# Using IATI data to enhance Bangladesh's AIMS

Methodological report

20th November 2015
Mark Brough and Matt Geddes

*This report follows on from the inception report concluded on 17th October 2015. It is informed by a mission to Dhaka in October 2015, which included ten meetings with development partners and five meetings with government staff, as well as numerous informal conversations in the Ministry of Finance.*

*For an introduction and overview of this work, please see the inception report.*

# 1. IATI: an opportunity and a challenge

## A global standard for aid data

The International Aid Transparency Initiative (IATI) was launched in September 2008, in order to make information about aid spending easier to find, use and compare. IATI is a voluntary, multi-stakeholder initiative that seeks to improve the transparency of aid in order to increase its effectiveness. IATI brings together development partners and developing countries, civil society organisations and other experts in aid information who are committed to working together to increase the transparency of aid.

IATI began its work by consulting developing country stakeholders on their information needs, and discussing with development partners the most efficient and effective way of meeting these. On the basis of these consultations, IATI developed and agreed a common, open, standard for the publication of aid information – the "IATI Standard". Development partners implement IATI by publishing their aid information following agreed common definitions, in IATI's agreed common electronic format. This usually happens on their website, with a link to a central registry, the IATI Registry.

IATI was born at the 2008 Accra Third High Level Forum on Aid Effectiveness with the promise of accelerating improvements in aid effectiveness through the provision of better information on development partners' aid activities. Now that much more data is available on development partners' aid activities, we can begin to use the data with the aim of meeting those aid effectiveness goals.

## Opportunities IATI presents

There are two broad groups of opportunities that IATI presents: improving data quality and reducing the burden of data collection.

Improving data quality has several dimensions. Coverage of development partners could be increased through the inclusion of development partners that are not resident or are unresponsive to the government's requests for data provision. Detail can be increased by including more granular financial data in the system or by including sub-national location data and project documents. Data can be updated more frequently, perhaps monthly or quarterly depending on the development partner. Finally, discrepancies between different systems can be addressed, which could be caused by manual keying errors or where project information is not shared by all parts of the organisation (for example, where regional projects are not known to a country office).

Many of the data quality improvements listed above could also be achieved without IATI data given sufficient resources. For example, development partners could be requested to provide as granular transactional data in the AIMS as they have available in their IATI data. However, the burden of providing ever more data manually at some point becomes such that development partners will be deterred from providing this additional data. In manual systems, there is therefore a balance between burden and quality.

IATI data can help overcome the burden issue by automating much of the process of entering basic fields into the AIMS – there is no need to look through project documents to find multiple dates or re-key project descriptions if this information can be pulled in from the IATI data. Similarly, frequently updating financial data should become much easier if projects can be periodically re-synced with IATI data.

A common problem in making AIMS sustainable is a vicious circle whereby development partners provide poor quality data, meaning the data is less useful and less likely to be used. When the data is not used it reduces incentives on the remaining development partners to continue providing data to a system that does not satisfy any user's demands. IATI data could help turn this process into a virtuous circle, whereby higher quality data encourages greater use of the data and encouraging development partners not in the system to begin providing their data.

## Challenges in using IATI data

While IATI data does present some clear opportunities, there are also a range of challenges which have limited take-up of IATI so far.

One of the largest barriers is the technical and complicated nature of IATI. Although the technical nature of IATI is essential to both accurately represent the complex nature of development cooperation and to ensure that data can be exchanged by machines, it also makes it harder for humans to understand the content of the data. This complexity has been compounded by the periodic release of new versions of the Standard. The complex nature of the data and the fact that many of these issues require a good understanding of both policy and technical issues has made it harder to establish guidance and procedures for using IATI data.

The complex nature of the data has also created misunderstandings. Among those trying to use data, there is a perception that IATI data is uniformly not of a sufficiently good quality to make it possible to use the data. While it is true that some data is of a very poor quality, much of the data is good enough to begin to be used. It is therefore important to approach the data in a nuanced way – aiming to use as much good data as possible while excluding the poor data. Analytics and rankings can be a useful guide to data quality, but DPs are also well placed to make such decisions if they have a strong understanding of their projects.

Among many DPs there is a belief that using IATI data should be a simple or straightforward task and that there is a problem with a country system if it is not able to automatically import a development partner's data. The slow take-up of IATI data in country systems can be attributed at least in part to the complicated and technical nature of IATI data and of software development. The lack of proven cases of successful and sustained integration of IATI data with an AIMS also has made it hard to make a business case for investing in software development.

## No "one-click" solution to IATI import

Using IATI data will continue to require time and effort on behalf of DPs and government. Import will need to handle a wide variety of issues. Depending on the DP, these may include:

- properly modelling trust fund contributions, pooled funds and co-financed projects;
- handling projects reported at multiple levels of the project cycle by different organisations;
- excluding DP administrative project components or components that are not reportable to the AIMS;
- excluding projects that have inaccurately been tagged as operational in Bangladesh, or where country-level staff do not recognise the projects reported by headquarters;
- handling projects where the title or description is not available in English;
- grouping very granular activities into larger project components resembling those available in the AIMS.

It is clear from the inception mission that "one-click" IATI import will not be possible for most DPs for a while, because only humans can handle some of these complex issues. Given the complex nature both of the data and of what it is trying to represent, taking time and care to get this right will pay dividends in future. In time, this approach may also suggest improvements to the IATI Standard or to the way in which it is interpreted and used.

Data is not available for certain development partners such as China. IATI data cannot solve this problem in itself. However, by offering a mechanism for obtaining high-quality data at lower burden it can reduce the cost to development partners of complying with government data requests. By increasing the quality of data in the system it can also help to build confidence and support for a system which is providing all stakeholders – development partners and government alike – with access to useful data. In time, it could therefore increase incentives for other development partners to begin providing data to the AIMS either through manually entering data or importing it via IATI.

Finally, IATI poses a challenge in disrupting existing processes. Some of this disruption may be positive - by providing more data and at lower cost than currently available. But care must be taken to avoid negative disruption which may do harm to existing systems of data collection - either by disrupting the process or the data quality that results.

# IATI data in Bangladesh

In Bangladesh, IATI data is of mixed quality. Some DPs reporting to the AIMS have not begun publishing to IATI, while others are publishing only very old data to IATI. On the other hand, there are many DPs with good quality data in IATI, and some DPs publishing to IATI are not accounted for at all in the AIMS.

Publish What You Fund's Aid Transparency Index provides a reasonable proxy for IATI data quality, given that IATI is the most highly-weighted component of the Index. Rather than repeating all of the analysis required for the Index, we use the most recent results across all organisations as a guide.

**Table 1**. Size of largest 20 DPs according to FY14 AIMS data and rating in Publish What You Fund's 2014 Aid Transparency Index[1]

| Development Partner | FY14 Disbursements (USD millions) | FY14 Rank | PWYF Rating |
|---|---|---|---|
| World Bank | 942.96 | 1 | Very good |
| AsDB | 704.32 | 2 | Very good |
| Japan, JICA | 375.75 | 3 | Poor |
| UK, DFID | 225.32 | 4 | Very good |
| Netherlands | 73.54 | 5 | Fair |
| Australia, AusAID | 45.40 | 6 | Fair |
| EC | 44.91 | 7 | Good |
| UNICEF | 43.48 | 8 | Good |
| GFATM | 39.17 | 9 | Good |
| Canada, DFATD | 34.36 | 10 | Good |
| UNFPA | 32.44 | 11 | *Not rated* |
| WFP | 30.45 | 12 | *Not rated* |
| Switzerland, SDC | 28.23 | 13 | Fair |
| IFAD | 27.92 | 14 | *Not rated* |
| Germany, GIZ | 19.49 | 15 | Fair |
| UNDP | 11.47 | 16 | Very good |
| Korea, KOICA | 10.29 | 17 | Poor |
| OFID | 10.11 | 18 | *Not rated* |
| Denmark, Danida | 8.11 | 19 | Fair |
| Sweden, Sida | 7.67 | 20 | Very good |

This table notably excludes some large DPs reporting to IATI that have not provided any data to Bangladesh's AIMS, including GAVI.

A 2014 study by USAID found that awareness of IATI in Bangladesh was limited. However, perhaps partly as a result of outreach involved in that study, we found fairly widespread awareness of IATI, though knowledge or understanding of DPs' own data and any data quality challenges was much more limited.

## DPs with specific challenges

A number of DPs face specific challenges with their data. These challenges are captured here not to criticise them, but to ensure that they are adequately taken into account in the course of this work. In some cases this may mean DPs changing the way they publish data at headquarters level, and in others it may mean that the importing tool needs to be flexible to deal with nuances in the way data is published. In time, there may be arguments for

---

[1] Disbursements refer to AIMS data as at 2015-11-17 for FY14, in millions of USD. PWYF Rating refers to Publish What You Fund's 2014 Aid Transparency Index: http://ati.publishwhatyoufund.org

changes to the IATI Standard or the way in which it is interpreted in order to make it easier for software to handle these specificities or differences. However, we suggest any changes to the Standard should be made only once different approaches have been thoroughly tested.

DFID publishes projects and sub-components in its IATI data. The approvals process in Bangladesh means that projects can only be reported after a certain stage, so certain of DFID's project components (preparatory work, M&E) should not be reported to the AIMS. The import interface will need to allow components to be deselected from projects before importing. It is possible that other DPs will face a similar challenge, and they may have a less clear division in their data.

Germany's data is published by BMZ and includes projects funded through GIZ and KfW. The data currently includes cumulative figures for each project, without any breakdown over time. This makes it difficult to know how much has been spent in any year. GIZ projects also contain multiple "phases", but it is unclear if these phases are or should be joined together to make a single project in the IATI data. These issues will be taken up with headquarters.

For the the Netherlands' projects, commitments are made in the local currency value - in Bangladesh, this is USD rather than BDT. The commitment value is stored in the Ministry of Foreign Affairs' internal project management system in EUR and USD. The EUR value is shown in IATI data. At the start of each year, the EUR commitment values are recalculated to ensure that the local project value in USD is accurately reflected in EUR, given currency fluctuations over the previous year. This should be fairly straightforward to handle given the mechanisms the IATI Standard has for dealing with currencies, but it is useful to be aware of this nuance. The Netherlands also spends much of its funds through large tranches of programmatic funding to implementing partners such as Dutch NGOs. Including IATI data published by Dutch NGOs may help to provide a fuller picture of Dutch development cooperation projects.

UNDP receives funds from headquarters and from DPs in country to fund particular projects. It is important to avoid double-counting the DP-reported projects with the UNDP-reported projects. We propose to handle this issue by allowing DPs to map their projects to the UNDP project, and then allow UNDP to decide which data it chooses to take - its own, or others' data. This issue of data being reported at multiple levels will certainly apply to other organisations and the user interface will need to handle it in a general way.

UNICEF has a unit of aid in its IATI data that shows results rather than projects. This issue cannot be solved in the user interface but would need to be taken up with headquarters to work out if the system could instead export project data.

USAID has a very granular unit of aid in its IATI data ("awards"), so activities will need to be grouped together to create meaningful projects. It appears that in Bangladesh, one project has one or many associated awards, so an interface to allow grouping of awards into projects will be useful. There may however be awards belonging to multiple projects - we propose that such awards should not be "split" for now, but should be handled manually. Systems improvements currently underway at USAID could in time help to improve this process.

There will undoubtedly be other issues specific to DPs that occur in the course of this work. Again, the point of listing these issues here is not to criticise the individual DPs; indeed, in many cases, there is nothing to criticise. However, being aware of and understanding these issues will be critical to successful IATI-AIMS integration. It will be important to work with DPs individually to understand the way their development cooperation works and to make adjustments to data published at headquarters level where problems are found.

# 2. Review of the Bangladesh AIMS

## Background

Launched in October 2014, the Bangladesh AIMS has been designed by local staff based on their needs and after reviewing several other AIMS systems. The design was implemented by Technovista, a local software development firm.

The AIMS collects project and financing agreement level data in a comparatively comprehensive system. The AIMS is able to record financial data for commitments, planned disbursements, actual disbursements and expenditure. The AIMS allows for reporting on trust funds and multi-donor projects, can handle multiple currencies for all transactions, sub-national location data, both thematic and sector priorities (according to the government's national development plan), aid effectiveness indicators, and project documents. Some fields are mandatory, but most are optional. There are a variety of pre-made reports and charts, all also downloadable and options to export custom data selections. Many fields are adjustable via the interface – thematic priorities, geographic areas, currency conversion rates and document types can all be modified without needing a developer to reprogram the AIMS.

Several aspects stand out with respect to IATI integration:
- consideration has already been given to providing an "API to ensure interoperability with other as such systems of the Government";
- the AIMS was designed with IATI data in mind, therefore wherever possible, taking IATI data definitions and structures as the starting point in an effort to make it simple to incorporate IATI format data;
- projects have an optional field for the IATI activity identifier, which could facilitate matching of projects;
- source code is owned by the government and can therefore be adjusted without having to request permission from software vendors;
- some of the mapping work between IATI codelists and local codelists has already been done.

DPs are responsible for entering data in the system. Since the launch, 47 DPs have provided data into the system[2]. Over time it is expected that more detailed project data is provided; there are a number of DPs that have yet to input any data to the system.

Data entry for a project involves filling some 40 fields even excluding project documents, results information or aid effectiveness indicators. Each financial transaction involves providing another 5-10 fields. Development partners report that this takes about 10 minutes per project. Some DPs have indicated that the data entry burden is partly responsible for the lower levels of data and lower data quality. These issues are likely to be compounded where

---

[2] Figure calculated by the number of Development Partners shown in the Development Partner Profile interface, 17th December 2015:
http://aims.erd.gov.bd/Areas/AIMS/ReportViews/DevelopmentPartnerProfile.aspx

DPs are non-resident, or where ERD has limited contact with them. As a result, ERD staff have also been carrying out significant data entry on behalf of DPs.

## Comparison between AIMS and IATI data

In October 2015 the AIMS underwent a substantial exercise in data collection in advance of the Bangladesh Development Partner Forum[3]. A snapshot was taken on the 17th November 2015. AIMS data was compared to IATI data from http://dportal.org taken on the same day. Both AIMS and IATI contain significant amounts of data, AIMS reporting FY14 disbursements of USD 2.841 billion and IATI reporting USD 3.156 billion for the same period. The aggregate numbers are broadly similar, but the breakdown tells a somewhat more complex story. The AIMS records 46 organisations and IATI records 88 organisations. 11 of the organisations reported to the AIMS are not reported to IATI[4]. AIMS data and IATI data therefore appear to be highly complementary. There are also significant variations between what is published to IATI and what is reported to the AIMS for specific DPs. Chart 1 shows all organisations for whom the difference in volume between the AIMS and IATI is greater than USD 50 million.

**Chart 1: Comparison of actual disbursements recorded for FY14 and FY15[5]**
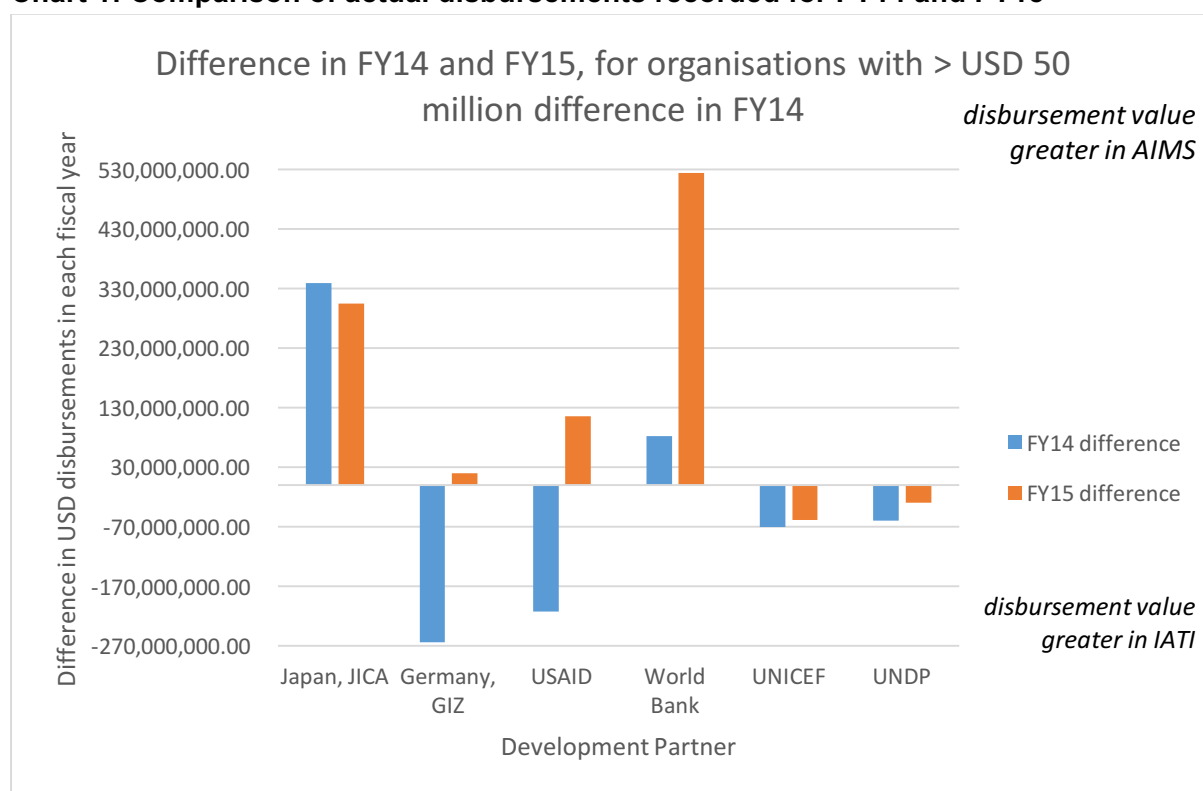


Where there is a difference between data in the AIMS, investigation is needed to determine the cause. The differences in GIZ are attributable to the issues earlier discussed regarding

---

[3] http://www.lcgbangladesh.org/bdf2015
[4] Many of the remaining organisations publishing data to IATI are implementing organisations.
[5] Data taken from AIMS and dportal.org on 17th November 2015. Note that FY14 in Bangladesh is 1st July - 30th June. d-portal allows you to specify the fiscal year for a country under the settings button.

cumulative financial data appearing in Germany's IATI data - therefore several years of expenditure are tagged on a single date in 2015. The "Germany" IATI data available through d-portal also includes projects not funded through GIZ (e.g. via KfW) so the numbers would not be expected to be identical – though the IATI data does allow the data to be distinguished by implementing partner. JICA publishes quite limited information to IATI; the latest data published to IATI stems from December 2014, according to the IATI Dashboard (it my also refer to an earlier period of spending)[6].

The reasons for the differences for the World Bank and USAID are less clear. For USAID, it could perhaps be an issue to do with spending reported to the AIMS for FY15 that actually occurred in FY14, but may also be due to a large amount non-AIMS-reportable spending from USAID in its IATI data. For the World Bank, the data is being published relatively infrequently, one quarter in arrears, with spending aggregated into six-month periods – so there could be an issue with spending from FY14 either not yet being reported or being attributed to the wrong fiscal year. We understand that there may also be an issue to do with the completeness of the World Bank's IATI reporting which is currently being addressed.

Below USD 50 million, there are still a number of DPs with large differences, but the numbers are more in line with what would be expected given differences between what is published in IATI data and what can be reported to the AIMS.

**Chart 2. Comparison of actual disbursements recorded for FY14**

---

[6] IATI Dashboard, accessed on 17th December 2015:
http://dashboard.iatistandard.org/timeliness_timelag.html

**Difference in FY14, for organisations with < USD 50 million difference in FY14**
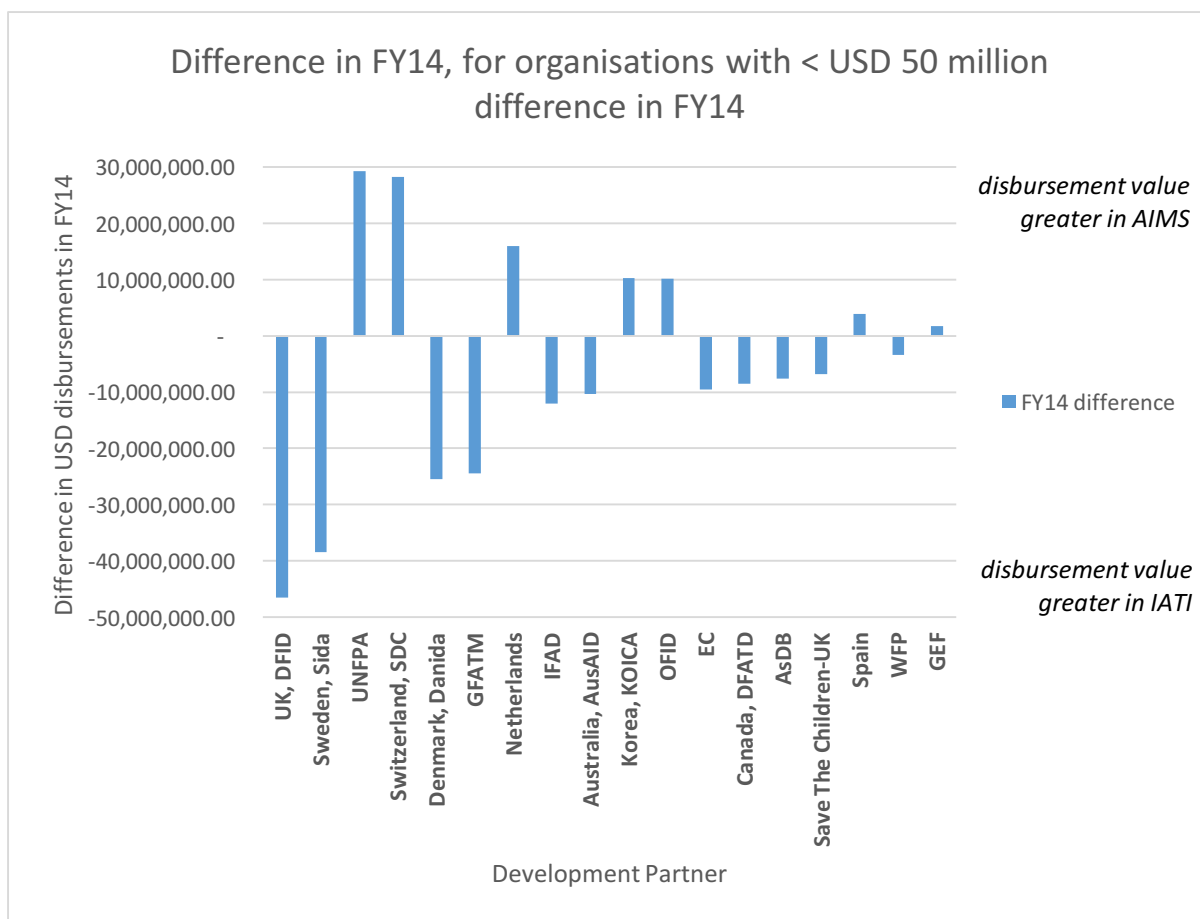


On this chart, we only show the difference in disbursement values for FY14 for space considerations. In general, there does seem to be a greater value reported to IATI than to the AIMS. However, it becomes more difficult to identify reasons for any discrepancies when looking across DPs at these smaller values. Some of these differences may also be due to the same spending being reported by different organisations at different stages of the project cycle.

The foregoing discussion underscores the importance of taking a DP-by-DP approach to importing IATI data to the AIMS, to ensure that discrepancies are identified, understood, and can be handled with care. It also shows the benefits of being able to see the numbers that are flowing direct out of headquarters systems – so IATI data can be used as a "sense check" even just by looking at the aggregate numbers.

We do not go into any further comparison of the quality of data available in different systems (e.g. detail of sector coding or geocoding); that is beyond the scope of this report. But it may be useful to conduct this kind of analysis in the later stages of this work in order to help users understand the nature of the data they are importing. It may also be useful as a way of evaluating the relative costs and benefits of using IATI data to supplement data captured by AIMS.

# Surrounding processes

The AIMS is hosted by the Development Effectiveness Wing within the Economic Relations Division (ERD) of the Ministry of Finance[7]. Within ERD, other data is also collected by the Foreign Aid Budget and Accounts (FABA) Unit, which is responsible for debt management and budgeting of foreign aid. The intention is that over time, the AIMS should provide all the data required by FABA and that there should be no need for a parallel data collection process.

We are keenly aware of the need to ensure the data provided by the AIMS to FABA satisfies the needs of FABA – which will be important if the link between the AIMS and FABA is to be strengthened. From our initial meetings, the key priority appears to be to ensure that there is detailed financial data (including actual rather than aggregated financials, because of the need for precise currency conversion), clarity around whether the funds are spent through government systems and a detailed mapping to government budget classifications (which is currently performed manually).

## Technology used in the AIMS

The AIMS uses a combination of Microsoft technology. It is hosted on a Windows server physically located inside the Aid Effectiveness Unit. The database is run on Microsoft SQL Server and the software runs on a .NET MVC framework with C# for the business logic. The front-end uses ASP.net and JQuery for some of the user interfaces. Charts are run through Highcharts and SAP reports are used in a few cases to create printable reports.

---

[7] For more on the organisational structure of ERD, see:
http://www.erd.gov.bd/site/page/e8cbce25-6478-4efd-b227-aec73ba416e2/Organizational-Structure

# 3. Approach and principles for IATI import

## Approach

As detailed in the inception report, we emphasise taking a careful and pragmatic approach to IATI import. The aim is to take data from as many DPs as possible while avoiding doing any harm to existing functioning processes – especially as the AIMS itself is relatively new. Our goal is a working, functional and sustainable implementation within six months. Given the short timeframe, the implementation will need to be somewhat simplified, but it should still be functional and sustainable in the vast majority of cases.

We think it is very important for users to understand conceptually what they are doing when they import project data – but we do not think they should ever have to look at XML or understand very much of the technical nature of IATI. The hardest part to get right will probably be matching and grouping projects, reconciling the same projects between IATI and the AIMS and making sure that multiple activities (from the same or different DPs) can be correctly matched together and modelled to avoid duplication. That is where we need to focus users' efforts.

## A separate module – more users, more maintainers

To avoid disrupting existing processes, as well as to widen the net of suppliers that may compete in a tender, we will develop module separated from the rest of the AIMS. Maintenance will also be easier if what is developed has wider applicability in other countries. This should also speed up development as we will have to understand less about the existing AIMS technology, structure and process.

Developing a separate module is also consistent with our goal of avoiding doing any harm to the existing AIMS. We can develop the module separately from the rest of the AIMS and only have to touch it towards the end when we write to the database, and only when we know exactly what we are writing. The existing AIMS interfaces and processes can therefore continue working in parallel, and we will not disturb the system until we have a solid working implementation and good user feedback.

## Principles for data input

DPs will remain responsible for their own data input. This means that DPs can decide and are accountable for the data that they choose to enter into the system. If a donor is happy with the way that their projects are represented in their IATI data then they can choose to use that information. However, we do not want to create a process which absolves DP focal points from any responsibility for the data that is entered, as this inevitably will reduce the quality of the data as well as damage the ownership that DPs feel about the data that is entered into the system.

The interface should also accept as much data as possible rather than setting a "minimum threshold" that will determine whether or not a DP's data should be accepted. Analytics are

helpful for gaining a general understanding of the nature of DPs' data across the board, as well as encouraging and incentivising improvements in individual DPs' data. But incentives also come in the form of increasing the accuracy with which a DP's efforts are being represented as well as reducing the burden of manual data entry. If a DP has some information available through IATI, and that information is accurate, then it should be accepted.

Nevertheless, many DPs will be missing some pieces of data for their projects in their IATI data – whether it is detailed enough sector coding, or geocoding, or forward data. In these cases, where data fields are required, DPs should be able to manually enrich or enhance their data to provide these additional pieces of information.

The interface should make reasonable effort to handle idiosyncrasies in DPs' data where these idiosyncrasies are fairly widely prevalent or easy to fix. However, DP-specific data problems – particularly where a DP's data does not conform to the IATI Standard – will remain the responsibility of that organisation. The issues that are found with individual DPs' data should be shared – both so that the DP can address data issues, but also so that others may shine a light on the problems or be able to consider how they would deal with the issues in future.

Finally, data input should automate those components that are possible and desirable to automate but retain manual input in other components where humans are required to interpret data. This will allow edge cases to be resolved in a sensible way but also reduce the labour intensity required for data entry. In practice, this means that, for example, the IATI Datastore should be queried behind the scenes to select data which is then displayed in a user-friendly way prior to import, rather than expecting a user to find, extract and upload an XML file. By contrast, users are needed to determine the (often complicated) relationship between projects published by different organisations, and asking computers to attempt to guess how to interpret these relationships will be likely to cause problems.

Manual data entry will always be required – either because data is unavailable, or because available data will contain inaccurate or insufficiently detailed data from the perspective of the AIMS. But the vast majority of copying and pasting or basic project information, as well as automatic updating of transactional data, should be greatly simplified through automation.

# 4. Technical integration of IATI data with AIMS

Importing IATI data to the AIMS will require the development of two components on the AIMS, as well as the development of a separate front-end module for handling the import process.

## Adjustments to the AIMS

In both of these cases, it would be strongly preferable to use a separate testing server set up a test server rather than pushing code to the live server before it has been fully tested. This does not have to be a physically separate server but could use a separate database schema on the same server for the database back-end, and the test application on the same server as the existing application. If this proves difficult then a new light-weight PC could be used as the server.

### Reading from the AIMS

In order to develop the user interface required for matching and grouping projects from the AIMS, it will be necessary to read data out of the existing database and provide access to it via an API.

The API will need to provide a list of project data according to a series of set queries, at least by DP, by sector, and for all data. It may also be necessary to add additional filters in time, so a flexible approach will be required for querying the data.

The data should be provided in the IATI-XML format so that it can be read in by the front end in the same way as IATI data published by donors. There is an additional benefit in doing this in providing IATI export functionality from the AIMS.

In order to do this, the software supplier will require access to the AIMS source code. The source code can be provided to the vendor providing they agree they will not disclose or use the source code for any other purpose.

### Writing to the AIMS

In order to implement the decisions taken in the IATI import tool, it will also be necessary to have the ability to write into the AIMS. This will be a question of updating the data in the AIMS for a particular project, by mapping project data from the IATI import tool to specific fields. The AIMS' notifications interface should also be considered for tying the IATI-AIMS import tool into the AIMS.

It will likely be necessary to make some adjustments to the database structure in order to record the provenance of data and allow for it to be automatically updated in the future. The nature of these adjustments will depend on the way the existing database is structured, as well as the effects of other business logic in the source code.

In order to do this, the software supplier will require access to the AIMS source code under the same conditions as stated above.

# IATI import module components

The front-end module should be developed as a separate component, and should only write to the main AIMS database once the effects of this have been tested and verified thoroughly.

There would be some advantages to the server-side components being written in Python, given the existing tools available for wrangling IATI data in Python. The fact that Python is also the dominant language in the IATI community provides an additional advantage in the form of increasing the likely sustainability of the IATI import module.

## a) Retrieve IATI data

The tool will need to automatically fetch relevant IATI data. In generally, this will mean selecting all available data for a particular DP, though there may be a couple of other queries required, so the query should be built in a flexible way (though the interface should emphasise simplicity).

> About the IATI Datastore[8]
>
> The IATI Datastore is the preferred data source for IATI data. It makes it possible to query for data across a range of different filters. The data is returned in IATI-XML format, though the activities may be in different versions of the IATI Standard for each donor.
>
> There are a several limitations of the IATI Datastore, including:
> i) it does not handle hierarchies very well. Donors can structure their data according to a hierarchy of activities (e.g., a donor may have a large water program with many project sub-components for individual activities). The datastore only returns activities that match the search parameters, and not any parent or child projects;
> ii) there are some other known bugs with the datastore.
>
> The IATI Datastore is an open-source Python project available on Github. In almost all cases there will be workarounds for the above issues. However, it may also be sensible and helpful to contribute to the Datastore to fix issues that arise, and pull requests would be welcome. If the IATI Datastore proves to be a significant barrier to development, then the data could be retrieved directly from the IATI Registry[9].

### Steps required for obtaining IATI data from the Datastore

On request, the IATI import module should:
i)      Retrieve the requested data from the IATI Datastore
ii)     Store the data locally
iii)     Make the data available to the the rest of the application, via the next component

---

[8] http://datastore.iatistandard.org
[9] http://iatiregistry.org

## b) Standardise IATI data; make it available to the rest of the application

IATI data retrieved from different donors may be available in different versions of the IATI Standard. A component to automatically convert all data to a single version of the IATI Standard (probably version 2.01 or 2.02) will make it significantly easier to handle the data later in the application. IATI-XML data from the AIMS should also be passed through this component. The import process of matching data from different sources will therefore become easier once all the data is in the same format.

This component should:
- Allow data from different IATI-XML versions to be imported and converted to a standard JSON representation, including nesting hierarchical activities
- Be available as a distinct module so that it can be used and maintained by other users of IATI data.

## c) Dashboard: show overview and summary statistics for selected data

Once data has been retrieved, an overview should be presented to the user to help them understand the nature of the data that has been downloaded (normally, for a particular DP).

The overview should contain some summary statistics including:
- The number of projects found
- Aggregate value of project commitments / disbursements / expenditures, per year
- Aggregate value of project commitments / disbursements, per sector
- Compare data found in IATI data with data available in the AIMS for that publisher – for example, to understand the total value of projects in each system in each year.

From this page, the user can choose to begin importing this data to the AIMS.

## d) Import stage 1: grouping and matching

As emphasised previously, establishing the relationship between projects in IATI and in the AIMS is the most important and difficult thing to get right. This component, as stage 1 of the import process, focuses on this relationship.

Users should be asked to group and match projects, from a list of unmatched projects in IATI data, to a list of similar projects in the AIMS.
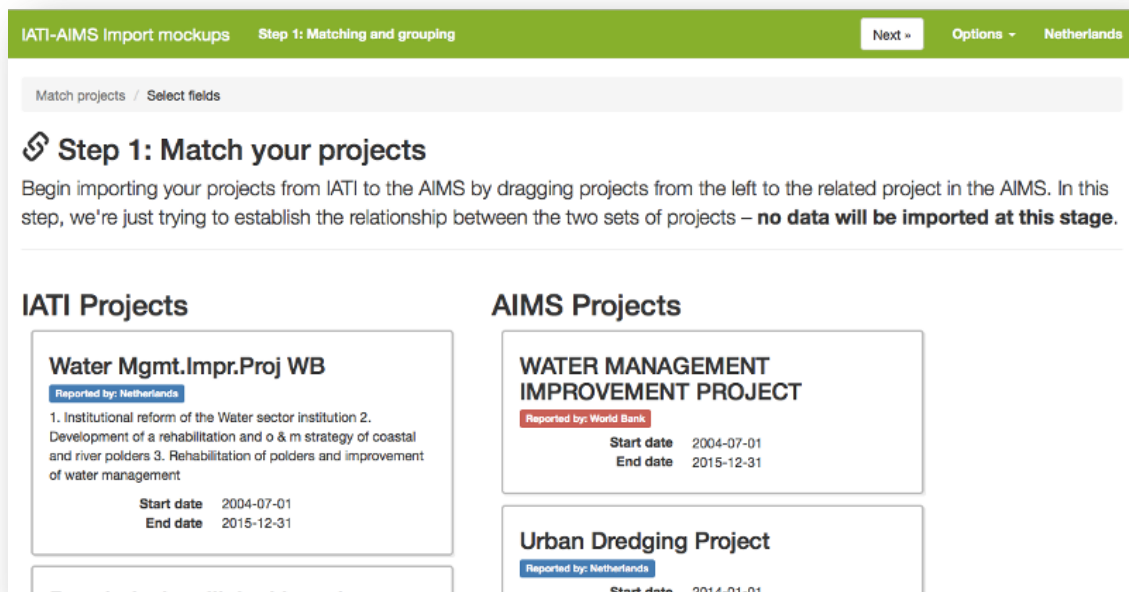
This implies several subcomponents:

i) *map components from IATI to the AIMS*: where a donor uses a hierarchy of activities, it should be possible to move the IATI activities according to their parent activity, to move individual activities independently, or to deselect certain components from the hierarchy of activities. For the activities that are moved, it should be possible to map them to the activities in the AIMS. This may take the form of dragging and dropping various IATI activities onto the AIMS, or another intuitive user interface.

ii) _find AIMS projects that are most similar to the IATI activities_: the AIMS' "most likely match" for an IATI project should be shown at the top of the list of AIMS projects. This should make the process of matching projects much faster for the user. The "most likely match" could be determined by simply filtering for the same sectors, or something more complex such as string comparison of titles. There will likely be many cases where a user would need to map their project to the project "owned" by another user (e.g. where the Netherlands finances a World Bank project, it should be possible to map the Netherlands project to the World Bank project as they refer to the same concept and it is important that they are joined together). It is very important that this sorting is fast if it is to be useful, so it would probably be preferable for the sorting to occur in the browser.

iii) _re-sort AIMS based on "most likely match"_: the list of AIMS projects should be re-sorted according to the most likely match. It should also be possible to disable this re-sorting.

The below mockups illustrate what this process could look like (though other UI suggestions would also be welcome):

**Stage 1: Matching and grouping – showing IATI projects on the left and AIMS projects on the right**

**Stage 1: Matching and grouping – dragging projects across to the AIMS**
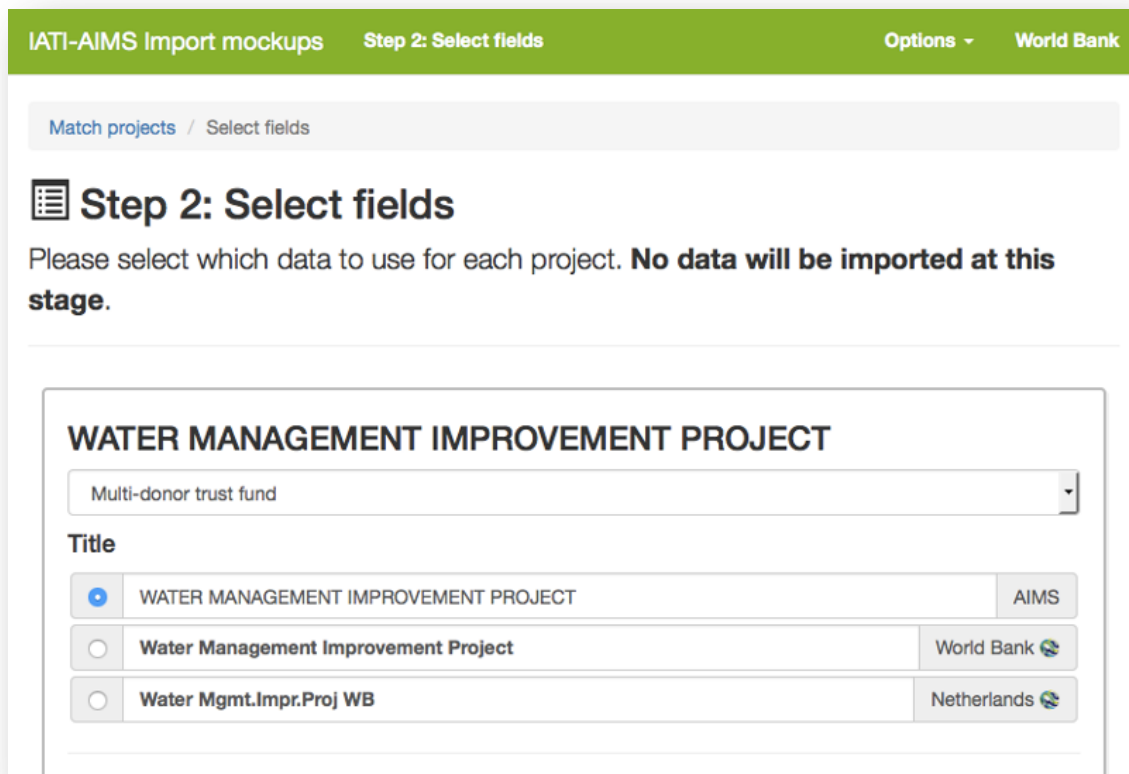


# e) Import stage 2: field level import - per project

Once activities have been mapped together, it is necessary to determine how this data (potentially from several different sources) should be imported to the AIMS. A user should be able to determine which data they would prefer to take from different sources.

Granular controls should be provided so that options can be selected for each field for each project. However, the tool should also make educated guesses based on user-determined defaults (e.g. "prefer my organisation's IATI data") so that a user can adjust and verify the selections rather than select every option.

Where multiple IATI activities have been grouped and matched to a single AIMS activity, the user will need to choose with data they prefer – from one of the IATI data sources or from the AIMS. The user should also be able to edit the values if they are unhappy with what has been suggested by either IATI or the AIMS – this would then be stamped as if it had been entered manually through the AIMS interface.

This process is again illustrated by the use of a mock-up, below:

**Stage 2: Field-level import, per project**



In this example, the user can choose to use data from the AIMS or from the World Bank or Netherlands IATI data.

# Updating IATI data

## Updating activities

Where a field has been tied to a specific IATI activity, it should be possible to update that field automatically as IATI data changes. IATI data should be checked for updates each night. To begin with, a user should be prompted when a relevant change has been identified. They should then be able to choose whether to accept the changes or reject them. Their choices should be persistent if the update is rejected, i.e. the following night, even though the IATI data differs from the AIMS data, the user should not be prompted to update. Notifications could be provided periodically be email (see "User settings and controls", below) or by using the AIMS' notifications interface (see "Writing to the AIMS", above).

We assume that users will generally want to update fields that have been sourced from IATI data, but not where fields have been sourced from AIMS data. However, at least to begin with, users should always have to choose to import data rather than for it to occur automatically without prompting. In time, as more analytics are collected on the way the import tool is being used, users could decide to allow the data to automatically flow in.

The following steps could be used to check for and manage updates:

i)      on a nightly basis, download relevant data from the IATI Registry;
ii)     record activities that are not matched at all in the AIMS. Flag those originating from DPs that have imported some data from IATI to the AIMS (they could be new activities);
iii)    where an IATI activity is linked to an AIMS activity, check to see if any of the matched fields have different values. This could be achieved by comparing values in the AIMS with the values in the most recent IATI data, or it could be achieved by comparing the most recent IATI data with the previously downloaded IATI data.

Different approaches should be taken depending on the nature of the change:

i)      if there are new activities, present the user with the option to begin importing those activities (following the methodology outlined above – beginning at stage 1);
ii)     if there are differences in matched fields, present the user with the option to update those fields with the new values;
iii)    if there are new fields that were previously unknown to the AIMS (e.g. there is now data on the location of activities), present the user with the option of importing that data to their activities (following the methodology outlined above – beginning at stage 2);
iv)     if activities are deleted, alert the user. Provide the option to remove the activities from the AIMS or unlink them from the IATI data. It is also important to state how many other activities were deleted – it could be a technical error that has caused the activities to be deleted. We should be particularly careful about deleting information and would suggest not doing so automatically for the foreseeable future.

It will probably be desirable to develop a distinct module that can compare and record differences in IATI data files, as well as expose that data in an intuitive way.

## Updating transactions

Financial transactions should be handled in a similar way to activities – using the same methodology for comparing and recording differences in IATI data files. However, given the potential large volume of changes to financial transactions, users will need to be presented with an aggregate overview of the nature of the changes, so that they can make an informed decision about the implications of any changes to the transactions, and whether they would like to import them.

Financial transactions should generally not be overwritten or deleted from the AIMS; IATI data should be assumed "append-only" in this respect. However, it is important to be aware of and handle potential exceptions to that rule (e.g. where an organisation changes the dates of transactions after the fact). We will consider the best approach to handling exceptions to this – whether it is better to try to adjust the data in the user interface or to work with donors to fix their data.

# Ancillary components and diagnostic tools

## User settings and controls

Initially, user authentication and authorization should be separate from the rest of the AIMS, though in time it should also be possible to share user management.

Users should be able to specify some default settings, including:

- Their IATI organisation identifier, which can be used to simplify the process of downloading data from the IATI Datastore
- Preferences about where to source data from – e.g. whether the system should generally suggest to use their own IATI data, others' IATI data about related projects, or the AIMS data. This should reduce the amount of clicking around in the above interfaces.

There should also be an interface to help users understand and diagnose errors in their IATI data or in the import process. This should provide warnings if more than a certain percentage of projects suddenly change, if invalid codes are used, if the XML could not be parsed correctly or if the file could not be found. Some of this information is already captured and exposed by the IATI Datastore, so it will partly be a question of showing those error messages in a friendlier way.

Users should also be able to receive email alerts (with the maximum frequency of alerts customisable) when new data is available for import.

## System-wide mappings

The tool should maintain mappings between fields as well as mapping tables to match IATI codes to their AIMS equivalents – for example, to convert IATI sectors to AIMS sectors. The mappings should be stored in a flexible way so that these mapping tables can be easily updated by non-technical users. Some mapping tables are already maintained in the AIMS.

## Logging and analytics

Detailed logging should be collected to record data changes. Detailed analytics on user behaviour should also be collected to understand the way the tool is being used and work out how to simplify the tool over time and thereby reduce the burden of data entry.

## Backups

The IATI-XML provided by the main AIMS database as well as the relevant data from the IATI Registry should be backed up periodically. This should occur up to once per day for a month, and then once per month going backwards. This will be particularly important to the development process as it may sometimes be difficult to anticipate everything that may happen when using live data and working with different live databases.

# Development process

The development process will need to be highly agile and iterative given the necessity to experiment with different approaches and techniques to handling the data, as well as responding to user feedback to ensure the end product is user friendly and intuitive. Some parts of the development process can happen in parallel – particularly back-end and front-end work – so that we can proceed more quickly as well as benefit from the specialisations of individual developers.

We will provide mock-ups of the front-end interface and work closely with the developers to fully understand what is being developed and how it relates to the data. All code will need to be on Github from the start, and in a public repository, so that we can benefit from collaboration with others (who also have their work on Github) as well as ensure that the final product will be useful in a wider range of environments and more likely to be maintained.

We will leverage existing tools and care strongly about maintenance and sustainability – so it may be preferable to use similar languages used in the rest of the IATI community (notably Python).

# Indicative development roadmap

In the **first phase**, data will be extracted from IATI and the AIMS, with projects then listed side by side on the same page. This will include a field-by-field cross-walk and mapping. *Requires AIMS to export data in IATI-XML format for each donor.*

In the **second phase**, donors will map their projects in IATI to those already contained in the AIMS. Establishing this project-level relationship is arguably the most important part to get right and we emphasise the challenge of this component - both in technical and conceptual terms.

The **third phase** will then test the import of a limited subset of fields - probably restricted to the title and description - into the AIMS. DPs can edit the IATI data before it is entered into the AIMS. *Requires AIMS to accept data from the IATI import module.*

In the **fourth phase**, we will begin to establish techniques to reconcile multi-donor projects, focusing on co-financed projects.

In the **fifth phase**, automatically updating fields with data from the IATI Registry will be established. This will need to provide a user-friendly interface for controlling updates, as well as development of logic for handling different data sources. *Requires, inter alia, AIMS to know whether a field comes from IATI or has been manually entered.*

In the **sixth phase**, we will move to encompass other fields - particularly financial data (transactions, including disbursements).

In the **seventh phase**, we will develop techniques required to handle multi-donor projects (parallel financing, co-financed projects, trust funds, etc.) and tools to mitigate against double-counting.

Again, various aspects of the six phases can be carried out in parallel. The implementation will be determined as a success if donors with high-quality IATI data are able to import data into the AIMS. Due to the short timeframe, the amount of time available for the later phases may be limited.