3

Integration and Interoperability – Standardisation is the Key

I n this chapter, we engage in the ongoing debate in the global HIS field, around issues of integration and interoperability – how these are different, how they may complement each other, and how integration and interoperability, respectively, could be achieved. We argue that, standardisation is the key to understand and to achieve integration and interoperability in HIS, and information systems more broadly. Furthermore, we argue that within the perspective of the management-oriented enterprise architecture, we develop in this book; the IHIA – as we have labelled it, makes it crucial for data standards.

In this chapter, we first position the debate(s) around integration and interoperability in HIS, taking examples from the development of the DHIS2 and its integration and/ or interoperability with medical records and human resource management systems. After discussing concepts of vertical and horizontal integration, in the last part of the chapter, data standardisation is linked to cases of development of the DHIS2 'data warehouse'.

3.1 Integration and Interoperability – Current Debate

Integration is a term with different meanings in different contexts – from being a very precise term in mathematics, as the art of calculating integrals, to a more ambiguous term in social sciences and everyday speech, as in cultural, ethnic, or religious integration. Similarly, in our field of HIS, meanings of integration will vary according to 'who you are', your role in the organisation, your relation to the health sector, and the perspectives on the HIS. While managers and health services providers would think that integration of information and routines are needed to perform certain tasks; ICT staff, software developers and more technically-oriented personnel would think that co-ordination of the software applications and protocols for communication are needed to achieve that.

During an electronic forum debate on integration, in the RHINO network meeting in Mexico, 2009, it turned out that, while some saw integration strictly as a technical term related to collating several software applications into 'one system,' regarded as a negative and an old fashioned approach; others used integration to describe approaches to get health programmes to collaborate and unify their data collection forms and routines. Those who understood the term integration as refering to a technical approach to make 'one big system', were against the notion of integration understood in this way. They would rather go for the concept of interoperability,

as they saw it more useful to make different systems to interoperate rather than to collapse them into one big system. In our view, however, integration, even at the level of software applications, cannot be reduced to an aim of constructing one big system. The definition of integration in the organisational context, takes the users needs of the HIS, the purpose of the systems, and the wider organisational perspectives as points of departure, and relate those to goals of better efficiency, effectiveness and co-ordination in organisations and enterprises. In such a perspective, interoperability is one of the means to achieve such ends – integration.

Interoperability is, contrary to integration, a term which may be given a more formal definition. In our context, interoperability refers to the ability of a system to use and share information or functionality, of another system by adhering to common standards. Without agreed standards shared by at least two systems, processes or other actors which are aiming at interoperability, interoperability is not possible.

In the last chapter, we described the example of the SDMX-HD standard for exchange of data. Two 'levels' of standards are needed in order to exchange data; a shared format or protocol, for example the SDMX-HD, and then, of course, one needs to agree on what data including their definition and meaning should be transferred. Later in this chapter, we discuss the distinction between syntactic and semantic standards, and elaborate on the challenges of developing data standards. But for now, we use the example of SDMX-HD to describe the essence of interoperability.

Box 3.1	An example from Sierra Leone

Building Interoperability

In late 2009, WHO had initiated a process to develop a data exchange standard for health metadata and statistics, building on an existing standard for financial transactions. This new protocol, SDMX-HD, was still not implemented by any application, far less in any real use-case setting by early 2010. Nevertheless, HMN, MoH in Sierra Leone and HISP, decided to initiate a pilot system that could both track patients on Antiretroviral Therapy (ART) and share this data with the data warehouse running on DHIS2, as stipulated in the original strategic plan for the country. OpenMRS, a medical record application, was set up in the main hospital in Freetown, the capital. Weeks before this took place in February 2010, major advancements were made to SDMX-HD, spurred by the fact that it was about to be implemented in two 'live' applications. The main idea was that OpenMRS has the functionalities needed for doctors and nurses to track patients over time, while the aggregate data from this system is used for general health management at hospital, district, and national levels. This two-level use applies to other domains and applications also, such as logistics management, human resource management, lab systems, health finances, and so on. With SDMX-HD set to go live in Sierra Leone at the given date, developers behind DHIS2 and OpenMRS not only pushed development forward by implementing this standard in the respective applications, they also contributed the necessary use-case to solve outstanding issues. The development of SDMX-HD made the HIS architecture of various applications for different business domains, a reality; anyone able to share data on this protocol could now be 'plugged in' in a country HIS.

Good Design

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 the West Africa Health Organisation (WAHO), to pilot open source human resource management software called iHRIS, in Ghana. Learning about the 'Sierra Leone architecture', WAHO took the initiative to include also the iHRIS suite of applications for human resource management in this architecture, which provided further impetus to the iHRIS implementation of SDMX-HD. These initiatives, together with the fact that many countries in the West African region were in the process of introducing the DHIS2 and/or the iHRIS, led to the organisation of a workshop on training in both these applications in Accra, Ghana, September 2010. This workshop was jointly supported by WAHO, HMN, WHO, Capacity*Plus* and HISP. The WHO organised a 'Connectathon' meeting for the SDMX-HD in parallel with the workshop. At the end of the workshop, the SDMX-HD standard was officially launched and implemented in the DHIS2, the iHRIS, OpenMRS, and the WHO Indicator Measurement Registry (IMR) integrated framework. This initiative was further consolidated in November 2010, where HIS staff from all the 15 WAHO member states were present. There, it was decided that HMN, WAHO, HISP, and Capacity*Plus* should collaborate to develop a centre of excellence at WAHO, for supporting member countries in adopting the interoperable solutions which had grown out of Sierra Leone.

Figure 3.1 depicts the three applications being part of piloting the SDMX-HD standard, depicting the relationship between integration and interoperability.

Interoperability is the ability to exchange data between two or more systems. In Figure 3.1, we see exchange of data:

▪ Between DHIS2 and OpenMRS.
▪ Between DHIS2 and iHRIS.

This is where, we have interoperability between, in this case, two systems. There is no interoperability between OpenMRS and IHRIS.

Integration here, can be understood as the process of joining distinct systems in such a way, that they appear as being a whole in a particular perspective. In this case, we see that patient data from OpenMRS and human resource data from iHRIS are 'joined' and integrated in the DHIS2 'data warehouse'. OpenMRS and iHIRS are integrated, but there is no interoperability between them. DHIS2 is interoperable with both OpenMRS and iHRIS.

Figure 3.1 emphasises a technically biased perspective on interoperability, since it is seen strictly as an interaction between two software applications. Such a technical perspective is quite usual and may seem the reason why the concept of interoperability has been mostly used and promoted by those with more technical and/or clinical medical backgrounds, who view the term as being mostly related to a medical records system. However, if we 'unpack' the SDMX-HD interoperability arrow in Figure 3.1, we see that it includes both the SDMX-HD standard for data exchange and the data which is to be exchanged. We may regard these two components and their internal relationship as being similar to the messenger and the message; the container and the content; and between syntax and semantics.

In Figure 3.2, we have depicted the overall integrated 'data warehouse' and the architecture concept: Integration across health programmes and services and across technologies. With Figure 3.1 and its discussion, we have extended the interoperability
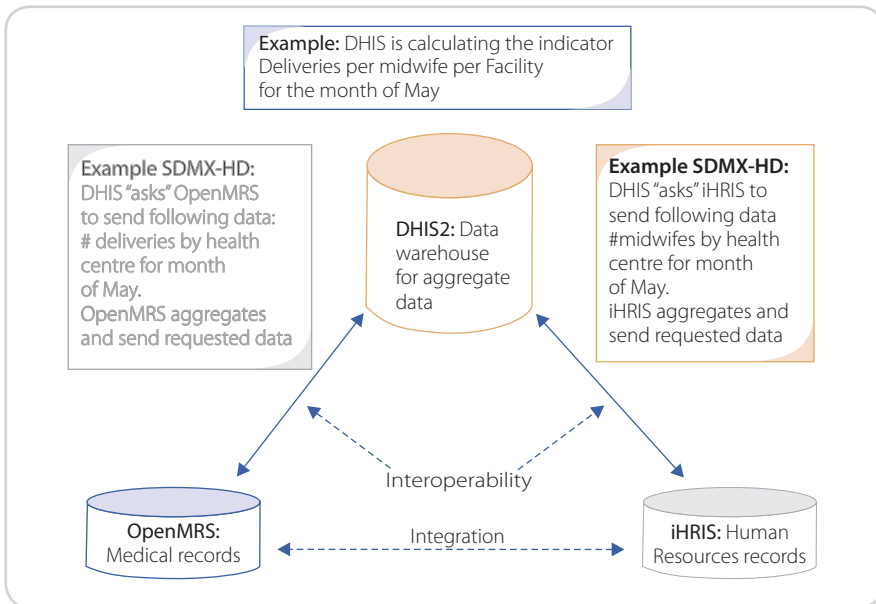
Good Design

Example: DHIS is calculating the indicator Deliveries per midwife per Facility for the month of May

Example SDMX-HD: DHIS "asks" OpenMRS to send following data: # deliveries by health centre for month of May. OpenMRS aggregates and send requested data

DHIS2: Data warehouse for aggregate data

Example SDMX-HD: DHIS "asks" iHRIS to send following data #midwifes by health centre for month of May. iHRIS aggregates and send requested data

Interoperability

Integration

OpenMRS: Medical records

iHRIS: Human Resources records

Figure 3.1 Examples of interoperability and integration

and integration concepts to also include the traditional routine HMIS data reporting, which may be paper-based or computer-based. Data may be reported using the Internet, USB memory sticks, as paper reports, or by mobile phones. The means of data transfer, whether it uses paper or electronic format, physical transport or the Internet, does not affect the use of the terms, as interoperability is not limited to exchange between computer-based systems. Figure 3.2 illustrates that data exchanged between the DHIS 'data warehouse', a range of health programmes and services, and data reported from different types of systems. For example, data generated and transferred from computer-based systems such as medical record systems, human resource management systems, logistics and drug management systems and vital registration; and data generated from paper-based systems such as register books used by health programmes, such as Expanded Programme on Immunisation (EPI), Reproductive and Child Health (RCH), and typically, from Outpatient Department (OPD). While some data is exchanged using SDMX-HD, other data are exchanged using paper-based reporting forms. In principle, however, the data standard remains the same, regardless of the means of exchange; electronic or paper-based. Distinguishing the technical level of data transfer from the data itself is therefore important, when using the concept of interoperability. Furthermore, Figure 3.2 also illustrates that interoperability and data exchange is also as much between people and organisational actors. Thus, the social system dimension is particularly relevant to the understanding of interoperability.

Given the many-faceted picture of interoperability, as discussed in Chapter 2, we have the three levels of organisational, semantic and technical which is analogous to our three local architecture framework.

Looking at the relationship between interoperability and integration, interoperability is a sub-set of integration, and more concrete and directly operational, than what is
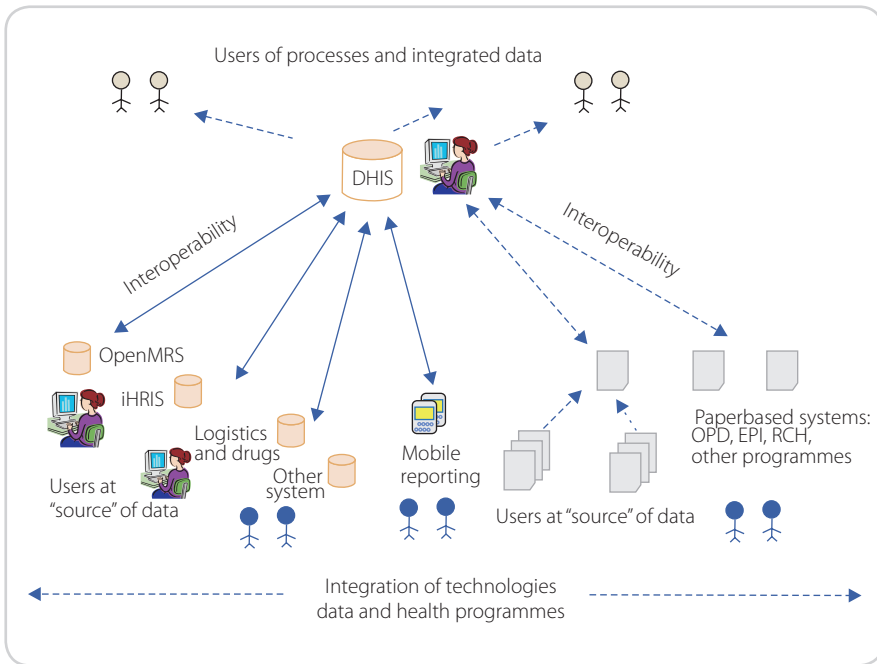
Figure 3.2 Using integrated data

often understood by the more open term of integration. The fact that integration may mean everything also covered by the term interoperability, and that it is also used in such overlapping ways, is a source for confusion. However, if we are explicit about what we mean, or want to achieve, when dealing with concrete system design and development, these differences between professional traditions should not be problematic to sort out.

3.2   Enterprise Integration

In this section, we approach the term integration from the view of the enterprise, and relate it to the overall view of enterprise architecture. *Enterprise integration,* is a topic of research and practice in the sphere of business schools, with many issues relevant to our enterprise architecture debate. For example, information systems and HIS practitioners and researchers will tend to see fragmented and uncoordinated computer-based systems as their problem domain, and integration as the solution through an approach to bring these systems together. At the same time, the Enterprise Integration folks would take the need for better co-ordination of the units and domains of the enterprise as their point of departure, and see integration as an approach to increase co-ordination and efficiency of the enterprise. Then, of course, with the needs of the enterprise as their point of departure, they would address the need for integration at the software application and technical level. These two perspectives might or might not be compatible, or even the same. The key issue is who is defining the integration needs; the 'business users' in the language of Enterprise approaches, or, as we have argued earlier in this book, that in our context would be the health

managers and service providers, who really need to define the integration needs, and not the IT people!

The networked enterprise facing global competition, operating in a context of changing economic conditions, mergers and new business strategies, is the model on which the enterprise integration approach is developing their concepts. The emerging strategy then, is towards enabling the flexibility needed by a networked enterprise or a network of enterprises, characterised by distributed control, inter-organisational business processes, shared information and knowledge, where nodes can be added or removed. For the enterprise then, integration means the development of a coherent information architecture, where the various administrative, business processes and systems are integrated, so that they appear in the most seamless way possible to the users, irrespective of where they are located.

The enterprise architecture model may as well be adapted to the organisational structures, objectives and 'business logic' of the health sector, as it aims for co-ordination, interoperability and efficiency. Since in the enterprise model, the health sector may be seen as being made up of a vast number of more or less interoperable units and sub-hierarchies, or by more or less independent enterprises, networks of enterprises, or sub-enterprises. Furthermore, and importantly, the health sector, as the typical enterprise, abounds in more or less uncoordinated information systems.

3.3   Horizontal and Vertical Integration

The two terms, horizontal and vertical integration, the contrast and the inter-dependence between them, are frequently used concepts in the Enterprise Integration domain, that are highly relevant for understanding integration in the health sector. *Horizontal integration* refers to integration across the various domains or business areas of an organisation, or across organisations. *Vertical integration* is integration typically 'military style' along the line of command from the top to bottom, in the hierarchy of an organisation, or the line of business from its top management down to operational levels. While our objective of developing an integrated architecture across the different health programmes and services, is by definition a quest for horizontal integration; our objective to enable a seamless flow of data between the patient and facility level and higher administrative levels is by definition vertical integration. At both the horizontal and vertical levels, we can employ strategies of *tight* or *loose* integration. A tight integration involves closely coupled systems, using a shared resource, where the failure of one could lead to the failure of the other. In a loose framework, there is more buffer and leeway between the systems, and the failure of one may not directly bring down the other.

Vertical integration is often labelled 'silos', thus emphasising the contrast or even opposition to horizontal integration, as in the HMN slogan 'From silos to integrated HIS', which advocates horizontal integration. Independent and uncoordinated silos, following a vertical integration approach, is the classic way to organise data reporting and information management in health programmes, according to donor agencies' requirements. 'Too much' vertical integration is, thereby, often seen as the main reason for the fragmented HIS in the health sector, as each health programme tends to focus on building their own HIS as silos with little or no interaction horizontally. While acknowledging these negative connotations, in our context, however, vertical integration is predominantly linked to the goal of achieving the most 'seamless' possible

flow of data between the local data sources. For example, the patient-health facility encounter and the various administrative levels above in the health administrative hierarchy. The Sierra Leone SDMX-HD is a perfect example of interoperability and integration, where the OpenMRS database for AIDS patients on ART and the DHIS2 enables a relatively seamless electronic flow of HIV/AIDS data and indicators from the patient and hospital level, to the national level, and further as national reports to WHO in Geneva.

Today, data management at the local facility is conventionally paper-based. In India, the reporting from the Sub Centres and PHCs is currently in the process of change, from paper to electronic reporting, and in some cases enabled through the use of mobile phones from Sub Centres and in others, by accessing the web-based application from the health facilities. In most developing countries, however, reporting from the facilities is currently typically paper-based. The paper-computer interface is therefore important in achieving 'seamless' vertical integration.

At a more general level, vertical integration may also be seen as being parallel to the aggregation of data from the patient or production or transaction levels, for information systems, where further up in the hierarchy of usage and management, only particular sub-sets of the information produced at the local and transaction levels of information hierarchy are useful, or make sense. At the local and operational levels, more granular information is needed for patient management and services delivery, while at higher managerial levels, less granular information or sub-sets of information 'produced' locally, would be needed. The smooth flow, or availability, of information between these levels of granularity is an important characteristic of vertical integration. The following Figure 3.3 contrasts vertical and horizontal integration.
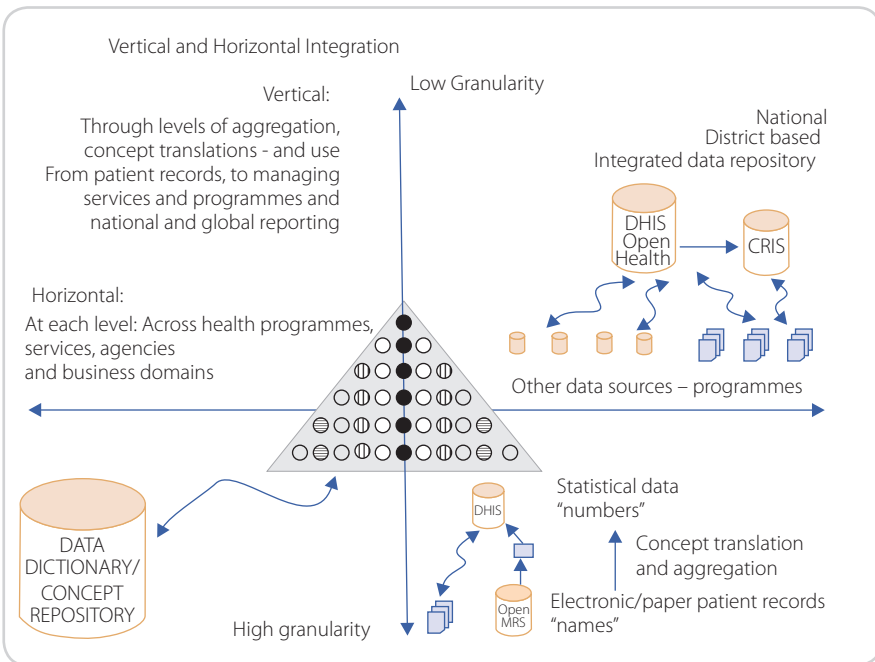

Figure 3.3 Vertical and Horizontal integration

Good Design

Horizontal Integration

Horizontal integration is understood as the integration of the various domains or business areas of an enterprise, or as in our case, the health sector – between the enterprise or the health sector and its environment and wider context of partners, community or public sector. In the health sector, horizontal integration would involve the co-ordination of information flows and information systems across the different health programmes, health services and other relevant data sources, so that all data can be accessed and analysed at 'one point'. The 'data warehouse' approach represents a typical horizontal integration, by the level of hierarchy; national, state, district and facility. The 'data warehouse' enables a horizontal view across the various sectors at each level, where again each sector is part of a vertical flow and integration of data, thus, manifesting a dynamic relationship between vertical and horizontal integration – up, down and sideways in the hierarchy. The Table 3.1 links the concepts of integration within the three-layered structure of the IHIA.

Such a three-layered perspective on integration, will help us guard against the view that integration can be easily achieved. Research points to the contrary – that most

Table 3.1 Horizontal and vertical integration in the health information architectures framework	
Three-level Architecture – Horizontal and Vertical Integration	
Level 1: The organisational, political level of information needs, users and usage	*Horizontal Integration* • Health management's need for information from across different health programmes and health services and health resources (for example, human resources), available at 'one point' of access, for co-ordinated management, M & E, correlation and analysis. • Hospital management's need for information from across wards, specialties, finances, and so on are available at one point.
	Vertical Integration • Managers, at all levels, need for 'seamless' flow of information between the peripheral levels of the hierarchy, the source of information, for example, patient-facility encounters, and district, state/province, national and even global levels.
Level 1 uses services from the level below (level 2)	
Level 2: Software applications and information systems	*Horizontal Integration* • Data warehouse for aggregate data and indicators, such as the DHIS2, integrating and managing data from different health programmes and service areas. Data imported electronically from other systems, or captured directly from paper reports. • Integrating patient-based medical records between different systems, typically using shared repository.
	Vertical Integration • Extracting aggregate data and indicators from medical record systems and loading them into the 'data warehouse' for aggregate data, as illustrated by the integration between OpenMRS and DHIS2, and the flow of these data upward, through the levels of the system. • The still dominant paper-based reporting routines, where aggregate data is extracted from register books and reported upwards through the levels of the health system.

Level 2 uses services from the level below (level 3)	
Level 3: Data exchange and interoperability standards	*Horizontal Integration* • Shared data standards prerequisite for capturing data from paper or importing data electronically to the 'data warehouse'. Harmonisation of datasets between health programmes and services to solve overlaps, inconsistencies and gaps is the first step. • Shared data standards, for example, SDMX-HD, represents a technical standard used for transferring aggregate and indicator data between systems.
	Vertical Integration • Shared data standards prerequisite for the flow of aggregate data reported up through the levels. Extracting aggregate data from medical record systems such as OpenMRS and loading them into the 'data warehouse', such as DHIS is based on the shared data standards and converted into the SDMX-HD technical standard for import into DHIS2.

integration efforts in the domain of HIS in developing countries do not often work, and the reason for this is mostly institutional. For example, donor projects often lead to the development of small fragmented systems, and when the donors leave, legacy of their systems remain. Donors and other stakeholders tend to be reluctant to integrate their systems with anything else, leaving an unhealthy trail of unintegrat(able) systems.

Integration, interoperability and standardisation are intimately linked. In the next section, we illustrate this through cases of developing data standards involving varying approaches. We will argue that standardisation of data and indicators is a pivotal part, and in fact, a precondition of integration efforts.

3.4   Standardisation and its Relation to Interoperability and Integration

Integration and interoperability both build on the use of standards. The alternative to shared standards is to build a custom gateway between each of the systems involved, while being of course possible, gets too complex as more systems are included and existing systems are replaced. Because, each new system will need to connect with all the other systems, and in total requires $n*(n-1)/2$ connecting gateways to be established and maintained. Looking above the level of technical standards for interoperability to the level of data standards, there is no alternative to shared standards. Or more precisely put, the alternative is *chaos* – fragmentation of systems and lack of co-ordination. Obviously therefore, data standards play a crucial role in HIS development generally, and in an architecture particularly.

We discuss here, the concepts of standards, and the process of standardisation as related to the integrated IHIA. In the example and Figure 3.2, standards of different types enabled the flow of data:

▪ The SDMX-HD enabled the technical exchange of data between systems.

▪ Data standards ensured that the sender and the receiver both understood the data being transferred.

▪ Finally, we may say that the standard institutional procedures and legislation ensured that data were collected, processed and reported.

Thus, the three types or levels of data are in line with the levels of interoperability listed above.

We can compare these three levels of standards with a cellular telephone system. The technical standards of transmission and handheld devices enable us to hear the sound of the voice at the other end, the one you are communicating with, and vice versa. The data and semantic standards is about, the meaning of what is being transferred; which can be likened with a telephonic dialogue, where if one speaks Hindi and the other Chinese, and none understands the others, language, not much would come out of the conversation. Therefore, it does not help if the technical standards are working well, unless there are standards for joint understanding, which we call semantic standards – a shared language, in the case of the telephonic conversation. And finally, we will also need standards for agreed procedures of communication and data sharing, at the organisational level. If, for example, two people understand each other perfectly well and are provided with a technically excellent telephone line, it does not help much if they do not want to talk, or have nothing to talk about. Similarly between two information systems, for example, the hospital system and the DHIS2 data warehouse, technical and data standards will not help much if the two organisational structures do not want to share and interchange data.

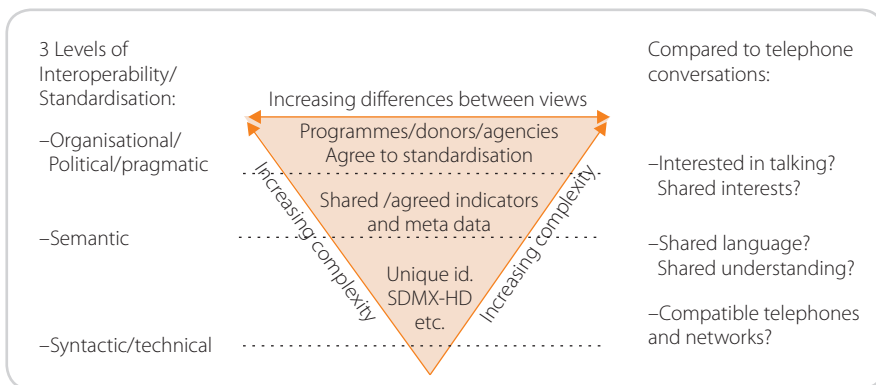In Figure 3.4, the three levels of standards are depicted.

Figure 3.4 Three levels of standardisation of the increasing differences and complexities

The inverted triangular model describes the three levels of standardisation and interoperability, adapted from Carlile (2004), who developed his model in the area of knowledge management to help conceptualise the management of knowledge across boundaries. In our context, the boundary is between those needing to share the standard; and who are engaged in negotiations on agreeing on the standard. The three levels are relating to what type of standards and agreements would be needed. At the syntactic and technical level, they need to agree on technical means of data exchange, as in the SDMX-HD example. At the semantic data level, they need to create shared understanding and agreement on the data to exchange, the datasets and concrete definitions. Finally, at the level of the social system and organisational politics, agreement is needed to share data and to change routines, systems and even ownership and control of data and information. The point with increasing differences

between views, as one move up through the levels, to arrive at a shared understanding and partial agreement, is only at the top level, the world of politics; with social and organisational dynamics. Having reached an agreement at this level, it would be easier to agree on some shared data standards and procedures for exchange. And having agreed on the need to share data and the data to be shared, it becomes relatively less complex to sort out how it can be carried out in practice technically.

Related to the three progressively, more complex boundaries of differences described above – syntactic, semantic and social customs boundaries, there are three progressively, more complex processes of managing these differences. And these processes, in our context, include implementing standards such as for transfer, translation and transformation. We summarise the levels of boundaries and processes needed to overcome them.

3.4.1 Syntactic-Technical Level

Data transfer and interoperability. This level is part of both paper and computer-based systems, and any combination, thereof. The technical level is relatively easy to comprehend. The SDMX-HD standard, enables an example of a syntactic description of data to be transferred, so that it can be understood both by the sender and the receiver. Paper-based systems and standards are often understood as something completely different. It is, therefore, important to note that in a manual system, paper-based registers and data reporting formats are representing the same syntactic level; the paper form is also a 'grammar' for coding the data to be transferred, so that the sender and the receiver understand it in a similar way. Also note, however, that the practical difficulties in changing paper-based reporting forms make up an important driving factor in the fragmentation of HIS and the problems faced by data standardisation. While data standards (like SDMX-HD) are *Software* based, and enables data transfer, implying that it is easy to change data standards, the paper formats are *Hardware*, or not changeable, and will need to be replaced!

At the syntactic-technical level: The process of agreeing on a shared 'grammar', or in our context 'protocols', as we call it, to able transfer of data and information. But, the meaning of what is being transferred is not in question here.

3.4.2 Data-Semantic Level

Meaning and shared understanding. This is the level of standards for data and indicators, data dictionaries and metadata; for example, on procedures for calculating indicators, health facility lists, and coding schemes like ICD10. Data standardisation is the focal point in most standardisation efforts, and in our context of integrated 'data warehouse' architecture.

At the data-semantic level: Processes of translation of interests and views, are involved in reaching shared meaning and understanding; for example, between different health programmes, on what should be the important and shared data and indicators.

3.4.3 Social System – Organisational-Political Level

This is the level of decision-making, using the power to decide standards at, mainly, the semantic level: the data and indicator standards. The standards for interoperability at the syntactic and semantic levels will also be reflected by 'softer' standards at the

Good Design

inter-organisational level, in terms of procedures, mandates, responsibilities and job-descriptions needed in order to effectuate the other standards.

At the social system and organisational level: Processes of transformation of, in fact, social and organisational structures, in order to align the interests between the actors, and in its' process; so as to accommodate new ways of sharing information across organisational boundaries, implementing new systems and new routines, and shared by several health programmes.

Boundary objects can be used by both sides of the boundary of differences, as a means of negotiating, arriving at, communicating and maintaining shared understanding. For example, they can mediate between two health programmes, working on developing shared routines for collecting and using information. In our context, the various information systems resources, such as the DHIS2, shared by different parts of the health system represents such a boundary object. The example of developing the SDMX-HD standard illustrates the role of DHIS as a boundary object. Demonstrating that data could be shared between DHIS2 'data warehouse' and the iHRIS human resource management system, allows the actors to agree on useful data and indicators to be shared; for example, the provision of indicators on number of patients served by staff for different services and areas. Concrete examples like this, again, make it possible for the different departments and health programmes to explore modes of collaboration, such as between the human resource department of WAHO and the general information unit there.

This example brings us to the most important part of the proposed framework, the need for processes of iterative cycling through the levels, to step-by-step develop a common understanding and alignment of interests. In the example of SDMX-HD, the first cycle has only been on the piloting stage, and for very limited datasets. Numerous cycles will be needed for standards, interoperability and integration in the area of information systems, shared between human resource management and the overall management of health services. And when segments of the interoperability have matured and become 'common sense', there will be new frontiers of piloting and development. Generally, repetitive cycles are needed because consequences cannot be seen and addressed with one try. New agreements and changes are developed through a process. Shared artefacts and methods play an important role as boundary objects, in providing the capacity to negotiate interests and transform knowledge. By cycling through the levels, the aim is to gradually negotiate and solve differences in interests and shared understanding, at the different levels.

An important characteristic with this cyclic prototyping approach is that, generally, problems at the 'higher' semantic and pragmatic levels will need solutions at the levels below in order to be solved. Adapting this strategy of 'cyclic evolution' of standards in the health sector, we may say that the differences in interests between, for example, two health programmes on content of data reporting and ownership to information can be addressed. This will involve the need to negotiate the particular data and indicators to be shared, and include it in the standard setting, at the semantic level, based on practical implementations of how to actually do it at the technical/syntactic level. To take the SDMX-HD example; the ability to solve the interoperability problem, caused the action of defining the data and indicators to be shared at the 'higher' data-semantic level, which triggered the processes of transformation at the 'highest' organisational level, and so the cycles continue.

This cycling between levels is illustrated through an example from Malawi (see Box 3.2). In Malawi, for quite some time, the Expanded Programme on Immunisation (EPI) and the Ministry of Health, were aware of the significant discrepancies in data quality and actual values, for immunised children between their two data collection systems. At the organisational level, they were eager to find a solution, but had no practical way to join or collaborate the systems. The first approach was only to regard the data from EPI to be of best quality, and therefore to be used for official purposes. Only when a practical solution was found for standardisation and co-ordination at the syntactic-technical level, which included the shared data warehouse,

Box 3.2	Designing and implementing an integrated HIS in Malawi

Standardising EPI datasets in Malawi

During 2000–2002, Malawi went through a HMIS reform, where, following the South African learning, the integration efforts focused on developing minimal datasets consisting of the most important data elements from all health programmes. Paper-based data collection forms were designed for reporting from the health facilities to the districts, and in the districts, the DHIS1.3, was implemented for data capturing and onwards reporting to the national level. Data collection forms and software were implemented country wide from 2002. Given the poor infrastructure in Malawi, data was only reported from the health facilities on a quarterly basis. However, since monthly frequency is regarded to be a better option in order to ensure data quality, the quarterly reports contained the data for each of the three months.

When the reformed HMIS was implemented, it was a good match between both the perceived and the real needs of the users and the information provided by the HMIS. As time passed by, however, the needs of the health programmes developed or changed, while the HMIS stood still, as if cast in stone. They could not adapt to changes over time, where new requirements were not included from users as the needs of the health programmes changed over time. The inability to 'change with the tide' caused users to regard the HMIS, as not fulfilling their needs.

The EPI, for example, re-established their own separate reporting system, as they claimed that the HMIS did not provide them the information they needed. While the HMIS collected only a sub-set of the vaccines provided, for example, DPT1 and DPT3, and not DPT2, the EPI collected all. While HMIS reported data on a quarterly basis, EPI did it on a monthly basis. Furthermore, the data reported through the two channels, EPI and HMIS, differed. While EPI was well resourced with staff and vehicles, and had established their own procedures for data collection from all points of vaccination including outreach clinics in their catchment area each month; the HMIS was less resourced and relied upon the health facilities own reporting, including of the outreach clinics, to the districts, every quarter. The HMIS routines for collecting vaccine data was, thus, less 'thorough' than those of the EPI. Consequently, EPI statistics showed better coverage of their data collection, with 'more data, higher values'. The variations between the figures reported by HMIS and those reported by EPI Programme, together, with the general assumption that the EPI data was of higher quality, resulted in the general sentiments expressed by a range of users, that the data quality was better in the vertical programmes on reporting structures of the EPI, than that of the HMIS.

joint data collection, paper forms and data capture. Only after this solution, could they arrive at shared information management, including the data, indicators and immunisation, which represented standardisation and co-ordination at the semantic level. Finally, co-ordinated and shared management of the information at the semantic level, enabled further co-ordination and sharing at the organisational level. This co-ordination manifested through shared and integrated procedures for various aspects of data collection, information management and shared co-ordination of the official immunisation statistics, for Malawi.

This example from Malawi indicates that practical and flexible solutions at the technical-syntactic level greatly enable harmonisation and standardisation of data, indicators and general information management at the semantic level, which again enable the harmonisation of organisational information procedures. The cycling between these three levels also represents a process of learning by doing.

General Aspects of Standardisation

The evolutionary 'rule of nature' approach to standardisation, advocated through this cyclic model, is in line with contemporary research on how standards actually develop. Standards are generally seen to evolve through use and by solutions, that over time become 'attractors,' that is, gaining momentum by attracting more and more users, and becoming more and more of 'global' standards. Bureaucratic committee based standard-making, is more or less futile, because the process is too slow, while the world is changing rapidly. Besides, emphasis is on organisational-political considerations rather than the flexibility of use. Standards need to be flexible, in order to adapt to a changing world, to remain 'attractors'.

Flexibility allows for standards to be useful, to gain momentum, to grow and to endure. Flexibility can be understood across two-dimensions, termed, use-flexibility and change-flexibility. How can a standard be flexible, is not a standard something that is fixed? Yes and No. For example, the new European Union standard requires all chargers for mobile phones to be interchangeable, that is to be similar in all practical ways, same plugs, connectors, cables, voltage, and so on, and would be regarded as a standard; fixed and rigid. While it has a high level of use-flexibility for all mobile phone users in all contexts; its change-flexibility is nil! In this case, the change-flexibility is low and its use-flexibility is high. In other cases, it may be the high change-flexibility that makes a standard and also, flexible to use. The SDMX-HD standard, for example, which provides the interoperability between the various systems and the data warehouse in our IHIA, is only useful for this purpose, because of its ease to change and adapt to variety of data standards needed to be transmitted or reported, from various systems. If we compare data standards for electronic reporting of data, with the traditional standard for data reporting; the paper form, the difference in terms of change-flexibility becomes obvious. The paper form is like 'cast in stone', in that it is form and content combined, the actual data and the view of the data tied together in the 'pen and the ink' and it is impossible to change. Contrary to the paper form, in the SDMX-HD standard (or any XML format really) the data is represented separate from any representation of it. Therefore, it is flexible in terms of changes in the data to be transferred, contrary to the paper form.

In the next section, we describe cases and practical approaches to data standardisation, using what we have labelled the 'data warehouse' approach.

3.5   Data Standardisation, Principles and Cases – A Pivotal Part of the Integrated Health Information Architectures

Developing data and indicator standards in health and related areas, represents complicated processes, as different areas of health services, health programmes, donors and governmental structures, such as states and provinces, will by definition have different views and interests. Standardisation seems to indicate more or less by default, more control being handed over to those higher up in the hierarchy, leaving less flexibility for those situated in the lower levels to pursue their own interests and views. Similarly, the push for integration will tend to imply increased co-ordination, and as for standardisation, more control to those higher up and less independence, for example, the individual health programmes.

The principle of *Hierarchy of Standards*, to be described next, represents an effort to address both:

- *The Paradox of Standardisation*: How to provide simultaneously both the flexibility and control needed to guide the standardisation processes.

  and

- *The Paradox of Integration*: How to provide both co-ordination and independence between actors, such as health programmes.

The rationale for the principle of *Hierarchy of Standards* is, that in the area of health data and information, the lower levels of the health services will generally need more detailed data – that is, larger datasets or longer 'lists' of data – in order to support their day-to-day work, than higher levels, in order to support their needs for co-ordination and overall management. For example, while the health facility needs a register of local Tuberculosis patients to ensure proper treatment of each individual patient, the district will only need, say, the number and percentage of patients fully cured or treatment defaulters, in their overall management of the Tuberculosis programme. The principle of the *Hierarchy of Standards,* schematically depicted in the Figure 3.5, as a pyramid, implies the lowest point of entry requires additional and more granular data (most disaggregated), whereas higher the level, more aggregated or lesser granular data.

The pyramid depicting the hierarchy of standards, also represents horizontal and vertical integration, as described in the earlier section on integration. While vertically, the levels of the pyramid represent a sub-set of the levels below and thereby illustrate vertical integration; horizontally, the pyramid consists of distinct datasets that maintain their independence, at the same time are (horizontally) integrated in a shared framework.

We describe the development of this hierarchy principle in practice, from our seminal experience in South Africa. This was developed by the HISP team, as a part of this process of developing the district-based HIS and the DHIS software application. The hierarchy principle is key, to the design on which the DHIS is also based and the general data warehouse design, described in this book. Flexibility is ensured by allowing for 'freedom' horizontally in the health system hierarchy, as long as standards are maintained vertically. The central core of the pyramid shows, the 'essential data and indicators' required by the national and global levels which all levels below must report, and more generally, every level below has the freedom to add to that core representing their local needs. Horizontally in the hierarchy, multiple domains of data, or datasets, may be envisioned. So, while a state can add some data, they cannot
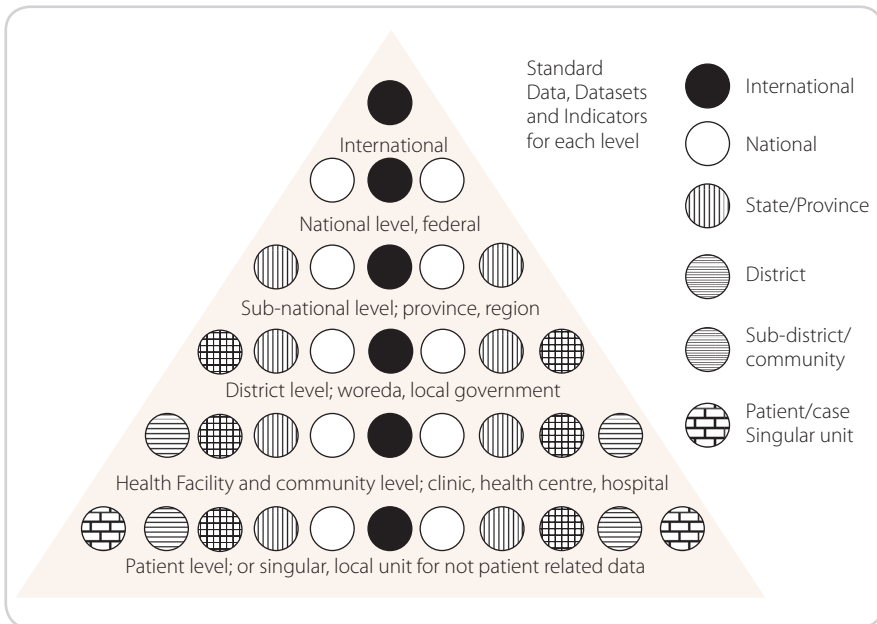
Good Design

Figure 3.5 Hierarchy of standards

delete what the national level needs, and the added data need not be sent to the national level, but only used for state specific purposes. Similarly, the level below the State (the district) must report on what the State and National levels wants, and in addition can incorporate other data, essential to the local needs. Similar logic holds for the levels below. This approach of 'flexible standards', allows for standards to evolve in a locally relevant manner. This mechanism for local empowerment enables users to view the HIS, as not a tool for top-down control facilitating upward reporting, but as a tool to facilitate local relevance and action. Following this principle, horizontal integration is enabled by, for example, multiple health programmes that can maintain their independence, while at the same time are integrated through the shared management of their datasets in the data warehouse and/or co-ordinating upward reporting.

Case Study 1: The South African Example

In 1994, as a legacy of apartheid, the new South Africa inherited one of the least equitable health care systems in the world, where 60% of the resources were used by the private sector; serving only about 20% of the population. In order to rectify the wrongs of the past, the new government launched the Reconstruction and Development Programme (RDP), with a title that clearly expresses its intent – The reconstruction and development of communities that suffered under apartheid. The restructuring of the health sector is based on a decentralised system of health districts. Development of a new HIS, to support the restructuring of the health sector was part of this plan.

During apartheid (1948–1993), the health services were extremely fragmented by race, type of service, and the system of 'homelands'. Until May 1994, there were 14 departments of health at the central level: the 'general' Department of National Health

and Population Development, 3 specific 'white', 'Asian,' and 'coloured' administrations, and 10 for 'blacks', 'homelands', and 'self-governing states'.

As a consequence of this fragmentation, there were no comprehensive national standards for data collection, and each province used different datasets, definitions, and standards. Furthermore, apartheid was characterised by extreme inequity in health services provision and health status, between populations and racial groups. Equity in health services provision and health status had therefore become a major political target, in post-apartheid South Africa. The important issue, in this case, was the striving for equity between geographical areas and racial groups, required a system of national standards to measure and monitor the extent to which this policy was achieved, to pinpoint areas where more resources and efforts were needed.

The Health Information Systems Programme (HISP), was initiated as part of the RDP in 1995, with the aim to develop integrated district-based information systems, supporting the new emerging decentralised structures in South Africa. Starting in 3 pilot districts in Cape Town, the focus was on strategising development of shared data standards as a way to counteract the extreme fragmentation of health services and health programmes, serving different population groups that made up the legacy of apartheid. Figure 3.6 illustrates the typical situation in a district in South Africa, where a multitude of different services, all reported to different head offices outside the district, with no information shared or co-ordinated within the district.

The strategy for moving from a fragmented centralised to a decentralised integrated information and management system, is depicted in Figure 3.7, from Atlantis sub-district outside Cape Town; creating a repository for all information being reported
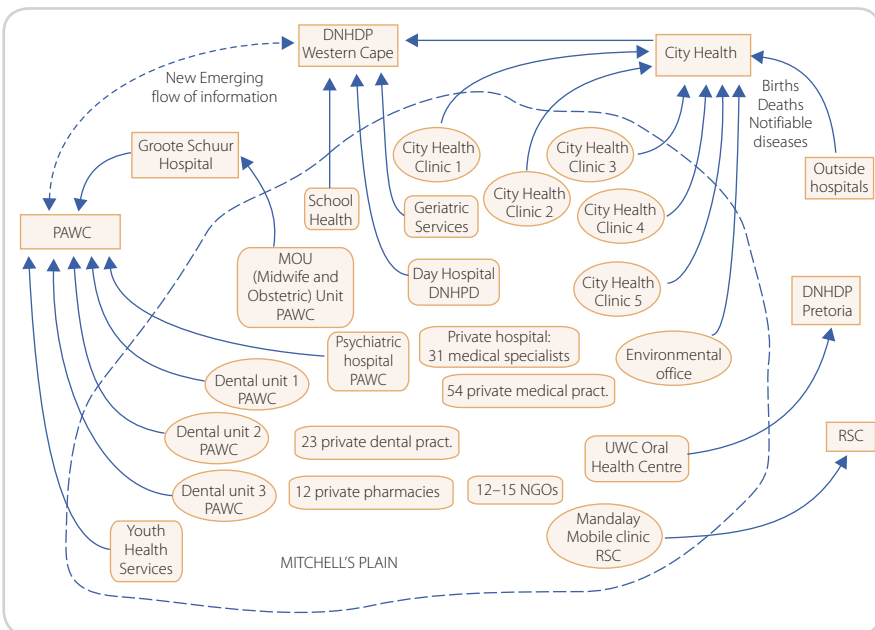

Figure 3.6   Legacy of apartheid; fragmented and centralised data flows, Mitchell's plain district in Cape Town. The figure featured in the document proposing to establish HISP written by the Western Cape Strategic Management Committee, 1994.
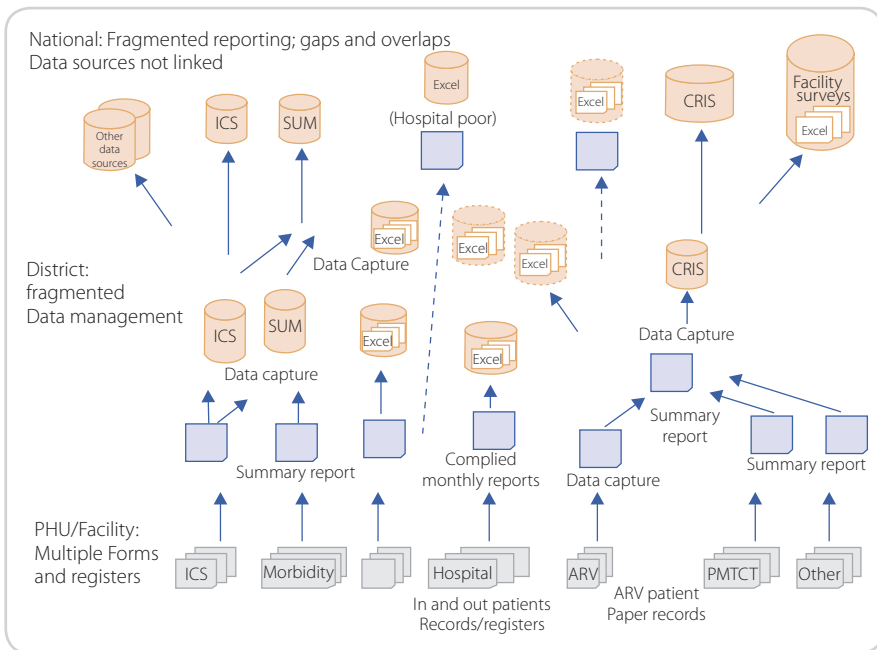
Good Design

National: Fragmented reporting; gaps and overlaps
Data sources not linked

Other data sources

ICS   SUM

Excel
(Hospital poor)

Excel

CRIS

Facility surveys
Excel

District: fragmented Data management

Data Capture

Excel   Excel

CRIS

ICS   SUM

Excel

Data capture

Excel

Data Capture

Summary report

Summary report

Complied monthly reports

Data capture

Summary report

PHU/Facility: Multiple Forms and registers

ICS   Morbidity   Hospital   ARV   PMTCT   Other

In and out patients Records/registers

ARV patient Paper records

Figure 3.7 Prior to interventions, fragmented helth information systems in Sierra Leone

from the district within the district, and shared so that it could be used by all the stakeholders. In this book, we label this as the 'data warehouse strategy'.

HISP was responsible for efforts towards standardisation and software development. Standardisation efforts aimed at developing minimal datasets – a uniform and minimal set of data elements (a dataset), with clear definitions to be reported by all health facilities, using simple paper reporting forms. These were essential so that the data elements could cover the key information needs across, and calculate the most important indicators. Implicit in this approach, was the reporting on minimal dataset to be established, in addition to the data already being collected by the various health programmes and organisational structures. Software development and prototyping efforts led to the creation of the first version of DHIS application, which supported the implementation and use of the minimum datasets.

In order to reach consensus on datasets, widespread negotiations and consultations with different health programmes and services were carried out, starting from the Western Cape Province. After about 9 months of intensive negotiations, driven by local managers in collaboration with the HISP team, the first essential dataset was implemented in all the local government health facilities in the Cape Metropole, in May 1997, and later in the whole of Western Cape. The first version of the DHIS was implemented and used, to capture and analyse monthly data at district, regional, and provincial levels, in Western Cape, from 1998. Simultaneously, in the Eastern Cape Province, a unified monthly minimal dataset was implemented as from January 1998, in all primary health care centres. The DHIS software application was used to manage the new minimum datasets in the two provinces. Interestingly, the two minimum

datasets resulted from two different processes in two different provinces, and were very different with only about 50% overlap. The DHIS had already, at that time, incorporated a flexible metadata structure, that could be used to manage multiple datasets; where the data elements that were the same in both datasets, could be seen as a 'shared' core dataset. The standards that where specific to each province could be seen as a locally defined extension of the shared core, and the concept of organising the datasets in a hierarchy, was born.

In the situation, where there were no shared national standards, and in provinces; no shared provincial standards, the two provinces had developed their own shared standards, followed similar approaches and were managed by the same software application, appearing to be a significant success. From 1999, the DHIS and HISP approaches, were endorsed as official national approaches, and projects to establish provincial datasets and agreed processes were started in all provinces. The first national essential dataset was agreed upon in June 2000. While all the provinces maintained their own extended datasets, the national dataset made up the shared core, where all the provinces needed to collect and report on. This dataset was revised several times since then. Over time, additional data elements were added, either as a process of expanding an existing programme dataset, or to accommodate new vertical programmes (for example, HIV/AIDS programmes). By 2005, it was expanded to become a National Indicator Dataset (NIDS), reflecting the increased focus on the use of indicators, as compared to the earlier focus on data elements.

While some health programmes, such as, within the HIV/AIDS segment, maintained their own full-fledged systems and only provided extracts to the NIDS, and was managed by the DHIS; others, such as the EPI, was fully included in the essential dataset and DHIS2 routines.

In 2000, EPI was collecting 44 data elements through their own system, whereas 5 data elements were collected through the minimal dataset. They then realised that the completeness and overall data quality was much higher for the few data elements being collected through the national 'minimal' system and DHIS, than for the much larger dataset they collected themselves. As a consequence, the EPI reduced their data elements from 44 to 13, and included them in the national reporting system. In this way, the national system grew in strength and quality, and was able to provide to 'everybody' with better data than they had, when each one of them maintained their own system.

Case Study 2: The Indian Example

From around 2000, the South African approach to integration by using unified dataset and reporting tools, was applied in other countries, typically ones with smaller scale settings. The example of India illustrates how these bottom-up approaches need to be modified and adapted to a more stable and entrenched bureaucratic structure, over a much larger geographical scope.

The HISP initiative from South Africa was adopted first in India, in December 1999, in the state of Andhra Pradesh in the South. Initially, the HISP team, inspired by the South African experience, and may be a little naïve about the 'heavy hand' of the Indian centre, tried to develop minimum datasets following a situation analysis of 12 pilot clinics in one district of the State. However, these attempts were largely unsuccessful as the national mandated datasets could not be modified by the levels below. Even

though health is a state subject in India, historically, the data collection formats have been largely governed by nationally defined vertical health programmes, in which international donors have a significant influence. Unable to make a dent in this process of modification, HISP India then adopted an approach of automating existing formats on an 'as is' basis, hoping that once HISP and DHIS is more deeply entrenched, they may have some leeway in influencing the design of the datasets.

However, the above opportunity did not come through easily, and after 5 years of struggle, HISP was terminated, as the state government in Andhra Pradesh had changed following the elections which also had the fall out of a change in the bureaucratic structures. HISP moved first to Kerala state, and then other states such as Gujarat, Jharkhand, and Madhya Pradesh. In each of the states, it was found that while the state health secretaries agreed to the logic of the 'Minimum Dataset' approach, when it actually came to making concrete changes in what data gets collected and why, they would back off saying that 'we cannot change as the national level wants this data.' So, in all the cases invariably, HISP India designed the respective state applications on a 'as is' basis implying the collection of nearly 2500–3000 data elements on a monthly basis. However, by making data from across programme areas and for the whole state available for analysis, through the DHIS application, important lessons on data quality and the eventual usefulness of the various data were drawn by a multitude of stakeholders; lessons that later provided valuable input in the national reform process.

In response to the grim state in the country of public health indicators, despite the achievements in the economic growth sphere, in 2005 the National Rural Health Mission (NRHM) was launched. The NRHM had the explicit aim of bringing about architectural corrections, within a health systems framework, in order to make health care, especially to the marginalised poor – more equitable, affordable and accountable. Amongst the health systems identified for introducing these architectural corrections was the area of HIS, including making these systems more integrated, decentralised and supporting of evidence-based decision-making.

While there had been previously (prior to 2008) various attempts both through the national level and various states to carry out HIS related reforms, systems remained fragmented and data driven. In 2008, another initiative was taken on re-design, which took technical inputs from the HISP India team, and also another public health expert who had been deeply involved in the HISP process in South Africa. The re-design started with a situation analysis of the HMIS of 4 to 5 states to identify the existing systemic problems. This analysis led to an inductively derived set of design principles, which could be applied to HIS re-design including to datasets, reporting formats, indicator lists, and feedback cycles, as well as, the DHIS 'data warehouse' application.

In carrying out the situation analysis, HMIS data from some states was analysed for 2007–2008. The problems identified across these states were quite uniform; a very high number of data elements being collected, of which a high number of zeros or blanks were reported (ranging from 50 to 60%). While a lot of data elements were being routinely collected, less than 5% of them were actually being used for the generation of indicators, reflecting very poor utilisation of information for action. Further, a lot of disaggregated data was being collected (for example, each data element being broken up by caste groups – Schedule Castes, Schedule Tribes, and Others) on a routine basis, which would arguably be better off being collected on an

annual basis through surveys. Another systemic challenge which existed in the HMIS, was the historical practice, of field nurses collecting 'area-based data', which involved them going to the houses and collecting information, such as, if a delivery has taken place, and recording it even though it may have taken place in a facility (like at a hospital or medical college) that may even be outside the jurisdiction of the nurse. A problem of duplication would then arise, as the same delivery noted by the nurse, would have also been recorded by the hospital or medical college, where the delivery physically took place.

To address these identified problems at a systemic level, HMIS guiding principles were inductively derived to help design, based on the overall philosophy of developing a HMIS that could support the broader NRHM agenda of decentralisation, integration and use of information for action. These principles included reducing duplication of reporting, use of surveys for collecting data not used routinely, build the HMIS on indicators rather than raw numbers, and establish a hierarchy of indicators for the different levels, as in South Africa.

After a long period of intensive negotiation across the different divisions, in the Ministry of Health, some decisions on re-design were taken, which led to a nearly 80 to 90% reduction of data elements to what existed earlier. Two of the national programmes (Immunisation and Blindness Control) were integrated into the existing routine HMIS. With other national programmes (for example, Tuberculosis and Malaria) basically adopted the strategy of 'wait and watch' to see how the new HMIS would fare, before deciding whether or not they would agree to integrate. Till then, it was decided to adopt a strategy of 'reverse integration', where the core indicators required from the disease programmes would be imported from their systems, electronically or manually, as the case may be into the national HMIS.

In this way, the national HMIS was re-designed with the aim of simplification, integration and also decentralisation. *Simplification* came in the form of reduction of data elements, the creation of facility specific datasets, and clearly distinguishing between recording and reporting formats. *Integration* came with the incorporation of datasets from two of the existing vertical programmes into the routine HMIS, and agreeing on a phased strategy for the other programmes. *Decentralisation* was designed by reducing the workload of field workers, with respect to the amount of data they had to collect; and setting up an environment where they could focus more on how they could use indicators to manage their everyday processes. The entire set of new recording and reporting formats were then under signature from the then Mission Director, NRHM, sent to all the states with the instructions that they should immediately switch over to the new formats, and all the existing formats (except those disease specific programmes, not yet integrated in the HMIS), would need to be discontinued with immediate effect.

The Indian example, in contrast to the South African case, was a top-down national level driven initiative. While in South Africa, the climate and political will was to create 'new systems'; in India, it was an attempt to make changes within a deeply entrenched system, in which various stakeholders preferred the 'status quo'. Standards established in India included the data elements, datasets, reporting formats and reporting relationships – a mix of technical and practice related. The case of standardisation in South Africa, and to some extent also in the case of India, emphasised flexibility in standard setting *vertically* in the hierarchy, analogue to what we have labelled vertical integration. For example, integrating the line of management from the policy setting and managerial

Good Design

levels, to the peripheral operational levels. In the standard setting, you are free to add your own standards, for your own domain, which may include levels below, as long as you adhere to the standards of the level above. Vertical integration is very much about managing levels of granularity and levels of aggregation. In the India case, what could not be effectively agreed was on the hierarchy of information, with the national statistics division insisting on receiving all the data that was being collected at the facility level. This insistence on receiving 'monitoring' instead of 'evaluation' data, which by design could not be useful to convert to action.

Case Study 3: The Sierra Leone case

In 2002, the civil war that had raged Sierra Leone since 1990 was declared over. Having cost 50,000 lives, and made many more refugees, the war had a devastating impact on the small West African country, which is currently ranked among the *last* on the United Nation's development index.

Sierra Leone is the HMN pilot country, for HMIS. The HMN formed a strong partnership with the Ministry of Health in 2005, thus, forming a solid political base for the integration project. After an initial HIS assessment revealed substantial challenges in relation to a myriad of data reporting structures, and weak ministerial data management, HISP got involved in developing a detailed plan of action during the last half of 2007, the implementation of which commenced in early 2008. Two efforts to integrate data reporting, had been carried out the previous two years, each of them consisted of the development of a new paper reporting form that included parts of other forms and datasets; and a computer database, to capture the data. The most recent integration effort was from early 2007, which focused on Mother and Child Health and immunisation services, and came with a customised database application based on EpiInfo. It was partly overlapping the preceding integration effort, which also included disease data, still in use, with the database application. In addition, came several other forms overlapping both of the initiatives. Numerous errors and inconsistencies were identified, but the developers had left the country, and the effort had no more funds. Furthermore, the UNAIDS CRIS application for HIV/AIDS data was implemented in all districts, but also overlapped with other systems on Mother and Child Health data.

The situation by early 2008 was, thus, that in a typical district information office, there were three software applications in little use, each capturing datasets overlapping with all the others, no communication between the applications, and many more data collection forms not captured by any software. 'Isn't it possible to get all this data into one database? Can DHIS do that?' were the initial questions, we were asked by the Ministry of Health, when we started the planning in 2007. In order to respond to this request, we started to set up the DHIS, as a data warehouse to be used at the district level, which could extract and load data from the three other systems. This rather complex approach, given the state of the technical infrastructure at the district level, was selected, in order not to offend the various system owners. However, as things developed during 2008, it turned out, that most districts stopped using the HIV/AIDS system CRIS; the HIV/AIDS staff employed by UNAIDS got on several occasions, the salary of the information officers, who then did not want to enter the data themselves; 'why do not they do it themselves, they do not have that much work to do!', was a statement made by one information officer. Also the EpiInfo project ran out of funding and stopped. As a consequence, the decision was taken to replace all applications with the DHIS2, as it was rolled out to the districts.

Given the fragmented, overlapping and inconsistent character of the existing datasets, a major effort went into designing a data structure in the DHIS2, that enabled the capturing of all the data from the paper forms in use through a computer user interface mimicking each of the forms, while at the same time solving the inconsistencies behind the scene, in the database. Overlapping data that had been entered for one paper form using the corresponding 'computer form' would then show up in the other computer forms it belonged to, without having to be entered into the database again. A consistent maximum dataset had been derived from the multitude of overlapping forms.

Following a 2 times 3 weeks training, of all district and national information officers in May and June, 2008, a major effort was directed to get all the data captured in the district, to be reported electronically to the national database. The most important result had been the documentation of the quality of data including the completeness. The most advanced district information officers were instrumental in spreading the lesson that, 'we can get all the data from all the forms into one database, why do we still have all these overlapping and complicated paper forms?' For the first time, all the data collected were available in one database, which again convinced all stakeholders that it was actually possible to harmonise and integrate the data reporting from all programmes. A revised set of data collection forms was drafted and circulated among stakeholders for feedback; alongside a drafted list of indicators with definitions of the data, needed to calculate them. In January 2009, a workshop with all the important stakeholders, programme managers and district representatives, agreed on revised and integrated data collection tools and data and indicator sets. Given that there were some refinements to these collection tools, and the time it took to print and distribute them, these new forms were in effect in all facilities, from January 2010.

As the new data collection forms were developed over the summer of 2009, the project started to approach the various international agencies, involved in health service delivery in Sierra Leone. The districts, which for the first time not just had most of their data available in one data warehouse, but also had electronic tools to ease the analysis and dissemination of this data, had been given extensive training in the use of information. Some best performing districts had started to engage civil society to improve their health indicators, a process that was fuelled by the national level publishing league tables, ranking the districts on key indicators. UNICEF was the first major international organisation that pledged to support the data warehouse structure, and drop their own parallel reporting systems, with others also following suit. This integration at the organisational level was only made possible by creating an attractor; a system that had proven results in data completeness and quality. By the second quarter of 2009, the percentage of facilities that had reported the main data collection form, reached above 90% for 7 out of 13 districts, up from 4 districts the quarter before. It was exactly this proven ability to increase data completeness that convinced the health development partners to support the DHIS2. This was not just the software by itself that increased completeness, it was supported by the whole enabling structure of capacity building, information products, and quite aggressive dissemination of feedback to districts to create an environment of friendly competition and peer pressure. All these efforts contributed to making the health development partners to accept the standards implemented in the software.

The Sierra Leone example shows the power of the whole 'data warehouse' approach to develop data standards in practice. With firstly, enabling the pulling in of all data

from different sources into the DHIS2, and then using its analytical capabilities to make visible the problems of redundancies and gaps; a stronger motivation and buy in, for the stakeholders to agree on data standards.

The three cases; all demonstrate that defining and implementing standards is a socio-political process, far from being a neat technical exercise. Standards cannot be dropped down from the top, but need to emerge in and through practice. The most effective form of practice seems to be the ability to make visible the problems, and allow the stakeholders themselves to come to an agreement on the need to standardise. This allows for ownership to develop, which is essential for the standards to get embedded in practice.

Good Design

Summary

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

1. Integration is not about creating one big technical system, as user perspectives will always vary. The definition of integration in the organisational context, takes the users needs of the HIS, the purposes of the HIS, and the wider organisational perspectives as points of departure, and relate those to goals of better efficiency, effectiveness and co-ordination in organisations and enterprises. In such a perspective, interoperability is one of the means to achieve such ends – integration.

2. Interoperability, contrary to integration, is a term with a more formal definition. In our context, interoperability refers to the ability of a system to use and share information or functionality, of another system by adhering to common standards.

3. The SDMX-HD standard being carried forward by WHO describes the essence of interoperability. Sharing data requires two 'levels' of standards, a shared format or protocol and data definition and meaning. SDMX.HD provides the protocols for sharing data.

4. A 'data warehouse' will enable both integration and interoperability of data. iHRIS and OpenMRS data interoperate with DHIS2, and the DHIS2 enables the integration of iHRIS and OpenMRS, by providing a common platform to join their data.

5. The European interoperability framework is described over three levels: Organisational, Semantic and Technical.

6. Integration can be conceptualised across horizontal and vertical dimensions. While horizontal involves integrating further business domains, vertical is concerned with levels of hierarchy.

7. Standardisation is fundamental and a pre-condition to both integration and interoperability.

8. Standardisation is a technical-institutional process, which takes place in everyday practice.

9. The principle of hierarchy of standards, provides a practical framework in which, standards can be created in the context of diversity, and provides a means of flexibility, as against central control, which standards imply.

Reference

Carlile, P. (2004) Transferring, translating and transforming: An integrative framework for managing knowledge across boundaries, *Organization Science,* 15, 555–568.