

InvestInternational

Questions submitted
Market consultation - Integrated Enterprise Reporting
7 December 2023

Nr.	Subject	Question	Answer
1	Prototype	Could you elaborate on the architecture of the prototype. Guiding questions are: a. Has the prototype been developed by your internal organization, or by an external party? b. Is the prototype hosted locally or in the cloud? If it is hosted in the cloud, what provider? c. What layers are currently included in the PostgreSQL database (e.g. staging, consumption) d. What modeling technique has been used for the data? (e.g. dimensional, normalization) e. What tool(s) is (are) used to schedule and perform the ingestion, data quality checks and transformation of the data? f. Is there already a DevOps/version control environment?	a. An external party. b. In the cloud. Hosted at Hasura. c. API, Staging, Datamodel, Reporting views. d. Normalization. e. This is done in the staging area or in Hasura. Relational datamodel. f. Yes, Github.
2	Prototype	Will there be any internal team members be involved in co-developing the datawarehouse, e.g. onboarding data sources, building dashboards, or is this under the full responsibility of the development party?	Internal team members are involved.
3	IER solution	What is the approximate data size and number of tables of the data that needs to be included in the data platform?	Input 30 tables, sizes up to 90 tables. We expect to go up to 50GB to 100GB. Not TB.
4	IER solution	Could you give a rough indication of the data quality per source (poor, medium, good)?	The data quality in Mambu is good, in nCino medium and in Exact medium/Good. Invest International will address data quality internally.
5	IER solution	Is PostgreSQL preferred as the datawarehouse, or would you be open to an OLAP solution if that would be more suitable for your situation (e.g. Clickhouse, Synapse)?	We are open to other solutions. Solutions should be cost efficient (in transition and hosting/maintenance). The current prototype is our preferred solution.
6	IER solution	Where should the datawarehouse be hosted, locally or in the cloud? If in the cloud, is there a preferred cloud provider?	We only want SAAS/Cloud solutions. Data should always stay in Europe. We prefer not to have servers to be managed by a separate cloud provider (we want a totally managed solution). If server management needs to be done by a separate cloud provider we prefer Microsoft/Azure servers in our tenant.
7	IER solution	Invest International is a semi-governmental agency (the Dutch Ministry of Finance (51%) and FMO (49%), which is also a 51% holding by the Dutch State). Are there technological and/or security prerequisites and prescriptions imposed from the government?	Yes, we need to be prepared for DORA and possibly NIS2.
8	PoC Architecture Design	It seems that the development of the POC is at an advanced stage and provides a solid foundation for further development towards the end-state. Where does prototype development stop and where does further development towards end-state start?	The further development to the end-state starts when we have a contract with a supplier for the end-state solution and the migration or handover from the prototype is done. This is where the prototype development stops.
9	IER solution	To what extent is the chosen technology of Hasura on PostgreSQL database and Auth0 coupled via JSON/REST with Mambu an established choice? Or is there possibility of a completely different choice in technology?	We are open to other solutions. Solutions should be cost efficient (in transition and hosting/maintenance). The current prototype is our preferred solution.
10	IER solution	There is no mention in the documentation about whether the solution can be built in the cloud or on premise. Is there a preference for that?	It should be built in the cloud.
11	Market consultation	In what language would you like the answers to the questionnaire and later the proposal to be delivered?	Please provide your answers in English.
12	IER solution	In slide 8, section "Outputs" a "IATI XML report*" is mentioned. What is the expectation or requirement here?	We need to report data according to the IATI standard: https://iatistandard.org/en/ . It is highly likely that we will deliver our data to an IATI tool which is predefined (and maintained) by an IATI expert vendor and which will convert our data to the IATI XML file. We do not wish to follow IATI reporting developments and maintain own developed software to create the XML. In line with our outsourcing policy, an IATI tool outsources this to an expert and changes to IER will be limited to data delivery to that tool.
13	IER solution	In slide 8, section "Outputs", "Integrated Reporting" is mentioned. Where should the reports be integrated in?	Integrated reports refers to combining data from different sources systems into an integrated report.
14	DWH	Why is it chosen to enforce referential integrity? In most modern data platforms (e.g. data lakehouse) this is not possible in this way and is sub-optimal.	To maintain data quality and because the source data is structured. Control and reliability of data in IER is considered fundamental in the technology choice. Specifically, we don't want to 'just' add data to our DW as would be the case with a data lake. Also our data sources are our own source systems, no external sources of data are currently in scope and if they are added, they will be very limited.
15	DWH	The PostgreSQL db is part of the back-end-as-a-service, is this also already decided as the main platform for the DWH?	We are open to other solutions. Solutions should be cost efficient (in transition and hosting/maintenance). The current prototype is our preferred solution.
16	DWH	Does the time traveling have to happen in the PostgreSQL db?	Yes, historical data needs to be included in the PostgreSQL database to facilitate time travel. Historical data is not maintained in our key source systems.
17	Datamodel	It is stated that the datamodel is 60%. Is this conceptual, logical or physical?	At all levels it will be 60% (80% conceptual).
18	Prototype	What does the current prototype exactly entail?	Please check slide 5 for information on the prototype.
19	Prototype	Is version control part of the prototype, if so how?	Yes, with Github.
20	Prototype	Which components use auth0 authentication?	Computer to computer and user login.
21	Prototype	To get an idea of the complexity of the current prototype can you provide us with some metrics for it? (for example; lines of code, database size).	Current database size is 1 GB but will grow. We expect to go up to 50GB to 100GB. Not TB.
22	Implementation	Does Invest International have any expectations concerning the duration and deadlines regarding the implementation phases?	Roughly, the expectation is as follows: Phase 2 needs to be completed during 2024. Subsequent phases will be confirmed in preparation for phase 3 & 4 although the intent is to finalise all phases as early as possible in 2025.
23	Implementation	In order to properly set-up self-service BI, a decent governance is important. Is this part of the tender?	Governance is definitely part of the overall project, it is not clear yet the extent to which this will be included in the tender. We expect this to be part of the internal organisation.
24	Outsourcing	Is in the scope of the project also the hosting of the platform or does Invest International have its own SaaS providers/contracts in place	The hosting of the platform is also in scope.
25	PoC Architecture Design	Which components have been used, can you list all the components / solutions used in the PoC environment and describe their role	Please check slide 5 for information on the prototype.
26	PoC Architecture Design	Which ETL tool is used and how is the data loaded from the source systems to PostgreSQL, and with what tool is the data transformed within the PostgreSQL database	API is ETL and by using the staging environment within Hasura.
27	PoC Architecture Design	Is this environment in the cloud or on-premise?	In the cloud.
28	PoC Architecture Design	Can you indicate why PostgreSQL/Hasura was chosen?	Please check slide 4 for this information.
29	PoC Architecture Design	What is the role of Hasura in this environment, our understanding is that Hasura can create API's on top of a database, is Hasura used to create API's between source systems and PostgreSQL or between PostgreSQL and consumer (tools)?	Both.
30	PoC Architecture Design	Has the PoC environment been implemented using infra as code and version control?	Version control with Github.
31	PoC Architecture Design	Have you thought about a CI/CD process? If so, what are your thoughts?	Yes. We work with CI/CD (not automated yet).
32	PoC Architecture Design	Which technology is used to determine the data lineage and which technology to make it insightful?	Via labeling source system data fields (with extended documentation).
33	PoC Architecture Design	Does the target architecture support data quality challenges and/or objectives? If so, what processes and technology were envisioned?	See slide 7 for this information.
34	PoC Architecture Design	Which arguments/criteria played a role in determining the flexibility, agility and scalability of the chosen solution?	It should be easy to add source systems with API's and add to the database with a few clicks. SAAS/Hasura is scalable enough to start small and cost efficient and grow as we go.
35	PoC Architecture Design	We understand that you no longer want to question the architecture choices made, is this still the case?	We are open to other solutions. Solutions should be cost efficient (in transition and hosting/maintenance). The current prototype is our preferred solution.

36	Business	What criteria are applied to cost control (OPEX and CAPEX)?	We are looking for a cost efficient solution.
37	Business	What business problem(s) will this data platform solve for Invest International?	Please refer to the lower half of slide 8 for the reports that need to be made that are now created manually or that are not yet created at all.
38	Architecture: Fit for Purpose principle	In order to validate if your architecture fits its current and expected purpose, can you share the required key data capabilities?	We believe this is not relevant during this stage of the project.
39	Business	Will this data platform be used for reporting to regulators? If so, Which regulators and which reports?	The data in the platform will be used to do all reporting and in the future that will most likely include regulators.
40	Business	Which sustainability insights are used in the primary processes?	We don't understand the relevance of this question.
41	Business	Can Invest International source meaningful ESG data, are there still gaps on datasets which you want to be able to report on?	Impact and ESG source systems will deliver data to the IER solution and the solution should provide for data enrichment.
42	Business	Does a functional business model of your key processes exist, which can later be the fundament for technical implementation?	Yes.