4 Strategies for Scaling of Integrated Health Information Architectures

In the earlier chapters, we presented concepts, principles and components of an IHIA, as a framework within which HIS may be designed, developed and implemented. Building on this discussion, here, in this chapter we address the next step – strategies, whereby these systems and architectures may be spread out and taken into use, in a country or a state, successively in all districts in a state, in all facilities in a district, and in all hospitals in a state or country, covering different user groups and technologies. We call this scaling.

The chapter is divided into four key sections. In the first, we describe the principles for scaling of IHIAs, arguing that this is more about the scaling of design principles than of 'boxes.' Since the IHIA is not 'fixed' and pre-defined at the outset, it needs to expand and develop organically in a paced manner, according to the local context and the opportunities available. The design principles outlined will help to understand how this scaling challenge can be approached.

The second section, describes the dimensions of scaling. Conceptually, we distinguish between vertical and horizontal processes. Vertically is 'down' the health hierarchy, from national and state to district, and the various sub-district levels of the health facilities and communities, and finally to the level of the patient or a beneficiary of a particular health programme. These different levels are associated with different granularities of data and information needs, and also the comprehensiveness of the solutions provided. The horizontal processes refers to the 'scope' of services and functional areas, included in the IHIA, as well as, the geographical scope of coverage; for example, from one pilot hospital, or pilot districts to all hospitals and districts in the state, or successively to include new programmes. Horizontal processes also include the incorporation of different technologies, as the IHIA increases in scope and functionalities.

The third section, describes the particular challenges of scaling ICT based solutions in developing countries, where the technical infrastructure and human resources are unevenly developed across various dimensions, such as urban and rural, rich and poor, politically advantaged and marginalised, and well funded and neglected health programmes. Such dichotomies, which are inherent in the context of developing countries, in general, and the health system in particular, makes scaling of uniform ICT solutions impossible and emphasises the need for context sensitive scaling approaches.

Finally, the fourth section focuses on strategies for scaling, describing the importance of creating attractors, for change as part of an evolutionary approach. This involves the creation of initial success and its demonstration, in order to rally scarce resources and political support for the process, and thereby to create more success and strengthen the attractor further. Case examples are provided to demonstrate the practical benefits of such an approach.

4.1 Scaling of IHIA – Scaling of Design Principles, Not Boxes

Scaling is often conceptualised in a rather static and stand-alone manner, concerning how a system used in one location can be replicated in other settings. We see such an approach to focus on the 'boxes' or the technologies being used, rather than the underlying processes through which such circulation takes place or not. Our focus in conceptualising scaling, is on the processes underlying the spreading out and expanding of the IHIA around the different dimensions of technology and institutional use of information. This will include aspects of the three layers of the IHIA, such as organisational work practices; for example, training materials and approaches, datasets and data collection tools, software applications, and other components of the HIS including the technologies and the capacities to use them. Scaling is generally used to denote the quantitative and mechanical expansion of the system in scope, size and number of users. However, we will, in line with our three levels IHIA, distinguish between the quantitative, structural and mechanical dimensions on the one hand and the more people centred qualitative, process-oriented and social system dimensions of scaling, on the other hand.

Quantitative, Mechanical, Structural Dimensions of Scaling: Here, we not only include the traditional expansion of the system in scope, size and number of users, but also the expansion of the system down in the hierarchy of the health system. Also, the inclusion of increasing number of health programmes and systems, in granularity of data, and in comprehensiveness of the architecture and systems taken into use.

Qualitative, Social System, Process-Oriented Dimension of Scaling: Here, we include the increase and expansion of quality of the use of information, the extent to which individual and institutional users are taking part in the development of the system, maturity of the users, and their feelings and sense of ownership to the system.

4.1.1 Cultivation of Social Systems and Improvisation

The concept of cultivation, in contrast to construction, denotes a way of shaping technology that is fundamentally different from rational planning, engineering methods and the construction of technology. A rational process will assume linearity in how system requirements are first frozen and then developed. In reality, requirements can never be frozen – they always change – and thus, need to be 'cultivated'. Cultivation is about interfering with, supporting and controlling natural organic processes that are in material; as the seeds sprout, they must be provided with proper cultivation; the soil must be prepared and the saplings cared for, nourished, watered and given appropriate sunlight. The term cultivation covers these processes of, and in our case, spreading the principles, tools and basic set-ups of the IHIA to a new place, and then help them to grow into place within the local social system of work practices and technologies, through processes of learning. The IHIA, may thus, be regarded as a socio-technical system, or an organism, with a life of its own, with its ability to learn and grow. The spread of technology is, therefore, better understood as a process of technology learning, rather than 'technology transfer', which metaphorically represents a 'design from nowhere'. Technology, like institutions, are also shaped through such processes of learning and growing into place. Methodologically, cultivation is characterised by incremental and evolutionary approaches, described in terms of 'piecemeal engineering'.

The 'bricoleur' or 'tinkerer' are alternatives to the rational engineer in shaping the design and change processes, following a cultivation approach. Typically, the engineer starts out by defining and specifying his or her goal, then choosing and constructing his tools in order to reach this goal. However, the bricoleur will take the possibilities given by the available tools, situation and other possessions, as the point of departure, and define the goals and way forward, according to that. The bricoleur will improvise out of the possibilities given, by what is in hand and the results might look like patchwork. Particular tools will be used for many purposes. The bricoleur thinks in association to concrete observable things, which helps to remember, while the engineer builds on abstract mathematics and technical drawings. The engineer builds abstractions on abstractions, where the bricoleur builds abstractions on observations. In developing the HIS and the IHIA, a bricolage approach allows and encourages tinkering by people close to the operational level, that is by combining and applying known tools and routines at hand to solve new problems. The value of tinkering lies in keeping the development, or 'cultivation' of the information system and the IHIA, close to the competencies of the organisation and its ongoing fluctuations in local practices. Evolutionary and participatory approaches have a bricolage approach, inherent by their flexibility and multiple ways to address cultivation and user participation. The bricoleur is the roadside 'bush' mechanic you meet in Africa who gets the car going by the way of 'magic' or by whatever is at hand in an improvised way. Due to contextual constraints, bricolage is particularly appropriate in developing countries, where local resources and potentials form the point of departure.

Given the extreme heterogeneity of the technical and human resource infrastructure in developing countries, scaling approaches need to be flexible and allow for different paces, according to the socio-technical and institutional contexts available. We argue against a perspective, which regards scaling as a property of complex systems, in which one part of the system reproduces the same structure and patterns that appear in other parts of the system. Broccoli is used as an example here, of scaling, in a natural system as branches and sub-branches have the same structure as the whole plant. This perspective on scaling is limited to the replication of the same, which implies in our context, a fully standardised package being replicated in a fully standardised context; consisting then of the same infrastructure, human resources situation, and sociocultural conditions across the various settings, in which scaling is taking place. Such replication is not possible.

4.1.2 Scaling and Poor Infrastructure

There are particular challenges related to the scaling of systems, in developing countries. The fact that technical infrastructure is very unevenly distributed between, for example, rural and urban areas, poses a particular challenge when scaling computer-based and networked solutions, for uniform implementation in a country. For example, currently while planning for the use of a General Packet Radio Service (GPRS) based mobile application in Himachal Pradesh in India, we had to consider, some of those remote districts that would not receive the GPRS connectivity, to then have the state application to cater to both GPRS and SMS based solutions. Similarly, when the HMN Technical Framework, advises countries to implement a central data repository; receiving data electronically from subsystems and the peripheral levels, it presupposes both uniform access to the Internet and that sub-systems are computer-based. This assumption is rarely valid, since data sources are often paper-based, as is the dominant tendency in Africa and even in various parts of India. Therefore, new ICT based solutions need to be established and interoperability created based on the available infrastructure, which ranges from a complete lack of power and connectivity through portable data systems, such as USB memory sticks all the way to mobile networks and the Internet. We give below, an example from a northern state in India.

Box 4.1 also emphasises the need to take a process perspective on scaling, rather than, treating it as a one time event of expansion. In this case, scaling decisions and outcomes were shaped by the institutionalisation of existing processes (such as the stability of existing reporting system), availability of infrastructure, and the intentions of the state to access disaggregated data. At another level, the hilly terrains of the state and the challenges of both physical and electronic access had an important bearing on shaping the scaling processes.

Box 4.1	Addressing the challenge of scaling in a context of uneven infrastructure

Uttarakhand – Scaling and Uneven Infrastructure

Uttarakhand is a state in Northern India, which has various districts nesting in the foothills of the Himalayas, making access, both physically and electronically, a challenge. The state started to implement DHIS2 for their state HMIS, in December 2008, initially based on a district model. This implied that each district would get their different facilities to submit on paper their monthly reports, to the district HMIS officer, who was then responsible for manually aggregating the report for the district, and then entering it into the DHIS2, which was deployed over the state server.

After about six months as the district-based process of reporting was stabilised, the state took the next step of scaling, by getting sub-district (called Block) based data. The list of all the blocks by districts was provided and these details were included in the organisation unit hierarchy in the database. Now, we had all the districts and within each, all the blocks were included. In the next step, each block now made a block consolidated report for the month (including all the data from the Primary Health Centres and Sub Centres in the block) on paper and sent it to the district office, where data was now entered electronically at the district level, but by blocks. The district consolidated report was generated electronically by the DHIS2, and in this way the state and district managers could drill down to the blocks, to identify performance issues.

Once this process was stabilised, the state took the bold decision to now scale the systems one step further: by now getting data disaggregated by particular facilities in the block. For this, it was important to do the data entry at the block itself, otherwise the load of data entry at the district, would be huge. However, while computers were available in the block level, internet connectivity did not allow online data entry at the block level. To deal with this, offline installations of DHIS2 were made for each of the 88 blocks in the state, which had the particular hierarchies (PHCs and Sub Centres) preconfigured in the installers for each block. These installers were then installed in the respective block computers, training was provided to the staff on how to enter data by their respective facilities and export the monthly data into a file, which would either be saved on a USB stick and sent physically to the corresponding district, or by e-mail using a dial-up connection. At the district office, this data would be imported into the online application, which would then be hosted on the server.

The possibility to drill down to the facility level helps to set in place various other processes, such as examining of data quality by facilities, establishing and strengthening of feedback and supervisory processes, and also allowing for the analysis of health status indicators by the facilities. This requires the maturing of user capabilities, and simultaneously, the scaling of the processes of capacity building – both in terms of the geographical coverage and the content of training. And as the state develops confidence in their HMIS related processes, they may consider including additional datasets in the application that would cater to the state reporting requirements in the future, in addition to the existing national reporting datasets currently included in the HMIS.

In this way, the challenge of very uneven ICT infrastructure, even within the same district and block, could be addressed using a mix of varying technological solutions, and synchronising it with different institutional practices to match these revised technical configurations.

So the challenge in front, is how to scale a 'whole' integrated architecture, the IHIA, in a setting that is characterised by differences in technical and human infrastructure, and where scaling of uniform solutions is not possible. The answer is, by focusing on the scaling of architecture, in terms of the underlying principles and not the boxes. If we combine scalability and comprehensiveness, meaning a long-term vision of the integrated architecture to be in place from the outset, it must allow for stepwise development that does not compromise further expansion. Therefore, the architecture is not cast in stone, it is never a finished product, it needs to develop according to institutional learning and changes in technologies and needs, and be open to leverage on opportunistic conditions. The need for scalable strategies applies also to the comprehensiveness of solutions. The pace of expanding the architecture along the vertical and horizontal axes, on one hand, will depend upon the combination of the evolving maturity and readiness of the 'target' system in terms of human resources, institutional learning and infrastructure. While, on the other hand, the development of the overall architecture, in terms of the design, particular open source components being used, and the more general technological development strategies being used.

The architecture will, in most cases, not develop uniformly in all areas and across the different dimensions. There will always be lead examples and best practices paving the way, but due to various constraints, the others may not be able to follow. For example, the HIV/AIDS area will typically take the lead on medical records, paving the way for other patient groups segments to be implemented later. However, HIV/AIDS programmes are usually better resourced, than others, such as, TB and Malaria control. Integrated datasets, and national and district data warehouses, will start with essential indicators on some, but not all relevant areas; geographical and administrative regions may develop at different paces depending on infrastructure and maturity. The key issue, however, is that the various scaling processes need to develop within an interoperable and integrated framework, which underlies the IHIA. The heart of this strategy is to not limit the focus of scaling to singular independent solutions, but on the design principles of the architecture and their embodiment, in simple but concrete instances. Based on our empirical experiences, we inductively propose design principles both around the two-dimensions of the IHIA, outlined in the beginning of this section. These are summarised in Box 4.2.

Box 4.2	Design principles for a scalable IHIA – quantitative and qualitative dimensions

Structure, Technical System, Quantitative Dimension: Components of the Scalable IHIA

- Agreed data standards and integrated data and indicator sets, represent the primary building block of the IHIA. Data standards need to develop, according to the user's needs and be flexible to adapt to a changing context, and expanded and changed with the addition of new data and datasets.
- Develop based on a data warehouse framework, or data repository for aggregate data, which provides the means to manage the data and integrate the various datasets and sub-systems. The data warehouse needs to be expandable and flexible, and cater to both data management and information usage.
- Integrate and manage datasets as they are emerging, changing and developing.
- Present and make data available according to domain knowledge and 'business intelligence', as user needs are developing and emerging.
- Establishing interoperability and data exchange using standards, and through the provision of gateways between the data warehouse and the sources of data, whether they are based on paper, computer or other electronic means. One such standard to enable data exchange is SDMX-HD, which helps to provide a 'plug-in' to sub-systems and data sources.
- Designing for modularity; since complexity is managed by enabling sub-systems and data sources to plug-in (or out) of the data warehouse, through the use of appropriately designed datasets and means for data exchange (paper or computer-based).
- The use of Free and Open Source Software (FOSS) and Open Standards. It is important that the system development is based on the use of mature FOSS components, and supports both web-based and offline deployments. FOSS contributes to the scalability; as there are no license transaction cost; and 'small letters', it makes for the threshold to: test, try and eventually implement the components. 'Open standards' are devoid of intellectual property constraints, and which enjoy legitimacy through open, fair and participatory development and maintenance arrangements, by not-for profit agencies.
- While the spread of replicable processes is enabled by standards; standards also constrain, as same solutions cannot easily be implemented uniformly across unevenly developed infrastructure in developing countries.

Process, Social System, Qualitative Dimension: Scaling and Cultivation Approaches

- User participation, learning and local empowerment. To get users at all levels committed, and foster a sense of ownership to information and system.
- 'Radical change through small steps'. Adopt an evolutionary strategy, start small and grow big(ger) as users learn and see the potentials.
- Going first, for the lower hanging fruits which are easy to reach, before moving up the tree. Getting bogged down with complicated problems from the outset will reduce user confidence.
- Scale the architecture gradually along the vertical and horizontal axes, depending on users, and institutional readiness and learning, according to available infrastructure.
- Create 'attractors' by establishing initial success, which serves as the best way to enroll more actors and user organisations, and thereby rallying scarce resources to increase the success. Focus on solving specific large problems shared by many, rather than addressing 'all problems,' or specific problems which do not affect many.
- Flexibility: Data standards, data warehouse and means of data exchange need to be flexible; to enable change according to redefinition of needs, infrastructure and overall context.
- The architecture can never be developed (or constructed) right from scratch, the existing installed base needs to be nurtured and cultivated, so that it evolves and scales over time taking advantage of existing capabilities.

4.2 Scaling as Horizontal Spreading of Replicable Vertical Processes

While in the last section, we presented a model of scaling of the IHIA, as the scaling of principles rather than boxes, in this section, we take this model one step further and present scaling, as the horizontal and vertical translation of replicable processes and the cultivation of these principles. We distinguish between scaling the architecture through the hierarchy of the health system and geography, on one hand, and scaling as increasing the comprehensiveness of the architecture and its technical solutions, on the other.

- *Scaling the Architecture Through Geography and Health Structures:* While horizontal denotes diffusion to new places at the same level of the hierarchy, vertical denotes the process of cultivating the principles and architecture components by growing vertically, translating them into context and creating the need of a complete social system present at one location. Vertical also illustrates, the movement down the hierarchy to new levels of the health system, such as from districts to sub-districts, where the horizontal and vertical processes are repeated; and the movement upwards, through the processes from facilities to sub-districts and districts.
- *Scaling as Increasing the Scope and Comprehensiveness of the Architecture and its Technical Solutions:* Horizontal and vertical scaling also represents a good model for increasing the comprehensiveness of the technical solutions of the architecture.

While vertical denotes increased granularity of the data being managed, such as, adding medical records and interoperability where before only aggregate data was reported; horizontal denotes the expansion of the systems horizontally, by including more health programmes and health services, or by simply, expanding the datasets. Scaling of technical solutions also involves development of new functionality to already existing data sources, such as, new analytical modules like GIS or data integrity checks.

- *Scaling as Increasing the Depth and Grounding of the System:* This involves increasing the quality of use of information – cultivating the system in the context of use. This is first about the appropriation of the system amongst users and grounding; to form it in the context of use. This can only be achieved through user participation and the creation of a sense of ownership to the information and the system, by the users. These processes, of translating the technical aspects of the IHIA into the dynamics of the social context of use; are the concrete local instances of translation, following the horizontal scaling from locality to locality. We focus on one dimension of this local translation – the quality of use. Here we apply the TALI tool (See Annexure 1), which helps to distinguish 3 levels of quality of information use:

Level 1: Achieve basic technical functionalities at each level of the system; if data flows, including feedback, are working; and, data quality and data completeness are acceptable.

Level 2: Achieve active analysis, use and dissemination of data. Graphs, reports, maps, and so on, presenting data is used; for example, on the walls in facilities and in meetings.

Level 3: Data used to influence decision-making, planning and evaluation of the performance.

4.2.1 Debating Dimensions of Scaling

Metaphorically, the vertical and horizontal expansions of the systems are taken from historic battlefields; where armies would expand into territories 'horizontally', and dig down to consolidate 'vertically.' Another way to look at this dichotomy, is to regard the horizontal movement as expansion in quantity, or width, and the vertical movement, in quality, or depth. The usefulness and relevance of using this dichotomy of horizontal – vertical, or quantity – quality, or width – depth dimensions, is to emphasise on the point that an expansion in scope will always need to be followed by efforts to sustain the system, through improved quality of use and heightened capacity development. When scaling an IHIA in scope, to new domains of use (for example, to include human resource management or logistics in the IHIA), or to new geographical or physical locations (for example, to new states, districts, facilities or wards in a hospital), consolidation and gradual improvement of the system will always need to be part of the overall process. Furthermore, improved quality of use, which also implies more use and increased learning about the potentials and limitations of the current system, will lead to gradual improvements and further expansion of the system. Along the qualitative social system axis, the key part of the vertical scaling is, therefore, through cultivation: to ground the system among the users, create a sense of ownership to the system, and to improve the quality of use of data.

There is also, however, a quantitative and technical dimension of vertical scaling. Vertical scaling also denotes an expansion of the systems, literarily downwards in the hierarchy of the health system, for example, in India the first computer-based HMIS was introduced at the state level, then successively in districts, sub-districts, PHCs, and finally, in the Sub Centres, as we included mobile phone reporting as part of the computer-based HMIS, or IHIA. Furthermore, vertical scaling of computer-based resources is also the movement down, the 'data hierarchy' of the IHIA, from aggregate data to individual or case based 'disaggregate' data; from aggregate data reporting to medical records systems, which is increased granularity of data managed in the databases.

It is important to note, however, that technical scaling, vertically in the health system hierarchy, is most of all about users' learning, capacity building and increased quality in data use. For example, vertical scaling is provided from the district to the Sub Centre, and in the data hierarchy, from aggregate HMIS data to medical records system, When a system is introduced, data quality and data use will generally be at the basic level, and a key part of the efforts to strengthen the system, will be directed towards improving quality and use. Capacity building and support, are the key issues in efforts to improve information usage, and will always make up the major part of budgets and resources, needed for scaling systems. What is important to note here, is that scaling a system vertically 'down' the hierarchy, exponentially increases the number of users, and with it the demand for training and capacity building. The key part of vertical scaling, is to ground the system among the users, create a sense of ownership to the system, and improve the quality of data use.

If the key components and design principles are in place, such as agreed datasets, FOSS based tools and procedures for collection or extraction of data, standards such as the SDMX-HD, ad hoc gateways to enable data exchange, and a data warehouse to support data management; the boxes or sub-systems may 'grow' from paper registers to electronic systems, and new boxes 'plug in' to grow into place. As an example of such a scaling strategy, the initial HMIS for routine data reporting serves as a starting point, or the backbone from which the architecture may expand both horizontally and vertically. The HMIS typically includes a number of data elements to be reported, which are organised into datasets, such as outpatient data, immunisation data, reproductive and child health data. Datasets are implemented in the data warehouse, like the DHIS2, which serves as an umbrella or framework, to further expand the architecture. Each one of the included datasets may start out as being relatively limited where the data sources will typically be paper-based registers. Scaling will involve the expansion of these datasets, using more robust technological solutions, and increasing the usability of datasets to a larger community of users.

Expansions of this basic architecture may be illustrated in the following way:

First, a new dataset on human resources is added, made-up by number and categories of staff. This dataset is customised in the DHIS2, and is also reported on paper. We call this the expansion of scope, as more datasets representing a horizontal scaling of the IHIA, are included.

Second, a human resource management database application is established, using the iHRIS software, which facilitates the management of data based on individual records.

Furthermore, a summary of human resource data is exported, using for example, the SDMX-HD standard, to the DHIS2 based data warehouse. We call this the 'downward' expansion of the architecture for vertical scaling; as computer-based management of human resource data is expanded one level down from only being based on aggregate data to also include individual data, making the data more granular and providing the service area (human resources management) with a more comprehensive solution. Introduction of medical records, for example, starting with HIV/AIDS and TB patients, and expanding it to more patient groups, will represent a similar horizontal/vertical scaling of the architecture, as will the introduction of new systems for logistics, drugs management, and laboratories.

In Box 4.3, we summarise the dimensions of scaling discussed.

Box 4.3 Horizontal and vertical scaling as involving processes of translation and cultivation

Scaling Dimension

Scaling is summarised, as taking place over horizontal and vertical dimensions – two complementary, yet contrasting 'directions' of the expansion.

Horizontal Scaling

Scaling in Geographical Scope: Scaling from state to state, from one district to more and all districts, from one hospital to all hospitals in a state, expands the system to all facilities in an area. For example, introducing mobile reporting to all outreach centres in a district. Components of the system, software and standards are also scaled from country to country – from India to Bangladesh, from Sierra Leone to other countries in West Africa.

Horizontal scaling will always involve local adaptation and translation of the specifics of the system and standards, to the local context. Within one district and one state, most systems may be standardised, and more of the same may be replicated. But when scaling to a new state, however, more changes will be necessary; and when moving to a new country, even more.

Standards are crucial, in scaling from one place to another (as without standards, there would have been nothing to replicate), but they will always have to be adapted and translated to the local conditions. The farther away from its origin; administratively, culturally, technically and work-flow-wise, the more changes will be necessary through translation of the new context.

Scaling in Scope of Services and Functionalities: Expanding the scope of integrated HIS and architectures, to include more health programmes or service areas, including TB, HIV/AIDS, and other disease specific programmes, in an integrated approach. And expanding the IHIA to include laboratory systems, drug logistics systems, financial systems and human resource management systems. In terms of medical records and name based registries, typical in many countries, starting with key patient groups, such as, AIDS patients on ART and TB patients, expanding to other groups of patients and clients or services. In India, in some states, processes are ongoingly expanding the architecture, to include medical records for hospitals, registration of pregnant women and the newborn for immunisation.

Typically in many countries, there already exists, different health programmes, which have their own existing systems that are not co-ordinated and integrated with other programme-specific systems. Initiatives to introduce integrated HIS and architectures, will ideally start with some core health programmes, such as Mother and Child Health and Vaccination, to gradually include more programme areas.

Vertical Scaling

Vertically 'Down' the Hierarchy of the Health System: A typical vertical scaling of HIS, is to first introduce a database system at a sub-national level in the hierarchy, to capture data submitted on the paper forms aggregated by the level below; for example, systems at the state level, in India capturing data by the district, and in Tanzania and Ethiopia, at the regional level. Thereafter, introducing the system at lower levels in the hierarchy, down to the peripheral unit and health facility, for example, to a ward within the hospital. What's important to note here, is that the number of users and the need for training and support, increases ten-folds for each level, as the system scales down the hierarchy.

Increasing Comprehensiveness of the System: Comprehensive refers to the need to provide solutions that meet the needs of each service area and level of the health system; from medical records for patient management, to aggregated data and indicator repositories supporting sub-district, district and higher level management, as well as, national and global monitoring and evaluation.

Increasing Data Granularity: Scaling to make database and system management increasingly disaggregated and granular, available to users. The typical situation is that at the global level (for example, WHO), data is analysed and used by the country, at the country level by the state, at the state by the district level, at the district level by the facility, and at the facility, data is managed by patients. Scaling, according to increased granularity is clasically then, to ensure data management in the HIS by the facility (and wards in hospitals), and then integrate 'seamlessly' transacting systems (for example, medical records, lab system) in the IHIA, as source of data for the HIS.

More data is needed at (more) local levels than at (more) central levels. The data at local levels is more granular and more disaggregated than at the central level. Data is generated through transaction systems, paper or electronic, at the local level, at the source of data. At higher levels in the hierarchy, more and more aggregated data is required and used.

Increasing the Depth and Grounding of the System by Increasing Quality of Use of Information and Users' Sense of Ownership to the System: Scaling as the appropriation of the system amongst the users and grounding the information system 'vertically' in the social system and forming it in the context of use. This can only be achieved through user participation and the creation of a sense of ownership to the information and the system by the users. The TALI tool (see later section) helps to measure the level of quality of information usage, which may be likened to the 'depth' and quality of the information system in the social system.

Box 4.4 Scaling in Himachal Pradesh

Scaling in Himachal Pradesh, India

The state of Himachal Pradesh signed an agreement with HISP India to design, develop and implement a hospital management system in all their district hospitals (20 in number) in the state. Ten modules were identified (Registration, Billing, Laboratory, Radiology, Outpatient Department (OPD), Inpatient Department (IPD), Finance, Blood Bank, Inventory and Pharmacy) as forming the core set of modules, to be included in the hospital management system. The plan was to design, develop and implement the system first in one hospital (in the state capital), stabilise it and then gradually also scale the system to the other 19 hospitals in the state. This involved scaling along different dimensions.

First, there was module-by-module scaling. HISP first developed the registration and billing modules, deployed it in the hospital, and started getting feedback through use, while simultaneously, worked on the development of other modules. Over a period of about 9 months, all the modules were deployed. This process of expansion can be understood as horizontal scaling, as the system gradually expanded in scope.

Second, a DHIS2 instance for the hospital was established to capture all the hospital semi-permanent data, such as infrastructure, human resources and government norms on infrastructure, human resources and equipment. Data captured through the patient level transactions (such as in registration, OPD and IPD encounters) would then be needed to be aggregated and exported to the DHIS2 hospital based data warehouse using the SDMX-HD standard. By combining the aggregated patient level data with the semi-permanent data hosted in DHIS2, allowed for the generation of indicators such relating to bed occupancy rates, hospital infection rates and other management indicators. Moving from the patient level data to its aggregated form, plugging into the data warehouse, processing and combining this data, can be conceptualised as a form of vertical scaling. The vertical expansion is represented by more comprehensive solutions for the hospital, increased granularity of the data and expansion of the architecture further down the hierarchy first, within the hospital itself, and subsequently the health system, as selected summary data would need to be exported from the hospital server to the state server that would host data for all the 20 hospitals. In the process described, the new datasets inscribed in the different modules in the hospital, represented horizontal expansion, while the development of the medical records system for the hospitals, from where the aggregate data for the hospital datasets were extracted, represented vertical expansion of the architecture.

Third, the scaling process involved a geographical expansion from one hospital to the others, in the state. The system developed in the first hospital serves as a frame of reference, or, the core standards, against which the systems for the other hospitals need to be developed. For example, the fields of data being captured in the hospital were listed down in a checklist, and form the basis of the requirements study in the other hospitals. The state agreed on the core standards being what the state wanted data on, and if other hospitals needed to collect any other data they could, but were not allowed to delete any item from the list of core standards. This process of requirement analysis was used, to scope out the system for each of the other hospitals, and for development to be carried out. The plan was to first try this scaling process in one nearby hospital, to test the process of customising the application to another setting; and to scope out the context free and context dependent elements of the application. Once the process was carried out in one hospital, it was possible to scale out to more than one hospital at a time (say a group of hospitals from adjoining districts).

This process of expansion represents a geographical scaling, which can also be labelled as horizontal scaling. That is, spreading particular solutions, as parts of the architecture, so that they cover a whole district and a whole state. Together, the horizontal and vertical scaling, make up a dialectic relationship where, first, a solution is scaled so that it covers the whole hospital, and then gradually spread to all hospitals in the state, and possibly to other states wanting to adopt the solution. For each administrative unit or location, the architecture is established as a result of horizontal scaling processes of vertical downward expansion.

We take the example of implementation of a hospital management system in Himachal Pradesh, to describe the horizontal and vertical processes of scaling, and how scaling is cultivated bottom up, but at the same time plugged into the 'top-down' vision of the IHIA (Box 4.4). There is a constant and mutual interaction between these bottom up and top down processes that help to ground the IHIA in reality, while giving it the flexibility to evolve as per the constantly changing realities.

As described in the example in Chapter 1, the development and spread of a hospital information system, or the IHIA including the OpenMRS for patient related transactions and the DHIS2 as a data warehouse for aggregate data, is one of the five applications, being implemented in the state, representing the overall IHIA. Vertical scaling with reference to increased granularity of data is illustrated by the gradual inclusion of medical and name-based records systems in the IHIA, and the registration of pregnancies and children for immunisation; first, in one hospital, and then more hospitals. Such expansions of the architecture vertically, in depth, lead to the management of increasingly granular data, providing an appropriate design; the ability to 'drill down' in the data from the district, to the facility to the individual. In each of the places or areas of use, where the new system or functionality was introduced, it needed for it to expand vertically in the hierarchy of the health system. For example, the new system had to expand vertically from the state to the districts, from the districts to the blocks, from the blocks to the PHCs, and finally to the Sub Centre.

Expansions of the IHIA along the vertical and horizontal axes are far from being mere mechanical or quantitative processes of development. However, as described above, as the IHIA is scaled horizontally, processes of translation are needed to 'ground' it, in the local context. Obviously, the 'further' the IHIA is scaled horizontally, the more the context would change; and if the context is changing, standards will need to change; and the more the change in standards, the more profound will be the required processes of translation. The further the IHIA is being scaled along the horizontal axis, the more the changes in the data standards and metadata, broadly speaking, as they would differ. Hence, there is the need for flexible standards. Changes in standards, however, are not only required when moving from one country to another; there are differences between even states and provinces within a country, and a local adaptation would be needed when moving even between districts within a state.

As the need for information, the reason that the IHIA is being evolved, data standards must be regarded as the most important part of the IHIA. In this perspective, the IHIA is a machinery being put in place, to serve the needs for information. The design principles described in the beginning of this chapter, take the data standards as a point of departure. In India, while the central data standards are shared by all states, each state has its own additional data standards and specific requirements. Scaling the IHIA in India, will therefore be based on a core of data standards shared by all states, with each state, having its adaptation by the way of additional or expanded datasets. If we 'move' beyond India, to Bangladesh, or to countries in Africa, the requirements will be very different. Focusing on the data, we see that the data standards and datasets are also different. But are they so different that we would need to develop a totally new IHIA as we scale further along the horizontal axis, or is it 'more or less the same'?

The integrated architecture framework focusing on management needs, called IHIA, presented in this book, is being developed over about 15 years through practical trials, errors and learnings in numerous countries in Africa and Asia. This varied nature (over geography and time) of inscribed experience, makes the framework robust and well grounded. There are two general reasons for this robustness. First, the context of global health is to some degree standardised across countries, among other things, due to the influence of and from international agencies such as the WHO. Second, the IHIA, while building on these standardised aspects of global health, provides design principles and a framework, rather than technical solutions. While the data warehouse, the DHIS2 in our case, is a technical solution in its various concrete instances and installations; at the outset, it simply is an empty shell that includes a toolkit with which building the metadata and functionalities of the data warehouse is possible, playing an essential part in the IHIA. In fact, the DHIS is, the toolkit and drawing board with which the IHIA is being evolved through successive cyclic instances in multiple countries over the last 15 years.

Scaling the IHIA within a country, state or district, will not involve many 'mechanical' changes in data standards as such. However, the differences in social, economic, cultural and other contexts will be significant. Continuous translation, rather than mere replication, will be needed on both, the horizontal axis of diffusion and the vertical axis of appropriation. In each location or node, as being part of the horizontal scaling, the IHIA will need to adapt to the local context and the users needs, to appropriate the system and build a sense of ownership, for development. This scaling, in scope and depth, are both required to reach the goals of using quality information to inform health management and strengthen service delivery.

4.3 Scaling Across Uneven Infrastructure and Heterogeneity

The need for scalable strategies, applies to the differences in infrastructure, and relate to the comprehensiveness of the architecture and solutions. For example, the recently mandated registration of all pregnant women and immunised children in India will only gradually be covered by computerised electronic records for tracking, as states and districts, rural and urban areas, will develop at different paces depending on their infrastructure and otherwise readiness. The pregnancy and immunisation register, for example, is starting universally as a paper-based registration system, but is gradually becoming computerised – with paper forms captured in the database, and primary registration of services also on mobile phones or other mobile devices, sending data to the database, will be able to access the database directly. The scaling model implies that the architecture and particular solutions need to be able to grow with time; from one to more places, from one level of the health system to the other levels (from district to the facility and then community level), to generally become more granular and comprehensive. This scaling process can be depicted as involving the following processes (See Figure 4.1).

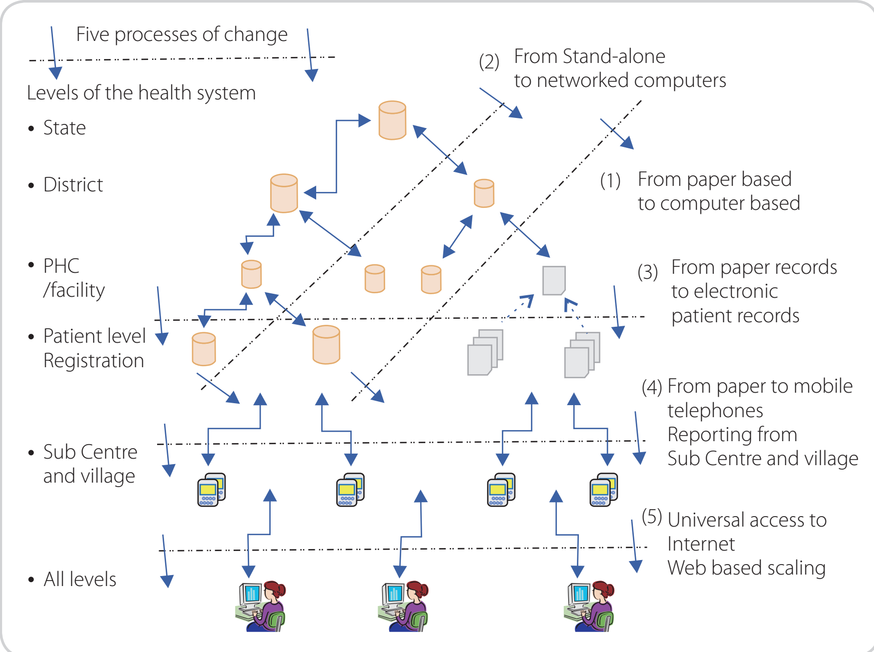

Figure 4.1 Scaling across different infrastructural contexts

Process 1 – From Paper to Computer: This is a general process that has been ongoing for a long-time, where paper-based registers and reporting procedures are gradually being computerised. Still, however, health workers all over the developing world are using large numbers of register books, typically one for each health programme, to register all activities related to individual patients and clients. These registers are yet far from being uniformly replaced by electronic record systems. In India, various states have taken significant steps towards computerisation with different paces. While nearly all are computerised at the district level and a high percentage also at the sub-district level of the Block, there are others (such as, Kerala, Karnataka, Gujarat and Tamil Nadu) also capturing data at the facility level of the PHC. In African countries, data capturing is, very often, limited to the district level.

Process 2 – From Stand-alone Computers to Internet and Networked Computers: In the beginning, all computerised HMIS systems were stand-alone, with for example, a database application installed in each district and using discs and later memory sticks for reporting upwards. Then, with Internet, data could be transferred as e-mail attachments, and finally, web-based systems are used in India and gradually also in various African countries. Taking the example of India, all states are networked at the district level, and a quarter of them at the sub-district level and may be a third of that at the facility level. In India, the transition from multiple installations reporting their data as e-mail attachments (networked, but not web-based), to direct access to a central server, has taken place just recently. As the network availability is enhanced, levels below the district are gradually being included in the network.

Process 3 – From Paper-based Patient Records to Electronic Medical Record Systems and Patient-based Tracking Systems: Introducing medical records systems in all health facilities and for all patient groups is a tremendous task and most developing countries have only just started, and then typically in hospitals and for priority patient groups, such as HIV/AIDS patients on Antiretroviral Treatment. In India, while initially, only a couple of states had initiated processes for introducing electronic patient records in district hospitals, other states like Himachal Pradesh, are starting similar projects. Success of this will potentially lead to a larger process of adoption in other states, since the application is being built on open source platforms which do not have license restrictions. The national Ministry of Health in India has now mandated all states to register, all cases of pregnancy and immunisation by names, which is being universally implemented on paper forms, but some states are initiating electronic community and name-based tracking systems. This registration scheme may trigger more rapid scaling of electronic patient or name-based registers.

Process 4 – From Paper-based Reporting to Reporting on Mobile Phones: As the mobile telephone network has been expanding to reach out to all corners of the world, the use of mobile telephones for data reporting and dissemination is gaining momentum. In India, some states such as Punjab, have initiated state-wide implementations of Sub Centre wise reporting of routine data, using a mobile phone application, directly to the DHIS2, being used in the state as the HMIS backbone. The mobile phone is small and makes it difficult to handle large datasets because of interface limitations. The paradox may be that the Sub Centres using mobile phones will be better networked than the larger PHCs, as, the much larger datasets reported by the PHCs are not well-suited for mobile phones. Through such efforts, the mobile applications are now a part of the IHIA. However, there is also a more negative trend in evidence, which is to establish vertical mobile reporting projects that are not part of the IHIA, but based on software services provided by mobile network providers. This may lead to disruptions in the architecture.

Process 5 – Web-based Scaling: Universal access to the Internet from multiple installations to central servers. While in India, the scaling of the IHIA as from 2009–2010 has been web-based with web access at the district, the block and the PHCs depending on the states; in Africa, scaling has been carried out on stand-alone installations, although typically with Internet access for submission of data. The scaling of the DHIS2 in India would not have been possible without web access to central servers, as support and maintenance of thousands of stand-alone applications would not have been possible. In Kenya, access is being enabled countrywide through mobile internet.

This brief overview of ongoing and often simultaneous processes, gives an idea of the various dimensions and aspects which are involved in scaling, including paper artefacts, computer terminals, servers and networks, mobile phone and mobile internet, and of course the various institutional conditions and practices that constitute the system and its use. This collective, which comprises the IHIA, is depicted in Figure 4.1.

Approaches for scaling of electronic patient records, from relatively simple name-based tracking systems to more comprehensive medical records systems, are being eagerly debated in many countries and within many agencies. A recurrent dichotomy is between those arguing that scaling of medical records will make the traditional HMIS approach superfluous, and those arguing that medical records systems will not scale fully in a foreseeable future in, for example, Africa, and therefore should not be prioritised. The Ministry of Health and Social Welfare in Tanzania once posed the question – Should we go for medical records systems or should we focus on the

HMIS? As argued in the first chapter of this book, this is a 'wrong' question. Within an enterprise architecture perspective, both medical records and information for decision-making and Monitoring and Evaluation (M & E) are part of the enterprise architecture for the health system. Within a scaling perspective, however, we may be more precise; for example, in Tanzania, it will take time to scale electronic patient record systems covering all patient and client groups to all the literally thousands of health facilities in the country. If the aim is to get full coverage data health programmes performance for management and M & E at different levels of the health system, it will not work to wait for full coverage of patient records. And the approach needs to be based on the IHIA data warehouse framework argued for in this book, where there are electronic patient records systems aggregated data needs to be generated and fed into the data warehouse, where there are no such systems, paper-based records must be used as the data source. The same holds for data on human resources, drug logistics, laboratories, and so on. More generic we may regard the IHIA as consisting, on the one hand, of more or less fine grained, paper-based or electronic transaction systems where patient-health service encounters are registered, staff recorded, lab tests registered, drug stocks updated, and so on; and on the other hand, the data warehouse where processed aggregate data from these primary data sources are managed.

The case of Himachal Pradesh described in Chapter 2, details how an integrated architecture approach enables a 'seamless' gradual scaling of the architecture, such as from paper records to electronic medical records first, then more hospitals and the similarly gradual expansion of the electronic tracking system for pregnancies and newborn. The key advantage with an enterprise architecture following a data warehouse approach in a scaling perspective is that it provides a framework within which computerisation, that is, new software applications, within all the various sub-areas may be scaled. When a new system for tracking pregnancies and immunisation is introduced in a facility, the data previously reported on paper will now be generated electronically and loaded into the data warehouse. New software solutions are 'plugged into' the architecture – the data warehouse – using, for example, the SDMX-HD. In this way, the data warehouse and IHIA are enabling smooth scaling of, for example, medical records systems.

4.4 Scaling Through Central Servers and 'Cloud Computing' – Universal Solutions in Heterogenic Infrastructure

Scaling of IHIAs in developing countries has so far been seriously hampered by poor Internet. While in industrialised countries, a HIS data warehouse may be installed at a central server with one support team and made accessible for all users in a country; in developing countries, until now, the HIS data warehouse application would need to be installed in literally hundreds of locations, data would need to be manually transferred and synchronised and a large number of skilled people would be needed to run and maintain the system. Data management in a system consisting of hundreds of stand alone databases with manual data flows to the central database is extremely complex and the maintenance and update of all the installations likewise. Scaling of such 'stand alone' systems in countries like Tanzania is difficult.

Given the improvements of the mobile telephone network, the 'standalone' problem may be overcome by using the mobile telephone to submit and receive data. In Punjab, India, a large project including 5000 Auxiliary Nurse and Midwives (ANMs)

Box 4.5	Scaling of mobile based health information systems in India

Scaling Using Mobile Telephones in Punjab, India

The first effort towards the use of mobile phones for supporting of reporting from the peripheral facilities was initiated in 2008, when the National Rural Health Misson (NRHM) started a pilot in 5 Blocks of 5 States in India covering a total of about 200 field nurses. HISP India developed the application for this based on GSM technology through which the report would be send via SMS to the DHIS2 application (used as the HMIS data warehouse application in the 5 pilot states) installed at three levels of the Block (Sub-district), District and State. In addition to building the application on a free and open source model, HISP India provided the required capacity building and support to the health workers on whose phones the application was installed. In this pilot, focused on Sub Centre reporting, the dataset included 77 data elements covering Antenatal Care (ANC), Child Immunisation and Family Planning. The pilot was evaluated as a success, and Punjab was one of the states that decided to go for a state wide implementation, covering 5000 health workers who would be given mobile phones from the state, and also CUG connections.

This initiative was a scaling challenge on various dimensions including functionality, geography and user support. With respect to functionality, while the initial application covered only 77 data elements, the State wanted to extend that to also include the reporting of mortality and morbidity details and other financial reports along with the basic Sub Centre report. This extended the data elements list to about 150, and these then needed to be split into 2 datasets as one was too long to be sent in a single SMS and resulted in memory errors. Further, the state also wanted to use the phone to carry out health services monitoring on a daily basis, and 10 data elements were identified in this dataset which mainly covered activities of a health worker on daily basis. From one dataset of 77 data elements, in this case the application was scaled to 160 data elements, which were then divided into 3 datasets – two monthly and one daily.

In terms of geography, while in the pilot only the field nurses from one Block (or sub-district facility) was covered in each of the 5 pilot sites (about 30 nurses each), the project now covered 5000 nurses from about 2500 Sub Centres covering the whole state (20 disticts). To deal with this massive expansion in scale, a shift was made in the model of reporting. Now all SMSs were only to be sent to a central server where the state DHIS2 application was hosted. This meant that the SMS would now also be sent to the Block and District databases, and users from this level would need to access the data from the state server. However, this raised challenges at the server end, especially in the use of the GSM modems for receiving the SMS, and for also sending out the acknowledgements back to the users whose SMS had been received. There were extreme challenges in managing this SMS traffic, especially the daily report covering 5000 messages. To deal with this, first the number of modems were increased to 3, where 2 were used for receiving the messages and one was exclusively dedicated to acknowledgements. Further, a change was also made to the queuing algorithm, wherein messages were processed in a batch mode rather than sequentially.

In terms of implementation support, HISP India had to enhance its team both in numbers and the mix of skills. While in the pilot phase, after the initial phase of training, the technical person would visit the sites may be once in 3 months for support, and the rest of the time remote support was provided. However, in Punjab given the scale of operations a total team strength of 10 members was created to cover all the training, and following that a team of 5 were based permanently in the State for 9 months to provide continuous support. The team now needed to include a dedicated server technician who needed to be at hand to deal with the constant problems related to the server and modem. Two of the five members were software developers who needed to be continuously engaged in making improvements to the existing application, but also create new enhancements as the demands of the state grew with time. For example, the state demanded a feature of Mass SMS to health workers in order to obtain feedback from health workers regarding any service, programme, stock requirement or grievance report. For example, the state wanted to be able to send a SMS to all 5000 health workers to get a confirmation whether or not their monthly stock of IFA or Iron tablets had been received. This facility of enabling 2 way communication scaled the use of the phone from not just reporting but to also co-ordination and communication.

The scaling of the Punjab state mobile project makes it, by far the largest implementation in the health sector in the world. It is interesting to also point out to how other states are in process of further developing pushing the boundaries of scaling. For example, the states of Himachal Pradesh and Orissa wants to scale the mobile based reporting for disease surveillance and control programme covering diseases like Polio, Malaria, HIV/AIDS, TB and other chronic diseases.

Various other states are planning to use the phone not only for number based reporting but also for reporting on names for tracking of pregnant woman care and child immunisation in order to provide timely care and monitoring. Further, though the application had only been used for reporting from Sub Centres, some other states also want to use it for reporting from PHCs and CHCs, where the datasets are much larger (about 150 data elements). This they plan to do because of the extremely poor internet connectivity, bad power situation, and also the difficult geography which makes travel complex.

using their telephones to report their routine databased on an application using SMS, is a good example of how the basic mobile network may be used to report data over distance to a central server. See Box 4.5 for details on this project.

Many projects in the developing world are exploring similar ways, or using the upgraded GPRS protocol, to use mobile phones for remote data reporting and communication. The global mobile network, however, is also opening up for new business models of 'software-based services', and what may be feared, the outsourcing of ICT and software services business from Africa to the West. What is happening in many countries in Africa now is that international mobile network providers are targeting donor funded health programmes and projects working in remote areas, typically in Africa, and providing end-to-end infrastructure and software solutions for data capturing, data management, and analysis at a cost, paid by the rich donor, and typically not integrated within a national Health Information Systems framework.

A typical example is a project by Pfizer and Vodafone in the Gambia, where Pfizer wants to monitor stock and distribution of Malaria drugs in the dispensaries and Vodafone is providing the entire infrastructure; SIM cards if needed (users have their own telephones), mobile network, air time, data management in their servers in the 'cloud' and provision of the data to the users. DHIS2 is running as the national HIS in the country and is including similar data. We have tried to argue the need for an integrated architecture approach, and suggested to feed the data reported by the

Vodafone mobiles into the DHIS2, that is the national HIS. We have also argued the need for pooling together the limited resources available for HIS in the Gambia. These efforts to integrate the approaches, however, have been in vain. Vodafone argued quite frankly that such integration efforts could only be included 'if there is a place for them in the value chain'.

It is our general view that business models that are locating value chains derived from Africa, outside Africa, are of no use for Africa. But more specifically, we will argue that the outsourcing of what may be labelled the 'ICT learning and innovation chain' from Africa to the West, as illustrated by the example of Vodafone and Pfizer, may be even more harmful.

The new situation of improved Internet connectivity and 'cloud computing', that is the availability of software and Internet-based services from anywhere in the world, represents, apart from some threats of further fragmentation of the HIS as illustrated above, and at the same time drastically improved opportunities for implementing and scaling integrated architectures. The rapid scaling of the IHIA in India has to a large extent been made possible by web-based access to the DHIS2 data warehouse, installed on central servers in each state. India does not have universal Internet access, but the network in most cases is extended to the block (sub-district) level. In most states, data is therefore captured at the block level, and in some states directly from the sub-block level of the PHCs. Experience from the previous stand-alone installations of DHIS2 in India, have shown that it is very complicated to support, maintain, update, and to some extent also synchronise hundreds of separate installations in PHCs, blocks and districts. This is of course following from the large scale of implementation in India. The case from Kerala in Box 4.6 describes the gradual transition from stand-alone installations to a central server solution and how that has enabled the scaling process.

Box 4.6 Scaling of DHIS2 in Kerala, India

Scaling of DHIS2 in Kerala – From Stand-alone to Central Server and Web Access

The very first instance of the DHIS2 started with a deployment in one Community Health Centre (CHC) in Trivandrum district of Kerala state in India. A CHC represents a sub-district level health unit in the state, and has below it PHCs and below which there are Sub Centres. A CHC caters to about 200,000 population. Kerala is the most advanced state in India with respect to public health indicators, which ranks with the best in the world. Furthermore, with its very high levels of literacy, Kerala was declared a 'Fully Literate State' in 1991, it becomes a very fertile and receptive place to initiate new innovations in the rural work force. And to the advantage of DHIS2, Kerala state has an explicit policy to promote the use of FOSS in the public sector, which provided us with the argument for this effort. Historically, the state has had a strong legacy of a left government which makes decentralisation a key political agenda, which was another argument made for deploying DHIS2 in a clinic level rather than within a centralised model.

HISP India, with whom the state has a technical partnership since 2005 and a formal MOU since 2008 is responsible for the DHIS2 deployment, and organising all processes of technical support and training. Since the stack of technologies used in DHIS2 was new to the HISP team (with prior experience in the Access-based DHIS1.4) and the DHIS2 itself was in a very nascent stage, it was an intensive and continuous period of learning for the team. There were constant requests for re-installation (due to virus attacks), more functionalities, and ongoing hand holding and support. The clinic was about 15 km away from the HISP office, and the field co-ordinators would go sometimes once or twice in a day to deal with these field problems. After about 3 months, the HISP team could demonstrate concrete results to the district and state office, and they gave the permission to extend the DHIS2 deployment to the 18 CHCs in the district. For this, the HISP team purchased the required hardware and got it installed in the CHCs through the vendor who were responsible for the hardware maintenance activities.

Extending to 18 CHCs started to expose the HISP team to the challenges of scaling. Now there were 18 such installations that we needed to take care of, and hence 8 field co-ordinators were hired and trained with each of them being made responsible for two to three CHCs for providing field level technical support. There were more technical problems to deal with, more training to be provided, and furthermore more databases to synchronise. For this, at the end of every month, all the paper-based data from all the facilities would be sent to the CHC data operators, who would consolidate that manually into a CHC wise aggregated report and enter it into the DHIS2. After the entry was done, the data would be exported into a file on a flash drive or CD, and this would manually be carried to the district office, where it would be imported into the district database to create the consolidated reports.

These scaling related challenges were only a preview of what the HISP team would need to do next, as the state was satisfied with the progress made by HISP in establishing the district systems and entered into a Memorandum of Understanding (MoU) with them to implement the DHIS2 across the state. The level of deployment was the PHC (a level below the CHC) and there were 858 such facilities in the state. This time around, the state procured the necessary hardware (computers, printers, UPS), and through the MoU required 2 of the HISP co-ordinators to be placed in each district (there are 14 districts in the state) for 1 year, and have an additional 2 in the second year to strengthen hand holding support. In this model, the entry at the PHC was to be done by each Sub Centre, and at the end of the month the PHC consolidated data (sum of the PHC and Sub Centre facilities data) would be aggregated and exported, and then sent manually to the CHC from where there would be internet access. Thus, the CHC would import their data PHC wise, plus enter data on services provided by their facility into the DHIS2 now deployed on a server hosted at the state IT centre.

Scaling challenges came across multiple dimensions which are briefly summarised:

The database had to be scaled to be able to handle data from 5391 Sub Centres, 858 PHCs, 181 CHCs, 86 District and Sub District hospitals, and also private institutions.

Server performance had to be constantly monitored and fine tuned to ensure, it could handle the load of may be 1000–1500 simultaneous users entering data in the last week of the month.

Capacity building programmes had to be carried out for more than 12,000 staff to ensure data entry and data quality management was done by the data collectors/compilers, that is the health staff themselves.

During the first year of statewide implementation all PHCs and CHCs were not provided with internet connection. Therefore, in PHCs and CHCs offline version of

DHIS2 had to be maintained. This involved ensuring stable installation and continuous trouble shooting in 858 PHCs, 181 CHCs and 86 Hospitals. The installations were quite a challenge and took about two to three months.

Data synchronisation every month involved the Sub Centres entering their data every month at the PHC from where data was exported manually or electronically where possible to the CHC where it was imported. At the CHC then sent up the data to the District. This way each of the 14 districts in the state had their respective database located at the district headquarters. This was a huge institutional and technical undertaking.

Providing technical handholding support to all the facilities in the state with the team of 56 co-ordinators and the state team.

With intensive implementation support, the HISP team was able to demonstrate very positive results, leading Kerala state to be ranked No. 1 in the country with respect to HMIS. The next phase of the scaling challenge came when the state now wanted the PHCs to access the server directly and carry out online entry. However, it was agreed that the data entry would be carried out for PHC consolidated (PHC plus Sub Centre data) rather than including data also by individual Sub Centres. This of course was good, in that now the problem of export-import which was being carried out manually would be stopped, but it led to serious performance issues with the server. There were constant crisis calls to the DHIS2 development team, especially at the end of the month, for trying to improve both server and DHIS2 performance. Slowly, this process was also stabilised.

A final challenge on scaling came when the state also wanted all Sub Centre data by individual facilities to be also placed on the server to enable the state and district officials to be able to drill down to the lowest level data when required. The HISP team tried to argue against this move saying that why should the state want to keep Sub Centre data and overload the server when for their planning and action, district and CHC level data would be sufficient. Instead, we could make the Sub Centre data available to them in an offline installation, and they could access that when necessary. However, the state was adamant on this, and insisted this was done. This was then done, providing huge amount of additional pressure on the server. This required again serious improvements to be made to the DHIS2, especially to improve the time taken in the loading of the screen. Simultaneously, the HISP team kept pointing out to the state the names of particular facilities where improvements in infrastructure needed to be carried. This process is now well-stabilised.

At this point, the HISP team with a great deal of pride can say that these scaling challenges have to a large extent been successfully dealt with, and significant improvements have been made in terms of data quality, data coverage and timeliness of reporting. In the next phase of the MoU, currently being negotiated between the State and HISP, the focus is shifting on strengthening processes of improving the use of information for local action, and improving the overall state ownership of systems.

The problem of maintaining numerous stand-alone installations represents a formidable obstacle for scaling the IHIA not only in India, but also in various countries in Africa. In Africa, the Internet has not been, and is still not perceived as robust or reliable enough to support web-based data warehouse solutions for routine data reporting. The norm, until today has been to capture routine data in stand-alone databases, and to report data electronically by e-mail attachments or physically to the level above. Significant human capacity on database, data management and system support is needed, in order to manage a national HIS based on numerous stand-alone database applications with fragile flows of data between them. Problems of data reporting, completeness and the maintenance of numerous stand-alone applications across the country, make it very complicated.

Building the IHIA on a web-based data warehouse on a central server is much simpler technically, in terms of data and database management. In terms of human resources too, ironically, it would be more appropriate for Africa, than their current reliance on stand-alone database applications. The paradox is that, currently, in Africa, more resources are needed to sustain a national data warehouse solution than, for example, in Europe, even though the resources required are fewer. The key challenge, therefore, is to explore how Africa may be able to optimise the use of Internet, including developing the required power back-up infrastructure on generators, since during power cuts, the fixed line Internet connectivity will typically be down. Internet over the mobile network is spreading rapidly, and may be a solution here, both as a back-up to fixed lines and as a solution on its own.

DHIS2 has been rolled out in 2011 in Kenya, based on a central server solution. The long-awaited sea cable has arrived in Mombasa, at the coast, and Internet over the mobile network is now covering the country. Kenya has become the first country in Africa implementing an online national data warehouse for universal capturing and use of routine data. In Chapter 6, we describe these new opportunities of online computing Kenya and Africa more generally and how offline data entry using the new HTML 5 standard may help managing the rather poor Internet connectivity found many places. Offline data analysis and use is another innovative approach developed to manage poor Internet connectivity; a small 'super lightweight DHIS2' application installed locally is used to download and update a 'datamart' including user specified data and indicators, which is then used to generate Excel pivot tables used for data analysis. Reports, charts and maps are generated when 'online' and downloaded in PDF format and archived in the offline application. As the Excel pivot tables are not easy to update, or 'refresh', online regardless of bandwidth, the offline local datamart is actually representing an improvement in some functionalities as compared with the web browsing.

The innovative approach to address the rather fragile Internet situation in Kenya illustrates how innovations are generated through implementation, use and problem solving. Innovation may thus be seen as being an embedded part of the scaling process. Furthermore, by using open source and shared software applications and enabling collaborative networks where countries and institutions share learning, best practices and solutions, innovations in Kenya is fed back to the entire community. The DHIS2 functionalities developed to enable the 'semi-offline' approach in Kenya have been included in the global core and thereby made available for everybody. As a next step, therefore, the semi-offline functionalities, including the ability to download data-marts and reports and to generate Excel pivot tables locally, are made available for users also in Kerala and other parts of the world.

Scaling may, therefore, be understood as a dialectic process that is enabling innovation; it is not the diffusion of the 'same', as the system and its components is changing and improving through the participative processes involved with implementation and use.

In scaling, while new things are gained, there are also always old things being lost representing a dialectical process of expansion. Feedback cycles are then bringing the innovations from one context, such as Kenya, to other contexts, such as Kerala, where new user interactions may lead to further innovations and improvements. In a next cycle then, the solutions may come back to Kenya, further improved. The FOSS approach is important in such a dialectic cycle of innovations as it enables free circulation of new solutions.

4.5 Attractors for Change – Drivers of Scaling Processes

Complex adaptive systems (CAS) is a research tradition, concerned with studying the dynamics with which complex systems, made up by other systems, networks and actors, evolve through adaptation. The health system illustrates not a singular but a complex system, made up of numerous more or less independent systems, organisations, structures and actors of different types. It typically is, not even 'one' system, and can be regarded differently, depending on what perspective and view point one is using to define the system, its components, or its point of gravity. Analogue to the health system, CAS is seen as being made up of semi-autonomous agents, with the inherent ability to change and adapt in response to other agents and the environment. It is important to understand how order within such systems is created without a 'designer', but out of emergence. Attractors are seen as 'points of gravity', which enable the emergence of order amongst processes that are divergent and apparently in a state of chaos. For example, in a situation where different health programmes, health authorities and other actors, are unable to come to agreement on the new shared data standards, an attractor in practical terms, would be the implementation of a simple but useful standard, which works; and on demonstrating its 'success', attracts attention and more followers, thereby enabling the implementation of even more effective data standards. Our case materials provides empirical evidence implying that such standard making attractors is the typical way in which new data standards can emerge and grow in use and importance, through attracting more followers.

Example of South Africa

In South Africa, as also discussed in Chapter 2, numerous actors, health programmes and health authorities were for a long period, after the liberation, unable to agree on shared data standards, basically because, as it turned out, they were unable to see practical ways of achieving it. A major problem that seemed difficult to overcome was the fact that the federal constitution made each province 'free' to make their own data standards, making it all the more difficult to see practical ways in unifying all these actors. The breakthrough came, when the HISP team finally managed both, to create and implement a minimal essential dataset and manage it close to the users through the first version of the DHIS application, at first one province, and then at a second. The third component in this package, in addition to the datasets and the DHIS application, was the principle of the hierarchy of standards, which was implemented in the DHIS. Important here, was that despite the fact that the two datasets were different, they shared only about 50% of the data elements, and were managed within the same unified framework. The DHIS framework used, was the early version of the IHIA discussed in this book. The DHIS, despite its simplicity in design, turned out to be quite powerful as a key attractor, driving the process of scaling along both the horizontal and vertical axis for many years in South Africa. When the various actors saw that the principles were working in two provinces, they understood that, they could be made to work in other provinces and the whole country, as well, motivating and driving cross country scaling. Since its emergence in the late 90s, the DHIS has been growing in strength and depth, and is today institutionalised as the national standard in South Africa.

Example of India

In India, the development of the new integrated health data standards, including data elements, indicators and reporting formats, was carried out within the framework of the NRHM, starting early 2008. Unlike in South Africa, where the development of standards took place through a bottom up process, or even in Himachal, where the hospital standards development has started from one hospital to be incrementally taken to other facilities; the NRHM was largely a top down and centralised process. In this top down process, the NRHM in consultation with the heads of different health programme divisions (like Child Health and Maternal Health), at the Ministry of Health and Family Welfare in Delhi, defined the standards and diffused them into the states for implementation. The probable arguments for adopting this top down rather than a bottom up (by states and districts) approach, could be as follows:

- The design was not being done from scratch but was modifying something already existing. It could thus, be argued that the requirements were already known.
- In a historically existing hierarchical and centralised structure, states will not take the initiative to re-design unless it has been mandated from the centre to do so.
- Given the large-scale and diversity of the different states in the country, a centralised and top down approach offers time and logistics expediencies.
- Since a number of the health programmes being implemented in the states are centrally driven, in terms of budgets and information systems, the states would not have the mandate to make any changes in design – and must comply to national directives.
- The problem on creation of standards was largely to streamline and reduce, rather than add new, so these decisions were best taken at the 'top' rather than by states who would not have the overall picture.

So, there were certain merits to the adoption of a top down process for the definition of standards and its scaling to the states. This approach was compatible to the existing context. The entire process of re-design of the national HMIS has been described in detail in Chapter 5, but in this section we describe the key features of the scaling strategies adopted. These are summarised in Box 4.7.

Box 4.7 Characteristics of scaling strategies

Characteristics of Scaling Strategies – India

1. The focus of the re-design process was to create the core standards, on what data should be reported by all districts, and all its facilities to the national level. In addition, the states had the freedom to add further local requirements, but they did not have the right to delete anything from the national core.

2. A set of design principles were created – rather than boxes of technologies and applications and these principles were scalable. The principles were articulated in a way such that was simple, based on common sense, was easily understood by all, and was accepted as being useful to address the common problems currently being faced. Examples of the design principles were:
   a. Data will only be entered on one format, and not more than one (to address the problem of duplication).
   b. Data will only be reported on the services provided by the facility (to address the problem of confusion between facility reporting and area based reporting).
   c. Only that data will be collected, which contributes the generation of at least one indicator (to improve the action ability of the HMIS).
   d. Disaggregations of data (for example, break ups by Scheduled Castes and Scheduled Tribes to be captured through surveys rather than routine data).
3. The revised datasets were designed in a scalable manner, which were flexible to adapt. For example, private hospitals could report using the CHC dataset as they were seen as equivalent facilities in terms of services being provided.
4. There was flexibility to add new datasets as new programmes were integrated into the HMIS dataset. In 2008, the immunisation data had been integrated. So, an incremental and flexible strategy was adopted, making it scalable.

Through this re-design process based on principles elaborated above, close to a 90% reduction on data elements to be collected was carried out. For the states that had agreed to adopt the DHIS2 as their state system, all the datasets, data entry screens and reporting formats were customised. Furthermore, many of the states requested for their local data elements or indicators to be added. These additions could easily be done in the DHIS2, which helped it become an attractor, since the users felt a sense of control over it and something they could tweak for their local ends. Another feature of the DHIS2 which helped make it a strong attractor, was the capability to customise organisational unit hierarchy to the sub-district level, if the user wanted it. This indeed helped to empower the users and enhance local control, something that was then absent in the existing national portal, which was based only on district level data entry and reporting. Further, DHIS2 was typically hosted on a local state server that implied easy accessibility to its users, which again was difficult with the 'black boxed' ministry portal, as it served primarily as a gateway for national, and not state reporting.

4.5.1 SDMX in Sierra Leone and West Africa

In this example from Sierra Leone and West Africa, we illustrate how attractors may drive scaling processes in one country, across countries; and in the scope of the IHIA, by including more components. We call this the creation of 'network of networks.' In the second chapter, we were told the story of how the IHIA was developed and scaled, in Sierra Leone, through initial success that created an attractor, which further led more health programmes and agencies to align their interests and join the scaling process. This is a development quite similar to what happened in South Africa.

Sierra Leone has been one of the HMN pilot countries, and by far the most successful. As HMN needed a success story, they have promoted Sierra Leone as that. While the success of the new IHIA in Sierra Leone has been real, it is clear that being declared a success by HMN and being published as a success several times in the HMN weekly bulletin, has significantly added to the reputation and knowledge of the project in Sierra Leone. Sierra Leone has thus, become an attractor for IHIA in particularly the West African region. In 2009, The Gambia started implementing DHIS2 on their agreed data standards, which previously had been implemented using Excel spreadsheets. Furthermore, during 2010, Burkina Faso, Ghana, Liberia and other countries in West Africa, have also decided to go for DHIS2, as the embodiment of the IHIA. Today, the West African Health Organization (WAHO) is co-ordinating technical support on DHIS2 to all West African countries.

In late 2009, WHO had initiated a process to develop a data exchange standard for health metadata and aggregate data, SDMX-HD. By early 2010, HISP and HMN, decided to pilot the standard in Sierra Leone by implementing the OpenMRS medical record system to track patients on Antiretroviral Therapy (ART), and to share this data with the DHIS2 data warehouse. This was the first implementation of the standard, and major advancements were made to SDMX-HD, spurred by the fact that it was now being implemented for real in two applications. The seemingly successful implementation of SDMX-HD for medical records data triggered interest among the many actors dealing with human resource management. While medical records system are slow in scaling, as health facility and patient groups need to be implemented one by one; human resource records are different, as in it the employees and salary registers are more centralised and complete. In fact, the problem in many African countries is that, employees and salary registers tend to be more than complete, sometimes containing unknown numbers of ghost workers. The integration and interoperability between the DHIS2 data warehouse and human resource records system was therefore seen as very attractive; as the number of staff by category could be analysed by facility and compared with other data on resource utilisation and consumption.

The development of SDMX-HD, made the IHIA architecture include 'any' application from different business domains, a reality. Anyone able to share data using the SDMX-HD could now be 'plugged in', in a country IHIA. In addition, as this development was associated with Sierra Leone, its perceived success accumulated, and, further events triggered. While HMN and HISP were collaborating in Sierra Leone, Capacity*Plus,* a partner of HISP and HMN specialising in strengthening health workforce information systems, was partnering with WAHO, to pilot the open source iHRIS application for human resource management in Ghana. Given the success of what was now seen as the 'integrated Sierra Leone architecture' in West Africa, WAHO organised a workshop in Accra, Ghana, in partnership with HMN, Capacity*Plus* and HISP, to promote the integrated architecture, and conduct training in DHIS2 and iHRIS. At the end of the Accra workshop, SDMX-HD implemented for interoperability of DHIS2, iHRIS, OpenMRS, was officially launched by a WHO representative pointing out the appropriateness of it being taken up among countries adopting the standard-supporting applications. In November 2010, WAHO invited HISP, Capacity*Plus*, and HMN to the first annual WAHO HIS strengthening workshop in Dakar, where HIS staff from all 15 WAHO member states were present. Here, it was decided that HMN, WAHO, HISP, and Capacity*Plus* should form a partnership to develop a centre of excellence at WAHO, for supporting member countries in adopting the IHIA, which had grown out of Sierra Leone.

The above discussion highlights the important role attractors play in enabling processes of scaling. In South Africa, the DHIS2 which enabled the implementation of the data standards and principle of hierarchy, served as an important attractor for other provinces who saw its value and took it up. In India, the principles developed for the re-design of the datasets, became an attractor in creating uniform standards, which could be scaled across geography and programmes. While in Sierra Leone, the DHIS2 which served as the IHIA became the attractor and the focal point, to expand the HIS to interoperate with human resources and patient based data. The IHIA initially represented by the DHIS2 and the supporting network, became the attractor to coalesce other networks – the OpenMRS and iHRIS – in both its technological and institutional constituents, leading to scaling through 'network of networks'. Thus, attractors are not only technological in nature, but well involve organisational and institutional constituents.

Summary

Key concepts that can be taken from this chapter are summarised below:

1. We emphasise that scaling should focus on design principles rather than on technological artefacts referred to as 'boxes'.
2. Scaling is around two sets of processes:
   a. Quantitative, mechanical, structural dimensions of scaling.
   b. Qualitative, social system, process-oriented dimension of scaling.
3. Scaling takes places across interrelated dimensions of vertical and horizontal axes representing depth and scope, respectively.
4. Scaling of systems involves processes of cultivation and improvisation.
5. Scaling in developing countries will always involve expanding across uneven and heterogeneous architectures.
6. Improving Internet and mobile networks are making it possible to scale universal solutions despite heterogenic infrastructures – through web and mobile access to central servers.
7. Scaling represents a dialectic process, where something is simultaneously lost and gained during the process of expansion.
8. Scaling along the vertical and horizontal dimensions goes hand in hand with the sustainability of the system.
9. Attractors – technical and institutional – are drivers for processes of change.
10. A current strategy for scaling is the creation of 'network of networks'.