Wageningen University - Department of Social Sciences

MSc Thesis Chair Group - Knowledge Technology and Innovation (KTI)

Scaling of user driven technological change

Case-based analysis of the scaling trajectories of M-PESA and iDART

April 2014

MSc program
Master International Development Studies (MID)

Specialisation
Communication, Technology and Policy (CTP)

Thesis code
CPT 80836

Name of student
Cees Peter van den Brink

Name of supervisor
dr. ir. SR Vellema

Word count: 18.870

Abstract

Keywords: user-driven technological change, inclusive innovation, developing markets, user-producer interaction, scaling, iDART, M-PESA

Processes of technological change in the context of developing markets are approached and researched, within both the academic- as well as the corporate environment, using different normative conceptualisations and strategies of how inclusive innovation or technological change comes about. Two essential directions can be discerned from this work. A first direction concerns a more top-down, strategic and firm-led exercise of developing and implementing products or services that impact the marginalised, whereas a second direction highlights the importance of technological change emerging from a more localised context. Few literature analyses how the interactions between these directions gives insight on the issues and processes that affect scaling of 'co-created technologies'. By means of two case studies, the scaling trajectories of two user-driven technologies (iDART and M-PESA) have been reconstructed in order to gain insight on how their scaling emerged and evolved over time. Scaling is found to be a gradual and emerging outcome that involves at least three processes, namely: processes of: 'technology adaptation and stabilisation', 'systemic feedback on organisational behaviour' and 'organisational closure'. Although all three processes affect scaling of user-driven technological change, it is concluded that only the first process comes forth from the interaction of user- and producer networks. Processes of systemic feedback come forth from scaling or organisational growth itself, whereas organisational closure originates from organisations actively executing a closure process. The latter process has been identified as a meta-theory of scaling because it organisational closure intertwines with-, gives interpretation to-, and directs processes of (i) technology stabilisation and (ii) systemic feedback. Future research can quantitatively test these six hypothesised causalities with scaling and can research if there are more theoretical processes involved in scaling than the three processes identified in this thesis. This thesis provides firms six 'control panels' that they can use for prospective analysis of scaling.

Table of Contents

Abstract iii

1. Introduction 1

1.1 Research background and research problem 1

1.2 Research objectives and research questions 2

1.3 Research design and research methods 4

1.4 Research assumptions 5

2. Literature Review 6

2.1 Conceptualisations of top-down and bottom-up directions of Inclusive Innovation & Scaling 6

2.1.1 Bottom-up approaches of Inclusive Innovation 6

2.1.2 Top-down approaches of Inclusive Innovation 7

2.1.3 Defining scaling 8

2.2 Insights on scaling of user-driven technological change 9

2.2.1 Insights from Innovation Studies 10

2.2.2 Insights from Systems Thinking 11

2.2.3 Insights from Organisational Studies 14

3. Different 'frames of theories' 17

4. M-PESA: a literature case study on scaling of user-driven technological change in the financial sector 19

4.1 Historic reconstruction of the development- and scaling trajectory of M-PESA in Kenya 19

4.1.1 Introducing M-PESA 19

4.1.2 Organisational- and technological developments over time 20

4.1.3 M-PESA's scaling over time in figures 24

5. iDART: A field case study on scaling of user-driven technological change in the healthcare sector 25

5.1 Historic reconstruction of the development- and scaling trajectory of iDART 25

5.1.1 Introducing iDART 25

5.1.2 Technological developments over time 26

5.1.3 Organisational developments over time 28

5.1.4 The scaling of iDART over time in figures 31

6. Cross-case analysis 32

6.1 Processes of technology adaptation 32

6.2 Processes of systemic feedback 34

6.3 Processes of organisational closure 35

7. Conclusion and Discussion 38

References.....41
Appendix A - Overview of interviewees IDART Case-Study 46
Appendix B - Overview of the constructive events of the scaling trajectory of iDART 47
Appendix C - (Incomplete) Overview of iDART implementation sites 48

1. Introduction

1.1 Research background and research problem

There are numerous barriers that prevent the poor, as consumers or producers, to actively participate in markets. For those who can access and successfully participate in them, markets can be an engine for individual human development and economic empowerment. Market failures, government failures, characteristic of poor business actors or consumer, the environment in which these actors operate and more other factors can however prevent this. Producers in low income markets are often excluded from labour- and product markets due to reasons such as: a lack of access to credit, geographical isolation or limited investment opportunities in human capital. One of the outstanding barriers for consumers in low-income settings is that many non-poor businesses are unfamiliar with the low-income environment and the possible opportunities these developing markets offer. Although these barriers for producers and consumers will not be discussed in detail now, one may distinguish two main avenues of making markets more inclusive for the poor. The first main avenue focusses on actions that remove these kind of market barriers by changing the environment of the poor; for example through institution-building or by extending legal empowerment to low-income populations. These types of investment however require significant amounts of time before they bear fruit. A possible alternative, a second main avenue or starting point, is to work within the present environment and to stimulate the development of market-based innovations that are designed to help the poor overcome some of the barriers hindering them from more market access or market participation (Mendoza & Thelen, 2008).

In the last couple of decades, many work has been carried out by scholars within this second main avenue. Processes of technological change in the context of developing markets have been researched with use of different normative conceptualisations and strategies that discuss how innovation can be inclusive to the poor. Recently, these processes are discussed under the label: 'Inclusive Innovation' (by amongst others: Foster, 2013a; Foster, 2013b; Suurs et al., 2013, Paunov et al., 2013). Central in this umbrella concept is the idea that innovation or technological change can incorporate the needs and interests of the poor, can provide the marginalised with benefits and is led by the private sector (Foster, 2013). Something less discussed and focused upon in innovation literature, is how Inclusive Innovations can be brought to scale. In order to impact the marginalised it is highly relevant to not only focus on how technologies can 'include' the needs and demands of the poor and how technology can 'address' their problems, but to focus as well on how to bring such technologies to a scale that matches the size of their problems and demands.

From the innovation literature that discusses 'innovation and inclusivity', two essential directions can be discerned. Within the management literature, thinking on Inclusive Innovation mainly revolves around the idea of 'Base of the Pyramid' (BoP) markets. Within this field of study, research looks at inclusive strategies of large firms to serve the demand of low income populations. Low income groups are seen as untapped markets that Multinational Corporations (MNCs) can profitably tap into. This literature field positions innovation mainly as a top-down, strategic and firm-led exercise of developing and implementing products or services that impact consumers or producers located at the BoP. These innovative products or services, that may be new technologies or adapted forms of existing technologies, aim to fit with the unique cultural, financial and social needs of poorer consumers, producers and entrepreneurs. A second direction of Inclusive Innovation highlights the importance of technological change emerging from a more localised context. These understandings of innovation in developing markets focus on 'innovative and localised processes of technology adaptation'.

Scientific work within this direction makes use of many different concepts and labels, such as: frugal-, indigenous-, local-, grassroots- or informal innovation. Although these concepts have their own specific meaning, they all shed light on how Inclusive Innovations emerges from a localised context. A local context where innovation emerges under the unique conditions, practices and constraints of a low income setting and where low income communities use their knowledge to solve problems and share solutions that are relevant to their local needs (Foster, 2013).

This apparent distinction of 'directions of innovation' may not be something totally new within literature. It is rooted in earlier work on 'push' and 'pull' factors of technological change. This distinction relates to the degree in which innovative activities are autonomous over short-run changes in the business environment and the extend in which innovation activities and technological changes are directed by market signals. From the demand-pull viewpoint (related to the bottom-up 'direction' of Inclusive Innovation), it is argued that the main determinants of technological change are rooted in market forces. It subscribes the importance of recognising technological needs. The technology-push viewpoint (related to the top-down 'direction' of Inclusive Innovation) sees technological change more as an institutional process of (scientific) discovery, technology development and technology implementation. It gives technology a more autonomous character (Danse & Vellema, 2007).

The two directions of Inclusive Innovation are often played off against each other within literature. Research that focusses on 'how to address the needs and demands of the marginalised' tends to either focus on top-down innovation or localised innovation flows. Few literature analyses how the interactions between these two directions can give insight on the issues and processes that affect scaling of inclusive and 'co-created' technologies. Research focussing on localised processes of Inclusive Innovation may have the weakness that it doesn't comprehend the external pressures or flows of innovation that are of influence. Such research runs the risk of missing the bigger picture on the opportunities and limitations of scaling Inclusive Innovations. Top-down innovations led by larger firms may have the advantage of having large distribution networks that are needed for scaling, but may have the weakness of designing and developing technologies that mismatch the local contexts in which these technologies are used (Foster, 2013). It seems likely that localised innovation processes can provide perspectives on innovation that larger firms can adopt and scale more widely to promote inclusivity. However, there has been carried out few analysis on how firms go about understanding and linking to rich localised processes of problem solving or user innovation, and on what issues and processes are present in the process of scaling co-created technologies. This is the main subject of interest in this thesis.

The discussed directions of innovation are not only subject of debate within the academic environment. Many organisations have come up the last decade developing their 'own' normative inclusive intervention method or strategy. Deeper understandings about the impact of user-producer interactions on scaling of technology in developing markets are relevant for these organisations too. Also from an organisational or business perspective, this thesis addresses a knowledge issue that needs further study.

1.2 Research objectives and research questions

This thesis will analyse what processes take place when top-down and bottom-up directions of Inclusive Innovation interact and will research how these processes affect scaling. To get a better understanding on how localised innovation processes link and interact with firms, this thesis will analyse cases where this actually has happened. Two scaling trajectories will be analysed whereby user-driven technological change has been picked up and included in the 'vertical logic' of a technology producing firm. The case studies will research how user involvement in technology development is supported and nurtured by market players during processes of

technology development and scaling.

So the main aim of this thesis is to determine how certain 'distinctive processes and conditions' (symbolised by the coloured balls A, C, D, P in figure 1), present in the interactions between 'horizontal' user networks and their 'vertical' relations (merchants, retailers, wholesalers, and producers), impact scaling trajectories. By theorising upon 'distinctive processes and conditions' this thesis hopes to provide more understanding on what processes influence and impact scaling of user-driven technological change in developing markets. The following research question are set up to work on this.

Main research question:	What distinctive processes and conditions, present in the interaction between user networks and vertical production chain actors, impact the scaling of technologies in developing markets?
Sub research questions:	i) How are these interactions conceptualised and theorised in literature and what does this imply for the process of scaling? ii) How have these interactions impacted the scaling of 'M-PESA' and what insights does this case study give on the phenomenon of scaling that are reproducible for other cases? iii) How have these interactions impacted the scaling of 'iDART' and what insights does this case study give on the phenomenon of scaling that are reproducible for other cases?

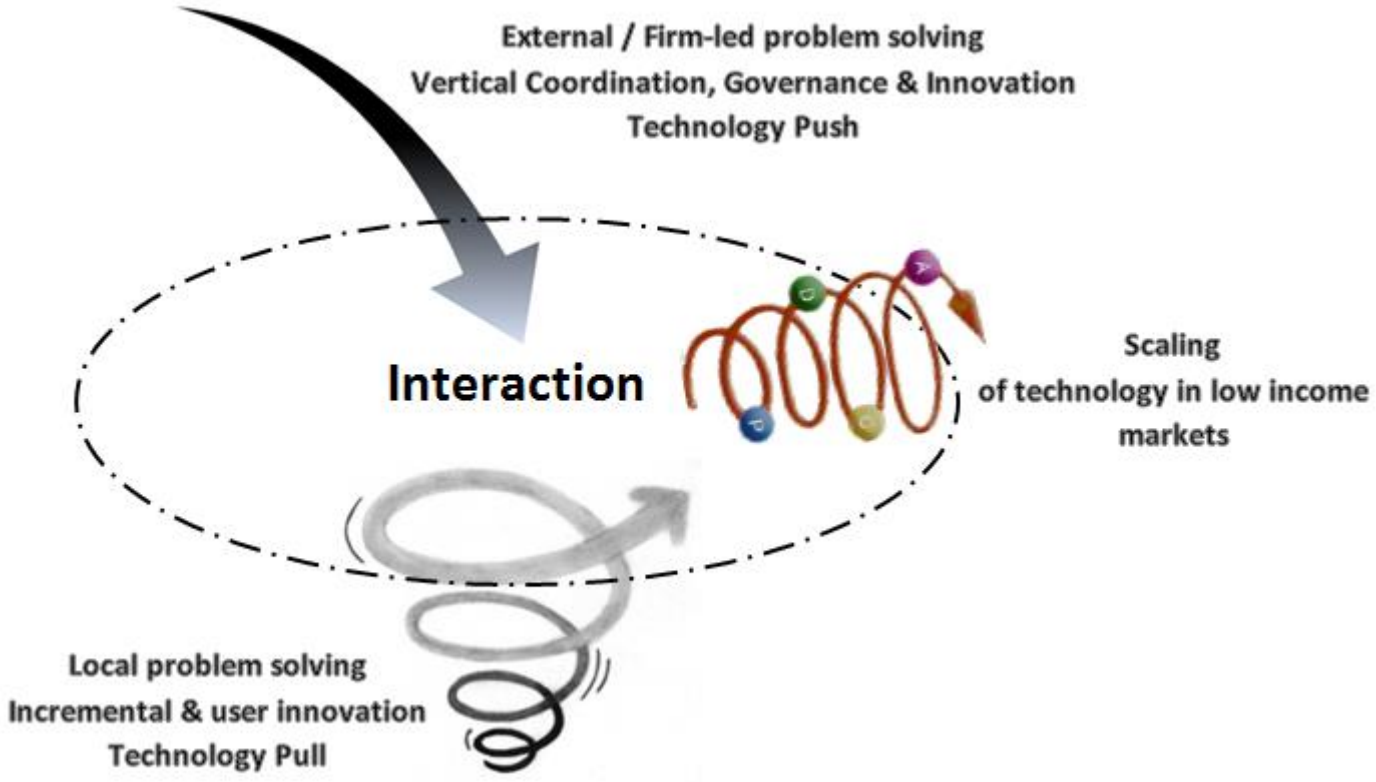


Figure 1 – Figurative representation of the research approach

1.3 Research design and research methods

The proposed research questions are addressed using multiple research methods during two stages of explorative research. Where chapter two starts off with a literature review, chapter three presents a 'frame of theories'; an overview and summary of the different understandings and perspectives gained so far on the phenomenon of scaling. In the second stage of this thesis, two empirical case studies are conducted. A first literature case study (chapter 4) looks at the scaling of M-PESA whereas a second field case study (chapter 5) looks at the scaling of iDART. These case studies are performed to test the usefulness of the different 'frames of theory' for analysing and interpreting technological scaling processes within two different sectors: the financial sector (M-PESA) and the health care sector (iDART). These case studies will be further discussed in chapter 6, where the main results of this thesis are presented. A final seventh chapter will discuss the main theoretical implications of this thesis and its implications for organisations.

First of all, chapter 2 investigates what different perspectives there are within literature on innovation and inclusivity. The review in chapter 2.1 will clarify what defines 'bottom-up' and 'top-down' directions of Inclusive Innovation and discuss what this implies for the interaction between them. Secondly, chapter 2.2 looks deeper into these interactions and theorises on how they impact the scaling of user-driven technologies. Academic literature on scaling is in an embryonic stage (Seelos & Mair, 2010). So there is not one particular delimited literature body useful for this. Therefore, several bodies of literature discussing the relation between innovation, inclusiveness and scaling will be reviewed. From Innovation Studies, theories or notions like: 'user innovation', 'innofusion' and 'user-producer interaction' are discussed. From a system thinking' perspective, innovation's systemic and evolutionary character is looked at. From organisational literature, the review discusses what theories and concepts like: 'resource-based view of the firm', 'organisational routines and practices' and 'organisational closure' teach us about the relation between innovation, inclusiveness and scaling. By reading a view core articles from innovation literature, more literature has been gathered using the snow-ball method.

A third chapter provides a summary of the different presented 'theoretical frames' that can be used to analyse scaling processes of user-driven technology. It translates the first findings into three main 'theoretical lenses' that will be used to analyse the following case studies.

The first case study is performed on the development- and scaling trajectory of 'M-PESA'. This literature case study contains of a thorough analysis of the development and scaling trajectory that this mobile banking application has gone through during the last decade. The analysis makes use of primary and secondary sources. Next to scientific papers, other sources are used like: company reports, policy documents and documents from NGO's and public-private organisations.

A next fifth chapter presents a field case study on iDART (an Intelligent Dispensing software for Antiretroviral Treatment). Cell-Life developed this stock management- and dispensing tool of Antiretroviral (ARV) drugs to support ARV clinics in treating HIV positive patients. Data for this case study has been gathered using two methods: document analyses and semi-structured interviews. For the period of 2000-2013, a reconstruction is made of the constructive events and processes that have been part of iDART's scaling trajectory. Interviews have been held with 12 people; 5 pharmacists, 4 employees of Cell-Life and 2 employees of the Desmond Tutu HIV Foundation (Cell-Life's first partner in implementing iDART) (see appendix A for an overview of the interviewees). Over 500 minutes of interview have been transcribed, and used as input for this case study. Next to qualitative data, also some quantitative data on the scaling of iDART has been gathered in

order to give some insight into the extent iDART has scaled during the selected time frame.

Using the three theoretical lenses (of chapter 3), the sixth chapter analyses the two scaling trajectories of M-PESA and iDART and distils from them the main distinctive processes that have affected their scaling. By means of a cross-case analysis, this chapter hypothesises on the causalities these distinctive processes have with scaling.

A final seventh chapters concludes and discusses how useful the different 'frames of theory' have been for analysing processes of scaling. It further elaborates on the implications of this thesis for a 'theory of scaling' and for management of organisations or firms.

1.4 Research assumptions

The main assumption underlying the approach of this thesis is that scaling of inclusive, user-driven technologies comes forth from the interaction between top-down and bottom-up processes; from the interaction between larger technology-producing firms and user-networks (figure 1). Due to this assumption, this thesis focusses on how processes within this interaction impact scaling trajectories. This assumption does not come out of the blue. In the last two centuries we have seen that production and consumption have increasingly grown apart because of mass-productions methods and low-cost transportation systems. Production and use of technology have increasingly been differentiated into separate clusters. Networks between production and consumption have lengthened, what led to an increase of the number of social groups involved in distribution (merchants, retailers, wholesalers and producers). Due to these lengthened networks, interaction between supply- and demand actors have become more complex, which causes misalignments between supply and demand, between technology producers and –users. Scholars have acknowledged this problem and have shifted their focus from a merely production-side perspective to a perspective that focusses as well on the user environment of technology. This resulted amongst others in theories of 'co-evolution'. These theories will be discussed more extensively later, but basically explain that implementation of new technology must be seen as a dynamic process with 'mutual adaptations and feedbacks between technology and user environment' (figure 2) (Geels, 2004).

Without alignment between technology and the user environment, productivity of technology is not optimal and less beneficial for both producers as consumers. The more technology and the user environment are aligned with each other, the more likely it will be that (i) the needs of producers and consumers are met, (ii) the demand for such technology will rise and (iii) the supply (or capacity) of such technology can be scaled (Leonard-Barton, 1988). This understanding supports the approach of this thesis to take the interactions between producer- and users networks as the main 'area' of analysis for gaining insight into scaling processes.

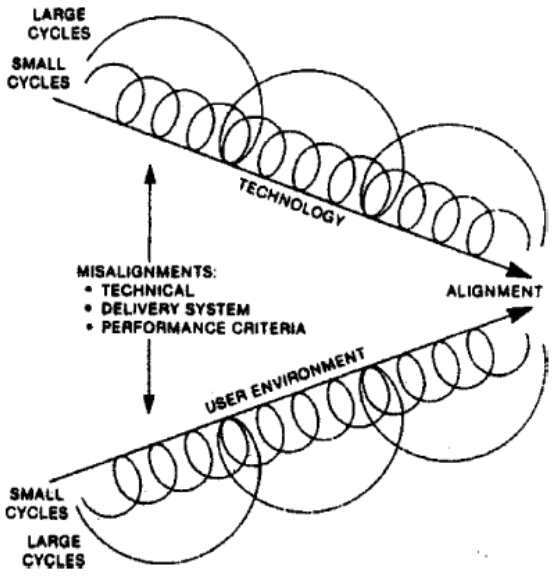


Figure 2 – Co-evolution of technology an user environment (Leonard-Barton, 1988:251)

2. Literature Review

2.1 Conceptualisations of top-down and bottom-up directions of Inclusive Innovation & Scaling

Schumpeter describes innovation as combinations of new or existing knowledge, resources, material equipment or other 'combinatory factors' (Schumpeter, 1934). Finding such successful combinations for developing markets is an activity that involves different types of actors: scientists, firms, regular users of technology and others. The plurality of actors involved in technological change in developing markets might explain why many different normative conceptualisations have emerged that discuss how technological change can benefit the marginalised. These conceptualisations are differently positioned in the 'bottom-up - top-down continuum' of directions of innovation. The following review (chapter 2.1.1 and 2.1.2) will clarify what defines 'bottom-up' and 'top-down' innovation processes. Two examples of organisations working in a top-down, or bottom-up way will be discussed. After more in depth understanding of the two directions of innovation have been presented, a first indication follows on what this implies for the interaction between them. After reviewing the different directions of Inclusive Innovation, chapter 2.1.3 will define the concept of scaling.

2.1.1 Bottom-up approaches of Inclusive Innovation

In the last three centuries, much of technological progress was directed to meet the needs of the rich. Most technologies therefore best-suited to operate in high-income environments. As a consequence, attempts to transfer new technology to low-income contexts was often inappropriate due to the fact that these technologies were highly capital intensive and operated only at large-scale. 'Intermediate technologies' were called for. Technologies that could operate at smaller scales, that were less capital intensive, more labour-intensive and simple to operate and repair (Kaplisnky, 2011). A not-for-profit "Appropriate Technology" (AT) movement arose in the 1970's. The movement emphasised the need to develop low wage and poor-infrastructure appropriate innovations for 'the South'. Many programmes emerged that focused on the implementation of such appropriate technical solutions and whether or not they had potential to be transferred to other developing economies. Scaling of these innovations focused merely on technical specificities (Suurs et al., 2013). Scaling of ATs was basically supported by funding and must be seen more as an 'act of charity' than a result of the pursuit of profit (Kaplisnky, 2011). More or less as a result of this top-down AT movement, a bottom-up movement with many normative conceptualisations have come up the last two decades that focus on localised, bottom-up processes of innovation in developing markets. These new bottom-up conceptualisations have been called different names (local-, grassroots-, frugal-, informal innovation and more) and all have their own particular focus.

The concept of 'Frugal Innovation' for example describes how producers create new, low-cost products or technologies with few resources. It often involves reducing the complexity of goods and their production process by removing nonessential features (Van der Boor et al., 2013). Another scarcity induced conceptualisation of innovation is 'Grassroots Innovation'. Poor technology users can creatively assemble available materials into functional technology (Cozzens and Sutz, 2012). Acting as innovative entrepreneurs, they can use existing or new technologies, ideas, concepts, both 'indigenous' and 'imported', to develop products and/or services that are uniquely differentiated to the needs of their local environment (Kaplinsky, 2011). What typically defines grassroots innovation is that it involves innovation 'by' and 'for' marginalised groups. It often takes place in informal settings, and therefore is also sometimes called 'Informal Innovation'(Cozzens and Sutz, 2012).

Many networks and organisations have come up the last two decades that operationalise these bottom-up conceptualisations of Inclusive Innovation. Some examples are: Ashoka, Practical Action (Peru), The National Grassroots Innovation Databank (Malaysia), the Social Technologies Network (Brazil), the Grassroots Innovation Augmentation Network (GIAN) and the Honey Bee Network (HBN) (India). This last organisation emerged in 1989 among a group of scientists, farmers and others' who were interested in documenting and disseminating traditional knowledge and local innovations in local languages. This has become the organisations main activity: to scout and document innovations and traditional knowledge. They do this through community visits, interviews and by organising competitions for 'undiscovered ingenuities'. In the last two decades, the HBN has reported over 100.000 innovative ideas, technologies and traditional knowledge practices. Many of these can be found at their online database (www.sristi.org). Besides documentation, the network also explores the commercial potential of some of the products and processes they have identified during scouting. They support local grassroots innovators in the process of patenting and offer them assistance in terms of prototyping, incubation and finding the first rounds of funding. On top of that they try to mobilise different public- as well as private actors at both grassroots level and levels higher-up. In this way they try to build up a broader enabling environment for grassroots innovation (Sone, 2012; www.sristi.org).

Paunov (2013) studied the socio-economic value of HBN grassroots innovations. It was found that grassroots innovations can make substantial contributions to-, and are valuable for specific low-income groups. It was however found that few HBN grassroots innovations have been patented or have become an export success. It was also observed that HBN innovations have hardly been adopted across other lower-income groups (Paunov, 2013). These findings indicate that grassroots innovations do not easily link up with higher-level actors and do not easily scale.

What becomes clear is that bottom-up conceptualisations and methods of Inclusive Innovation underscore the idea that the marginalised, as users of technology, are capable 'innovation agents'. From a bottom-up perspective, Inclusive Innovation therefore stresses 'the active inclusion and involvement of the marginalised in innovation processes' (Foster and Heeks, 2013b). In order to address the demands and interests of the poor (inclusivity in output), the importance of 'inclusivity of process' is stressed. This has implications for processes of technological change. First of all, it captures the idea that interactive social learning processes between producers and marginalised users play an important role in (re)engineering and (re)shaping innovations or technologies (Foster and Heeks, 2013b, Nahuis et al., 2012). Secondly, 'inclusivity of process' does not only appeal to participation of the marginalised in processes of technology development, but also stresses their involvement in the business processes and strategies established around it (Martinez-Gutierrez and Paunov, 2013). So when bottom-up approaches of Inclusive Innovation interact with top-down approaches, interactions may occur around processes of technology development as well as around strategic organisational processes.

2.1.2 Top-down approaches of Inclusive Innovation

An example of a more top-down directed movement of Inclusive Innovation is known under the term: 'Base of the Pyramid' (BoP) markets. Organisations that work with the 'BoP vision' develop, learn about, and accelerate Inclusive Innovations that serve the demand of low-income groups. By redesigning business processes and by making products more accessible to low-income consumers, larger companies can profitably tap into developing markets. This is the argument of Prahalad and Hart (2004) that was introduced in their book: 'The fortune at the bottom of the pyramid'. The movement primarily focusses on enabling innovation 'for'

marginalised groups and to a lesser extend involves innovation 'by' marginalised groups. Under stewardship and leadership of Stuart Hart, this movement started to get more attention and the BoP vision got more intensively researched. Over time, several 'BoP Learning Labs' were initiated. Such learning labs are now stationed in 18 countries around the world. They facilitate 'incubator spaces' for the development of new innovations and focus on knowledge creation and dissemination on the theory and practice of creating sustainable businesses at the BoP. One of the things that the BoP movement learned over time is that many BoP ventures fail because of reasons like: 'product misfire, low sales penetration, high distribution costs and inability to scale' (Klein et al., 2012a). Over time it was learned that co-creation is critical to BoP business success. Co-creation requires deep dialogue; two-way-communication with the target group that must be seen as partners and colleagues rather than 'passive' clients or consumers (Klein et al., 2012a). Not only is co-creation found to be important for processes of product design, organisations also need to *"develop a new 'native capability' that focuses on co-creating business concepts and business models with the poor"* (Klein et al., 2012a: 3).

One of the organisations working with the BoP vision is the Base of the Pyramid Innovation Center (BoPInc), located in the Netherlands. Publications from BoPInc (Klein et al., 2012a; 2012b; 2013; Chevrollier, 2013) stress that deep interaction between producers and users are important for Inclusive Innovation, from a technology development-, as well as a business development point of view. The publications touch upon a range of different issues (e.g. co-creation, innovation cycles, partnership development, inclusive marketing and more) but in-depth content on (i) user producer interactions and (ii) scaling of technology is not provided or still very superficial.

This first review of innovation literature that discusses different directions of Inclusive Innovation clarifies two things. Both directions stress the importance of more interaction between bottom- and top-levelled actors. Secondly it clarifies that neither this interactions, nor the process of scaling, are well conceptualised. The review does give direction on what interactive processes to further look at, namely: (i) processes of interactive technology development and (ii) interactions regarding development of sustainable business.

2.1.3 Defining scaling

Scaling is a concept that is often applied to different types of processes. It is associated with efforts related to intensification and/or multiplication of existing technologies or 'good' practices. It however also includes efforts to align technical opportunities with policy, governance, or organisational processes; efforts with a more 'catalytic' character (Berg et al., 2012). How the term scaling is used depends very much on the specific innovation practice and the context in which this practice takes place. Berg et al. (2012) therefore conceptualise scaling as *"an outcome of the interaction between how the innovation is implemented and the context in which it is embedded"* (Berg et al., 2012:3).

There are different ways in which scaling can occur (figure 6). 'Scaling up', sometimes called 'vertical scaling', refers to initiatives that attain larger volumes or higher capacity of products and services while continuing to use the same processes and structures. Hereby, these initiatives can reach larger numbers of customers, or can better satisfy their needs through a higher capacity of their product or service. A second type of scaling is called 'scaling out' or 'horizontal scaling'. This takes place when ideas and business models are used and copied to other initiatives. It's about the replication of interventions or practices; doing the same thing in a different context (Berg et al., 2012; Woodhill et al., 2012). Woodhill defines a third route of scaling: 'scaling adaptively'. This more or less is a combination of the first two types of scaling. It involves innovation

and adaptation. A form of scaling which occurs through the rapid development of many new ideas and initiatives based on inspiration from initial successes (Woodhill et al., 2012).

These definitions make clear that scaling is not only about 'gettin g bigger'. Scaling is all about the challenge in which what works successfully at small scale in one context, will not necessarily work at a larger scale in a different situation and context. This puts the focus of scaling on 'bringing about change at a sufficiently large scale'; a scale that equals the scale of addressed problems. Positioning scaling in this way makes it a highly complex phenomenon. Better understandings are needed on how to support, incentivise and strategically guide good initiatives towards growth. One way of getting closer to such an understanding is by monitoring how scaling emerges and how it evolves (Berg et al., 2012; Woodhill et al., 2012). This approach is used for the following two case studies.

Figure 3 - Different ways of scaling (Woodhill et al., 2012:7)

2.2 Insights on scaling of user-driven technological change

How do the interactions between user networks and production chain actors impact scaling of technology? In order to gain more insight into this 'black box' of interactions, three research angles are used for this review in order to theorise upon the phenomenon of scaling (figure 4). The former chapter identified two of these. Interactions concerning organisational processes of sustainable business development will be theorised by reviewing organisational / management literature. Interactions around processes of technology development will be theorised by reviewing innovation literature. Reviewing Innovation literature led also to a third perspective that was found to be of importance, namely: 'Systems thinking'.

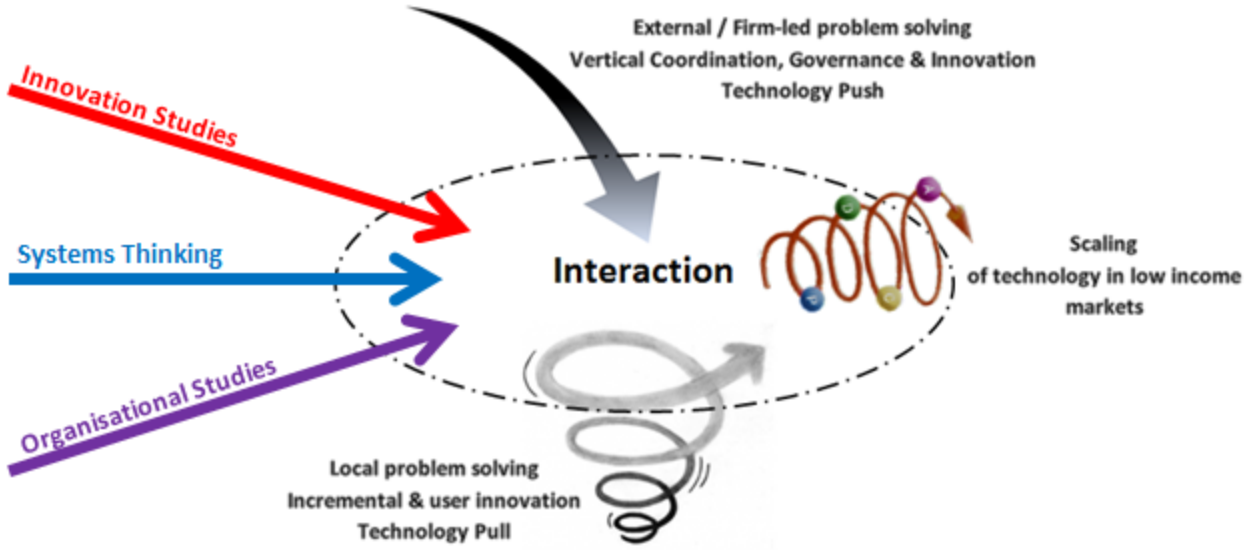


Figure 4 - Different research angles used for conceptualisation of user-producer interactions that impact scaling

2.2.1 Insights from Innovation Studies

According to conventional linear models of innovation, processes of invention, production, and diffusion (distribution and implementation) are separate processes. Scholars have agreed that these models bear little relation to the realities of innovation. Production and diffusion are not separable processes, they are inextricably intertwined. Invention, production, distribution and implementation involve continuous innovation and feedback. Some scholars have therefore named this process 'innofusion', an intertwined process of innovation and diffusion (Foster & Heeks, 2013b). When users in such processes of innofusion deliberately adjust and adapt technology, this can be called user innovation (Nahuis et la., 2012). Users can exploit the flexibility in technology designs, and customize technologies to meet their specific needs.

Users not only appropriate or domesticate technology for themselves, they might also exchange their user-innovation related information with producers through UPIs. Especially in developing markets, where the portfolio of commercially offered products and services is smaller, it is more likely to find a wide range of unmet needs as well as customers and users who create their own solutions to meet these unmet needs (van der Boor et al., 2013). Little literature is however available on the diffusion of user innovations in developing markets. In an attempt to discover what happens with user innovations in 'the north', it was found that users often lack incentives to diffuse their innovations. When they are shared, this often happens for free and through non-market mechanism, such as peer-to-peer diffusion or through communities (von Hippel, 2012 according to van der Boor et al., 2013). Producers seem designated to further develop and commercialise user innovations in the broader market. They however often stumble upon the fact that information about unmet needs are 'sticky' to the user. It is costly to acquire and transfer such sticky information (Von Hippel, 1994). Debates on how firms can capture the role of users as sources of innovation are, amongst others, discussed under the notion of User-Producer Interaction (UPI) (Nahuis et la., 2012, Lundvall, 1988).

The work of Nahuis et al. (2012) gives deeper understanding on interactive processes between users and producers. They define UPIs as *"Interactive learning processes between users and/or producers leading to or aiming at the reduction of uncertainty about the relation between product and demand characteristics"* (Nahuis et al., 2012: 1122). In their report they distinguish users as being either 'co-producers' or 'users as implicated actors'. Where the former are active agents in technology development, the latter are just affected by technology and rather invisible during its development. Whether or not users are included as active co-producer depends much on two 'modes' of UPI. Interactive learning might resolve around 'feedforward-' or 'feedback loops' (Nahuis et al., 2012).

Through anticipatory activities like*: '*broadening', 'user characterisation', 'upstream involvement' and *'*first user enrolment', firms can involve users in technology development before their product or service is released. In early phases of development, firms might broaden their perspective to identify and characterise the range of potential users and other affected actors. After 'zooming out', when all potential users are within the firm's scope, 'user characterisation' can give the firm insight and specification for whom, and how exactly, to design the technology. In this design process, firms might choose to be open for 'upstream involvement'. This is as much to say that firms let users participate actively in the process of research, design and development of the technology. In this way users can articulate their demand and articulate when a technology is acceptable. Often when coming to the end of the design process a pilot project is carried out among lead users to test its usability in a real-world setting and to make the product or service ready for its release. These feedforward learning activities influence how a sufficiently malleable technology gets steered and shaped

towards a specific use and user-context (Nahuis et al., 2012).

The second 'mode' of UPI resolves around feedback loops. Two distinct forms of feedback are 'Learning by using' and 'Configuring the user' (Nahuis et al., 2012). 'Learning by using' is concerned with technical characteristics. Not all performance characteristics of a technology can be understood on forehand, and often actual experience and use of technology leads to redesigned technologies. 'Configuring the user' is a form of feedback that is not concerned with technological characteristics, but instead with the characteristics and capacities of users. Configuring the user is defined as *"a process of defining the identity of users, and setting constraints upon their likely future actions"* (Woolgar, 1991 according to Nahuis et al., 2012: 1124). Innovating agents observe user's confusion, wrong use or alternative response to innovative technology, and find ways to teach and (re)direct them on how to use the technology (Nahuis et al., 2012). In more plain words: producers constrain users towards a certain usage of technology. These two forms of feedback give insight on how misalignments between technology-producing firms and users can be corrected (figure 2). Either by altering the technology (learning by using), changing the user environment (configuring the user), or both (Leonard-Barton, 1988).

2.2.2 Insights from Systems Thinking

Three contextual factors may have influence on UPIs: types of technology, phases of technology development and the heterogeneity of the user population (Nahuis et al., 2012).

Every design of a technology consists of a 'set of affordances and limitations' which prescribes or suggest a particular use of the product or service. However, the degree to which this set of 'affordances and limitations' is specified, the extent to which a technology is 'flexible', differs for different types of technology. Fleck (1989) distinguishes three types: discrete technologies, system technologies and configurational technologies. These types of technology differ in their flexibility. Discrete technologies are products or services that can be seen as almost completely 'autonomous'. Its design specifications are highly standardised and its user prescription is clear and stable. Users can 'take it or leave it'; they don't have much opportunity to participate in the development and further redesign of such technology. System technologies are more complex. They consist of several technological components that condition and constrain each other. The more interrelated and interdepend these components become, the more stabilisation and standardisation takes place. The technology becomes less flexible and leaves less space for user-customisation or -involvement. Finally, configurational technologies are products or services that consist of both technical and social components. These technologies lack tight standards for it use or performance during the time before their implementation. The affordances and limitations for such technology become however specified during their implementation processes, when users have opportunity to configure and customize them according to their preferences (Fleck, 1989). Not all types of technologies have the same potential for user-involvement. In term of analysing scaling trajectories, users have most influence on configurational technologies that are still flexible, and not yet standardised.

The idea that there are different types of technology that differ in their extend of flexibility is rooted in evolutionary thinking of technological change. An 'evolutionary systems approach' conceptualises innovation in a way that is similar to the evolution process found in biology; a process of variation, selection and retention (Smits, 2002). Different technology-producing firms altogether introduce a wide variety of innovative products and services. These compete and enter a 'selection environment', which Geels (2004) calls 'the ruling social-technical regime'. This ruling regime is formed through continuous processes of co-evolution between several continuously changing sub-systems (science-, policy-, socio-cultural-, technology-, and user & market system).

From the variety of new innovative products or services entering this regime, some 'perish' and some 'survive'. The ones that survive become part of the existent ruling socio-technical regime (figure 5) (Geels, 2004). This short and simplistic explanation of the evolutionary character of innovation is sufficient to explain that technologies can go through different phases of development. In early phases of development, a new technology can bring various opportunities. The direction in which it will do this successfully is however hard to predict. New 'configurational technology' is still very much flexible and malleable. Not only because it still can be customised by users, also because it will be shaped under the influence of other regimes that are part of the Social-Technical (ST) system (Geels, 2004). It is in this way that during the process of 'being shaped by the selection environment', standardisation of the technology can set in and a configurational technology can become path-dependent. During such a process, configurational technologies might evolve into 'system technologies'. Even further in the process of stabilisation and standardisation, system technologies might evolve into 'discrete technologies' (Nahuis et al., 2012). Such technological trajectories, wherein flexible and malleable technology stabilise and standardise over time, are a common practice. Scholars have used notions like 'path-dependence' or 'lock-in' to describe how technology stabilises when it is taken up-, or becoming part of existing systems (Geels, 2004). When researching the scaling trajectory of user-driven technologies, it seems important to understand how processes stabilisation and standardisation play a role in this.

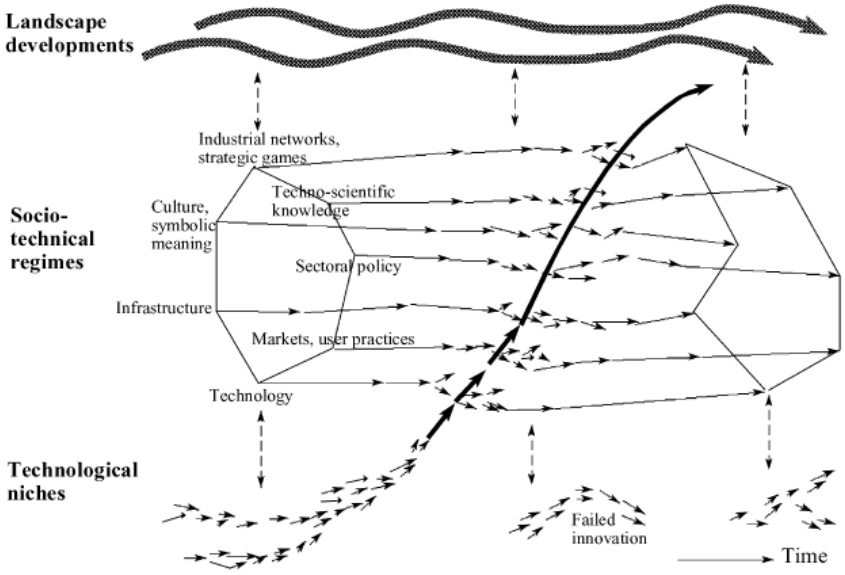


Figure 5 - A dynamic, multi-level perspective of system innovation (Geels, 2004: 915)

A third contextual factor that influences UPIs is the notion of 'heterogeneity of the user population'. It gives extra understanding on how path dependency emerges. The more heterogeneous users are, the more complicated it is for a technology-producing firm to align with all the different user demands. Complete alignment with all demand seems impossible, so what occurs is that alignment is realised by 'closure of controversy'; standardisation or convergence of demand. By exception, alignment of demand of heterogeneous users is not pursued through practices of closure, but by provision of differentiated products. Application of the latter option is not common for discrete- or system technologies, and primarily something that can be applied to configurational technology (Nahuis et al., 2012). In terms of developing and scaling technologies, these processes of 'closure' and 'differentiated product provision' might be influential.

In search of a practical model for scaling of family health innovations in low-income settings, Bradley et al. (2012) performed research on its disabling and enabling factors. They identified five important interrelated components for scaling processes (figure 6) (Bradley et al., 2012). By applying work from Paina and Peters (2012) about Complex Adaptive Systems (CASs) (see table 1) unto the AIDED model (figure 6), some more systemic understandings underlying processes of scaling become clear.

Firms understanding of their consumers' specific context and demand is often very poor. This is a major barrier for scaling of innovative products or services (Bradley et al., 2012). It is caused by continuous interactions within systems that lead to continuous creation of spontaneous orders. New emergent behaviour arises of the actors facing these changing orders (Paina and Peters, 2012). Because emergent behaviour is hard to predict, firms that scale a particular product or service need to intensively and continuously engage and assess their environment during their technology development and scaling efforts (Bradeley et al., 2012).

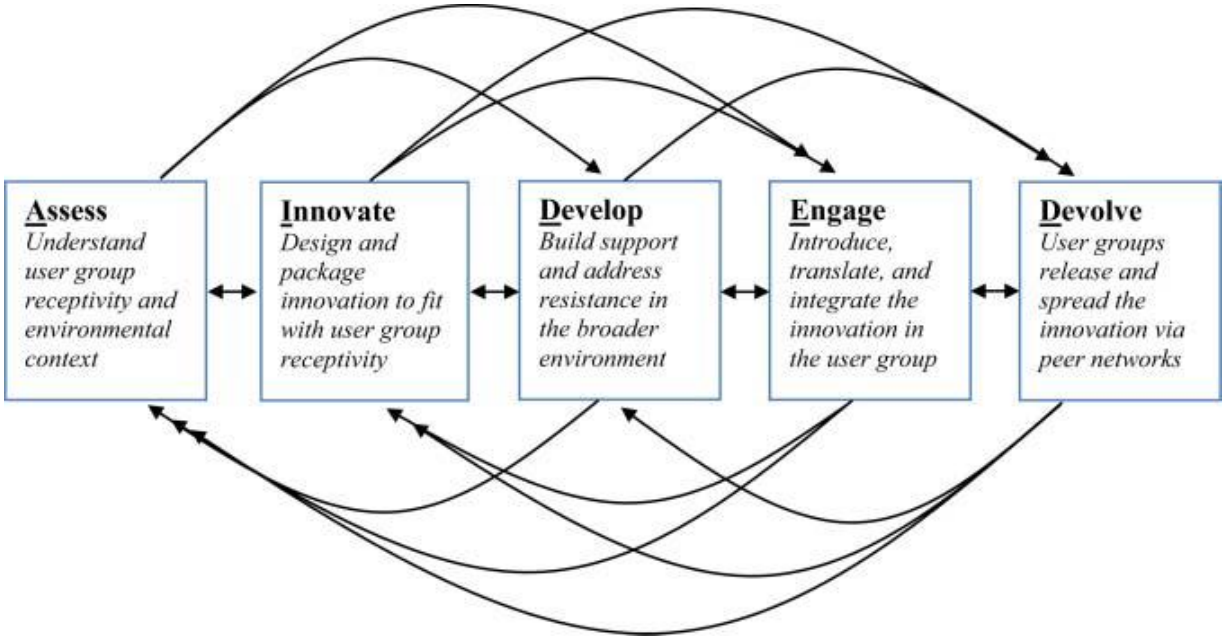
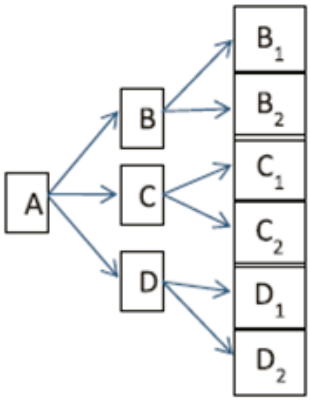


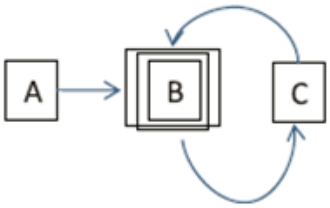
Figure 6 - AIDED model: interrelated components of a scaling process (Bradley et al., 2012:6)

Another thing that firms often stumble upon is that they have to deal with resistance from other actors in their system. These actors might be threatened and make losses because of their competitors development and scaling activities. To 'build support' is therefore crucial for bringing potential scalable products towards higher acceptance levels, towards certain 'points of transition' or 'tipping points'. When reaching such points, technologies may quickly become dominantly accepted (Bradley et al., 2012). Scale-free networks, such as peer groups, are found to be of key importance for building support and for the process of releasing and spreading health innovations to new user groups. Using scale-free networks however requires organisation to relinquish control, which involves certain risks. Release and spread through peer networks can cause positive-, but also negative unintended consequences and feedback effects (Bradley et al., 2012). By closely engaging user groups, organisations might be able to observe these positive and negative feedback loops and take action to 'enforce and trigger' or 'disable and supress' these (Seelos and Mair, 2012).

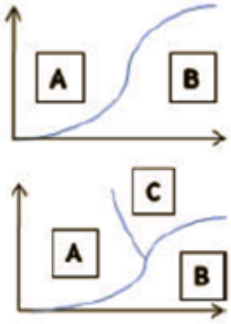
Table 1- Complex Adaptive Systems phenomena (adapted from Bradley et al., 2012)

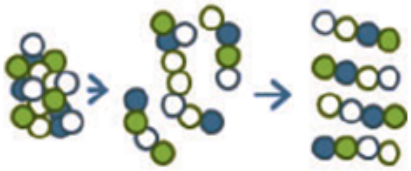


Left: *Path dependence*. Non-reversible processes have similar starting points yet lead to different outcomes through choices made along the way. Right: *Feedback*. Positive feedback increases the rate of change of a factor towards an extreme in one direction. Negative feedback modulates the direction of change.



Left: *Phase transitions*. Events that occur when radical changes take place in the features of system parameters as they reach certain critical points.
Right: *Scale-free networks*. Structures which are dominated by a few focal points or hubs with an unlimited number of links, following a power-law distribution
Down: *Emergent behaviour*. The spontaneous creation of order, which appears when smaller entities contribute to organized behaviours as a collective.





2.2.3 Insights from Organisational Studies

To bring an innovation at a scale were it matches the size of the addressed problem is typically an organisational activity. Organisations need to scale and benefit from successful innovations of the past to make an impact on 'todays' needs and problems. They however also need to have the organisational capacity to continuously innovate (OCCI) and explore new knowledge for 'tomorrow' in order keep up with competing organisations (figure 7). So organisations have to deal with two important dual and counterintuitive processes. The following elaboration takes a closer look at the issues present at both sides of this organisational balancing act that relate to the matter of scaling .

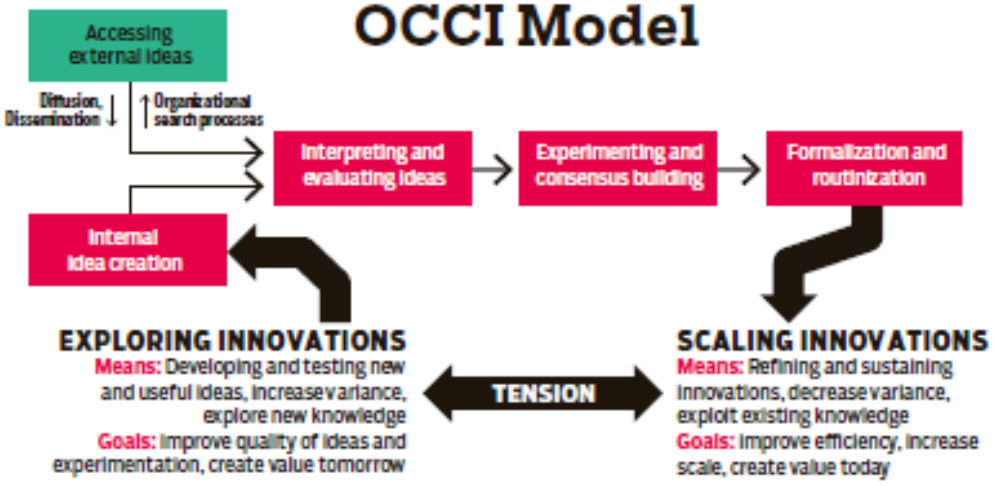


Figure 7 - Organisational tension between exploring & scaling innovations (Seelos and Mair, 2013:1)

On the one hand of this balancing act, organisations are required to focus on exploitation of existing knowledge or products within their current operating model. Fine-tuning of existing innovations and products often involves value-creating activities like: incremental improvement, standardisation and deep organisational routinization (Seelos and Mair, 2013).

According to Seelos and Mair (2010), the extend in which organisations can scale their outcomes is determined by something called: 'organisational closure'. In their 'closure theory', scaling is argued to be about 'event regularities'; the idea 'that doing more of A, or doing A better, leads to a higher expected outcome of B'. In this process, organisational closure is the prerequisite for 'sustained outcome regularities'. organisational closure can be defined as a 'distinct set of organisation knowledge that affects to what extend organisations can achieve and scale their organisational objectives'. This set of organisational knowledge contains of understanding on (i) the organisations' functional mechanisms and the way to trigger and enable these, and (ii) the organisations' dysfunctional mechanisms and the way to disable and supress these. As it is a form of knowledge, there exist different 'degrees of closure'. High degrees of closure correspond to great ability of organisations to identify, specify, enable and trigger its most functional mechanisms in a regularly way. High degrees of closure can also be achieved when organizations learn how to identify, specify and disable or supress their dysfunctional mechanisms. Closure is argued to be important for understanding inclusive growth. When working in resource constrained environments, it is often too challenging to replicate organisational activities. There are often not enough resources for this. Low productivity levels, often present in businesses in developing markets, make it imperative to first gain knowledge on how to enable higher closure levels, in order to obtain achieve higher productivity (Seelos & Mair, 2010).

Taking this a step closer to practice, closure often implies stabilisation and standardisation of organisational processes. We discussed how these processes affect technologies, but they also affect the organisations that develop and scale them (Seelos and Mair, 2013). In order to scale technologies, improved organisational efficiency is needed and often pursued by organisations through creation of different kinds of rules and routines. Sometimes these rules are not specifically created, but emerge as a sort of feedback. Cognitive rules are on example. They influence an organisations 'belief system'; its expectations and perceptions of the future. It determines how firms focus on particular technological directions. The available competences, skills and knowledge within an organisation often determine the setting of these cognitive rules. Also normative- (e.g. on 'proper behaviour in the organisation') and formal/legal rules enable scaling of technologies (Geels, 2004). Next to rules and routines, scaling often implies organisations to work together with-, or outsource certain activities to other actors in their network. You can think of partnerships with research laboratories, manufacturing corporations, regulatory bodies and more. Intensification of production involves new established mutual dependencies with partners, which increases formalisation of organisations (Geels, 2004).

One the other hand of the balancing act, organisation need to focus on exploring new innovations. This asks organisations to commit themselves to increase the variety of-, and experiment with new ideas (Seelos and Mair, 2013). Within the theory of the 'knowledge-based view of the firm' (KBV), a great deal of work looks at 'knowledge transfers' or 'knowledge flows'. Work from Lee and McNamee (2013) discusses how 3 types of knowledge flows (figure 8) can enable or constrain 'reverse knowledge transfer' or 'reverse innovation'. The concerned knowledge flows are: (i) a knowledge flow (red arrow) from the Multi-national Enterprise (MNE) to its 'subsidiary' (its daughter organisations, a local entrepreneur or first user of technology), (ii) a knowledge

flow (black arrow) from 'the local' to the subsidiary agent and (iii) a knowledge flow (blue arrow) from the subsidiary agent to the MNE ('reverse knowledge transfer'). It is argued that the intensity of these three knowledge flows impacts the efforts and activities of the MNE to access external knowledge for organisational innovation (Lee and McNamee, 2013).

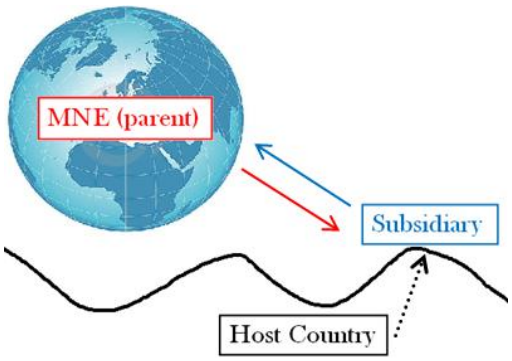


Figure 8 - Three types of knowledge flows (source: developed by the author)

Two notions are useful for explaining how the intensity of these knowledge flows can influence the innovation exploring activities of MNEs: 'knowledge rigidity' and 'subsidiary's *autonomy'*. The first notion captures the idea that when within a firm there is high dependence on some specific organisational knowledge, this might reinforce a 'rigid knowledge structure'. When looking at the first knowledge flow (red arrow) this means that when there is a high level of knowledge inflow from the parent to the subsidiary, the later can become restricted in its 'knowledge exploring' and 'value-creating' activity. If a subsidiary is overly constrained by the perspectives of its parent regarding market needs or product features, it will never be able to develop knowledge or products that have the potential to flow back to the MNE as 'reverse innovations'. True knowledge creation can only occur when a subsidiary is given enough autonomy to explore new ideas. Autonomous subsidiaries are more likely to react appropriately to local conditions as they are given the freedom to experiment with novel solutions.

This stresses however another point. Because besides giving autonomy it is at the same time crucial that a MNE provides a subsidiary with a sufficient level of knowledge inflow. A subsidiary with low levels of MNE knowledge inflow and high levels of 'local' knowledge inflow may have so much autonomy, freedom and flexibility to experiment with new ideas that this might results in too radically new knowledge and/or products that can't reverse to the MNE. When new knowledge differs too much from the MNE's 'dominant logic', current practices and technical and managerial resources, it can become irrelevant for the MNE. It is therefore critical for MNEs to ensure both a sufficient level of MNE knowledge inflow as well as a sufficient level of autonomy to their subsidiary agents (Lee and McNamee, 2013). Only then these agents are able to really adopt new external ideas and innovations into their operating system and work towards scaling them.

3. Different 'frames of theories'

This chapter presents an overview of the main understandings gained so far on the interactions between technology-producers and -users and how these affect processes of scaling. This overview must not be seen as a finished and comprehensive theoretical framework on the practice of scaling, but rather as an overview of the different 'frames of theories' or 'theoretical lenses' that can be used to look at this phenomenon. The three presented theoretical lenses focus on different types of processes that make sense of scaling trajectories differently (table 2). Table 3 provides a more detailed overview of what distinctive scaling processes and conditions belong to these three different sense making 'lenses' of scaling.

Table 2 – Different theoretical lenses that help to make sense of scaling trajectories

The innovation perspective mainly looks at scaling trajectories as *processes of technology adaptation*. Technologies are shaped and manufactured through user-producer interactions.

The systems perspective looks at *feedback mechanisms* in scaling trajectories. It reveals how these feedback mechanisms come about, how they impact the behaviour of the involved actors and how this influences processes of scaling.

The organisational perspective looks how organisations actively perform certain strategic activities in their period of scaling to *enhance* organisational *efficiency* and to maximally *exploit the functioning of technology*.

The following two case studies are carried out to see what deepened understandings these perspectives give on the phenomenon of scaling . These lenses will not only be used as a tools that give guidance in analysing the scaling trajectories of the following case studies on M-PESA and iDART. It will also be used to research whether the different processes that they focus on have different causal effects with scaling .

Table 3 - Overview of distinctive scaling processes and conditions found in literature

	The concept	Distinctive Scaling Processes	Distinctive Scaling Conditions
Processes of technology adaptation	User Innovation Innofusion	(a) Emergent behaviour of users that adapt technology (b) Continuous emerging feedback loops between processes of invention, production, distribution and implementation	(i) Existence of a 'configurational platform technology' that is still flexible and malleable (ii) Enough freedom / autonomy for users to act as 'adaptive innovators'
	User-Producer Interaction	(c1) Feedforward learning (broadening, user characterization, first user enrolment) (c2) Feedback learning (learning by using or configuring the user)	(iii) Involvement of users in early phases of technology design (iv) Firmly established linkages and engagement between users and producers

Processes of Systemic Feedback	The flexibility-stability evolution of a technology	(d) Path dependency: standardisation and stabilisation of a technology (e) Alignment of user demand. Through: (1) convergence of demand or (2) provision of differentiated products	
	Characteristics of Adaptive Systems	(f) Building support: breaking through the resistance and path dependency of important partners to come closer to a 'phase transition' (g) User network effects: targeting central hubs within scale-free networks to roll out a technology, monitor emergent behaviour of users and identify feedback effects	(v) A sufficient level of conformity and compatibility of the new technology with existing dominant platform technologies
Processes of Organisational Closure	exploiting old- and exploring new innovations	Organisations need to: (h) balance the dual pressures of finding new knowledge and technology for the long term and exploit and scale up existing knowledge and technology for the short term (i) balance the amount of autonomy they give to innovating subsidiary agents	
	Formalisation	(j) create rules and routines for efficient exploitation and functioning of existing technology (k) Outsource specific production or service tasks to specialised agencies for efficient production or service delivery	
	Organisation-al Closure	(L) Trigger and enable their functional mechanisms. (m) Disable and suppress their dysfunctional mechanisms	(vi) The capacity, competence, commitment and resources available within an organisation to perform and enforce closure activities

4. M-PESA: a literature case study on scaling of user-driven technological change in the financial sector

The telecommunication industry is currently the most rapidly growing developing market. Initially the mobile phone was conceived as a basic tool for communication. It however quickly evolved into a multipurpose platform that is used for all sorts of services (Van der Boor, 2013). Mobile banking is one of the examples of how users have been creative in addressing their needs by adapting the mobile phone technology. The enormous uptake of mobile phones in developing countries have enabled the rapid and widespread adoption of mobile banking applications. There are nearly six billion mobile subscriptions worldwide, of which more than 4.5 billion are in the developing world. Roughly 1.5 billion of these people do not have access to a bank account (Van der Boor, 2013). As the following two examples show, the need or demand for mobile banking existed before private companies introduced formal mobile banking services.

In the Philippines , around 1998, pre-paid mobile phone customers realised that they could use the functionality of 'top-up scratch cards' to transfer airtime load to each other by SMS. People started to send these codes to relatives across the country, who would top up their phones with them. Scratch-card credits were basically used for sending remittances and these Filipino customers invented this way of mobile airtime transfer (Van der Boor, 2013). Top-up credits were also passed along in exchange for services (Lallana, 2004). The Philippine telecom operator 'Smart' initially wasn't aware that its customers had identified an effective way of sending mobile money across long distance. But soon private companies realised the potential of these developments. In 2003, 'Smart' launched 'PasaLoad'. Shortly after, 'Globe Telecom' followed with the launch of 'Share-a-Load' (Van der Boor, 2013).

A similar innovation was observed in Uganda under the name: 'Sente'. People started to transfer money to family members by calling 'local village phone operator'. These small businesses make a profit by offering others to use their communication service. So people would call a village phone operator located nearby there family member and would read out their airtime code. The operator would use the code and its unlocked airtime for his/her own 'business' phone and would complete the money transfer by giving the same amount of money, minus a small commission fee, to the family member in cash (Chipchase, 2009 according to Van der Boor, 2013).

There are now around 120 mobile money ventures in the world (Heyer & Mas, 2009 according to Van der Boor, 2013). One of the best known mobile banking services is: 'M-PESA'. The following reconstruction of the development path and scaling trajectory of M-PESA reveals how both larger corporations as well as technology users explore the potential functionalities of a mobile phone, and how they connect and interact with each other when doing so. A view core articles have documented M-PESA's development path, which are used to make the following reconstruction of events (Hughes and Lonie, 2007; Gebregziabher, 2011; Mas and Radcliffe, 2011 and Vaughan et al., 2013).

4.1 Historic reconstruction of the development- and scaling trajectory of M-PESA in Kenya

4.1.1 Introducing M-PESA

In 2000, the U.K. government's Department for International Development (DFID) established the Financial Deepening Challenge Fund (FDCF). In mid-2003, M-PESA's journey starts when Nick Hughes (a Vodafone executive) decides to write a proposal for this competition fund. The proposal focusses on how mobile phones can be used to deepen financial access of the marginalised. It was basically a preliminary needs assessment.

Meanwhile, Hughes build support at two sides within Vodafone; support from senior executives within his company and from colleagues located in East Africa. The proposal of Hughes was awarded a funding of nearly 1 million British pounds, which was matched by Vodafone.

In 2004, Vodafone conducted a series of open workshops in Nairobi and Dar es Salaam to brainstorm with banks, microfinance institutions, technology service suppliers, NGOs involved in microcredit, and telecommunication- and financial sector regulators about ways to increase access to financial services. It resulted in a pilot partnership between Vodafone, Safaricom (Kenya's largest mobile telecommunication company), Faulu (a Kenyan microfinance institution) and the Commercial Bank of Africa. The partnership enabled developed a platform to allow customers of Faulu to receive and repay microfinance loans using a mobile phone, with payments made via the Safaricom distribution network of airtime resellers. Besides customer convenience, the service aimed to also bring business efficiencies to Faulu, to allow them to grow its business to more remote locations. The service was given the name M-PESA, with 'M' for mobile and 'Pesa' meaning money in Swahili (Hughes and Lonie, 2007; Vaughan et al., 2013).

4.1.2 Organisational- and technological developments over time

In preparation of the pilot program, three target locations were identified: the city centre of Nairobi, Mathari (a slum near the city centre) and a town called Thika. Faulu had recommended a few customer groups in these areas that were most likely to understand how to use a cell phone and embrace the service. Within these areas, a Vodafone sales team identified the most appropriate airtime dealer stores where customers of Faulu would regularly meet. The head offices of these stores were visited, and the potential benefits of M-PESA for were explained to them. Many shop owners quickly agreed to become M-PESA agents. Airtime dealers could have been rightfully worried about entering the mobile money business. Becoming an agent initially didn't appear to be lucrative because commissions on mobile money were a little lower than commissions on airtime (Cobert et al., 2012). Because Vodafone and Safaricom were such big, dominant and respected players, airtime dealers however quite easily decided to become a M-PESA agent. The sales teams of Vodacom developed 'operating instructions', 'frequently asked questions', and 'escalation procedures' and prepared training sessions that would educate M-PESA agents on the basic principles by which M-PESA worked.

Safaricom had two main preparation tasks. First of all they created a special 'M-PESA customer line' (234) to take calls from customers or agents having issues with the technology. Secondly, Safaricom was responsible for the management of cash flow to and from the M-PESA bank account. They prepared how they would accurately reflect this into the M-PESA system.

Faulu had to deal with some difficulties in their preparation to the pilot. At the time, Faulu used a paper-based system. They collected hand-written records from the field every week. This data was manually put into their central micro banker system. Faulu thought it would have difficulties in simultaneously managing this manual procedure and the automatic real time data entry system of M-PESA. They decided that their best solution was to instruct M-PESA agents to keep written records of the payments made by entering these data unto clerks print outs. These M-PESA records would then as well manually be put in their central system (Hughes and Lonie, 2007; Vaughan et al., 2013)

The M-PESA pilot started on 11 October 2005, with eight agent stores that were given a M-PESA phone. Around 500 microfinance clients of Faulu were given free mobile handsets, a few dollars in their M-PESA accounts as extra incentive and instructions on how to use M-PESA to repay their loans. These enrolled customers of M-PESA were able to perform three main functions. Customers could (i) withdraw their loans in

cash at participating M-PESA outlets. The M-PESA agent would then receive a text with the assignment to give the borrower a specific amount of money from the shops' till. Secondly, customers could (ii) repay their loans. By depositing cash at the outlets they could top-up their M-PESA account and return their borrowed money electronically to Faulu by SMS. Loans were however often given to groups of people instead of individuals. So customers were also enabled to (iii) send their mobile money person-to-person, to other members of their loan group. Mobile money would be send first to a 'group treasurer'. When all the borrowed money of the group was collected in his/her M-PESA account, he or she would transfer it to Faulu (Hughes and Lonie, 2007).

The first major problem during the pilot was that M-PESA agents hesitated to pay out loan withdrawals. Although training had told them they should, shop assistants often felt too insecure to obey the text message that told them to hand out cash from their employer's till. This problem was overcome by giving agents a separate M-PESA cash float. With such a 'second' till, shop assistants had more confidence to actually pay out the loans to M-PESA customers (Hughes and Lonie, 2007).

Another learning point had to do with the SIM cards used. Every customer was given a phone, but half of them already owned one. So they carried their M-PESA SIM in their wallets and swapped them with their own SIM when they wanted to do a transaction. Many customers lost their M-PESA SIM cards because of this. Swapping the SIMS also hindered clients from making 'spontaneous' transactions. To solve these problems, Safaricom enabled their customers to use their old phone number and their new M-PESA account on one and the same SIM, by using SIMEX cards. Over time, when M-PESA was nationally launched, Safaricom ended up fully absorbing the SIM replacement of around 6 million customers (at over $2 per customer) that needed to switch to a SIMEX card (Mas and Radcliffe, 2011).

By attending Faulu group meetings, by spending time with M-PESA retail outlets, by observing customer behaviour and by asking feedback it became clear that before launching a national service for millions of customers the process had to be simplified. So after two months into the pilot, the so far gained experience and feedback resulted in some technical modifications. It was made possible for customers to receive their loan via M-PESA, rather than receiving it in cash. Customers were also enabled to buy prepaid airtime with their M-PESA e-money. This need was highlighted by user who felt it would increase the utility of the service if they could use their M-PESA account to buy airtime without having to withdraw the cash first (Vaughan et al., 2013). So Vodafone created a new menu item, and connected M-PESA to the Safaricom billing system (Hughes and Lonie, 2007).

It was after these modifications that M-PESA transactions were taking place for a wide variety of purposes that were not intended. These transactions patterns were closely watched by Vodafone, until after some time a few researcher were given the assignment to explain them (Hughes and Lonie, 2007). Aside from standard loan repayments, the following alternative transactions were observed: (i) repayment of loans to others in return for services or cash, (ii) trading between businesses, (iii) larger businesses that were using M-PESA as an overnight safe (because banks would close before closing times of these businesses) (iv) People who would secure money while journeying (depositing cash at one end and withdrawing it at the other), (v) people who started to send their M-PESA purchased airtime directly to their relatives in rural areas as a form of informal remittance, (vi) Poor people that started using M-PESA as a way to build very small levels of savings, and other alternative uses were discovered (Hughes and Lonie, 2007; Gebregziabher, 2011; Vaughan et al., 2013). These innovative uses of M-PESA that customers introduced during the pilot program, convinced Vodafone and Safaricom that there was a market for M-PESA far beyond its use for microfinance purposes.

Evaluating the pilot, Faulu found integration of their accounts with those of M-PESA a too big obstacle. Loan disbursement through M-PESA also caused for significant over-riding of Faulu's existing routines. Faulu was used to organise weekly group meetings with clients to stimulate social relationships among members and to collect repayments of loans. A significant amount of clients felt however no longer the need to attend to these meetings, since they had M-PESA to repay their loans. Faulu saw these group meetings as beneficial and important and found the drop off in attendance disturbing. When the pilot officially ended on 1 May 2006 (although in practice it ran until October since people didn't stop using M-PESA) it became clear for Faulu that the service was more user-friendly for its clients than for themselves. On top of this, the pilot partnership decided that a mass market launch with an MFI would be too complex. So Faulu stepped out the partnership after the pilot, and Safaricom and Vodafone continued together (Hughes and Lonie, 2007; Gebregziabher, 2011; Vaughan et al., 2013).

Although the dropout of Faulu had mainly to do with organisational issues, practice also revealed that the technical flexibility of M-PESA wasn't appropriate for MFI clients. When a borrower for example would repay an average 20 week loan to its MFI using M-PESA, this would cost him or her around 600 KSh (roughly $10). You could say that this corresponds to a raise of interest of 8 of 9 percent. These added cost for repayment are costly for people that already are in lack of money. For sending remittances however, the commission costs of M-PESA were nothing in comparison with what one had to pay using alternatives (Foster, 2013a). So the design of the M-PESA technology itself also steered towards a certain usage, where other uses (like loan repayment) stumbled upon certain service barriers that limited such a 'user tactic'.

End 2006, Vodafone and Safaricom prepared a national launch. They signed a service agreement under which each party covered the costs over which it had control. They as well signed a corresponding revenue-sharing agreement. In order to increase the familiarity of customers with M-PESA, Safaricom and Vodafone decided upon the following things: (i) deposits of cash (for E-money) at M-PESA outlets were made free of charge, (ii) M-PESA customer were enabled to send money to non-M-PESA customers that used the Safaricom, Zain, Orange, or YU network in Kenya (customers could basically send money to anyone with a mobile phone) and (iii) pricing was made very simple, transparent and predictable. Agents couldn't charge direct fees, because transaction fees were subtracted straight from the customer's account. This reduced the opportunity for M-PESA agents to defraud. Transaction fees were also specified in fixed currency terms rather than as a percentage of the transaction. This would make it easier for customers to calculate the exact absolute value of money they would transfer to someone else (Mas and Radcliffe, 2011).

M-PESA was launched nationally in March 2007, with the key proposition: 'Send Money Home'. In order to provide simple functionality, the launch service was limited to three features. User could (i) deposit or withdraw cash to or from their M-PESA accounts at agent stores, (ii) transfer money person-to-person and (iii) buy prepaid airtime (Hughes and Lonie, 2007). In addition, some basic applications were enabled that allowed users to manage their accounts (balance checking, changing pin codes etc.) (Gebregziabher, 2011). Recognizing that M-PESA wouldn't be adopted unless customers had enough trust in the M-PESA retail network that transferred their money, Safaricom employed several measures during the first two years of scaling that improved the service of M-PESA agents. Firstly, when popularity grew and Safaricom was flooded with request for agency, the requirements for becoming an agent were tightened (Mas and Radcliffe, 2011). Due to rapid expansion, M-PESA agent training was outsourced to a sales training agency named 'Top Image'. They were now responsible for the start-up and further running of a M-PESA outlet. They offered set-up kits to new agents, helped these agents use their handsets, branded their shops and trained their staff. They visited agents

on rotation, at least weakly, to disseminate information, retrain the agent, gather information and keep an eye on things (Hughes and Lonie, 2007; Mas and Radcliffe, 2011; Vaughan et al., 2013). Their guidance was quite strong. An example of this is when some M-PESA agents in slums started to adapt the service and began to allow trusted customers to transact without showing identification. Peoples in slums find it often too insecure to carry around ID cards, while some of them do not even have one. When this development was noticed, Safaricom reacted by closing down these M-PESA businesses for one month (Foster, 2013a).

Once a critical mass started to use M-PESA, new functionalities were added alongside the M-PESA system. In December 2007 the service 'Pay Bill' was provided and in March 2008 Safaricom launched the 'Bulk Payment' application. The former application enables businesses to open organisational M-PESA accounts and collect money from their customers (e.g. for paying bills, or school fees). 'Bulk Payments' enables organisations to make one time bulk payments to customers (e.g. for salaries). In September 2008, Safaricom together with PesaPoint (an ATM network company) launched a service that enabled customers to withdraw money from ATMs using their M-PESA account. The introduction of this service made cash more accessible to M-PESA users. In December 2008, M-PESA introduced international remittance and in 2010 'M-KESHO' (meaning: savings account) was launched. M-KESHO subscribers have an actual bank account that is linked to their M-PESA account. This enables them with some benefits, like: earning interest on savings and obtaining micro loans and insurances. In October 2010, Safaricom introduced another application called 'Buy Goods'. This point of sale application enables users to make payments from their M-PESA account at till points in supermarkets and at other retailers. All these additional services were a result of further interaction of M-PESA with its environment, that desired to use M-PESA for other purposes than person-to-person money transfers (Gebregziabher, 2011).

The biggest issue during the whole scaling process of M-PESA was, and still is: 'float balancing'. If customers make too many cash deposits, M-PESA stores find themselves running out of e-float. The other way around, If customers do too many withdrawals, the stores run out of cash. During the scaling trajectory there were three main ways in which liquidity management was taken care of.

A first method involved the shop owner and its head office. Several M-PESA outlet agents would often work for one and the same head office. Money would never balance exactly in one outlet, but 'colleagues' from other outlets could then help out. Moving e-float and cash between different stores was the first way to solve the liquidity problem. The head office of these stores would send 'cash runners' to do the job, or shop owners would come to the head office and the transactions would be carried out over there. Vodafone created a web tool for the heads office's corporate accountants to support this method. For head offices without internet access, the same tool was enabled on 'head office mobile handsets' (Hughes and Lonie, 2007).

Because this first method was time consuming, risk intensive and quite expensive for the head office, a second liquidity management method was introduced. It made use of 'sub-agents' (most often banks). When having excess cash, an M-PESA agent could buy M-PESA e-float by depositing cash into the account of its head office at the nearest bank branch or ATM. Once the head office confirmed receipt of the funds into its account, it transferred e-float to the store's M-PESA account. When the store had excess e-float, the process would be carried out vice versa. This method was still quite ineffective, especially for the outlet agent. Money transfers from and to the bank account would take one or two days during which the outlets would be out of business. It also required shop-owners to leave their shops during working time, because banks have limited visiting hours. A better system was called for.

With the third method, M-PESA stores are able to interact directly with a bank that has registered itself as an M-PESA 'Super-agent'. M-PESA stores would open an account at such a super-agent bank. At the nearest branch or ATM of such a bank, stores could electronically buy and sell e-float against this bank account in real time (Mas and Radcliffe, 2011; Vaughan et al., 2013).

4.1.3 M-PESA's scaling over time in figures

At launch Safaricom had 750 M-PESA agents stores that covered all of Kenya's 69 district headquarters (Mas and Radcliffe,2011). Within the first month of operation (March 2007), 20.000 customers signed up. Within one year, M -PESA had more than 2 million customers. Within three years there were almost 10 million customers. According to the latest data publication of Safaricom, more than 14 million people have a M-PESA account and nearly 28 thousand M-PESA outlets have spread throughout the country by April 2011 (figure 9) (Safaricom, 2011).

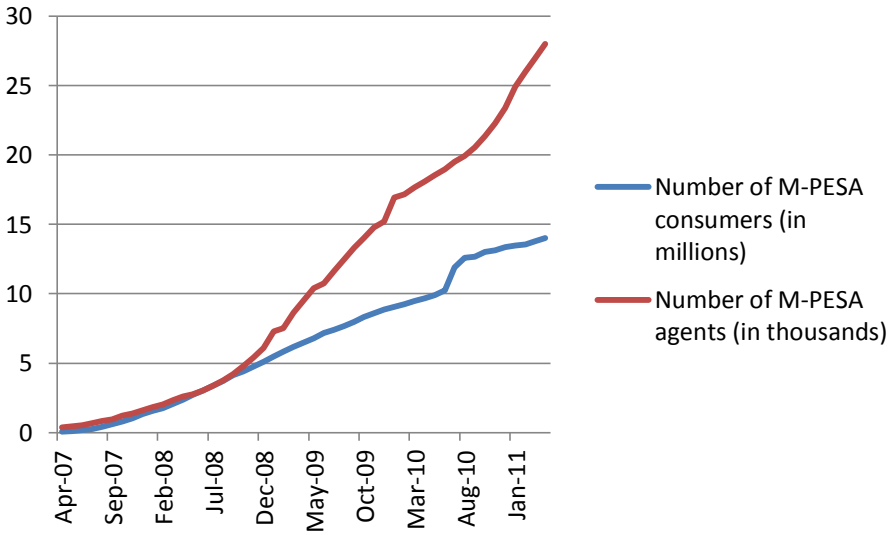


Figure 9 – The scaling trajectory of M-PESA (Safaricom, 2011)

Not only M-PESA's membership rates skyrocketed, also the volume of mobile money transfers rose enormously over time (figure 10). The following facts help to interpret these numbers. Throughout the whole year 2011, M-PESA transacted a total estimated amount of 7 billion dollar. During this year they transferred substantially more money than the Western Union Bank did that year in the whole continent Africa. The 7 billion transferred dollars are equivalent to 30 percent of Kenya's GDP (Vaughan et al., 2013). M-PESA now provides mobile-banking facilities to more than 70 percent of the country's adult population (Cobert et al., 2012). The enormous growth of M-PESA reveals that there was a strong latent demand for a technology that could transfer domestic remittances in a convenient and safe way (Mas and Radcliffe, 2011).

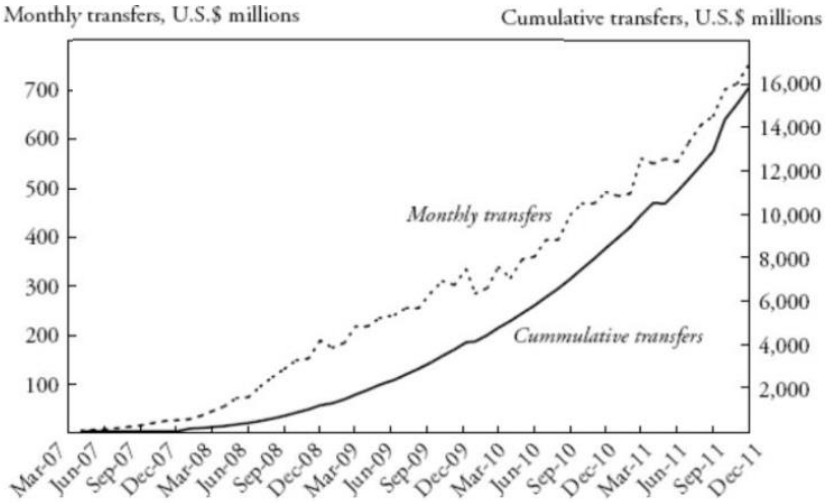


Figure 10 - M-PESA Person-to-Person Transfers, 2007-2011 (Figure adopted from Vaughan at el., 2013:162)

5. iDART: A field case study on scaling of user-driven technological change in the healthcare sector

This field case study presents how the organisation 'Cell-Life' has developed and scaled electronic technologies that help to monitor and support HIV patients' adherence to antiretroviral therapy (ART). The case study specifically focusses on an electronic pharmacy system called: Intelligent Dispensing of Antiretroviral treatment (iDART). The total number of persons living with HIV in South Africa in 2013 was estimated at 5,26 million (Statistics South Africa, 2013). The enormous HIV pandemic in South Africa, that came up during the 1990's, urged for the setup of a massive ART program. However, the ART programs in South Africa were rolled out very slowly. The initial lack of political will of the national government to provide the necessary resources for the expansion of the number of Antiretroviral(ARV) treatment sites, plus the slow and long accreditation processes that were involved in opening new ARV pharmacies have for many years put great pressure on existing treatment sites to expand their patient load. Without comprising the quality of care, efficiency became a central concern for these clinics (Wessels, Nattrass, Rivett, 2006). Because the national government in South Africa did not push for rapid expansion of ART, many NGO's and research organisations took initiative in founding and supporting ARV clinics. In this context, also Cell-Life came into existence.

5.1 Historic reconstruction of the development- and scaling trajectory of iDART

5.1.1 Introducing iDART

Cell-Life began as a research collaboration between the University of Cape Town and the Cape Peninsula University of Technology. Researchers and students from the faculties Engineering, Computer Science and Health Science participated. Soon after, also researchers from other institutes and interest groups (TAC, DTHF) got involved. Since then, Cell-Life has come a long way (see appendix 4 for an overview).

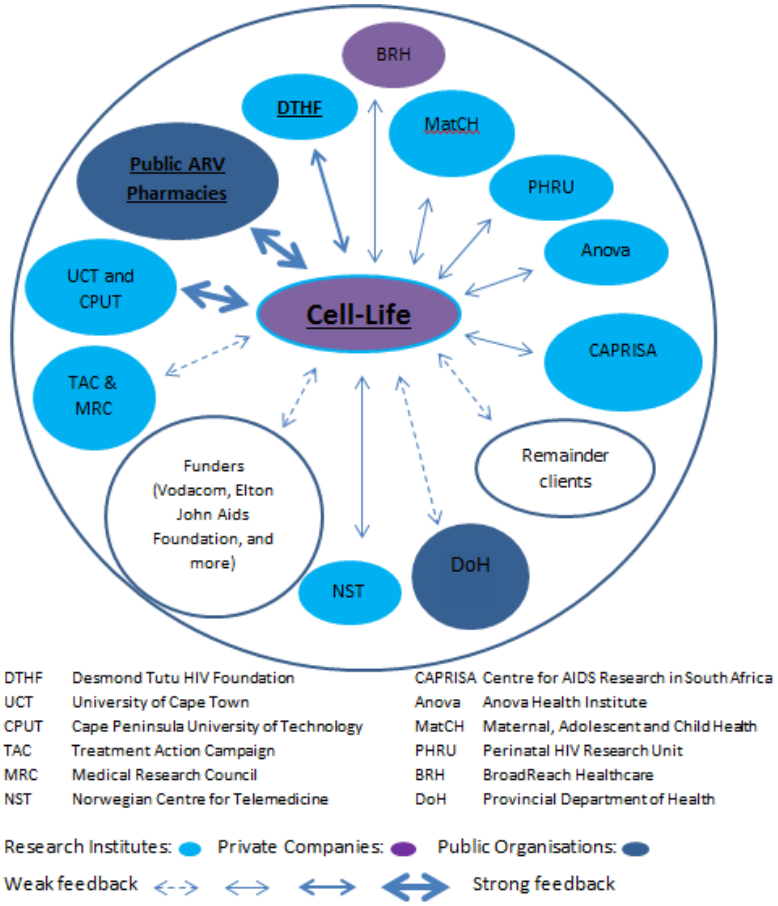


Figure 11 – Overview of the involved stakeholders during the beginnings of Cell-LIfe

The main technologies that have been developed by Cell-Life are: 'iDART' (an Intelligent Dispensing software for ART), 'Capture' (a mobile and web-based data collection system) and 'Communicate' (a communication platform enabling information and communication services to HIV/AIDS patients and organisations). Figure 11 shows the first group of actors involved during the beginnings of Cell-Life. Cell-Life's first clients consisted of a group of research institutes (DTHF, MatCH, PHRU, Anova, CAPRISA). The Desmond Tutu HIV Foundation (DTHF) was the first client. It founded the Hannan-Crusaid Treatment Centre (HCTC) in Gugulethu and another treatment site in Masiphumelele, where the first versions of iDART and Capture were implemented during the period 2004-2006. Communicate was introduced by Cell-Life a couple of years later.

In order to be able to obtain rich data in a limited amount of 8 weeks of doing field work, this case study limited itself to conducting interviews with people from Cell-Life, the DTHF and the HCTC and merely focused on the development path of iDART. Cell-Life's developments are well documented (Wessels et al., 2007; Wood et al., 2008; Rivett and Tapson, 2009; Tolly and Alexander, 2008; Loundon and Rivett, 2011). Besides interviews, more data has been obtained through document analyses of, amongst others, these articles.

5.1.2 Technological developments over time

iDART was initially developed as an IT solution for a new 'down-referral dispensing model' (figure 12) that was introduced by the DTHF. Before, ARVs were prescribed and pre-packed by a central pharmacist. Incoming shipments of ARVs would be pre-packed for specific patients, labelled with barcodes, scanned into iDART, and scanned out when these packages were sent to the dispensing sites. When the packages arrived, they would be scanned into iDART again and scanned out when patients would come to pick up their medication. iDART's functionality was twofold: it was used as (i) a stock management tool and (ii) as a dispensing tool. It gave pharmacists at the dispensing sites accurate stock control. They would know when their medication would come in and 'when, what and how many' new ARV medication to order. It gave them dispensing-related information as well, like: patients' dispensing history, dates of when patients were supposed to come back to pick up new ARVs, what patients were defaulting treatment and other relevant information. Although the down-referral model is still used at clinics located in remote rural areas, over time clients started to request iDART to support 'direct dispensing'. The numbers of patients at these treatment sites started to increase rapidly and down-referral dispensing became impractical. iDART was adapted to support direct dispensing. It became less stock management oriented and focussed more on supporting dispensing functionalities.

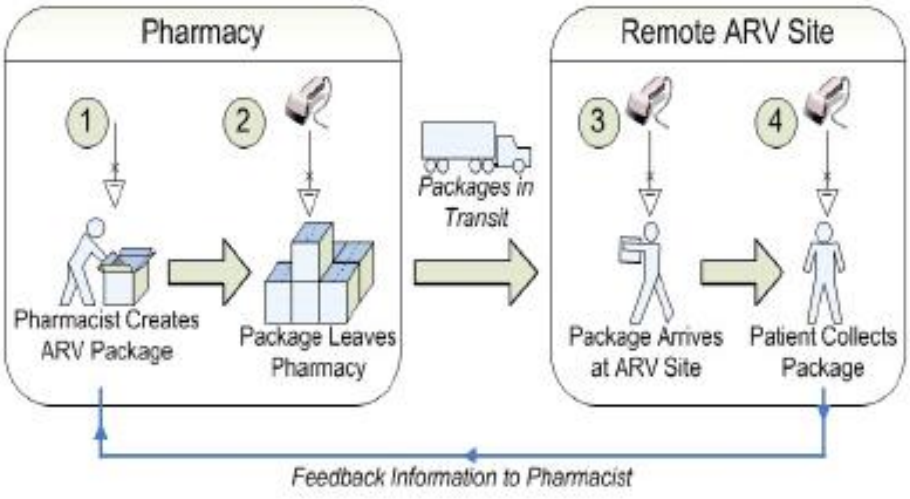


Figure 12 –The down-referral dispensing model (Cell-Life, n.d.: 4)

From the 1.0 version to the current 3.8 version of iDART, the software has undergone more substantial changes. The DTHF requested iDART to be 'cheap, simple to use and compatible with low capacity computers'. The first prototype of iDART was 'just a fancy label printer'. Over time, iDART extended with functionalities that enabled clinics to do more advanced things, for example to: (i) capture the return of packages into stock, (ii) dispense multiple packages of different sizes of the same drugs, (iii) dispense additional medication alongside ARVs (e.g. vitamins), (iv) capture additional patient data (e.g. the number of patients' episodes, why patients stopped treatment, data if the patient was down referred and more data), (v) create different types of statistical reports (that clinics were obliged to hand in to their funding agencies) and other extensions.

These kind of new functionalities were added, depending on the needs of the users. *"Users were constantly involved in 'evolutionary growth'. [..] We placed the prototype out there, and it evolved"* (S.N. Mahomed, personal communication, January 21, 2014). Business analyst Sarah Brown confirms this: *"We developed our processes around the requirements that we gathered on our way. When we would go into these sites, they would be like "Well, we like your system, but can't you do 'this' and 'this'?" Our budgets were always really tight, so with me being up in the North West in a rural town the instructions were: 'Don't come back until it is working up there'. So there were many crazy nights when I would be in a hotel room and I would be phoning Simon: 'Simon, I need a pill count feature now. It must do this, and must do that'. And then he would make something up and send it over the next day"* (S. Brown, personal communication, February 17, 2014). Most of the software development of iDART happened quite haphazardly. Although it was an open-source software, most development was done in-house; within Cell-Life. Cell-Life did look for partners to contribute code but didn't succeed in that. The only actor that contributed code was 'Anova'. They created an additional module called 'Patient Visits and Statistics', which contains patients statistics (CD4 counts, Viral Loads, ALT) and patient visits reports (wellness tracking); information that is vital for patient care. Most iDART implementation sites were updated with this module.

With more and more iDART projects coming in, and more and more request for extensions of iDART, the architecture of iDART slowly becomes like a 'Jenga tower'. In the development process of continually adding bits and pieces that enable additional requested features, the developers didn't focus so much on iDART's roadmap and didn't work on building a strong code base. *"And that is fine for a while, but if you really wanted to take something out or change something, the development time starts increasing exponentially because the changes would brake everything else"* (S. Brown, personal communication, February 17, 2014). Simon Kelly elaborates on this: *"iDART kind of grew out of its original design. The original architecture of iDART is now limiting where iDART can be used and what you can do with it."* (S. Kelly, personal communication, February 13, 2014). Cell-Life always wanted to be able to dispense any drugs from iDART, not just ARVs. But because the initial design of the system was built around chronic communication and prescriptions of medication, it became too complicated and expensive for Cell-Life to redesign it into a system that functions as a general pharmacy system. That is why Cell-Life two times tried to rewrite iDART from scratch. The first attempt failed because the employee that worked on the new version left Cell-Life before it was finished. Other developers didn't know enough about this design to be able to either finish it, or to integrate it into the existing version of iDART. The second time, the new design was about 30% through when funding for this work ran out. As a result of continuing to work within a 'Jenga tower architecture', there were various important features that were too expensive to integrate into iDART. Another consequence was that iDART became very slow. For example the loading of a patients profile or "*When you need to generate a report, it takes too long. It takes 5 to 10 minutes for 1 patient."* (E. Brant [pharmacy assistant], personal communication, February 17, 2014).

Discussing how iDART came to such a situation, Simon says: *"In most of the projects that we did, clients*

would have request that either didn't make sense or were just not really necessary. So we always were doing some negotiation. […] But I can't think of anything major that we turned down. Mostly we didn't say 'no' to anyone's requests, because that wouldn't be an answer" (S. Kelly, personal communication, February 13, 2014). Kieran more bluntly explains why 'Saying no' was most not an option: *"Cell-Life was very small. We made the changes the people pay for. Less so we made the changes that we would think were strategically important in those days. If a programmer or a director would think that something was strategically important for the product, you would have to do that at night"* (K. Sharpey-Schafer, personal communication, February 13, 2014). New projects, the ones that brought in good money, would often imply requirements that didn't fit into the current existing product. So additional development needed to be done to extend the product to cater for new requirements. An example of this is how Communicate was integrated with iDART for a new research project and how a function was added to iDART that enabled to make randomized selections of patients. *"None of that work wouldn't be done if we hadn't this project. The projects had a big influence on the direction of the technology"* (S. Kelly, personal communication, February 13, 2014).

Figure 13 - Impression of use of iDART by the Hannan-Crusaid Treatment Centre

5.1.3 Organisational developments over time

The technological developments over time can't been seen separately from the organisational developments of Cell-Life. One Cell-Life employee speaks of three phases that Cell-Life has gone through over time: (i) a student 'mickey-mouse' period, (ii) a growth period and (iii) a period of becoming a mature and sustainable business.

In the first phase (2002-2006), Cell-Life had all the dynamics of a research group. There were more students involved than staff members. Everything was thought of in an academic way. Processes took very long and involved a lot of: research papers, discussions and communication. The focus lay on learning and much time

was taken for launching the first prototypes of iDART and Emit (later Capture). During 2005-2006, this slowly started to change. More software developers were hired. Besides doing research, Cell-Life became a software development house. At the time Cell-Life would make software, deliver, and disappear. Clients would however often change their minds. During a conflict with a client about the specifics of a product delivery, Kieran remembers one of his client ones said: *"I don't want what I asked for, I want what I need"* (K. Sharpey-Schafer, personal communication, February 13, 2014). During 2006 the following important lesson was learnt that changed Cell-Life's way of doing software development. The old 'waterfall model' was replaced by a development process where they would: 'built something early, show it to the customer quickly and then basically iterate the waterfall model'. The new development processes helped overcoming the pains experienced during the old way of development. In 2006, Cell-Life officially became a non-profit organisation. The number of stakeholders and funders increased after this.

During the period 2007-2009, Cell-Life became much less a research and development house and started to focus more on project management and implementation. Cell Life employees don't recall this being an explicit decision. It was a gradual shift that was caused by the fact that their activities were dependent on incoming funding, which was easier to obtain for implementation- than for engineering activities. *"There was still a focus on research, learning from the technologies and instituting them, but besides a lot of pilot projects were coming in. They scattered all over the place."* (S.N. Mahomed, personal communication, January 21, 2014). Part of this timeline change was the coming of a new director: Peter Benjamin. He was experienced in working within the 'NGO world' where he had a big network. This allowed Cell-Life to write grant proposals to big funders and sell iDART (and Emit) to NGO's. This led to *"the point where it was like: 'deliver, deliver, deliver'. Cell-Life ran of the energy and engine of individuals and grew because of what Sarah did: 'grow, grow, push, push'. Emit would have gone nowhere if I didn't spend five nights coding. Even then it was still not a product, more a prototype. I was just making it work. But at Cell-Life we went and sell that prototype to this NGO. […] Our turnover grew with millions of Rands and we had to hire new staff for 'today'. So it put a lot of pressure on you to hire people. And I was 24 or 25 years old, what do I know about hiring people? Now at that point you have so many salaries, you have to take the contract for the features that would pay us money, not the contracts that the product perhaps needs. So then you get stock in a cycle of delivery. We didn't do any product management, until you got a big wave of money from someone,"* (K. Sharpey-Schafer, personal communication, February 13, 2014).

Not only did increasing staff numbers enforce a 'cycle of delivery', It also created a need for more and better processes. This started late 2009 when a consultant joined Cell-Life for a few weeks to work on restructuring the organisation. Cell-Life started to work with Product Managers, Business Analysts, and a Support Team (ST) (responsible for implementation). Cell-Life started to build 'a knowledge base'. Because the ST was taking over operational activities, knowledge started being documented. Training materials were developed and an escalation process was set up with guidelines for when the ST had to report back to their managers. Teamwork was still very unstructured and inefficient at that time. Business analysts worked some days on 'iDART' and some days on 'Cellphones4HIV' (later: Communicate) in cooperation with different developers. *"This woman gave us more of a framework to do simple things like: 'if you start a project you must have a project manager. The project manager sources resources from the head of development. You have to get agreement from them before you start the project'. It gave us a framework for dealing and processing all the demands that were coming in"* (K. Sharpey-Schafer, personal communication, February 13, 2014).

Most of the 'formalisation proces' takes however place in the period 2010-2014. During this time Cell-Life starts: (i) doing Human Resource Management (HRM), (ii) using SCRUM (a software development

methodology), (iii) working with Quality Assurance processes (QA) and (iv) researching and integrating the use of a 'pay-as-you-use model' for Emit. For some of the iDART projects, Cell-Life starts to partner with other organisations on the implementation side (BroadReach Healthcare, Anova, PHRU, technical teams in Nigeria). Cell-Life trains these teams to enable them to give training sessions and to manage further support to the client.

A major event that stimulates the formalisation development is the initiation of Cell-Life's commercial subsidiary: 'mHealth Solutions'. *"Over the years Cell-Life developed the reputation of being the leaders in mobile health communication, regarding: content, behavioural science, what technology works, how you reach the patient, what best communication platforms to utilise and how you use multiple platforms."* (S.N. Mahomed, personal communication, January 21, 2014). Cell-Life realises that it is able to compete in the commercial market. So in 2010, Emit is rebranded into Capture and the technologies developed as part of the 'Cellphones4Hiv' program are now sold under 'Communicate'. In comparison to Capture and Communicate, iDART didn't have a clear business model for commercial use. So commercialisation didn't apply to iDART. With mHealth Solutions, Cell-Life is able to keep funding contracts on the one side while also having commercial contracts. Profit form the commercial contracts help fund Cell-Life's research and development activities. Cell-Life is not anymore bound to specific requirements of funders, but can invests more freely into strategic development of their technologies. It enables Cell-Life to move iDART away from a stand-alone system to a web-based system. The web-based version of iDART has the advantage that it operates interactively. It also enables integration with other patient record systems. The development of the web-based version of iDART is now finished and currently in its 'testing phase'.

For some clinics, like the HCTC, this development seems to have come too late. iDART is planned to be taken out of the HCTC during April 2014. The provincial government has contracted 'JAC' (another pharmacy system) to be rolled out in all the government supported clinics throughout the Western Cape. This roll out started in 2011. Heli Moung and Eugene Brant agree that the government has chosen for JAC as the 'government pharmacy system' because of 'consistency and uniformity reasons'. The JAC system, for many years, already allows interactive operation and communication. Commenting on this development, Sarah Brown says: *"If you want to put iDART in every clinic in SA, you need to have a tender with the government. This is top-down and takes ten years to do, where we went bottom-up. We came in with a free IT system that made life so much easier for many. Everyone loved it. The problem is, if you are not top-down you won't do that expansion that you are looking for. It's about politics. At the end of the day it is not who has the nicer system, it is about who can get in"* (S. Brown, personal communication, February 17, 2014).

From 2012 onwards Cell-Life slowly starts to move away from communicating about, and working within, three separate 'silos' of products (iDART, Capture, Communicate). It start to work in a less supply-driven and more integrated and solution-driven way. Besides doing more marketing and branding, Cell-Life gets more efficient in serving their clientele. Project managers start using standardised application forms for new clients that capture their demand. They hand these over to the development teams, who make estimates of the work that need to be done together with the costs involved. The project manager than return these forms to the client.

During this period in which Cell-Life becomes more efficient, the culture of the organisation changes. *"It was a challenge moving from that passionate NGO space to a scaled up corporate environment. The biggest challenge was the people coming in from corporate. People came for different reasons and had different motivations. With hiring more people it is the idea of handing over, but I feel like, now looking back at my time, I didn't hand over enough. I kept too much to myself. Maybe it was a pride thing, an ego thing, or an*

combination of things. By not hiring the same competent and passionate people, it is hard to give up on that dream" (S. Brown, personal communication, February 17, 2014).

5.1.4 The scaling of iDART over time in figures

The way that iDART scaled, in terms of its growth in technical capacity and functionality, has now extensively been discussed. We have seen that it's scaling hit certain limits. iDART nevertheless has impacted, and still impacts, ARV clinics in their ability to cope with increasing patient numbers. iDART was installed into 18 clinics during 2009. During the following two and half years, Cell-Life monitored these sites to determine the increase in capacity as a result of the use of iDART. Monitoring the growth of the number of patients receiving antiretroviral treatment showed that all the sites increased substantially over this period. The average number of patients in these 18 clinics in 2009 was 775. Two and half years later this average was estimated at 2706, which is an increase of 249%. Although a portion of this increase must be attributed to capacity enhancing activities by the Government and the natural growth of the clinics, these figures give an indication of how iDART impacted the patient capacity of ARV public health facilities (Cell-Life, 2014).

iDART is now being used in over 120 district clinics and over 500 down-referral sites. Since 2012, Cell-Life began to work with more larger international projects and gets more experienced with implementing their systems internationally. Continual scaling may therefore be expected. Figure 14 gives an indication of the scaling trajectory of iDART during the period 2004-2014. These figure exclude numbers on down-referral implementation sites. The graph is based on figures presented in appendix C, which gives an overview of Cell-Life's implementation sites.

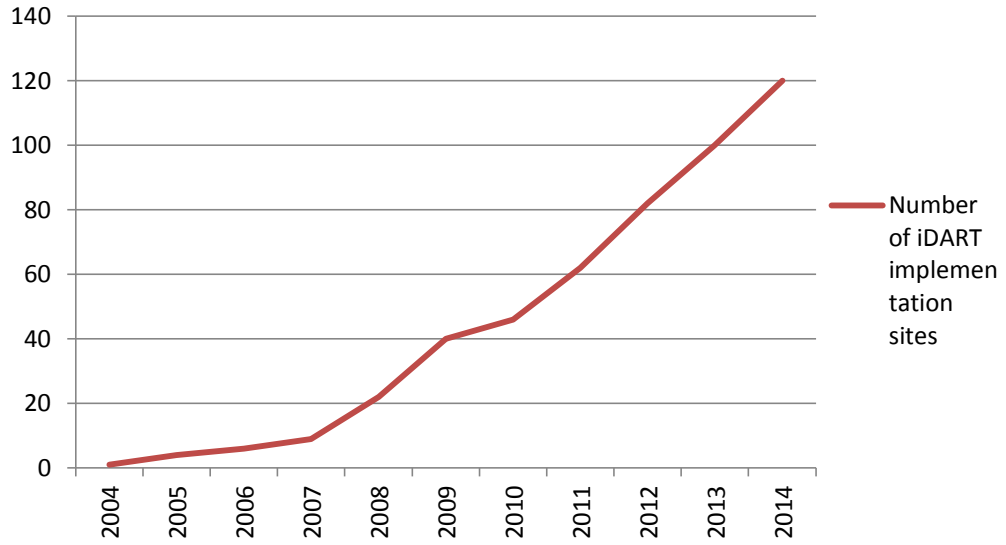


Figure 14 – The scaling trajectory of iDART (Source: Cell-Life, n.d; Cell-Life, 2008 and Cell-Life, 2014)

6. Cross-case analysis

Two scaling trajectories have now extensively been reported. The presented data contains insights on different scaling processes. These processes will now be distilled through application of the three theoretical 'frames' or 'lenses' introduced in chapter 3. Table 4 gives an overview of the results of this distillation process and summarises the main results of this thesis. It seems unquestionable that the processes presented in table 4 have impacted the scaling trajectory of M-PESA and iDART. However, in order to gain a deeper understanding on the phenomenon of scaling, knowing *what* distinctive processes are part of scaling trajectories is not most interesting. It is more relevant to know '*how?*' and '*to what extend?*' these processes have impacted scaling. Because scaling consists of a mix of processes, causal relationships are complex and hard to determine. The quantitative data gathered on the scaling trajectories of M-PESA and iDART don't provide sufficient detail to proof causality nor to determine the extent of these causal relations. The aim of the following elaboration is therefore too qualitatively describe these causal relationships, through a cross-case analysis. Based on the empirical data, several hypotheses concerning causality with scaling are proposed (table 5).

6.1 Processes of technology adaptation

The iDART case study reveals that hypersensitivity to user need may lead to development of unstable architectures of technology. Such 'Jenga' architectures limit further product development and refinement. Hypersensitivity to user needs seems to negatively impact scaling.

By analysing the M-PESA case study, it can however also be argued that a lack of sensitivity to user demand negatively impacts scaling. M-PESA's scaling would not have been so enormous if Vodafone and Safaricom would have continued to serve the initial targeted demand: the need for more financial inclusion and access of micro credit clients. If they would have kept tight control over the initial micro credit edition of M-PESA and would not have noticed, allowed or reacted to the alternative uses of M-PESA initiated by its first users, M-PESA would not have seen such enormous scaling figures. Too strong guidance of producers on users and their technology-related behaviour also limits firms to exploit the full capacity of their products and/or services.

This thesis therefor argues that balanced levels of say and autonomy of users and producers in processes of technology adaptation are crucial for scaling. Such balance positively impact the scaling of user-driven technological change (hypothesis a). 'Sensitivity to user demand' may be seen as one of the 'control panels' that determines the emergence of scaling (figure 15).

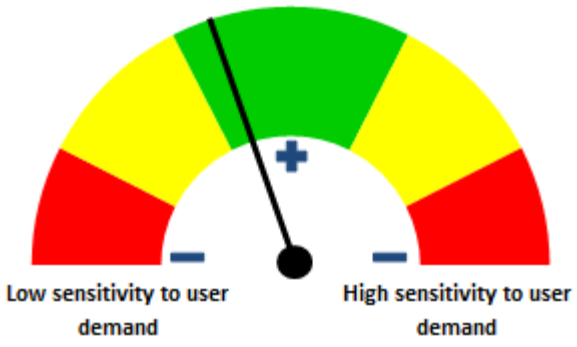


Figure 15 – Control panel (A): sensitivity to user demand

The iDART case study teaches us as that it is extremely difficult to preserve the customisability; the multi-purpose employability of a technology, during times of implementation. While implementing, Cell-Life adapted iDART through additional software development in order to address various upcoming user requests. This

resulted in quick stabilisation and standardisation of the technology. Because of a lack of demand selection or 'convergence of demand', the technology stabilised into a final architecture that could not perfectly support or serve the different demands that it over time tried to address. Once stabilised, iDART did not fully functioned as an ARV dispensing software (iDART became very slow, couldn't provide certain crucial functionalities) while it wasn't either useful anymore for development of a general pharmacy system. When looking at the end of the scaling trajectory of iDART, it can be observed that its scaling got a boost after Cell-Life moved away from 'making their products customisable' towards 'offering multiple custom-made products'. When mHealth solutions was formed in 2010, Cell-Life still worked in user-driven way. They now would however would serve heterogeneous user demands by developing and offering differentiated products, instead of adaptation of one and the same 'customisable' product.

The case study on M-PESA reveals how it was really important for Vodafone and Safaricom to first focus on only one of the many potential functionalities that M-PESA could offer, namely: enabling person-to-person mobile money transfer. Vodafone and Safaricom converged their technology development- and implementation efforts and focused them on this very specific user demand: 'to send remittances home easily, safely and affordably '. After adequately serving this demand, additional products and services were developed alongside M-PESA that were compatible with the existing M-PESA platform. These products and services served the other discovered user demands. It seems that 'convergence of demand' followed by 'differentiation of product delivery' have been successful methods for Vodafone and Safaricom in dealing with the many heterogeneous user needs they faced.

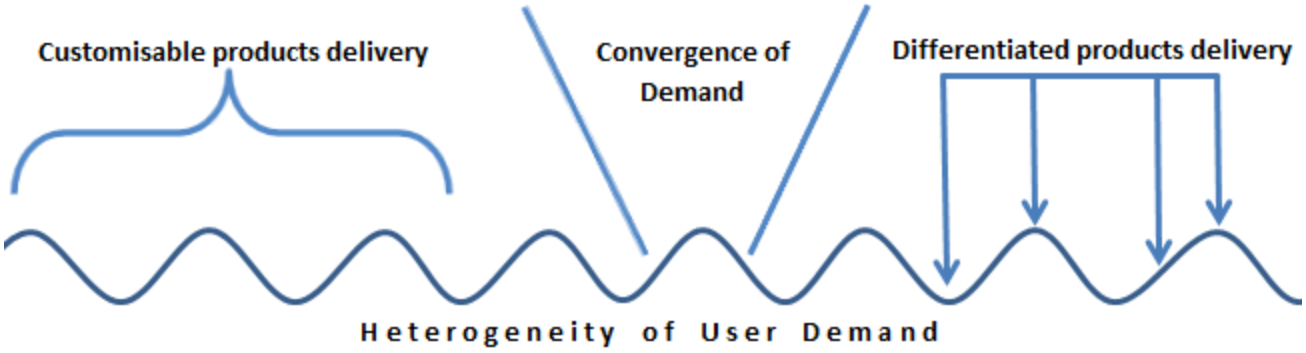


Figure 16 - Three methods of demand selection (source: developed by the author)

Both case studies give the understanding that guidance of the direction of technology development involves strategic organisational selection of user demand. Three methods have been observed that organisations use to come to such a selection (figure 16). Because technology stabilises and standardises towards a certain direction and certain usage, it is important for organisations to strategically select what demand (not) to serve. If organisations do not accurately select one particular demand, they might end up serving multiple demands poorly. It is more likely that producers can accurately fulfil demand and can scale their products or services when they put effort into serving heterogeneous user needs through 'convergence of demand' and 'product differentiation'. This thesis proposes that higher levels of demand selection, through 'convergence of demand' or 'differentiated product delivery', positively correlate with scaling (hypotheses b). An important note hereby is that convergence of demand implies that certain user demands are focussed upon, while others demands are left out. 'Inclusive scaling' can therefore only take place when convergence of demand is followed up by 'differentiated products delivery'. The degree in which organisations' select 'what demand to serve?' is a second important control panel that determines the emergence of scaling (figure 17).

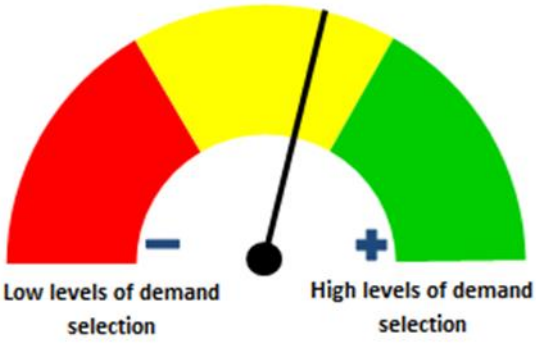


Figure 17 – Control panel (B): levels of demand selection

6.2 Processes of systemic feedback

iDART's implementation activities have increased due to, amongst others, initiation of more and better organisational processes within Cell-Life (for examples, see table 4). Likewise, M-PESA has been able to scale so quickly because Vodafone and Safaricom actively guided and coordinated its development process. In both the iDART case study as well as the M-PESA case study it becomes clear that the technology-producing organisations needed to formalise their organisational procedures in order to be able to keep up with the expansion of their productivity level and/or product- or service capacity. Instead of personal and ad hoc responses to systemic feedback (involved in iDARTs development process), increased product delivery requires organisations to respond in a more procedural and coordinated way. These different ways of responding to system feedback forms another parameter or 'control panel' influencing scaling (figure 18). The hypotheses is proposed that formalisation of organisational processes positively influences organisations' capability to scale their products and services (hypotheses c).

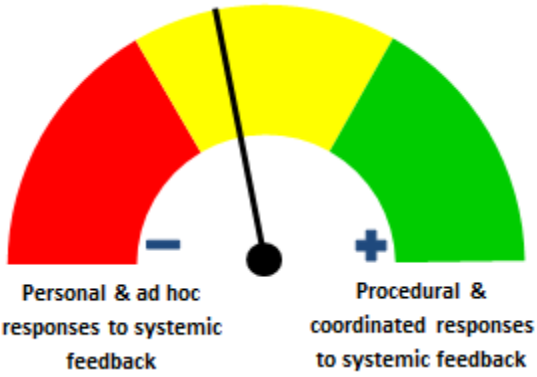


Figure 18 – Control panel (C): types of response to systemic feedback

Figure 19 – Control panel (D): short- vs. long term management

When looking how scaling trajectories are affected by changed organisational behaviour of producing organisations, two phases of systemic feedback can be identified. During its period of rapid growth, Cell-Life got to a point where it merely focused on project management and implementation of its technologies. As the organisation grew because of this, there was more to be taken care of for the longer term (extra personal, a more complex technology, relationships with partners, etc.). Cell-Life however hardly performed any long term management or long term strategic planning and got stuck in a 'cycle of delivery'. This had its consequences. iDART quite quickly stumbled upon certain technical limits. Another consequence was how Cell-Life had to lay off some employees because there was a lack of sustainable income. Only after 2010, when Cell-Life became

more financially independent, Cell-Life started to focus on long term strategic management and started with 'sustainable business development'.

We see that during a first phase of systemic feedback, scaling causes Cell-Life to focus on short term, implementation oriented activities. During the second phase of systemic feedback, further scaling forces Cell-Life to think about-, and act on how to make their organisation more sustainable. When organisations don't respond to this urge or feedback to work with long-term sustainable management, they might get stuck in the first feedback phase. Organisations can become short-sighted, caught up in 'cycles of delivery and implementation' and may neglect long term strategic product- and process management. On the one hand the first phase of feedback is important and of positive influence on scaling. On the other hand it might negatively affect scaling when organisations get stuck in this short-term focus and neglect to deal with the second phase of systemic feedback. A balance between these two focuses therefore seems crucial for scaling (hypothesis d) (figure 19).

6.3 Processes of organisational closure

The case studies have taught us that knowledge of organisational functionalities and dysfunctionalities are often gained through learning by doing. The case studies also reveal that the level of organisational closure competence depends on the organisations ability to enforce and act on this knowledge; the extend in which organisations deliberately supress dysfuncionalities or trigger functionalities. This requires human- and financial resources.

These resources were clearly not in place within Cell-Life. Many times Cell-Life simply didn't have the financial and human resources to exploit the full potential of their technologies through organisational closure. The examples presented in table 4 clarify that organisational closure didn't really contribute to the scaling of iDART, but did stimulate the scaling of M-PESA. This thesis proposes that closure competence is a stimulating factor that stirs scaling (hypotheses e) (figure 20).

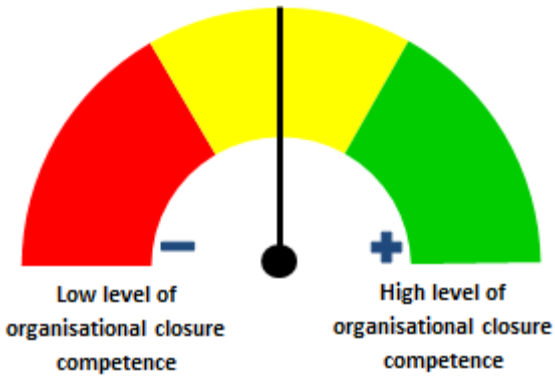


Figure 20 – Control panel E: low vs. high levels of closure competence

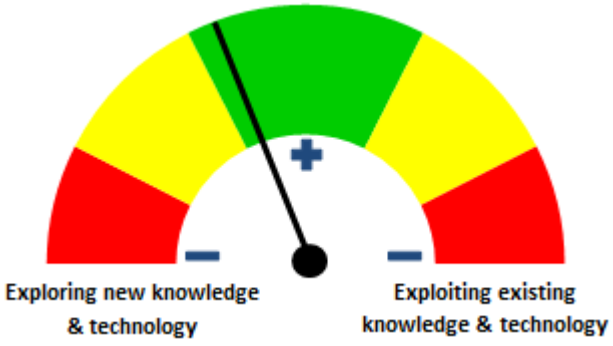


Figure 21 – Control panel F: exploring new- vs. exploiting existing knowledge and technology

From M-PESA's scaling trajectory it can be observed that M-PESA really benefited from Vodafone's and Safaricom's efforts to explore new knowledge and technology. Vodafone and Safaricom gradually investigated and enabled new applications that could be offered alongside M-PESA's main service of sending 'mobile remittances' (2007: Pay Bill, 2008: Bulk Payment, 2010: M-KESHO and Buy Goods).

Cell-Life's efforts of exploring new knowledge and technology on the other hand, did more harm than

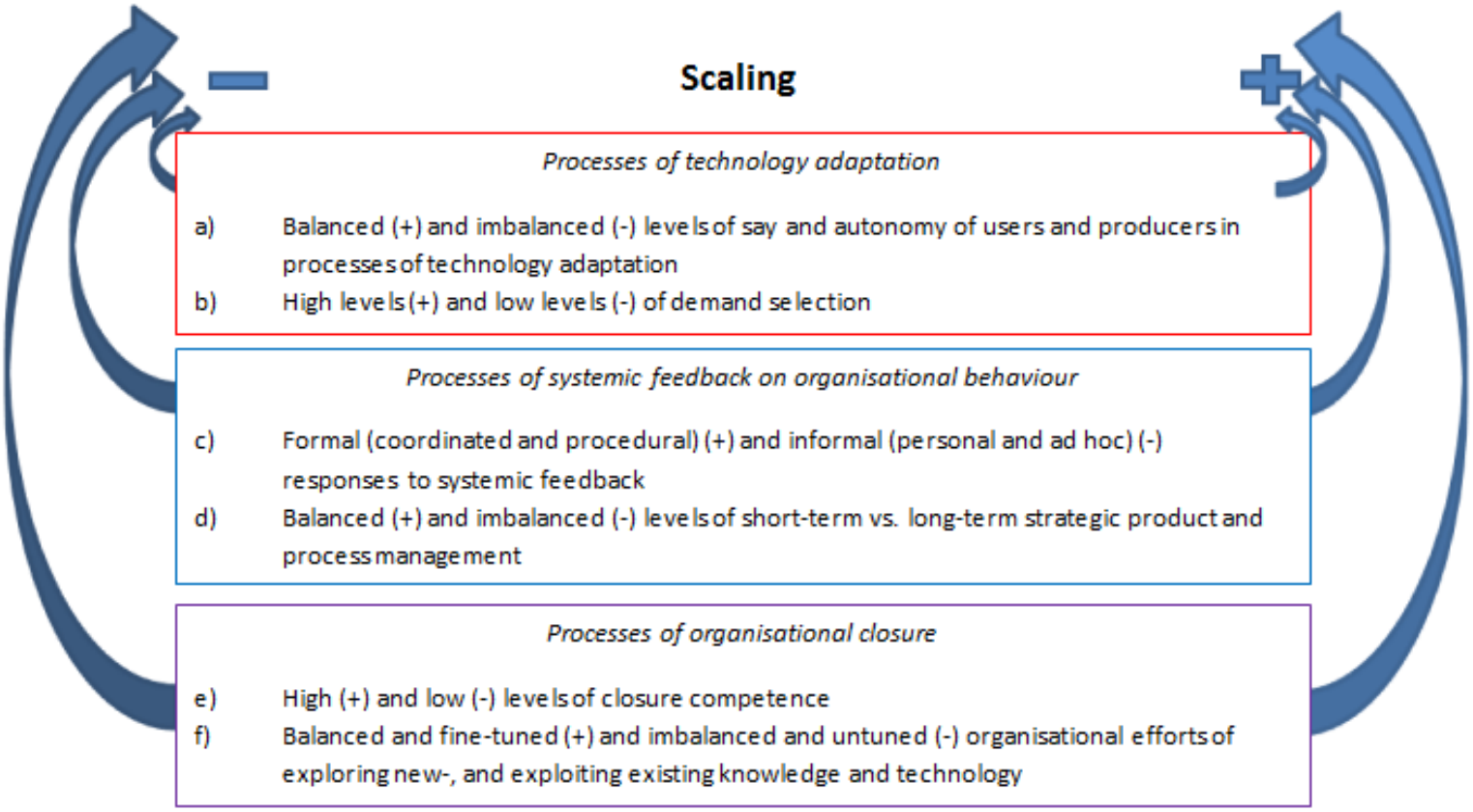
Table 4 – Cross-case analysis: distilling distinctive scaling processes and conditions

	Distinctive scaling processes and conditions	Examples from the M-PESA case study	Examples from the iDART case study
Processes of technology adaptation	Ai) Strong guidance of firms on users limit their ability to sense the potential of user innovations. Aii) Hypersensitivity to user needs leads to fast stabilisation of technology. B) The more organisational demand selection, the more the architecture of a technology will develop and stabilise in a way that is appropriate for the selected demand.	Ai) M-PESA's first target group consisted of Faulu and their clients. Vodafone then discovered several promising user innovations and unaddressed user demands. They altered their course, and served these upcoming demands. B) Vodafone addressed the discovered user demands one by one, through delivery of differentiated products and services (Pay Bill, 2007 - Bulk Payments, 2008 -M-KESHO, 2010, Buy Goods, 2010).	Aii) During its time of implementation, iDART was continually adapted in order to meet requirements of users and funding projects. iDART evolved into an unstable architecture that hindered further development. iDART became unsuited for general pharmacy dispensing. B) For a long time, Cell-Life didn't strategically select demand. It selected their projects on the basis of profit maximisation.
Processes of systemic feedback	C) As organisations scale, personal & ad hoc responses to systemic feedback are replaced or 'formalised into' procedural & coordinated responses. Di) Scaling requires firms to focus on implementation and short term management. Dii) Further scaling requires firms to focus on long-term management and sustainable business development.	C) The setting up of M-PESA agent stores and the quality assurance over their service are activities that over time have been outsourced to-, and coordinated by 'Top Image'. Dii) Vodafone followed a technology roadmap. They seem to have strategic planned and managed the coming of new applications alongside M-PESA.	C) Over time, Cell-Life differentiated its internal tasks (business analysts, product managers, a support team) and improved its efficiency through initiation of more processes and procedures (regarding project management, customer support, quality assurance, R&D, HRM). Di) Cell-Life operations were led by project requirements and product/service delivery. It brought them into a 'cycle of delivery' which supressed doing long term product management. Dii) For a long time, Cell-Life didn't work on sustainable business development.
Processes of organisational closure	E) The closure competence of firms determines the extend in which they: (i) trigger their functional mechanisms (ii) suppress their dysfunctional mechanisms. F) Balancing and fine-tuning effort to find new-, and exploit existing knowledge and technology	Ei) Vodafone provided customers with certain advantages (enabling: free deposits, deposits to non-customers, a simple pricing system) Eii) Vodafone dealt with: (a) distrust of M-PESA agents to pay out cash withdrawals, by providing them second 'M-PESA tails', (b) the SIM to SIMEX transition, by replacing these themselves and fully absorbing its costs, (c) the issue of float balancing, by developing new liquidity management methods. (F) Over time, Vodafone gradually researched, implemented and scaled up new technology opportunities alongside M-PESA.	Ei) Cell-Life initiated mHealth Solutions and shifted their operations from customisable product delivery to differentiated product delivery. Eii) Cell-Life deliberately surprised dysfunctionalities by: (a) Developing a web-based version of iDART and integrating iDART with other pharmacy and medical record systems. (b) hiring a consultant to help with restructuring the organisation. (F) Except for the last three years, there was hardly any strategic coordination in the research-, development- and implementation activities of Cell-Life.

good to the development and scaling of iDART. Cell-Life worked on many (new) projects at the same time. Because of this, iDART suffered a lack of strategic focus.

An important prerequisite for the beneficence of organisational closure seems therefore to be the ability of an organisation to balance and fine-tune their efforts of exploring new-, and exploiting existing knowledge and technologies (figure 21) (hypotheses f).

Table 5 – Overview of the hypothesised causailities with scaling of user-driven technological change



7. Conclusion and Discussion

What has become clear throughout this thesis is that scaling is in essence a mixture of processes. Scaling is a gradual and emerging outcome of different processes that can't be explained by one theory, but must be interpreted using multiple theories. The three theoretical processes that have been used in this thesis to make sense of the phenomenon of scaling will first evaluated and weighted. Is there on process that best explains and defines the phenomenon of scaling? Thereafter, the implications of this thesis for further research and for organisational practice will be discussed.

In the introduction of this thesis it was proposed that scaling is affected by certain distinctive processes present in the interaction between top-down and bottom-up directions of technological change (figure 1). The results of this thesis only partly confirm this. It was found that interaction between users- and producer networks do affect scaling through interactive processes of technology adaptation. Interactive adaptation of technology cause a technological 'flexibility-stability evolution'; a stabilisation process of technology. The case studies illustrate that such stabilisation processes impact processes of scaling. Although processes of systemic feedback and organisational closure also affect scaling, these processes do not come forth from the interaction between user- and producer networks. The cases reveal that systemic feedback effects come forth from a process of scaling and organisational growth itself. And although interactions between user- and producer networks might coincidentally trigger organisations to execute a process of organisational closure, this process does not arise spontaneously and ultimately but emerges because organisations actively executes a closure process.

Seelos and Mair (2010) have introduced 'organisational closure' as the main theoretical process that defines and explains the phenomenon of scaling. Based on the results of the case studies, two remarks can be made their work. First of a al the organisational closure perspective was found to be useful for retrospective analysis of two scaling processes that occurred within two different sectors; the health care sector and the financial sector. However, the 'closure theory' provides too superficial understandings to support organisations to prospectively interpret and analyse scaling. Secondly, when analysing processes of organisational closure in the case studies it was noticed that the two technology-producing organisations did not intentionally work on 'triggering their functionalities and suppressing their 'dysfuncionalities'. They stumbled upon them and then tried to deal with them in a most profitable way through learning by doing. Organisational closure processes occurred in both cases when the technology-producing organisations responded to processes of 'systemic feedback' and 'technology adaptation'. So rather than a stand-alone process that stimulates scaling, this thesis concludes that organisational closure intertwines with these two other processes and thereby positively impacts scaling. Organisational closure can be seen as a meta-theory of scaling that interprets and directs two processes that affect scaling, namely: processes of (i) technology stabilisation and (ii) systemic feedback.

This former is a result of processes of interactive technology adaptation. Technology stabilisation is however not always constructive for scaling and does not per definition cause scaling. The iDART case study illustrates that the stabilisation process can also lead to development of unstable architectures of technology with the result that these technologies are not functioning optimally and are not scalable. The organisational closure theory provides an interpretation of how stabilisation processes of technology can be guided and aligned by firms in such way that the stabilisation leads towards scalable technologies. The same goes for processes of systemic feedback. The organisational closure theory gives interpretation how feedback effects, that emerge when organisations grow and scale their operations, can be dealt with in such a way that enables

continuous scaling.

So this thesis concludes that besides organisational closure, scaling can be defined and explained by processes of systemic feedback and technology adaptation. It has further worked out the organisational closure theory of Seelos and Mair (2010) by giving it more substance. Figure 22 gives an overview of the three different theoretical processes and how they are positioned in relation to scaling.

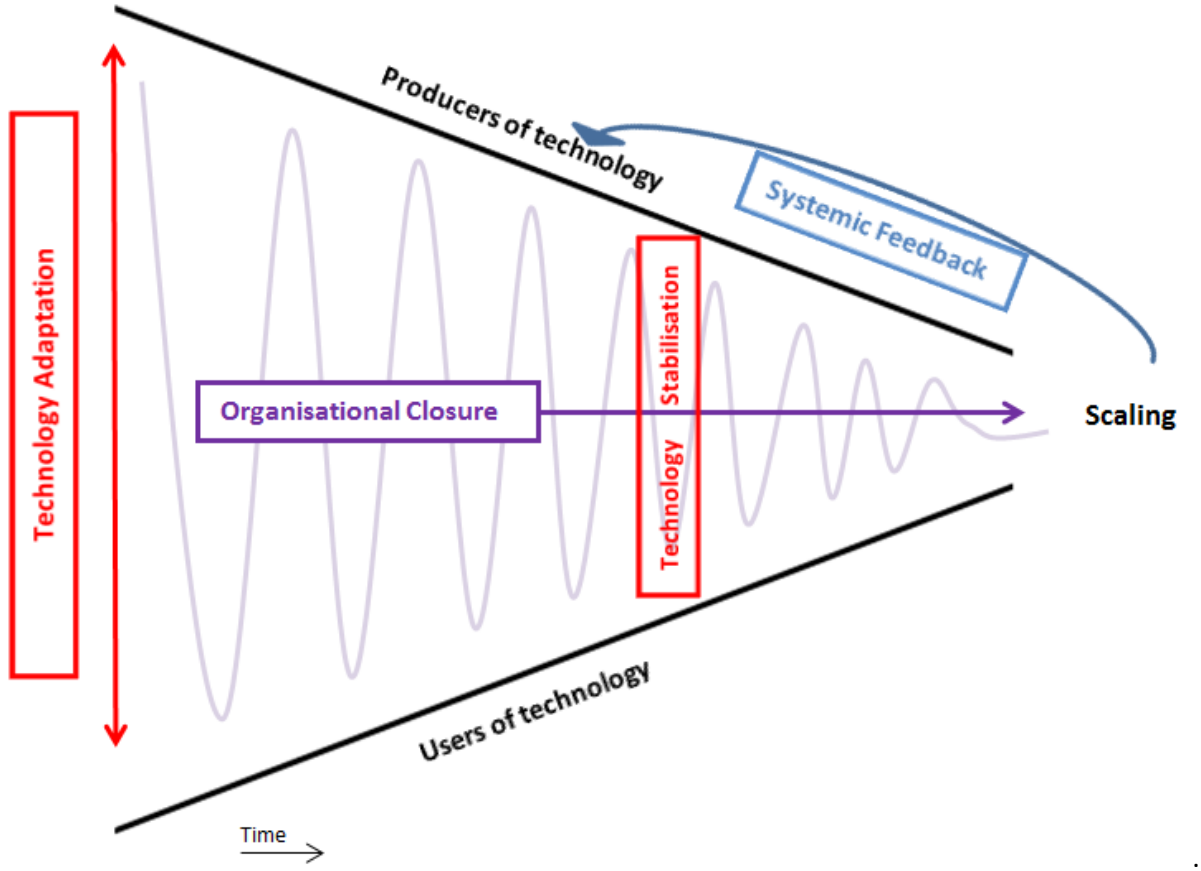


Figure 22 – Positioning of the different theorethical processes in relation to scaling

Because scaling is an emerging outcome of different processes, it is a hard and complex phenomenon to predict. It is not only hard to predict whether scaling will occur or not, it is also difficult what exactly determines 'scaling up' and 'scaling out' (the different forms of scaling introduced in chapter 2). When firms connect with user-driven technological change, when new ideas and adaptations merge with existing technologies, these types of scaling intertwine and therefor occurs 'adaptively'. This thesis found that such 'adaptive scaling' (AS) is an emergent outcome of a process of organisational closure (C), that intertwine with-, give interpretation to-, and directs processes of technology adaptation (TA) and systemic feedback (SF). Future research might first of all try to qualitatively determine if there are 'other important and influential processes' that impact scaling and that can be added to the formula: AS= C * (TA + SF + ?). The case studies give some indication that there may be other interactions than user-producer interactions that might be influential. Effects of 'culture-producer interactions' or 'politics-producer interactions' might be worth looking into. Secondly, and even more challenging, future research might try to quantitatively test the proposed hypothesis (*a* to *f*) of scaling (table 5). Empirically testing the proposed causal relationships with scaling seems however only relevant after more explorative research has been performed on what other processes are part of scaling.

Although scaling is an emerging outcome that is hard to predict, this thesis provides businesses and organisations the following understandings and tools for prospective analysis of scaling. Six 'control panels' have been presented that give organisations more in depth insight in what kind of processes exactly affect scaling. In order to scale, this thesis stresses organisations the importance of: (A) balanced levels of sensitivity to user needs, (B) high levels of demand selection, (C) procedural and coordinated responses to processes of systemic feedback, (D) balanced and fine-tuned levels of short vs. long term management, (E) high levels of organisational closure competence and (F) balanced levels of exploring new-, and exploiting existing knowledge and technology. So the control panels help organisations to indicate 'what scaling processes to monitor' and 'whether these processes positively/negatively affect scaling'. Businesses can experiment with these six control in order to find out what control panels need alteration in order to scale their products and/or services.

In sum, this thesis has contributed to 'a theory of scaling' by providing more substance to the organisational closure (meta-)theory of scaling. It has also provided organisations tools that may be used for prospective analysis of scaling.

References

Boor, P. van den, Oliveira, P. and Veloso, F.M. (2013). Unusual Suspects? Innovation by users in developing countries: evidence from mobile banking Services. *Working paper series*, 1-41. Retrieved from: http://dx.doi.org/10.2139/ssrn.2009926 [on 1 October 2013]

Berg, J. van den, Vellema, S., Wigboldus, S.A., Lee, J. van der, Giani, A., Wolf, P.L. de, and Potters, J.I. (2012). How are innovations in the agrofood sector going to scale? Towards a methodology for comparative analysis of transitions towards sustainable food provision in Europe and Africa. *Centre for Development Innovation*, Wageningen UR (University & Research Centre), Wageningen.

Bradley, E.H., Curry, L.A., Taylor, L.A., Wood Pallas, S., Talbert-Slagle, K., Yuan, C., Fox, A., Minhas, D., Ciccone, D.K., Berg, D. and Pérez-Escamilla, R. (2012). A model for scale up of family health innovations in low-income and middle-income settings: a mixed methods study. *BMJ Open,* retrieved from: http://dx.doi.org/10.1136/bmjopen-2012-000987 [on 7 October 2013]

Cell-Life. (n.d.). *Capabilities Document: Cell-Life's background, how iDART, C4H and EMIT work, and what services are offered as part of Cell-Life's technological solutions*. Cell-Life publication, 1-14. Retrieved from: http://allafrica.com/download/resource/main/main/idatcs/00021001:484d034 11543547c492b96c41e bec99f.pdf [on 17 January 2014].

Cell-Life. (2008). *iDART: The Intelligent Dispensing of Antiretroviral Treatment.* Cell-Life publication, 1-13. Retrieved from: http://unpan1.un.org/intradoc/groups/public/documents/unpan/unpan037372.pdf [on 17 January 2014].

Cell-Life. (2014). *iDART: iDART in Action*. Retrieved from: http://www.cell-life.org/systems/dispense-idart [Accessed on January 17 2014].

Chevrollier, N., Klein, W. van den, and Collée, L. (2013). *Co-creating BoP ventures: A new form of interaction to develop shared value among equal stakeholders.* Base of the Pyramid Innovation Center (BoPInc), Utrecht.

Cobert, B., Helms, B. and Parker, D. (2012). Mobile Money: Getting to scale in emerging markets. *McKinsey Quarterly,* February, 1-13.

Cozzens, S. and Sutz, J. (2012). *Innovation in informal settings: a research agenda*. Ottawa: IDRC.

Danse, M. and Vellema, S. (2007). Small-scale farmer access to international agri-food Chains: A BOP-based reflection on the need for socially embedded innovation in the coffee and flower sector. *Greener Management International*, (51), 39-52.

Fleck, J. (1989). The development of information-integration: beyond CIM?. In: J. Ranta (Ed.), *Trends and Impacts of Computer Integrated Manufacturing* (pp. 721–745), IISA, Laxenburg, Austria.

Foster, C. & Heeks, R. (2013a). Analysing policy for inclusive innovation: the mobile sector and base-of-the-pyramid markets in Kenya. *Innovation and Development, 3*(1), 103-119.

Foster, C. & Heeks, R. (2013b). Conceptualising Inclusive Innovation: Modifying Systems of Innovation frameworks to understand diffusion of new technology to low-income consumers. *European Journal of Development Research*, *25*(3)*,* 333-355.

Foster, C. (2013). Nurturing user-producer interaction: Innovation flows in a low income mobile phone market. *Workshop paper.* Retrieved from: https://www.google.nl/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CC8QFjAA&url=http%3A%2F%2Fwww.cdi.manchester.ac.uk%2Fnewsandevents%2Fdocuments%2FFosterPreWorkshopPaper.doc&ei=lsVWUs2YGoqZtQaqxICQCA&usg=AFQjCNE33NFKRxrPCpCav3RJGpJyzSWWDg&sig2=UO8wDxYmpS-zplDEK8Rrnw&bvm=bv.53899372,d.Yms [on 1 October 2013]

Gebregziabher, S. A. (2011). *The "Bi-Directional" Influence between technology and society: How M-PESA is shaping and being shaped by society in Kenya* (Master's Dissertation). Retrieved from: http://upetd.up.ac.za/thesis/available/etd-02172012-172140/

Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. *Research Policy,* 33(6-7), 897-920.

Heiskanen, E., Hysalo, S., Kotro, T. & Repo, P. (2010). Constructing innovative users and user-inclusive innovation communities. *Technology Analysis & Strategic Management*, *22*(4)*,* 495-511.

Hughes, N. and Lonie, S. (2007). M-PESA: Mobile Money for the "unbanked" turning cell phones into 24-hour tellers in Kenya. *Innovations*, 2(1-2), 63-81

Hippel, v. E. (1994). Sticky Information and the locus of problem solving: Implications for innovation. *Management Science, 40*(4), 429–439.

Kaplinksy, R. (2011). Schumacher meets Schumpeter: Appropriate technology below the radar, *Research Policy, 40*(2)*,* 193-203.

Klein, v.d. W., Chevrollier, N., and Collée, L. (2012a). *Inclusive Innovation: Shared value at the Base of the Pyramid.* Base of the Pyramid Innovation Center (BoPInc), Utrecht.

Klein, v.d. W., Mancheron, H., Wertheim-Heck., and Collée, L. (2012b). *BoP Insights: Inclusive marketing research*. Base of the Pyramid Innovation Center (BoPInc), Utrecht.

Klein, v.d. W., Sprenger, T., Collée, L., Brouwer, H., Tubergen, v. K., and Bults, R. (2013). *Developing BoP partnerships: towards collective impact at the Base of the Pyramid*. Base of the Pyramid Innovation Center (BoPInc), Utrecht.

Lallana, E. (2004). SMS in business and government in the Philippines. *ICT4D Monograph Series.* Retrieved from: http://unpan1.un.org/intradoc/groups/public/documents/other/unpan024834.pdf [on 25 October 2013]

Lee, A. and McNamee, R. (2013, June 17-19). *In Search of a Theoretical Framework for Reverse Innovation.* Paper presented at the 35th Druid Celebration Conference, Barcelona, Spain.

Loundon, M., Rivett, U. (2011). Enacting Openness in ICT4D Research. *Information Technologies & International Development*, 7(1), 33-46

Lundvall, B.A. (1988). Innovation as an interactive process: From User-Producer Interaction to the National System of Innovation. In G. Dosi, C. Freeman, R. R. Nelson, et al. (eds), *Technical Change and Economic Theory* (pp. 349–370), Pinter, London, UK.

Mas, I. and Radcliffe, D. (2011). Mobile payments go viral: M-PESA in Kenya. *The Capco Institute Journal of Financial Transformation*, 32: 169-182. Retrieved from: http://ssrn.com/abstract=1593388 [on 20 November 2013]

Mendoza, R. U. & Thelen, N. (2008). Innovation to make markets more inclusive for the poor. *Development policy review*, *26(4)*, 427-458.

Nahuis, R., Moors, E. H. M. and Smits, R. E. H. M. (2012). User producer interaction in context. *Technological Forecasting & Social Change*, *79(6)*, 1121–1134.

Paina, L. and Peters, D. H. (2012). Understanding pathways for scaling up health services through the lens of complex adaptive systems. *Health Policy and Planning,* 27, 365-373.

Paunov, C. (2013). Innovation and inclusive development: A discussion of the main policy issues., *Technology and Industry Working Papers*, OECD Science. Retrieved from: http://www10.iadb.org/intal/intalcdi/PE/2013/11458.pdf [on 23 October 2013]

Prahalad, C. K. (2012). Bottom of the Pyramid as a source of breakthrough innovations. *Journal of Product Innovation Management, 29(*1), 6-12.

Rangan, V.K., Chu, M. and Petkoski , D. (2011). Segmenting the Base of the Pyramid. *Harvard Business Review*, June.

Rivett, U. and Tapson, J. (2009). The Cell-Life Project: Converging technologies in the context of HIV/AIDS. *International Journal of Community Research and Engagement*, 2, 82-97.

Safaricom. (2011). *M-PESA customer and agent numbers.* Safaricom publication. retrieved from: http://www.safaricom.co.ke/personal/m-pesa/m-pesa-resource-centre/statistics [on 17 October 2014].

Schumpeter, J. A. (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle*. Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=1496199 [on 25 October 2013]

Seelos, C. & Mair, J. (2010). Organizational mechanisms of inclusive growth: A critical realist perspective on scaling. *IESE Business School Working Paper,* 1-24. Retrieved from: http://dx.doi.org/10.2139/ssrn.1550719 [on 7 October]

Seelos, C. and Mair, J. (2013). Innovate and scale: A tough balancing act. *Stanford Social Innovation Review* (in press).

Smits, R. (2002). Innovation studies in the 21st century: Questions from a user's perspective. *Technological Forecasting & Social Change*, *69*(9), 861–883.

Sone, L. (2012). Innovative initiatives supporting inclusive innovation in India: Social business incubation and micro venture capital. *Technological Forecasting & Social Change*, *79(4),* 638–647.

Statistics South Africa. (2013). *Mid-year Populations Estimates: Annual report 2013*. Retrieved from http://beta2.statssa.gov.za/publications/P0302/P03022013.pdf [on 20 February 2014]

Suurs, R., Diaz Lopez, F. J., Boer, J. de, Miedema, M., Kamp, L. de, and Mashelkar, R. A. (2013). A Systems-based model for the successful scaling up of sustainable innovation at the Bottom of the Pyramid, *Workshop paper.* Retrieved from: https://www.google.nl/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&ved=0CCwQFjAA&url=http%3A%2F%2Fwww.cdi.manchester.ac.uk%2Fnewsandevents%2FSuursEtAlPreWorkshopPaper.docx&ei=ycdWUvDwJdHHswbhrYHABA&usg=AFQjCNFX2jgdWqayYvxLTZh4qqi6xiAufg&sig2=NmbMIWbCcDAIFgQ0-ZuRmA&bvm=bv.53899372,d.Yms [on 1 October 2013]

Tolly, J. and Alexander, H. (2009, April 1-2). *Innovative use of cellphone technology for HIV/AIDS behaviour change communications: 3 pilot projects*. Paper presented at the W3C workshop on "The Africa perspective on the role of mobile technologies in fostering social development", Maputo, Mozambique. Retrieved from: http://www.w3.org/2008/10/MW4D_WS/papers/kdetolly.pdf [on 20 February 2014]

Vaughan, P., Fengler, W. and Joseph, M. (2013). Scaling up through disruptive business models: The inside story of mobile money in Kenya. In Chandy, L. Hosono, A., Kharas, H. and Linn, J. (Eds.), *Getting to Scale: How to Bring Development Solutions to millions of Poor People* (pp. 154-181). Washington: The Brookings Institution.

Wessels, X., Nattrass, N., Rivett, U. (2007). Improving the efficiency of monitoring adherence to antiretroviral

therapy: A case study of the introduction of electronic technologies in Gugulethu, South Africa. *Development Southern Africa,* 8(4), 607-621

Williams, L. D. A. and Woodson, T. S. (2012). The future of Innovation Studies in less economically developed countries. *Minerva,* 50(2), 221-237.

Wood, R., Bekker, L., Brown, S., Rivet, U. (2008). The utility of pharmacy dispensing data for ART programme evaluation and early identification of patient loss to follow-up. The Southern African Journal of HIV Medicine, Autumn, 44-48.

Woodhill, J., Guijt, J., Wegner, L., and Sopov, M. (2012). From islands of success to seas of change: a report on scaling inclusive agri-food markets. *Centre for Development Innovation*, Wageningen UR (University & Research Centre), Wageningen.

Appendix A - Overview of interviewees IDART Case-Study

Name:	Date of interview:	Employee of:	Function:	Over the period:
Sikander Noor Mahomed	January 21,2014	Cell-Life	Product Manager (iDART) New Business Development for mHealth Solutions	July 2009 - present January 2011 - present
Sarah Brown	February 17, 2014	Cell-Life	Software Developer Business Analyst	2005 - 2006 2006 - 2011
Simon Kelly	February 13, 2014	Cell-Life	Software Developer CTO Cell-Life	January 2006 - April 2007 April 2007 - March 2012
Kieran Sharpey-Schafer	February 13, 2014	Cell-Life	Mobile Team Leader	March 2006 - April 2010
Pamela Jacobs Raymond Eugene Brant	February 17, 2014	Hannan-Crusaid Treatment Centre	Pharmacist	2005 - present unknown - present unknown - present
Heli Moeng	February 17, 2014	Hannan-Crusaid Treatment Centre	Pharmacist	March 2005 - December 2007
Emeritus Professor Robin Woods	January 24, 2014	Desmond Tutu HIV Foundation	CEO of the Desmond Tutu HIV Foundation	1990 - present
Nicola Kelly	January 24, 2014	Desmond Tutu HIV Foundation	Pharmacists Researcher	Unknown - present
Doctor Richard Kaplan	January 24, 2014	Desmond Tutu HIV Foundation	Doctor and working with DTHF for Monitoring & Evaluation purposes	Unknown - present

Appendix B - Overview of the constructive events of the scaling trajectory of iDART

Stage	Organisational events	Technological events
Initiation of a research group: Cell-Life (2000-2001)	• Visionary discussion between Ulrike Rivett and Jon Tapson (UCT) • Networking and initiation of a Cell-Life research team, with students and researchers from UCT, CPUT, TAC, DTHF.	• Rough ideas are made about using mobile technology and internet to monitor adherence of HIV+ patients to ARV treatment.
Cell-Life going public: First funding and first prototypes (2002-2005)	• Cell-Life is present at the 2002 World Summit on Sustainable Development. It receives funding from Vodafone for implementation of 'Adherence'. • The Hannan-Crusaid and Masiphumelele clinic are hitting a 'treatment ceiling', the DTHF decides to initiate a 'down-referral' dispensing model. iDART is designed with the aim to assist the monitoring of ARV stock and support the core functions of ARV dispensing to HIV+ patients. • Further donations from Vodacom enable Cell-Life to register as a not-for-profit organisation. In 2005, Cell-Life swaps it's CPUT offices for new offices at UCT.	• 'Adherence' is a first prototype of a data capturing mobile application (later developed into 'Emit' and rebranded into 'Capture'). Using a simple WIG (Wireless Internet Gateway) menu loaded on a mobile phone, patient information is collected (by therapeutic counsellors from the DTHF) and via SMS send to administration offices that synchronize the data onto PC's, where doctors can access it. • A first prototype of iDART is made in 'Visio Basic' by Stephen Rausch, an undergraduate honours student, who works on iDART for 4 months for his thesis. During 2004 the prototype is already used by the HCTC.
Cell-Life in a period of first growth (2006-2009)	• In 2006, staff grows from 5 to 9 people, while 6 postgraduates and honours students are part of the team. In 2009 staff grows up to 25 people. • Cell-Life starts to work and interact with users in a more early stage of (software) development. • iDART is implemented in around 35 clinics. • Emit is developed and gathers momentum, many Emit projects are coming in. • From 2009 onwards, Cell-Life researches the opportunities of communication services for HIV health care (under the name: 'Cellphones4HIV'). • A consultant joins Cell-Life for a few weeks to work on (re)structuring the organisation. • Cell-Life's focus shifts gradually away from research and development to implementation and project management.	• Sarah Brown rewrites iDART (from Visio Basic) into Java. • The down-referral model is not often applied anymore, clinics start to dispense ARVs directly to patients. iDART is adapted for direct dispensing. • User feedback over time leads to a number of extensions that are designed in interaction with users and integrated into iDART. Almost all software development takes place in-house. • Adherence is developed into Emit. Data is now captured by mobile phones using a Java application and loaded onto a central database in real time using mobile internet. The system is made more customisable, enabling different kinds of data to be captured, when clients requested so. • The architecture of iDART is found to be like a 'Jenga Tower'; too instable and inflexible. Two times Cell-Life tries to rewrite iDART from scratch, both attempts fail.
Major scaling of Cell-Life's implementing activities (2010-2014)	• mHealth solutions is formed as Cell-Life's wholly-owned commercial subsidiary. • For iDART, Cell-Life starts partnering with some organisations on the implementation side (Anova, BroadReach, technical teams operating abroad). • Cell-Life grows more mature and becomes a more professional and corporate organisation with more efficient processes. • iDART is now implemented at over 100 clinics. • Cell-Life is trying to get rid of external communication and an internal structure around product lines (iDART, Capture and Communicate), and aims to become more (mHealth) solution driven.	• The Cellphones4HIV program gets momentum, and from 2010 onwards Cell Life starts selling communication services under a new 'Communicate' product line. • For a research project for the Canadian International Development Agency, Cell-Life enables iDART to interact with Communicate. It sends automatic clinic appointment reminders and daily positive living SMSs to (randomized selected) patients. • From scratch, Cell-Life develops a new server for their mobile data capturing application. Emit is rebranded into 'Capture'. • Cell-Life is moving iDART away from a standalone system, to a web-based system which enables the use of a centralised database. This version is now in the Quality Assessment phase, tests are planned for end April 2014.

Appendix C - (Incomplete) Overview of iDART implementation sites

(Data obtained from Cell-Life, n.d; Cell-Life, 2008 and Cell-Life, 2014)

Timeframe:	Clients:	Implementation sites:
March 2005 - January 2006	Desmond Tutu HIV Foundation (DTHF)	4 sites in the Western Cape (Hannan Crusaid Treatment Center, Crossroads Research Site, UCT Medical School-central pharmacy, Masiphumelele Treatment site)
2006 -2008	Department of Health (DoH) of West Province and Northern Cape, funding provided by the Elton John AIDS Foundation.	5 sites: 4 in the North West Province and 1 in the Northern Cape (Tuang treatment centre, Ganyesa Hospital, Klerksdorp paediatrics Hospital, Tshepong Wellness centre and Galeshewe Day Clinic)
June 2008 - January 2009	Anova Health Institute	11 Sites in Gauteng, Western Cape, Kimpopo and Mpumalanga.
May 2008 - August 2009	Wits Reproductive Health and HIV institute (WRHI)	6 sites in Johannesburg (Johannesburg General Hospital, Hillbrow Hospital, South Rand Hospital, Tambo Memorial Hospital)
During 2009	Eastern Cape Department of Health with funding from Vodacom Foundation	5 sites in the Western Cape, and 1 in the Eastern Cape (Zithulele Hospital, Molteno Hospital, IKwezi Wellness clinic, Canzibe Hospital and G.F Jooste Hospital)
September 2009	CAPRISA	2 sites in KwaZulu-Natal (eThekwini Clinical Research Site, Vulindlela Clinic)
January 2010 - June 2011	Maternal, Adolescent and Child Health (MatCH)	4 sites in KwaZulu Natal (R.K. Khan's Hospital, Charles James Hospital, Cato Manor Hospital, Kwa Mashu Community Health Centre)
March 2010 - June 2011 September 2010	The DoH of the different provinces. Funding provided by Vodacom Foundation.	8 sites in Eastern Cape (Cofimvaba Hospital, Thornhill CHC, Mt Ayliff, Port Alfred Hospital, Glen Grey Hospital ,Andries Vosloo Hospital, Madzikane Ka Zulu Memorial Hospital, Bisho Hospital) 6 more sites in the Eastern Cape (Aliwal North Hospital, Burgersdorp Hospital, Cloete Joubert Hospital, Empilisweni Hospital, Umlamli Hospital, and Taylor Bequest Hospital) 6 sites at Kwazulu-Natal Natal Umzinyathi District (Church of Scotland Hospital, Dundee Hospital, Charles Johnson Memorial Hospital, Greytown Hospital, HLabisa Hospital,Kwa Msane Clinic)
December 2010	South African Clothing and Textiles Workers Union (SACTWU)	1 site in Durban
January 2011	Ubuntu Education Fund	1 site in the Eastern Cape (Ubuntu Centre at Zwide Township)
November 2011	The Institute of Human Virology (IVHN)	31 sites at Nasarawa District, Nigeria.
July 2013	Médicins Sans Frontières	1 (pilot) site in Moputo, Mozambique (Da Maio Health Centre)