques nous donnent les produits gratuitement. Ca fait partie du *business model* aussi.

Oui, tout à fait. A terme, je pense qu'on arrivera à monétiser ça davantage mais aujourd'hui on est déjà content de notre idée des cadeaux gratuits sans avoir à les payer nous-même. Et on commence à avoir des choses chouettes, on a eu Zalando qui donnait des bons de 20€ quand même donc ça commence à apparaître.

Donc la monétisation pour l'instant ce n'est pas quelque chose qui freine les réseaux d'effet pour l'instant ?

Non, je pense que ce n'est pas ça qui freine. Mais parce qu'on prend des montants très faibles aussi, on n'est pas sur des montants où l'on prend 20% de la transaction comme pleins de plateformes le font. Si on fait ça, là oui on seraient hyper *bypassé*. Je suis devant le bouquin du *Platform Revolution* là, c'est exactement ce que tu fais pour le *premium service*, essayer de pousser les gens et augmenter la qualité globale de la plateforme.

Je suis aussi en train de voir un peu la notion de *critical mass*, et ils définissent *critical mass* comme un point où la plateforme devient auto-suffisante, c'est-à-dire qu'on ne la booste plus et qu'elle s'autogère avec le nombre d'utilisateurs dessus. Est-ce que tu penses que vous avez atteint une *critical mass* ?

Sur certaines villes oui, et puis sur d'autres non. Chez nous c'est hyper local, presque chaque code postal à sa *critical mass*, parce que tu ne vas pas avoir une baby-sitter de Nice qui va faire ton baby-sitting à Tourquois. Nous en fait chaque fois c'est de la *critical mass* et de la *critical activity* de manière super locale.

Mais pourquoi certaines villes de marchent pas alors ? C'est la confiance encore ?

Oui et après c'est aussi culturel, en Belgique les gens font plus facilement confiance qu'en France par exemple. Et je me rends compte aussi que maintenant en Belgique Bsit à une pénétration intéressante : de plus en plus de gens vont sur Bsit et directement ils savent que c'est une *app* de baby-sitting qui fonctionne bien parce qu'ils en ont entendu parler par leurs amis, du coup tu as un effet aussi de *word of mouth* qui s'accélère, avec la crédibilité de ta marque. Tu ne l'as pas tout au début quand tu commences, évidemment personne ne connaît ta marque, ça prend du temps, les gens doivent changer leurs habitudes. Il y a de toute façon un *timing* pour commencer à utiliser une plateforme. Enfin moi je prends mon expérience personnelle avec Uber, entre la première fois que je l'ai téléchargé et la première fois que je l'ai utilisé, j'ai peut-être mis 9 mois, 12 mois.

Et eux sont toujours dans le rouge niveau revenus ?

- Oui en effet, pas en revenus mais en profit. Ce qui se passe c'est que ce sont des *business* dans lesquels tu veux faire de l'acquisition et ça prend un certain temps avant de le rentabiliser. Maintenant s'ils arrêtaient tout de suite de faire de la nouvelle acquisition, ils seraient en profit. Mais c'est juste que vu qu'ils veulent grandir, être dans plus de ville et être plus présent, le taux d'investissement pour l'acquisition c'est ça qui coûte cher. C'est un peu comme Pitsetting qui a acheté son concurrent, qui a acheté Roger. Oui il y en a pleins qui rachètent leur concurrent. Airbnb a racheté pleins de concurrents à l'époque, leur concurrent anglais juste avant les Jeux Olympiques, leur concurrent allemand... etc. Ça fait partie du truc, ça te permet d'avoir une *critical mass* en un coup et ça accélère ton modèle.

Donc pour toi qu'est-ce qui augmente les réseaux d'effet ? Quels sont les principaux acteurs qui poussent ces réseaux d'effet à devenir puissants ?

Je te dirais de nouveau que chez nous ce qu'il faut c'est la confiance, parce que plus il y a des profils de *sitters* de confiance, plus les parents sont en confiance et créent du coup des baby-sittings. Et s'il y a plus de baby-sittings, il y a plus de baby-sitters et c'est comme ça que se crée le *network effect*. Il faut que tout la *value unit* de la *core interaction*, par rapport au bouquin, chez nous c'est vraiment le profil de la baby-sitter qui va donner confiance. Maintenant on se *focus* vraiment là-dessus, on a compris ça.

Et l'interaction pousse aussi énormément les gens à rester scotché à la plateforme ?

Oui, tout à fait. Et après on va lancer ça maintenant, dans le *network effect* tu prends d'un côté baby-sitters et parents, pour moi il y en a quatre : Il y a des baby-sitters qui attirent plus de baby-sitters, il y a des baby-sitters qui attirent des parents, des parents plus de parents et les parents les baby-sitters. Donc sur ces quatre parties-là, on a à chaque fois des actions différentes. Et effectivement on va pousser les baby-sitters à utiliser l'*app* avec des parents chez qui ils faisaient déjà du baby-sitting avant. On va dire aux baby-sitters que chaque parent pourra choisir un baby-sitter avec il ne paiera jamais de *fees* par exemple. Ça c'est pour essayer de pousser ce truc-là. On va envoyer des emails aux parents en disant « Cette baby-sitter et cette baby-sitter habitent près de chez vous » ou « Elle est prêt de chez vous et a déjà été utilisée par vos amis ». Maintenant ce qu'on fait c'est de demander aux baby-sitters de recommander trois autres baby-sitters à qui elles font confiance. Donc elles savent que ça crée de la confiance, de voir telle baby-sitter recommandée par 25 autres baby-sitters. Et en plus ça crée un effet de réseaux parce que ça a attiré d'autres baby-sitters. Donc vous créez aussi une base de données avec plusieurs réseaux de personnes. Oui, là vue que les parents partagent leur numéro de téléphone, on a au total 7 millions et demi de connections, de gens qui sont connectés entre eux. Moi par exemple j'ai 170 personnes connectées sur Bsit mais bon c'est normal... Il s'agit vraiment de créer l'attraction entre les baby-sitters et les parents en fait, créer cette colle, comme on pourrait dire, à travers la plateforme. Et c'est la plateforme qui donne l'outil. Exactement. Moi je crois vraiment à l'économie contributive et c'est ça que j'aimerais mettre en avant, c'est de faire comprendre à tout le monde que plus ils utilisent la plateforme, plus ils vont apporter de la confiance à la plateforme qui va du coup leur donner un meilleur service. Si des personnes arrivent à comprendre ça, il y aura moins de *bypassing*, de *lickage*. Ça fait partie du réseau d'effet ça aussi oui. C'est du travail de longue haleine. Ça s'accélère mais ça ne vient pas en un jour non plus, il n'y a pas de secrets. Et le réseau tu dois le créer chaque fois un peu localement aussi. Tu vois que culturellement c'est clairement entre la France, la Belgique, la Hollande, le nord et le sud de la France, les gens n'ont pas la même notion de confiance. Oui, et c'est pour ça que maintenant on a créé un système, enfin c'est vers là qu'on va, de labels. Les personnes locales peuvent créer leurs labels, c'est en plein développement maintenant ce sera pour la fin de l'année, mais on va permettre à n'importe qui de créer sa petite agence de baby-sitting ou de *self-care* dans Bsit. C'est vraiment notre vision de comment ça va fonctionner. On a bien compris que la confiance est la clé mais qu'elle est différente d'un endroit à l'autre et que l'on ne peut pas avoir déjà nous-mêmes des Bsit dans toutes les villes du monde ou les villages du monde comme Uber fait. Il y a moyen mais ça te demande des capitaux qui sont délirants donc on essaye toujours de trouver des solutions collaboratives. Tous les challenges chez nous sont résolus en se disant « Comment est-ce que l'on pourrait y arriver en étant plus ouvert, plus collaboratifs ? ».

On assiste quand même aujourd'hui à un marché de monopole avec les plateformes car il y a tellement de données sur ces plateformes qu'il y a des grosses barrières pour les concurrents, pour rentrer sur le marché. Si on

compare Facebook, Google+ ou d'autres plateformes, dans le monde des plateformes il y a souvent deux-trois grands leaders sur le marché. Est-ce que tu vois Bsit obtenir ce marché de monopole ?

Aux Etats-Unis il y a quelques boîtes assez grandes mais elles ne sont vraiment encore que là-bas, et en Europe de ce que je connais il y a un Français qui s'appelle Yoopies et qui est en train de grandir, ils ont racheté une boîte anglaise d'ailleurs. Ils sont plus vieux que nous de quatre ou cinq ans, donc eux ont un peu d'avance. Mais ils ne fonctionnent pas d'une manière collaborative comme nous et ils s'appuient très fort sur le B2B. Parce que dans la monétisation aussi, ça je ne te l'ai pas dit, on commence à vendre notre service d'abonnement aux entreprises pour qu'elles l'offrent à leurs employés. Et ça je pense que ça va vraiment représenter une bonne partie de nos revenus, qui vont être par les entreprises qui veulent offrir à leurs employés un service complémentaire pour s'occuper de leur enfants, trouver des gardes ils sont malades, des choses comme ça.

- Donc sous forme d'abonnement alors ?

Ecoute on a lancé ça en même temps que notre service d'abonnement il y a deux mois et là on a déjà signé avec cinq-six boîtes. On est vraiment en contact avec des grosses boîtes qui sont sur le point de signer. Aux Etats-Unis là-dessus je crois vraiment que ces boîtes-là, enfin l'équivalent de nos collègues américains, c'est comme ça qu'ils gagnent beaucoup leur vie.

Donc vous faites du B2B, du B2C et du C2C un peu ?

Oui. Du B2B2C un peu. Il y a un peu tous les différents axes en fait. Parce que là je suis en train de travailler sur le chapitre « Monétiser » mais c'est assez compliqué parce qu'on voit que le monde réel va plus vite que la théorie. Il y a des *business model* qui arrivent chaque jour et donc c'est un peu dur d'expliquer. Beaucoup d'économiciens on fait des théories il y a cinq-six ans avec des cas très spécifiques, dans les cas de Mais compétition avec trois agents... etc. Mais je ne trouve pas ça hyper intéressant, je crois que l'UNIF voudrait qu'on aille dans cette direction-là mais moi je préfère me concentrer sur ce qui se passe réellement maintenant. moi c'est ce que je te disais, je pense que chaque cas, chaque industrie est différente. Et dans chaque industrie tu pourras apprendre de manière différente. Tu vois quelque chose, tu tâtonnes, tu ajustes, ça a toujours été comme ça. Tu as l'impression que le modèle est clair mais pour arriver à ce modèle-là, pendant des années et des années ils ont testé pleins de choses. Ils sont obligés de faire ça. Tu commences par quelque chose mais tu ne te lèves pas un matin et puis c'est bon pour toute ta vie. Sur Airbnb par exemple, il y a une

partie pour les hôtes, une partie pour les voyageurs, et tout ça est un fin mélange. Et puis tu peux faire payer quelqu'un pour qu'il se liste dessus. Nous on pourrait dire que maintenant on peut mettre en avant certaines baby-sitters si on vérifie leur casier judiciaire, mais elles nous payent quelque chose pour qu'on les mette en avant. Il y a pleins de solutions différentes. Et puis maintenant tout ce qui est abonnements est en train de monter de plus en plus. Ça je vais faire un chapitre « Abonnements », je trouve ça intéressant. Parce que je crois que ça commence. Amazon et autres sont un bon exemple de systèmes d'abonnements dont on sait que ça marche super bien ; tu prends Amazon Prime, tu ne payes plus jamais la livraison. Je pourrais aussi un peu essayer de mettre en parallèle transactions et abonnements, ce sont deux modèles différents mais... Nous quelque chose comme Bsit si déjà ils prennent l'abonnement, enfin si on s'assure qu'ils continuent à utiliser la plateforme sinon ils ne renouvelleront pas l'abonnement, ça nous donne une stabilité de revenus qui est évidemment assez sympathique.

Et le taux de renouvellement vous ne l'avez pas encore ?

Non, parce que ce sont des abonnements de trois mois et on les a lancés il ya moins de trois mois donc. Ça ça va être hyper intéressant comme *metrix*. Oui, c'est le challenge.

C'est marrant ce chiffre de B2B, B2C, chez SOPHiA ils ont eu le même cas un peu aussi. C'est intéressant de voir le parallèle. Eux ont eu un « *chicken and egg* » de problèmes au début : pour attirer une communauté, ils ont dû vraiment solidifier un des côtés de la plateforme pour attirer l'autre, donc ils sont passés du B2C à du B2B, ils ont un peu *switché* leur modèle pour attirer encore plus de réseau d'effet.

Et bien nous avant que ça ne se décide j'avais lancé une autre start-up de *patrol*, c'était Sidebox, qui connectait les gens qui veulent envoyer quelque chose quelque part via des gens qui font déjà le trajet. C'est Haitchers, c'est deux jeunes qui ont lancé ça il y a deux ans. Non, ça n'a pas marché. C'est super compliqué comme *business*. Mais eux ont fait un partenariat avec Total et les pompes pour mettre les colis dedans. A faire marcher je trouve ça super compliqué parce qu'en fait une des choses qui doit pour moi fonctionner sur une plateforme c'est qu'il doit y avoir énormément d'échanges. Et nous au début on était partis sur des envois de privés, ce sont des privés qui voulaient envoyer des choses, et concrètement en tant que privé tu vas peut-être envoyer deux fois par an un colis, plus ou moins. Et du coup si tu n'as pas d'utilisation plus récurrente tu n'en a pas l'utilité, donc je pense qu'il faut trouver plus des professionnels, des ou plus deset c'est vraiment chaud. Nous on était vraiment loin. Et puis tu as tous les problèmes de concurrence aussi, ou de savoir ce qu'il y a dans le colis que tu transportes parce que tout le monde pense tout de suite qu'il peut y avoir de la drogue... J'ai vu aux Etats-Unis un énorme truc qui s'appelle et qui a été lancé en même temps que nous. Ils ont déjà mis 10 millions, puis ils ont remis 12, mais je n'ai pas l'impression que ça explose. Je crois qu'à terme ça va marcher, et ils avaient aussi

fait des partenariats avec des marques de restaurants, une franchise qui est un peu partout, pour déposer les colis. Mais je trouve ça vraiment difficile comme *business* parce qu'il n'y a pas assez d'envois en fait. Ce n'est pas comme la nourriture que tu manges tous les jours, ou du baby-sitting que tu fais quand même toute les deux semaines. Envoyer des colis, on ne fait vraiment pas ça souvent, donc avec le B2B même le B2C du coup ils ont d'autres problématiques. Maintenant on te demande tellement de livrer dans des délais hyper courts que tu ne veux pas prendre de risques. Ce n'est pas une demande de particuliers. C'est chaud, je trouve ça chaud. Je crois que ça va marcher un jour mais dans longtemps et peut-être qu'entre temps il y aura peut-être des voitures automatiques ou des choses comme ça.

Donc c'est vraiment le nombre d'échange qui est essentiel sur la plateforme ?

Oui. Ça dépend lesquelles après, des plateformes comme les nôtres oui, mais si ce sont des plateformes comme Immoweb ou des trucs immobiliers c'est différent. Par contre tu prends du coup un taux beaucoup plus élevé, parce que forcément tu vends ta maison une fois tous les sept ans. Donc c'est différent car il y a moins de transactions mais il faut que ce soit plus entier et prendre plus par transaction. Mais quand tu es prêt à vendre ta maison, la valeur est importante donc tu n'as pas trop de problèmes à payer 500€ si tu trouves l'acheteur qui va acheter ta maison au juste prix. C'est comme sur les plateformes de recrutement, pour poster c'est vite 500, 700€. Donc la valeur de la transaction en tant que telle a aussi de l'importance.

Ha oui ça c'est hyper important donc. Je le mets ici. Parce qu'ils disaient que dans la monétisation ce qui est intéressant c'est de passer d'un modèle qui est gratuit à un modèle qui devient payant. C'est cette transition qui est intéressante, les réseaux d'effet ne sont au départ pas chargés mais on passe d'un modèle non-chargé à un modèle chargé.

Oui, ça c'est très américain tu vois. Aux États-Unis ils ont réussi à faire ça avec pleins de plateformes c'est extraordinaire, tu vois par exemple Facebook, LinkedIn les choses comme ça. Mais c'est vraiment chaud de faire ça. Je crois que c'était faisable avant mais... Les choses vraiment gratuites comme ça aux États-Unis ça passe mieux parce que les gens sont hyper ouverts à tester des nouveautés technologiques, et que les investisseurs sont prêts à investir dedans même s'il n'y a pas de revenus pendant quelques années. En France, les gens sont beaucoup plus frileux et c'est beaucoup plus complexe. Il n'y a plus beaucoup de nouvelles plateformes, à part les réseaux sociaux, qui se lancent en étant gratuites.

Mais c'est un peu le modèle *freemium* qui marche bien aussi, si on voit un peu les jeux sur l'AppStore tout ça...Oui mais bon, ça marche bien pour du *gaming* des choses

comme ça, mais c'est encore autre chose. C'est encore autre chose oui. Mais c'est vrai que c'est un modèle encore à la mode pour l'instant. Je vais essayer de développer aussi ce modèle *freemium*. Mais tu vois à un moment il y a pleins de modèles et soit tu fais quelque chose d'assez théorique et tu peux écrire ce que tu veux, soit tu fais un ou deux cas pratiques il faut creuser les options, se demander pourquoi ça marcherait, pourquoi ça ne marcherait pas... etc.

Appendix 3.8. Thomas Grivegne

What is your role at Bordet institute?

I am business analyst in big data applications. My job is to prepare the development of operational applications that could be use by Bordet institute. I am in the business side and don't have direct contacts with doctors. We act like business consultant for the hospital and see if a specific application could be financially attractive for the institute. We elaborate a precise decisional process for the operational side. We are in management side, less research. We give proposition to directors and they make the decision.

What do you think about big data in medicine application and NGS data analytics platform?

I don't know specifically this type of platform, but this is true, we are working on big data to find new insights all the time. Genetics are crucial in the precision medicine side. We worked also on how to follow a patient treatment life but it's still complicate to coordinate. We don't have issue with big data, it's more about data mining. It's not huge amount of data finally, it's not technology issue, we huge database, the technology is ready and the capacity to treat big data! But how to treat this data, how to get insights? The interpretation depends on the human decision process. At our hospital everyone is talking about Big data, big data, big data as a new revolution. However, the challenge is not the big data itself but the human cooperation behind to understand. Technology is moving faster than human cooperation. Currently, we are more now in coopetition system where each hospital has his own interests and compete. To go further in big data we need to share some data to have bigger insights. There is bad collaboration now. Big data is a tool to support the human knowledge. Technology is complex but this is not the challenge, it is a human challenge! I was working before in scan company managing medical image with very high volume, the technology was not the issue, but the human organization. Humans don't want to collaborate for now, hospital try to monetize on their own side. I read an article about the CERN, it is the same, human interactions are so complex that the challenge is to cooperate, not the technology. But there is always a winner, how to collaborate? Human wants to win, specifically with data, with this race to data. They need to work together.

Do you think that platform strategy helps this collaboration between human?

We need to understand, how a tool can help the human behaviour but sociologically it is very difficult. It's idealistic to consider that human can collaborate because a platform helped us.

Do you think people are afraid about medical community?

Again human... Hospitals need to see what is the cost/benefit from this community. Data is a new currency, data is considered as gold, Google is doing all the money with data. Genetics is the same. It's very complex and idealistic to gather all the data, to do huge data base, there are too much interests in data itself. Could be interesting to analyse big European projects about doctors who participate from different institutions to collaborate. Europe forced to collaborate with industrial company, but

finally the industrial company left with all the collaboration research done. There were not enough incentives for doctors to collaborate in those case and collaboration was not respected. In Belgium there is a case with the platform e-health that try to gather patients medical registers. However, since 20 years, the project is still on hold. I think that these kinds of collaborative projects need to be managed in a public environment to avoid private interest. It is very complex when there is private stakes in the game to build a collaborative project. They call this coopetition again. -Interesting to understand, why we don't trust each other. It's the human mind and behaviour.

How to convince institutions to buy Sophia?

Doctors don't do the final decisions in hospital, it is the financial administration. So you need to convince them that there is a financial benefit behind the use of Sophia DDM.

How to convince doctors to be more interactive, curated the data on the platform?

It's always complicate when you are an external company to the hospital, to establish trust, trust is extremely important. The most important is too establish a trustful relationship. They are scared that you are going to use your data commercially. You ask for a contribution and they pay for it, it's a paradox!

Do you think the fact of being a private company block Sophia expansion?

Yes! They are scared. Medical data is very sensitive. They can be sued, they need to respect all the medical rules, becoming even more stringent. There is a new European rules being voted about data sharing that is even more strict. Don't be surprised that hospitals are not willing to share data, because they are scared that you use it. However, your system looks very protected, data protected and encrypted. As I said, I work for an imaging company before and people need trust!

If there is a new technology that is proposed to Bordet institute, do you take the decision to implement it?

No. I'm doing the business analysis, the cost-benefit strategy. I just present conclusions to the board and they take decisions about it. There is a huge gap between doctors and their administrator. Sophia needs to convince both, quality and financially attractive!

How is Bordet with other hospitals, do they collaborate?

Ah it's very complicated; each hospital have their own standards. So Bordet is very strict and it is difficult to cooperate.

What do you think about monetization of NGS data analytics platforms?

It is also complicated, because there is the approach of open source and institutions want to contribute. Doctors maybe don't want to pay because they already contribute. This is difficult to convince doctors. They don't want to work for NGS platform by contributing and paying. Everybody needs to contribute but it's delicate. Genomics data is so sensitive that monetization is a big challenge.

What would happen if the platform was free?

Yes it would be the best. But how to cover the costs. You could be subsided, be an

association and analyse which business model to choose. The business model is everything. For example, Google search engine and ads. It is crucial to choose the good one. How to create the most fluid experience on the platform. If price is cheap, if juridical issues are solved, if there is transparency, your business model could work.

Do you think we could add another group of agent on the platform such as Pharma?

Yes you could try. Why not? But always, you need to put the customer experience first.

What do you think about the business model of platforms?

Yes, this is the thing. Business models are evolving each day. Freemium is interesting in your case. Odoo is an example; the integration in different systems is interesting.

What do you think about the model pay per use?

4 different levels of payment in your case: Payment per utilisation, membership, freemium 100% free with exploitation of the data and complete open source.

You could charge specific users?

Yes it is possible. There is a lot of different ideas, customers need to see the value!! If doctors upload the data and pay for it, that means they really need the product. The challenge is not the business model, it is the human challenge, cooperate.

What do you think about the price governance?

Yes it could be interesting, people in a platform could manage their own prices and this can push people to do more transactions. Google succeeded to do that, by putting advertisers in competition to offer the best price for the best place.

Appendix 3.9.Dr.Stephane Gerart

Est-ce que tu peux juste nous rappeler ta position dans SOPHiA GENETICS pour l'interview ?

Je suis responsable des ventes au Canada pour SOPHiA. Et je suis le seul pour l'instant. Je ne sais pas, pas bientôt, mais l'idée c'est qu'un jour oui, potentiellement plutôt vers l'ouest parce qu'il n'y a personne pour l'instant. Mais bon, c'est compliqué le Canada.

Et tu as bien réussi pour l'instant ?

Ça va, ça décolle cette année. Ça prend du temps, c'est beaucoup moins dynamique que le reste de l'Amérique du Nord. C'est plus long, c'est plus frileux, mais ça commence.

Et par rapport à DDM, quelle est la valeur principale que ça rapporte au client, tu dirais ?

Je dirais que la valeur principale c'est le fait d'avoir un *workflow* qui est très simple. En gros, ils ont une plateforme où ils font tout, par rapport à la situation du Canada où aujourd'hui ils ont leurs propres outils, ils ont bricolé depuis quelques années mais au final ils utilisent plusieurs *software*, plusieurs solutions, ils sont obligés d'avoir plusieurs sites web. L'idée c'est que dans SOPHiA tout est intégré, je dirais que ça c'est la première valeur ajoutée. La deuxième c'est la qualité des analyses, qui est meilleure que la concurrence. Et je mettrais en troisième l'effet communauté.

L'effet communauté oui, exact, ça c'est intéressant. J'ai parlé aussi, petite parenthèse, à Cam, le nouveau de Washington, qui lui aussi parlait de l'intégration hyper importante, avec le système LIMS, il dit que c'est essentiel.

Pas encore justement, là on a un peu de mal, enfin on dit qu'on cherche. Mais concrètement je crois que c'est assez laborieux. Au États-Unis par exemple, ça va forcément être pas un problème mais quelque chose qu'ils vont devoir faire, c'est l'intégration avec Epic, qui est le LIMS qui cartonne partout en ce moment, ils font un raz-de-marée. Quasiment tous les labos américains l'utilisent et au Canada c'est un peu le cas aussi. Mais il faudra faire cette intégration parce que tout le monde est en train de se jeter sur ce LIMS-là.

Et donc ça c'est LIMS, c'est ça ?

LIMS ça veut dire Laboratory Information Management System je crois. En gros c'est un système qui te permet de gérer tes données dans un laboratoire ou dans un hôpital en l'occurrence. C'est toutes les données sur ton patient, tous les tests, tous les résultats à thème que ce soit génétique comme SOPHiA, que ce soit les radios, les scanners, enfin tout quoi. En fait l'idée c'est que tu as tout sous le même logiciel et comme ça ton médecin accède à ça de chez lui.

Et ce serait aussi une belle plus-value pour SOPHiA si elle arrivait à l'intégrer ?

Oui carrément. Justement on en avait parlé l'année dernière un peu au Portugal, ils pensent que c'est vraiment LE truc. Si tu arrives et que tu dis "Voilà on travaille avec

Epic", parce que Epic tout le monde connaît, ou "On travaille avec d'autres trucs" parce qu'en Europe il y a peut-être d'autres trucs qui sont utilisés, c'est clairement "Oui on est partenaires de machin et en gros on vous assure une intégration complète avec votre LIMS". Ça les gens ils aiment bien parce qu'aujourd'hui sinon ils sont obligés de le faire un peu à la main. C'est vrai que c'est souvent le cas aux États-Unis, surtout qu'ils ont besoin de Epic, l'intégration serait beaucoup plus facile avec, et si on disait que ça vient avec Epic ils seraient beaucoup plus content, je pense. Je ne sais pas, en tout cas ce ne serait pas un frein. Je pense clairement que c'est une valeur ajoutée.

C'est vrai. Et ça demande un pilote ? Tu ne sais pas les prochaines étapes ?

Je ne crois pas que cette année il y ait encore quelque chose de concret mais ce serait bien un jour. En fait, je pense qu'on a besoin de s'y mettre effectivement, donc un pilote ce serait bien, concrètement on saura le faire, c'est de l'informatique, c'est du codage quoi. Mais il faut qu'on soit remis en face du logiciel, ce qui n'a pas été le cas aujourd'hui, et qu'on comprenne sous quel format et qu'est-ce qui doit être apporté. Oui, c'est vrai que ce serait beaucoup plus utile pour les labos en fait. C'est vrai qu'on n'a pas parlé de ça pendant six mois.

Par rapport à la communauté, est-ce que tu vois dans ton territoire quelles sont les réactions par rapport à, quand tu arrives vers un client, si tu dis "communauté" est-ce qu'ils appréhendent ou est-ce qu'ils sont attirés par ce mot ? Quel est ton ressenti par rapport à ça ?

Au Canada, je dirais que c'est un peu mixte, tu as un peu les deux aspects. En fait ça dépend, au Canada en tout cas, c'est peut-être pas transposable à tout le monde, mais au Canada quand tu arrives dans un labo qui est assez expérimenté, ce n'est pas qu'ils ont peur mais ça les gêne un petit peu on va dire, ils n'ont pas forcément envie de partager leurs résultats avec la communauté. Et quand tu arrives dans un labo justement qui débute, ce qui est le cas de beaucoup de labos en diagnostique au Canada, quand ils débutent en NGS ils sont contents parce que ça les rassure en se disant "Voilà, on aura de l'aide" ; ils ne sont pas forcément experts en analyse donc ça peut être un atout d'avoir ces données sur la communauté, ça nous aide nous à faire nos rapports.

Et pour ceux qui ont peur tu fais comment alors, qu'est-ce que tu leur dis ?

Ceux qui ont peur, on leur explique qu'il y a plus d'avantages que de désavantages, que ça reste anonyme. Au final, ils comprennent bien que l'on ne va pas partager leur base de données spécifiquement mais qu'eux aussi bénéficient de la communauté.

A travers le *flagging* ?

Oui c'est ça, et au-delà du *flagging* il y a aussi l'aspect communautaire sur le fait qu'il y a éventuellement un *feedback* positif de la communauté en termes d'analyse. C'est-à-dire qu'au-delà du *flagging*, on partage aussi avec la communauté des améliorations

sur *software* ou des choses comme ça. Donc c'est important d'avoir une communauté pour ça. C'est pas mal aussi ça. Ce que disait Shirine en France, c'est qu'ils ont créé des petits *clusters* de communauté, où ils sont beaucoup plus, ce sont des petits centres où ils sont quatre-cinq partenaires ensemble et donc ils partagent vraiment bien leur données, et on voit que l'intensité des *flagging* est beaucoup plus forte dans cette communauté.

Est-ce que tu penses que ça peut être un modèle au Canada aussi ou on peut se rapprocher de petits *clusters* plus local, de communauté ?

Ça pourrait, je pense effectivement qu'ils aiment bien, enfin en tout cas on commence à avoir ça. Les Canadiens communiquent beaucoup entre les labos, ils partagent volontiers des informations entre eux donc oui, mais il y a des initiatives nationales, notamment le panel canadien de TSC. L'idée c'est que tous les labos l'utilisent et qu'ils communiquent autour de ce kit-là. Donc je pense que ça marcherait, mais nous ce n'est pas forcément notre intérêt. Mais pour l'instant on n'en a pas vraiment parlé justement, on a évité d'en parler pour ne pas qu'au final ils se cantonnent à ne faire que ça, mais si un jour c'est un *dealbreaker* pourquoi pas, on pourra parler des clubs.

Oui, parce que par exemple avec la communauté AstraZeneca, les sept labos qui sont dans Validation Program, ils ne communiquent pas entre eux ?

Non pas encore, enfin c'est prévu, on en parle, mais il n'y a rien de concret encore.

D'accord. Et dans la plateforme tu as deux types de réseaux : tu as les réseaux qui sont créés direct, c'est-à-dire que plus il y a d'utilisateurs sur la plateforme, plus la valeur globale augmente parce que l'algorithme devient plus précis et plus performant, donc le prochain utilisateur qui rentre sur la plateforme va avoir une valeur plus élevée qu'avant, qu'il y a un an par exemple ; et tu as l'effet indirect où là c'est plutôt inter-type utilisateurs,. Qu'est-ce que tu penses de ces effets de réseau par rapport à SOPHiA ? Penses-tu que l'effet direct est assez visible en SOPHiA ?

Oui, l'autre effet je pense qu'il est un peu moins visible parce que justement on ne partage pas et on est clair là-dessus, qu'on préfère dire "Vous devriez venir parce que je ne sais pas, n'importe quoi, parce que tu utilises Sophia DDM du coup vous allez pouvoir bénéficier de leur *flag*", mais en fait ce n'est pas vrai parce que dans SOPHiA il n'y a aucun moyen de savoir qui a *flaggé*. Et la version directe oui, parce que plus il y a de monde et mieux c'est, et puis ils ont l'effet communauté de toute façon sur le *flagging*, mais ce n'est pas comme s'ils pouvaient choisir en gros qui ils *flaggent* et à qui ils l'envoient.

Mais est-ce que tu penses qu'on pourrait augmenter l'effet indirect, ce serait vraiment augmenter le type d'interactions entre utilisateurs donc installer peut-être un forum ou des types de communication entre utilisateurs pour qu'ils interagissent entre eux et qu'ils s'attirent l'un et l'autre finalement ?

-Je pense que oui. Alors aujourd'hui ce n'est pas le business, mais moi en tout cas

quand j'en parle aux gens, les gens se disent "moi j'aimerais bien". Alors peut-être je suis d'accord qu'il ne faut pas que ce soit public dès le début, c'est-à-dire qu'il n'y ait pas le nom de qui a *flaggé* ça, mais les gens aimeraient bien pouvoir contacter le labo qui a fait ce *flag*-là par exemple.

Oui c'est ça. Et ça ils le font déjà un peu en France. En France il y a d'autres logiciels qui permettent de faire ça, où justement c'est public celui qui a fait le flag. Après dans les bases de données publiques c'est public justement, tu sais quel labo a *flaggé* donc éventuellement tu peux le contacter. Mais bon, je pense que ce serait un bel avancement de dire "Celui-là a *flaggé* un truc différent de moi, est-ce que je ne peux pas avoir un contact direct avec lui ?". Et peut-être que ça peut être validé par le labo avant que le contact soit révélé : le labo peut recevoir un message comme "Il y a un labo qui veut vous parler de ce *flag*-là, est-ce que vous êtes d'accord ?" et s'il dit oui je ne sais pas après il y a un forum qui se forme. On doit améliorer le flagging, et l'échange entre utilisateur. Pour le moment, ils doivent nous contacter et demander la permission. Y a des moyens de pousser ces *flaggings* comme des moyens financiers.

Qu'est ce que tu penses du transaction fee et pas le membership fee ?

Question intéressante, je sais pas trop. Peut-être que c'est pour améliorer la transparence des coûts. Avec le système de transaction, il est plus facile de voir où les coûts proviennent. Deuxièmement, les clients ont l'air content de ne peut être engagé directement pour une année par exemple avec une licence, comme ça ils peuvent tester la plateforme.

Est-ce que tu penses qu'on puisse donner l'accès gratuit à la plateforme ?

Non pas pour aujourd'hui, c'est toujours très compliqué. De nos jours, même les plateformes gratuites ne marchent pas très bien, la qualité n'est pas au rendez-vous. Le fait de mettre un prix, permet de garantir une certaine qualité dans le traitement des données. Chacune des plateformes, protège leurs intérêts.

Est-ce que tu penses qu'on pourrait mettre en avant un modèle de monétisation via la pub ?

Très compliqué. Règles médicales sont très strictes. Je verrai mal de mettre en avant un type de médicaments d'une certaine marque sous prétexte qu'ils ont payé. Donc à réfléchir. Mais ce modèle de monétisation me semble très compliqué avec les data génétiques qu'on a aujourd'hui.

Est-ce que tu pense que les produits bundle ou IDS renforcent les effets de réseaux ?

Oui, bien sûr, l'offre nous permet d'avoir accès à des nouveaux clients et d'en attirés d'autres. Cependant certains veulent juste la chimie et pas le dry lab. L'intégration du système est hyper importante dans les labos, appelés lims système, Pour la collecte des données, c'est une méthode qui pourrait grandement aider SOPHiA. Faudra encore

172.

pousser les interactions entre docteurs.

Est-ce que tu penses qu'on pourrait rajouter des nouveaux agents sur la plateforme ?

Pourquoi pas mais faudra toujours impliqué les docteurs dedans pour qu'ils puissent remplir les données, tels que les résultats des traitements, ou données phénotype.

Qu'est-ce tu penses du marché dans le futur ? Ca sera un marché d'oligopole je pense, entre plusieurs plateformes principales, avec des bases de données fortes qui auront un pouvoir.

Appendix 4:

Observations' data

Observations notes from internship:

Blood Cancer Wisconsin meeting (February)

Interested about the product. Surprised about the quality we could reach. Need more explications. Think price is too expensive for them. Interested about the community aspect. Didn't know about Sophia before, harder to convince. Were enthusiastic about the bundle solution, myeloid solution. Liked the transaction fee aspect. Concerned about genomic data security issues. Very price sensitive, private lab, looking for cost-beneficial of product for them.

Vanderbilt medical center meeting (February)

Good timing, they decided to move in-house all the genetic tests. Loved Sophia DDM's design, interface, integration, rapidity of results. The community aspect was a plus, liked the flagging. Had a lot of questions about the platform pre-classification. Competitive price for them. Liked transaction fee to test first the product, in very small volumes. Were attracted by the rapidity of Sophia to get them into routine quickly by offering a validation program. The boss and the assistants were motivated for the integration of Sophia into their system. Much more cheaper to do in-house compare to send it out to private labs they say. Liked the European side of Sophia Genetics, the "swiss" quality.

Walter military medical center meeting (March)

Obviously concerned about genomic data security. Want a data center on campus. Interested by the fact we optimised number of sample on a MiSeq from Illumina for TSM of 16 samples. It's cheaper to do more samples at the same time on the sequencer. No strong follow up, looking for sequencing important volume of samples in the future with other military hospitals. Not interested in the community aspect. Was particular attentive about the cost-benefit analysis of using Sophia Genetics' products.

John Hopkins medical center meeting (March)

Interesting meeting. The bioinformatician and her assistant were there. We did all the first call deck presentation, and a demonstration of the platform. Assistant was enthusiastic about the product, and saw a real plus-value and a time-saving to our solution she liked also the community aspect. However, the bioinformatician was more perplex and hesitant. She asked technical questions she said " You platform looks great but I want to understand how does it work". We answered to the questions, but at the end she seems enthusiastic but never called back afterwards. It seems that her job was to do exactly what our solution provides so this could threaten her job. They were also interested by ACMG guidelines. Does Sophia follow ACMG guideline? They to trust us

but still complicate.

University of North Carolina meeting (March)

The head of lab was directly involved in this meeting. Ready to collaborate with us, having issues with analysing specific genes in their panels (ILLUMINA TST 170) so Sophia could be interesting for them. Not a lot of samples though to analyse CNVs. Maybe interesting to work with their myeloid panel. Impressed by the platform interface, attracted by transaction fee, successfully analysed their data. Future collaboration possible. They looked happy about our meeting.

Dartmouth hospital university meeting (February)

Huge issues with interpreting the results from their sequencers. Data too much in disorders, hard to get useful insights. Used to work with old labs that missed critical variants, they were sued. Missing variants already happen few times in this clinical lab and it's critical to miss variant in the diagnostic world. Looking for a reliable solution that ensure 100 % sensivity and specificity to confirmed all the variants, no place for missing variants anymore. We talked to the head of the lab, Dr. Tsongalis, looked enthousiastic about the solution. Afterwards, we talked with the bioinformaticians who looked also impressed by the interface. Interested by the community But no move forward, maybe a threat for the bioinformaticians.

Cedars Sinai hospital meeting (June)

Meeting with the two chefs of the lab who were interested about the solution. They want to move fast into routine and miss any variants. However, mostly interested by the oncoportal, we said that it was not totally ready for us. They miked our honesty. Interested by our huge client base in Europe, money is not a probem for them. Used to work with PerianDX but they missed variants, want to change. Interested by the transaction fee so they can simply test. bThey are highly motivated by our undle solutions apporaoch where the price seems to be competitive for them. They liked to buy everything so it's easier for them. Interested also by the IDS solution we can provided, sending samples to another lab and doing all the process for them.

Cincinatti Children's hospital meeting (April)

Looking for a bioIT solution. Work principally with whole exomes. We talked to two persons; one was the head of genetic lab and the other one head of the cardio genetic lab. They emphasis on how ohio was strict with genomic data. Both look also interested and impressed by the interface. Very big volumes for paediatrics. We need to convince Brian who is head of the direct users of the platform. They work on MiSeq and Hiseq from illumina.

James Cancer Center hospital meeting (April)

Meeting with pathologic lab. They all said how was integration important. They all are using EPIC integration system and say that it is still weak in NGS diagnostic analysis. However, it is becoming better. Sophia Genetics to find solution to be better intergrated they say. Not interested by the community. Not interested to bring Sophia in the hospital because too much work to integrated it with Epic. Epic tries now to do also to bioinformatics. Ohio rules they say are aslo very strict for genomic data sharing, need to check ith the hospital if they can use a cloud to put data.

University of Wisconsin, Milwaukee meeting (March)

Meeting with germline genetic head of Lab. Specialized in whole exome. She was very dubitative and did not trust our community. Saying that “ How can I trust your community that it is only expert with high quality backgrounds”. The community aspect was not an advantage for her. He claimed that her hospital was one of the most reputed in whole exome. Scared to share her work on this.

Broad Institute meeting (April)

It was the Eye genetic lab, we talked with the head of bioinformaticians team. He said” If you came with this solution 8 years ago to our lab, we would definitely adopt it, but now it’s too late, we already worked 8 years on collecting the data, finding new insights about it such as you did with Sophia DDM” We cannot switch now, it would be 8 years wasting time, specifically because Sophia DDM is a black box, we don’t know which technology you are using, which rules do you apply etc... He also talked about how delicate was genetic data sharing. They worked 8 years on building a strong base of genetic data with relevant interpretation about it. They don’t want to share it to a whole community of people without any incentives. The data is huge value they say, they worked on it, they cannot just share it like that. Each hospital has his own interests to protect their data, they worked on it since years... Sophia Genetics is doing exactly the same for him, collecting the data and create a community, it is the same a hospital but with a global approach. He understands why hospital to share their data to this community.

Cleveland Hospital meeting (March)

Meeting of head of the lab, very interested by first the drylab analysis and afterward the bundle. Still complicate to see what is the best solution for them and what is the difference with IDTs products. He was surprised of our very cheap price for the whole exome bundle solution. Need a meeting with the big boss to convince, he is not the one will directly use DDM. For somatic, he as already convince with archer, big competitor, price cheap and they offer a lot of promotions.

Virginia commonwealth university meeting (June)

Meeting with all the doctors of the lab. they absolutely need a BioIT solution. However, the second head of the lab was not convinced about the term AI, machine learning and how variants were curated. We were not enough precise in technical questions. Not convinced by the community of experts, because they don’t who they are. Skeptical about data sharing. Had huge issue with their Ion panels maybe want to switch to another brand. In the meeting there was the bioinformatician, but he was not very enthusiastic about our solution.

Columbia Medical Center meeting (April)

Very advanced already in genetic analysis. The bio it team didn’t want to work with us. Don’t need the community they say because already too much advanced. Huge volume, don’t see how Sophia can compete in quality or too scared to share their precious data and results. New york restrictions on genomic data is the most stringent in the US, intense paperwork, how Sophia can be better with those rules.

U Miami Medical Center meeting (April)

Interested in bundle solutions. Already strong BioIT team behind, but interested about our chemistry and price competitiveness. They want to buy everything together, the chemistry and the dry-lab analysis. They were very price sensitive and were attracted by the price we offered. Liked the panel HCS and want more informations about the other panels. Strong separation between the lab and the pathologists, need to convince both!!

Moffitt cancer center meeting (June)

Already worked with PerianDX, with annual fee. Were interested by all our different offers and all different application we can provide. Interested in the bundle solutions and our panels. Looked happy about prices, which were surprisingly cheap for them with all the chemistry. Indeed, the bundle aroused curiosity during meeting, because we were not missing FLT3-ITDs and have a nice coverage of CEBPa with this chemistry. They were already using PerianDX for the bioinformatics, but for chemistry they were still open to a solution. Also looked for a solution in solid tumor, a new panel. Our STS is perfect for them. Everyone in the meeting was involved, the administration and the doctors. Lets see now the results in the future.

Drexel University (March)

Small lab. Really need a solution as the oncoportal. Interested about the transaction fee, no big engagement for small volumes they have. Liked the community aspect, they use small panel so they don't really need a bioinformatics support, because no difficulty to classify variant for the TST 15 and 26. However, they saw the Oncoportal which exactly fit their needs, it's gain of time for them they say, because no need so search for publication, which treatment is the best.... Oncoportal to be more efficient in time gaining.

