Obstacles to Integration

annex 9

Introduction

The use of mHealth and eHealth tools is rapidly expanding, and is anticipated to accelerate into the foreseeable future, creating a significant opportunity to improve health outcomes. When effectively implemented, integration facilitates improved data analysis, data sharing, and data interactions between systems. Increased data utilization and sharing enhances public health surveillance and monitoring and evaluation, which in turn augment routine healthcare and interventions to improve outcomes. Integration also promotes a coordinated strategy for information, communication, and technology (ICT) that can increase system efficiency and reduce costs for an implementing agency, country, or program. Despite these benefits, challenges exist in the ability to successfully integrate multiple mHealth and eHealth tools. The goal of this annex is to provide an understanding of the obstacles to integration and to offer an actionable framework to guide conversations between partners on systems integration.

Initially, it is important to understand the terms “interoperability” and “integration”, which are used loosely and in a variety of ways, often creating more confusion than clarity. A common definition of these terms is helpful in addressing the challenges and opportunities that exist and to successfully integrate diverse mHealth and eHealth systems.

Integration is the process of physically or functionally linking multiple sub-systems or disparate data to create a combined system or unified solution. Among health information system projects and professionals, integration typically refers to data integration. Additional types of integration, apart from those pertaining to data systems, may also help meet organizational goals and objectives. (Trowbridge et al., 2004)

A group of community health workers in Bihar, India

Types of integration include:

1. **Data Integration:** the combination or exchange of data from multiple sources into a tool or platform that uses it for transactional or analytical purposes. For instance, using an application programming interface (API) to share de-identified transactional data of disease incidence that are obtained from multiple facility-based health information systems, is a type of data integration related to public health surveillance. Another example could be the propagation of reference data such as lists of facilities, health providers and terminology codes so that other systems don't have to independently maintain those datasets. Common approaches to data integration involve the creation of data marts and warehouses, data synchronization, and data publish-subscribe services. (MSDN Data Integration, 2004)

2. **Process Integration:** the understanding, management, and coordination of individual processes among multiple systems to achieve complex objectives. For example, a laboratory identifies a positive test result; consequently, processes within separate but integrated systems -- such as the physician's electronic medical record (EMR), a frontline health worker (FHW) based mHealth platform, and an SMS reminder platform for the patient -- work in coordination with each other. Typically, process integration requires the definition of clear business rules and a system's operational processes that can then be executed by workflow engines and service buses.[2] These process integration tools can be deployed through either a centralized or decentralized approach. Centralized tools implies that the workflows are controlled by one process orchestration service, whereas decentralized processes are executed externally by the participating services (MSDN Process Integration, 2004)

3. **User Experience Integration:** simplifies the user experience, enabling user interaction with multiple back-end systems or data sources in a consistent user interface. An example of this type of integration is a user portal that encompasses user-specific reports, tools, interactions, and access to underlying data from multiple systems. For example, a "Patient Portal" could allow doctors to see a patient's entire historical data, aggregated from a combination of EMR and lab systems. Another example could be a FHW mHealth toolkit that delivers scheduling support and digital job aides that are sourced from different backend systems into a single user experience. Common approaches to user experience integration involve the creation of portals, dashboards, and plug-in-based clients in which functionality and data can be added as modules. (MSDN Portal Integration, 2004)

4. **Infrastructure Integration (Entity Aggregation):** combines infrastructure used by multiple systems into a common asset. Infrastructure integration is typically needed when a critical mass of deployments exposes inherent inefficiencies and excess infrastructure. One example is the process of maintaining multiple data centers with separate servers, each with its own maintenance procedures and applications with disparate error logging, authentication, and authorization approaches. Such inefficiencies significantly increase costs. Centralization of infrastructure components can lead to more effective and efficient operations. Moving eHealth systems into cloud computing environments, or into national data centers, are examples of infrastructure integration. In the area of mobile health, infrastructure consolidation can also yield technical and business advantages, as gateways for voice and SMS and other mobile operator channels are consolidated and shared across systems. (MSDN Entity Aggregation, 2004)

Interoperability is defined by HIMSS as the extent to which systems and devices can exchange data, and interpret that shared data. For two systems to be interoperable, they must be able to exchange data and subsequently understand that data so that that it can be utilized by a user or other systems without changing the data's semantics. (HIMSS Link) While the HIMSS definition for interoperability focuses on data, the definition can also be extended to other types of integration, such as process and user experience. While interoperability is commonly stressed as an important quality for systems, it should be noted that interoperability is also a means to integration. While the terms are related, their distinction is clear. Integration is the process of linking multiple systems, whereas interoperability is an intrinsic property of the systems themselves, which describes their ability to exchange and interpret data. System interoperability suggests a potential for cheaper, lower-risk integration, even though it is not an a priori requirement for integration or a guarantee that the integration won't encounter challenges or problematic side effects.

Typically, interoperability is achieved in two ways:

1. Through a process of prior integration, during which all known compatibility issues in making the systems work together are fixed.
2. Through compliance with a predetermined standard or convention. Such standards and conventions specify the behavior of each system with enough detail to enable both systems to work cohesively together. An "interoperable" system has typically undergone a process of specific compliance.

[1] *Workflow engines and service buses are software tools that can interpret, reconcile, manage, and direct complex processes between integrated systems.*

HIMSS also organizes levels of health interoperability hierarchically into three categories; basic, functional and semantic. The basic and functional levels of interoperability can also be described as syntactic interoperability, meaning that the systems are able to successfully exchange data. Alternately, semantic interoperability is the ability for systems to understand the data that was physically exchanged.

Basic interoperability is when a message from one system can be received -but not interpreted- by another system. The next level of interoperability is functional interoperability, in which two systems can share information from a data element, recognize the type of element, and disseminate the information in a comparable element to another system. Current technologies and development approaches make basic and functional interoperability easily achievable from a technical standpoint, regardless of which database, programming language or development framework is used to build the application. Finally, semantic interoperability describes systems that can understand and interpret the data being distributed among common data elements.

For each level of interoperability, standards play an important role in facilitating communication and interaction processing of information between systems. Standards are established norms or requirements used by systems and governed by a recognized consortium. Systems implement recognized specifications to rapidly achieve both syntactic and semantic interoperability by maintaining a common set of expectations and approaches. A more in-depth discussion of standards and their role in health information systems integration can be found in Annex 10.

To summarize, successfully integrated systems is the goal. Interoperability is a means to an end, and standards are semantic and syntactic norms or tools that aim to simplify achieving interoperability.

FHW demonstrating mHealth application in Bihar, India.

Obstacles to Integration

Considerable investments of time, money, human resources, and physical infrastructure are required to integrate systems, whether or not the integration is ad hoc or whether generic interoperability is sought through the use of standards.

Prior to making such a commitment, the concerned parties should answer a series of questions related to the requirements needed to achieve predetermined goals. The discussion frequently begins with high-level questions and becomes more specific and implementation-focused as the conversation continues. Objections to integration are commonly raised at more specific, technical levels; however, these challenges often stem from higher-order questions that remain unresolved or unaddressed. For this reason, it is important to resolve such first or second order questions before addressing the specific and technical concerns. We present five general steps that provide a roadmap to successfully navigate the major obstacles to integration.

The hierarchically organized questions include:

1. Is there any value to be derived from the integration?
2. Are there organizational incentives and alignment to invest in integration?
3. What is the regulatory, contractual and policy framework for data sharing?
4. Is integration feasible at the semantic (data meaning) level?
5. Is integration feasible at the technological level?

Figure 1: Pragmatic Approaches for Obstacles to Integration

1. Is there any value to be derived from the integration?

At the onset of a potential integration project, the initial obstacle that must be addressed is to clearly ***define the value to be derived*** from the integration. These are best expressed as goals with objectives. Simply put, what will be achieved through integration that cannot be achieved in the current state of each system? This type of analysis may be completed through use of a gap or SWOT analysis that takes into account the current operational state and a proposed future state. Also, the reasons to integrate may vary among stakeholders as each group has different interests and requirements. This may include goals dictated by implementers, programs, funders, governments, policies, caregivers, or even individuals who interact with each system. While the process of determining goals across stakeholder levels can complicate the evaluation of those goals, creating a common vision for what will be achieved is critical in the early stages of the process.

2. Are there organizational incentives and alignment to invest in integration?

After the goals and objectives for integration have been sufficiently defined, the next step is to determine the ***organizational incentive and alignment in achieving those goals***. While an abstract integrated system may seem appropriate, the costs and benefits of integration must also be cross-referenced with each organization's mission, mandate, policies, and capacity. This question has to be evaluated in terms of upfront (capital) costs and ongoing (operational) costs. Three examples of the interaction between an organization's incentives and alignment to integrate systems -- and the subsequent costs and benefits -- are described below.

1. The use of standards can sometimes lower the cost for integration, but not always. For example, the up-front investment in supporting a complex standard can be interpreted as an untenable imposition if it surpasses initial goals and budget of a specific project. A simpler path to integrate systems on an ad-hoc basis may better fit the organizations' constraints in the short term, while still achieving the overall goal of the integrated system.
2. Maintaining integrated systems requires an ongoing effort. Similar to bridges that need to be continuously maintained and monitored for safe use, ICT systems need to be maintained to continue running over time. Sometimes, the ongoing cost of the integration falls on the organization that has the technical capacity to keep things running, not the organization that has the mandate and resources, or derives the most value from the integration. This scenario highlights the potential divergence in incentives and alignment in successfully achieving and maintaining integration.
3. In some instances, centralized integration infrastructure, such as when using central enterprise service buses or data integration through unified data warehouses, may increase the costs because of a new requirement to coordinate with a third party with its additional charter, agenda and constraints.

Trust plays an important factor when different teams or organizations are integrating systems; transparency at the outset in sharing incentives and disincentives to integration is a highly efficient and effective tool to help groups solve first order concerns with regard to desirability of data integration across organizations (or departments). A common symptom of failure to resolve the lack of trust and misalignment of incentives is that these issues may taint the second order questions (e.g., #3, #4, and #5) as being ‘insurmountable problems’.

Public health surveillance activities tend to encounter problems at this level of the discussion, as surveillance programs often assume an approach of ‘we want to extract your data’ under the assumption that doing so is both advantageous and cost-effective. A better approach could be to ask program implementers what surveillance information they could utilize to improve their performance, and help them acquire it.

3. What is the regulatory, contractual, and policy framework for data sharing?

After the collaborating organizations each have sufficient incentives and alignment to pursue integration, the next consideration is whether there is a ***data sharing framework*** in place that details the regulatory, contractual and policy requirements for integration. Ambiguity in data sharing frameworks is a stronger deterrent to integration than a clearly defined but poorly designed framework. A data sharing framework should clearly define considerations such as:

1 Are there data sharing policies that affect each organization? Which ones are they? Who imposes them? Are the policies compatible?
2 What are the constraints with respect to access rights and physical placement of data? Does it matter when and where data is stored, who has access to it and how, for how long?
3 Are there different guidelines for provided, observed, or derived information? Information that may be given explicitly by patients, information observed by a health provider, and information derived from analysis of seemingly unrelated or exhaustive data may have different frameworks of rights and restrictions. For example, a health provider may treat the data showing that a patient visited an HIV clinic with a different set of concerns than a mobile operator who has derived the same information from signal strength logs. Herein lies an opportunity and a challenge in the area of public health surveillance, as the combination of observed and derived information has the potential to make obsolete the information that is directly requested and provided.

Unfortunately, the most common data sharing framework at hand is a composition of vague country guidelines, implementers' attitudes of fear or defiance, and subjective interpretations of off-hand official remarks, and too little contractual specification from funders.

The general lack of capacity and professional grade competence in the areas of information security, information access rights, and legal frameworks for licensing information makes the dialogue between public health stakeholders even more difficult.[2]

4. Is integration feasible at the semantic (data meaning) level?

Following development of cohesive and supportive goals, objectives, incentives and policies for integration, the next step is to determine ***if integration is even possible at the semantic, (or data meaning) level.*** While syntactic interoperability, or how data is formatted, is a prerequisite for semantic interoperability, it has largely become a non-issue given the current widely used syntactic markup languages and schema standards such as XML, RDF, JSON and standards such as ISO/IEC 9075 or SQL-DDL. Therefore, achieving semantic interoperability is the next major obstacle to integration once the aforementioned non-technology related barriers to integration have been resolved.

Metadata repositories and data dictionaries can play a crucial role in documenting, adhering to, and sharing the accurate meaning of a data element within a system. As documented in Annex 5, data dictionaries and metadata describe the definition, location, format, quality, limitations, and other requirements to successfully understand and interpret the data in the context of integrating it into a larger system.

One example of this is counting cases of 'malaria' with different approaches -- one being a diagnostic test, the other a questionnaire-based FHW assessment. Both approaches may track cases of 'malaria', but these two datasets cannot be simply integrated while also retaining the original meanings because of the divergent case definitions used. Similarly, FHWs working in two separate provinces and using checklists containing different data elements for 'suspected malaria', or using the same checklist with different levels of training, may generate skewed or biased information, that should not be naïvely compared or combined. In a scenario in which these two systems are to be integrated, appropriate considerations would be required to align the differences and understand the resulting limitations in sensitivity and specificity that will occur.

[2] One idea InSTEDD has put forward is using licensing terms that build on copyright law so data owners can control others' rights to the flow and uses of information, and data consumers or service providers can make their terms of service explicit in terms of those same rights. More information.

5. Is integration feasible at the technological level?

Finally, the last obstacle in the process is to consider whether integration is ***feasible at the technological level.*** Exchanging numbers, dates and strings and how to arrange them into more complex documents is fortunately not the bottleneck to 21st century systems integration.

For almost every scenario imaginable, there are many proven industry solutions that resolve technical interoperability needs within even the most exceptional constraints. Technology considerations relevant to health systems integration include issues involving deployment topology, credentials and permissions, and live vs. batch integration. Specifically, integration at the technological level must overcome challenges pertaining to:

1. Having a clear agreement around whether the integration desired is of data, process, or user experience (as described in the beginning of this annex). This in turn will suggest approaches that are more or less adequate for the problem -- e.g., an enterprise service bus is not a good tool for data integration, data marts are not good for process integration, etc.

2. System deployment and connectivity: which systems are installed where, with access to what, and under whose control? For example, a computer running in an office of a non-governmental organization may access the internet, but itself may not be addressable as a web service. This creates practical constraints such as 'who calls who', or 'push vs. pull' notifications in a given integration scenario. Also, inconsistent connectivity from mobile devices, facilities, and general internet availability in low and middle income countries (LMICs) make it necessary to cache data and queue messages for reliability. It also may be required to put processes in place to resolve conflicting updates or lost messages.

3. Accessing data securely typically implies authentication (securely identifying users and systems), authorization (limiting who can see certain data or perform a particular function), and auditing (tracing what was done). Sometimes in LMICs, an organizational obstacle to integration emerges due to the lack of consensus about who manages these overall permissions and how, often resulting in access that has to be maintained point-to-point.

While the technological considerations all need to be specifically addressed, these are in general the easiest challenges to resolve in the process of integration. Among health care and other fields the technological obstacles and resulting solutions have been well documented and can be reapplied in the future. (Trowbridge et al., 2004) & (Hohpe, Woolf. 2012)

Pragmatic Approaches to Resolving Obstacles

Figure 2: Pragmatic Approaches for Obstacles to Integration

a. Countries and funders can declare specific datasets to be of common interest to overarching health goals e.g. MDGs
b. Workshops and regular meetings can help discover integration opportunities

a. Funders can make the production and maintenance of high quality data a program goal.
b. Organizations can explore quid-pro-quo data exchanges
c. Implement a centralized health system where integration is mandatory.
d. Explicitly budget for systems integration and maintenance

a. Assist countries in developing, publishing and updating data sharing policies
b. Include data sharing obligations and rights as part of grants and contracts

a. Support the creation and regular maintenance of data dictionaries and metadata repositories
b. Include QA processes around data produced by projects
c. Choose & Localize ontologies for use in the country and maintain mappings between system data and these standards

a. Use interoperability profiles of standards to reduce cost and risk
b. Take an iterative approach to integrating first, standardizing later
c. Provide de-facto standardized platforms, endpoints and integration tools

While the obstacles to integration present a challenge, pragmatic approaches exist to mitigate the obstacles at each stage of the process. A brief breakdown of each approach is offered below.

1. **Goals and Objectives of Integration**
    a. Certain data sets pertaining to national and international indicators can be useful tools for achieving priority health outcomes. (e.g., MDGs or other organizational goals) thus creating intrinsic incentives for their utilization.
    b. Systems integration workshops can help implementers tackle common problems to achieve common goals.
2. **Organizational Incentives and Alignment to Integrate**
    a. Funders can include the goal to produce and share data that can be used in other programs as part of their agreements.
    b. Assess implementer landscape for pair-wise or quid pro quo approaches to data sharing, where an implementer A has information that helps implementer B and vice versa. This approach is described in greater depth below, in example 2a.
    c. Reduce the risk and cost of integration by providing a value-add health information system, where the "default" option is to contribute and utilize data from a pre-existing integrated system.
    d. Ensure costs of integration are explicitly budgeted as a value-add activity and not expected to be absorbed in overhead.
3. **Data Sharing Framework**
    a. Ensure that the country has a data sharing framework in place.
    b. The funder should have clear policy guidelines and contractual language that define the obligations of implementers who will share data with the funder and the country.
4. **Semantic Feasibility**
    a. Maintain up-to-date data dictionaries and metadata associated for information systems (see Annex 5)
    b. Include QA efforts as part of any project that produces useful data to validate assumptions about its quality.
    c. Maintain ontology mappings between terminology standards and project datasets.
5. **Technological Feasibility**
    a. Use interoperability profiles of standards and interoperability specifications to reduce the surface area, cost, and complexity of implementing standards. This approach is described in greater depth below, in example 5.a.
    b. Integrate first and standardize later. Successfully integrate systems before solving the generic technological challenges; this will increase efficiency and ensure that the standardization process is better informed. This approach is described in greater depth below, in example 5.b.

The following proven techniques are explained in more detail due to their usefulness in helping overcome barriers to integration or the relevance to the field realities observed in the Bill and Melinda Gates Foundation's (BMGF) projects in Bihar.

2.a Creating Scenarios that Support Mutual Benefits

Stakeholders typically hope to achieve varying objectives and mandates. The resulting differences between organizational incentives and alignment can potentially discourage data sharing and integration. One strategy to overcome this obstacle is to find common or synergetic benefits to sharing data. The varying types of stakeholders described above in the framework collect different types of data to support their organizational activities and incentives.

For example, consider two health information systems operating in a similar geographic area. One is a clinical eHealth system operated by organization A and the second is a separate FHW mHealth system, focusing on maternal and child health, operated by organization B. Given the different levels of interaction and focus of each system, they may be hesitant to share strategic data. Discovering which data is produced by each partner organizations (e.g., departments, agencies, institutions) would have clear benefits, particularly to help identify common data sharing use-cases and aligning organizational incentives. With respect to the example above, this could be achieved by identifying data needed at health facility levels that FHWs can provide and vice versa. Examples of this could include:

1. Clinical facilities that wish to understand population level health metrics, which are collected by FHWs, to inform program development, resource allocation, training requirements, and a broader understanding of local and regional health trends.
2. Clinics may also wish to provide referral information to FHWs to improve health-seeking behavior and to more efficiently and effectively link communities to their surrounding formal healthcare structures.
3. FHW programs may wish to improve the effectiveness and efficiencies of their services by receiving information on confirmed cases of specific diagnoses that require routine clinic follow-up visits, long-term treatments, longitudinal laboratory monitoring, in addition to peer advocacy, providing practical/emotional support and health education, (e.g., pregnancy, chronic non-communicable diseases, HIV, TB, & malaria, etc...)
4. Surveillance and public health programs can benefit from having routine access to clinical site data to monitor changes in disease patterns. Outbreaks often materialize first in health care facilities (e.g., cluster of severe acute respiratory infections (SARI) presenting to emergency rooms, clinical laboratory notifies public health department of positive results for reportable diseases). Sharing information about known cases of incident disease, discovered in the community through surveillance departments, can improve public health response and delivery of clinical services
5. Resource allocation and training can be directed using bi-directional data on observed trends. Both systems may also benefit from a data source to verify trends found in their system, such as a potential outbreak, before taking action.

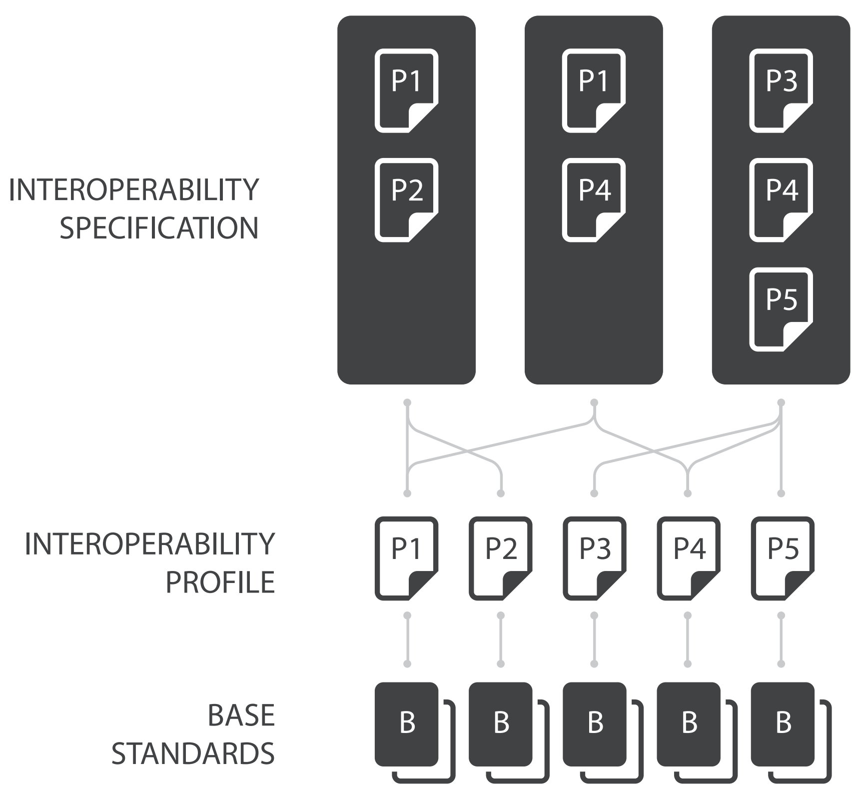
5.a Use profiles to simplify standards adoption

A simple framework of interoperability can be described by three types of requirements: - interoperability specifications, interoperability profiles, and base standards (Reynolds, 2009; Ritz, 2012). Base standards are generic guidelines to define the structures, protocols, and terminologies that describe aspects of what and how two systems communicate with each other. There are various types of standards ranging from semantics, syntactic, vocabulary, and grammar.

Profiles act as an intermediate and flexible layer to mediate the use of base standards to different technical use cases. Simply put, a 'profile' is a pragmatic subset of a standard that sufficiently achieves interoperability for a smaller number of scenarios than the original standard specification envisioned. This can make it feasible to claim interoperability with the profile when a claim of general interoperability is outside the scope or resources of a project.

Using combinations of different profiles, interoperability specifications describe how multiple technical use cases are used in coordination to achieve business use cases or program interoperability objectives. (Reynolds, 2009; Ritz, 2012) For example, in support of public health surveillance, an interoperability specification may be used to describe how to collect indicators that utilize data from multiple sources, each of which use separate profiles.

Figure 3: eHealth Standards Building Blocks Approach



5.b Iterative approaches to choosing standards

Complex standards require an up-front investment in accessing those standards, training on how to use them, and testing for integration. When parties seek to integrate two well-known systems, rather than open-ended interoperability, the return on investment (ROI) of implementing a standard may seem low in the short term. This is especially true when the cost of adding support for it later is lower than the cost of adding support at the present. The disadvantage is that this provides an opportunity for ICT health projects to externalize the costs of being part of a system, where tomorrow's risk is weighed against today's benefit.
This is particularly a hard tradeoff in LMICs, where the eHealth ecosystem remains nascent and today's measurable costs have more guiding power than tomorrow's potential savings.

Much of the cost associated with standards implementation in open-ended scenarios accrues from:

- Trying to adopt large standards that have a lot of 'surface area' beyond the integration goal - for example, HL7 is a family of standards -- it is false to assume that any product implementing HL7 can communicate with any other product implementing HL7.
- Trying to force an architecture for something it's not designed to do because 'it's the norm' - e.g. trying to perform data integration through a workflow engine or through event-based message exchanges.

These try to solve the generic interoperability problem before validating that the systems can be integrated as intended, or that the integration does indeed create value.

The goal of integration and interoperability can commonly be achieved by:

1. Integrating the desired systems in a way that is pragmatic and follows good technical practices.
2. Validating that the integration is possible and yields the desired value to the parties.
3. Analyzing the results of the integration work critically and choosing standards that simplify this otherwise ad-hoc work.
4. Refactoring the systems to implement the desired standards within whatever resource constraints exist.