

## **Request for Proposals for a Business Intelligence and Analytics Platform – Q&A**

- 1. What's the timeframe on the estimated "50-100 GB of data"? Is that day 1 starting volume? What is it expected to look like in 5, 10 years?**

Day 1 starting volume is expected to be about 25 GB from our legacy database, which is only used by a few countries anymore. All countries used that database to house data from 2000 – 2013, when some countries stopped using it and began building separate systems. The volume of additional legacy data from the other systems is currently unknown, so this is an estimate based on historical volumes.

- 2. Is it expected that the "dynamic ETL" concept be able to consume data through any possible change to the source systems, or would there be a list of "supported" formats (specific in-country DB schemas, DHIS2, an array of flat files definitions) that the system would be able to "pick" from for each source country/organization?**

Some of our country offices use Access databases or Excel files to collect data. The problem we run into is that the data they collect may change on a quarterly basis, so indicator names will change, new indicators will be added, and some will be deleted. In those cases, we want to avoid having a rigid ETL process that needs to be modified every quarter to manage those changes. We understand that we may need to manage those changes, but we want that process to be as simple as possible. We also expect that if we do provide a list of supported formats that we will inevitably have data that is not provided in those formats, and we would want to have some kind of alert system to let us know if the provided data is not in the specified format.

- 3. Would it be acceptable to have to know ahead of time, for any given source organization, which possible data structure should be expected from them? IE, if a country has changed from a flat file delivery to a DHIS2 system, this would be configured in the ETL system before the extract happens that month.**

Yes – we could work with county offices on a process of communication so we would know in advance what format to expect. We expect there could be cases where we aren't notified, however, so we would want to be alerted by the new system if there are errors.

- 4. The RFP didn't seem to have Appendix I that contained the list of Global Priority Indicators... Can that be provided?**

It is now posted.

- 5. Does the dynamic correlation of country program metrics to Global Priority Indicators trigger re-calculating existing measures with the new rules every month or do those changes only apply to the current (and future) months?**

In most cases changes would only apply to current and future months. We would most likely only need to recalculate historic values if an error in the calculation was found.

6. **Is the currently-licensed SQL Enterprise Edition core-based? Is it expected that all on-prem SQL resources would be installed on the same server, or could more licenses be procured to utilize more than one piece of hardware?**

Yes, it is core-based. We currently have one enterprise license available for this, but expect that we would need to procure another for a test environment. We do our licensing through TechSoup, which also limits the number of discounted licenses we are able to obtain. We are also open to using Azure instead of On-Premises, but pricing will need to be clear.

7. **Is it expected that all reporting for all audiences will utilize the same tools? IE, if some reports for internal EGPAF Global users are delivered via SSRS/Power BI Report Server, will those reports also need to be available to the public audience (of course, based on the "public" role's data security) also via SSRS/PBIRS?**

No, different audiences can use different tools.

8. **Could more details be given on the desired "data warehouse metadata auditing and report tracking for access and consumption metrics"?**

We would like to know who is submitting data and when. We would also like to know which reports/dashboards are being used and which aren't. We need to track errors as well.

9. **Is there intended to be permanent, dedicated non-production development and/or QA environments created and available to support future ongoing changes to the system?**

Yes

10. **What is the current and comprehensive number count of countries where data is being collected via the DHIS2 system/process?**

There are 3 different DHIS2 systems related to this project:

- 1) The national system. Most, if not all, *countries* (as opposed to EGPAF country programs) have a **national** DHIS2 instance which is owned and run by the respective ministry of health and used throughout the country.
- 2) The donor reporting system. Our main donor, PEPFAR, has its own reporting system which is built on DHIS2, and most EGPAF country programs need to use that system for their quarterly reporting
- 3) EGPAF systems. Currently, only one EGPAF country program has built its own instance of DHIS2 which they use as their main data warehouse. A second country program is in the development process of building their own instance as well.

At this time, the platform described in the RFP should prioritize on interfacing with the donor system and any EGPAF systems that exist or will exist in the future (2 & 3 above).

**a. What is current count of countries NOT using the DHIS2 system/process?**

Roughly 13 EGPAF country programs do not have their own instance of DHIS2. DHIS2 is becoming a standard used by many different countries and programs, however, so we expect more countries to begin implementing DHIS2 at some point in the near future.

**11. For countries not using DHIS2, are they scheduled to be converted or will they be before or during this planned project?**

It is hard to say at the moment. There certainly is no plan to build a DHIS2 instance in all country programs but some may decide to do so.

**a. Is the process different for each of these countries not using DHIS?**

Yes. The country programs have a variety of data management systems. Some use a MySQL database, others use Access, and many use Excel.

**b. How many different file formats do each of those countries use: Ex. 1 per country/clinic, etc....**

Each country may implement several different programs, which could each have a different format. One country could be implementing 2 - 8 different programs, and each could have a different format. The formats will be standard across all of the facilities for that program, though may change over time.

**12. What amount of development work does EGPAF expect their internal teams to develop and be responsible for, if any?**

Internal teams will do the report development using Power BI, and would like to be involved in the ETL development to ensure continued success after the end of the project. There are 2-3 EGPAF developers who will have time to dedicate to the project. EGPAF staff will be developing the reports and dashboards based on the new system.

**13. Currently EGPAF utilizes Power BI for reporting:**

**a. Is there an expectation of current reports being migrated, and if so....what is the approximate count of those reports?**

We also use SSRS and we plan to migrate about 12-17 reports to the new environment.

**14. Are there internal EGPAF project management (PM) resources that will provide oversight and project coordination, or will this need to be provided by the Consultant?**

We are flexible to different models of service but EGPAF's Sr. Data Architecture Manager will oversee all aspects of the project. She will not be providing day-to-day coordination of tasks. We would expect that to be handled by the Consultant.

**15. What project management methodologies are currently used or preferred by EGPAF?**

**a. Waterfall, Agile, Adaptive, etc...?**

We use more of Agile development approach, but we do not have a formal required approach.

**16. What types (Power BI, SSRS, SSIS, etc..), materials, and delivery method of training is EGPAF expecting as part of this work effort?**

We would expect to get training on the different tools that are used to build the system, particularly specifics on how to manage changes to source formats.

**17. From a documentation perspective, EGPAF is requesting documentation on mappings, roadmaps, and data dictionaries. Are there any other specific artifacts to add to this list, or is this it?**

Documentation on how to manage changes to source systems and data models for ongoing maintenance.

**18. Can 3rd party products be included in the proposed solution and how should those be accounted for in the official RFP response?**

Yes, we are open to 3<sup>rd</sup> party products, but cost will need to be clear.

**19. Should ongoing support or maintenance services be accounted for in the RFP response?**

Yes, please include any proposed ongoing support and maintenance anticipated/recommended

**20. What specific version control tools are currently in place if any?**

We have used Microsoft Visual SourceSafe 2005 on our 2005/2008 SQL Servers.

**21. Will EGPAF be providing necessary laptops or other technical tools necessary for successful delivery?**

Please include technical requests in the proposal.

**a. Will VM's be provided?**

Laptops and/or virtual machines will not be provided unless specifically requested in the proposal.

**b. Will remote work be available?**

Yes. In fact, we expect a lot of work to be remote, with the exception of training and face to face interviews.

**22. Is there any objection by EGPAF to utilize offshore resource for delivery?**

No objection but please specify details in the proposal.

**23. Will EGPAF be providing necessary Testing and QA resources for this project?**

EGPAF will conduct its own testing and QA however a) it is expected that the Consultant would do enough testing to ensure the deliverables are working according to expectations, and b) the Consultant will provide EGPAF guidance to do its own testing and QA.

**24. What is your current legacy platform built on? Microsoft – what version?**

Microsoft SQL Server 2008 R2.

**25. Does your current legacy platform allow users to export of data outside in other formats – CSV, XML, JSON? Do we need to migrate all your old data into new platform?**

The legacy system uses SQL Server 2008 R2. We would like to migrate old data to the new platform.

**26. Please provide the data formats / metadata of information collected across multiple sources and country programs.**

Data formats are listed in the RFP. Existing data collection tools use DHIS2, Access, Excel, MySQL, and SQL Server.

**27. We are assuming data integration and collection as part of the scope apart from the reporting needs. Please confirm if this is correct.**

This project does not include developing data collection tools. Data is already collected by each country program using their own tools. This project aims to integrate the data that is already being collected.

**28. Do you have any cubes, tabular reports defined in the legacy platform? If so, can you share them for scope assessment?**

No cubes at the moment. Tabular models currently only exist within Power BI reports – no existing SSAS models.

**29. Could you give a sample mapping on country program's indicators map to the respective global indicator?**

We do not have a sample mapping to share.

**30. How much amount of data is pulled per pull between country program and EGPAF?**

Each country program supports anywhere between 60 and 2,000 health facilities. A program could collect between 150 and 1,000 indicators. Data is aggregated to the facility level for each indicator and is generally collected once per quarter or once per month.

**31. Do you use currently DataZen capabilities for mobile reporting?**

We experimented with DataZen in the past but currently only use Power BI. We are open to other technologies.

**32. Do you use Active directory for managing user access and permission?**

Yes – Azure Active Directory.

**33. Does new platform mobile support include iOS, Android, and Windows based smart phone support? Or just tablets?**

This project is mainly focused on the infrastructure to support mobile reporting, and will not require the development of specific reports. The EGPAF team will develop reports and dashboards based on the new BI platform.

**34. Could you tell how many resources within Global Informatics Unit are involved in maintaining the current legacy platform?**

The Global Informatics Unit currently has 2-3 staff members whose responsibilities include (but are not limited to) the current and future platform. Because of the nature of the project, the IT department may also provide support as necessary.

**35. It is mentioned “Data volumes for the data warehouse are expected to fall between approximately 50-100 GB of data”. Is this projected based on annual data volume or historical?**

This is based on historical volumes. The volume of data in the other sources is not yet known.

**36. Do you have any existing integrations with REST API/DHIS2?**

Not yet, but we are in the process of developing a Power BI Connector for DHIS2, so we are gaining experience in working with the DHIS2 REST API.

**37. List the different type of database integrations we have to support for supporting different country program integrations? We see MySQL listed, anything else?**

MySQL, Access, SQL Server

**38. Since we don’t have access to your current environment, could you provide some idea around the estimated reusability of the existing legacy platform modules in the new platform.**

We will be starting from scratch. The existing legacy platform is based on data from only one SQL data source, which is no longer in use by most countries. We will incorporate this data into the new system, though.

**39. Given you are a nonprofit organization, any specific budget numbers we should be mindful of?**

We are a nonprofit with a small budget for indirect country support (which this project entails), and therefore often receive “nonprofit pricing”. For example, we can get discounts on software through TechSoup, and we qualify for Microsoft's discounts of Power BI Pro, which is a 70% discount on commercial pricing.

**40. What countries do we have to support other than Africa, India? We primarily want to understand any country specific regulatory, privacy requirements that needs to be complied with.**

Currently, all EGPAF country programs are in Africa and India. We do not expect that to change in the near future. See our website for the list of all countries we support. <http://pedaids.org/dashboard>

**41. Any language support given users are global?**

Not at this time, but additional languages used in our supported countries are French, Portuguese, and Swahili.