

Roadmap for Data Citizenship @PHARMCO's R&D Team

Rationale for this document

This template is for an R&D project as it entails the following aspects:

1. Introducing a new way of working for both the IT support team and the R&D data science users;
2. Proving the enhanced capabilities of the team, based on its expectations which will be formulated as use cases and user requirements;
3. Combining the individual freedom of data citizenship and the underlying self-service data acquisition platform with process and data governance.

After further conversations, other aspects may emerge but these three fundamental observations provide sufficient base for a draft roadmap.

We advise you to consider this as a consulting project as we can clearly distinguish project phases and their deliverables. Per deliverable we can make an estimate, based on the knowledge acquired in the previous phase. Needless to say, the first phase estimate will have a larger margin of error than the next one(s).

All deliverables will be documented in our Enterprise Architecture(EA) tool, Sparx which can deliver both Word documents and .xmi files for exchange with other EA tools PHARMCO may use. Where appropriate, each phase description is illustrated with the relevant meta models we use in our EA development method, based on TOGAF 9.

Overview of the phases

These 10 phases can be deployed in an agile way as of phase three. Depending on the first intakes, we can determine the scope of the epics, e.g. based on domains, persons, algorithms,...

Then we can also decide on the shape of each epic: end to end from data to analysis delivered or layer by layer in the overall data and Business Hub architecture.

The first two phases lay the foundation of the project:

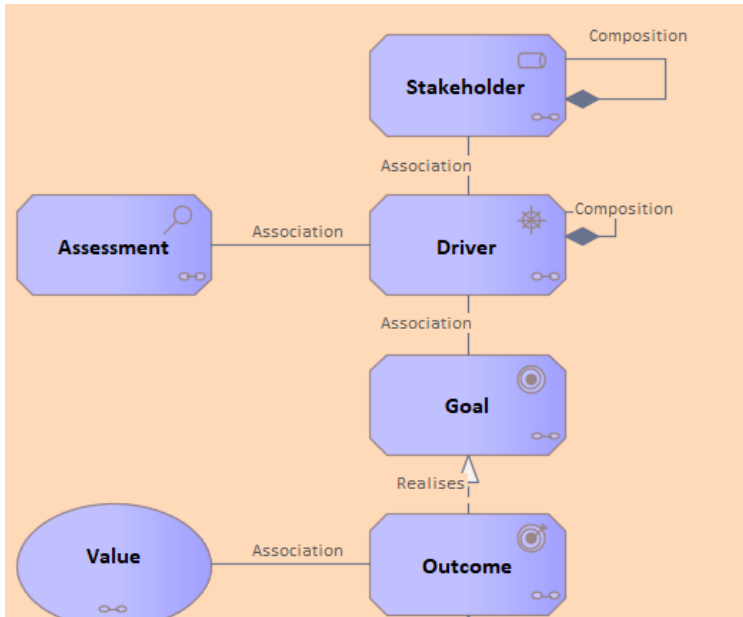
- Phase 1: stakeholder mapping and project value statement;
- Phase 2: high level project brief

Then we can start working in iterations:

- Phase 3: In-depth analysis where the epics and their user stories are defined;
- Phase 4: Business Hub design where we make sure we can always reiterate if a better configuration is found;
- Phase 5: Business Hub configuration and testing;
- Phase 6: Deployment;
- Phase 7: Documentation and training development and delivery;
- Phase 8: Go-Live and Go-Live support;
- Phase 9: Post implementation review;
- Phase 10: Continuous improvement, maintenance and user support.

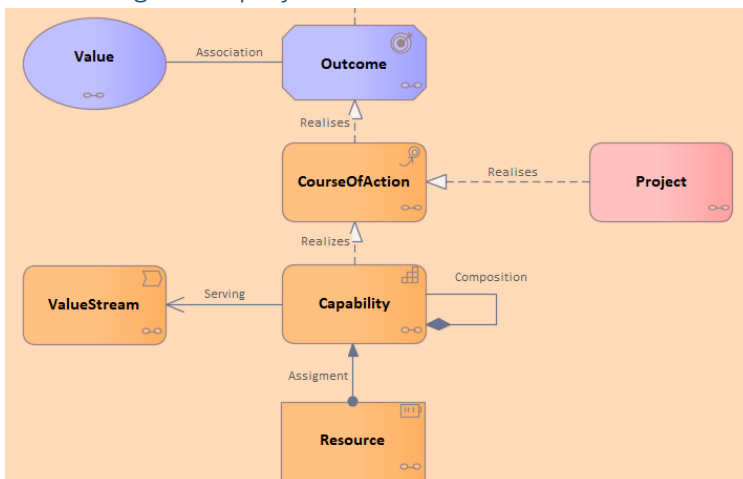
High level roadmap

Phase 1: Stakeholder mapping and project value statement



We will establish a stakeholder map, including PHARMCO's assessment of the situation, the objectives of the project and the desired outcomes and the value they produce. We estimate this work at 1 day intake with the IT Manager or an appointed contact and 4 days to produce the report.

Phase 2: High level project brief



All stakeholders being known, their high level requirements need to be refined in to use cases which serve as input for a project charter and – in the next phase for user stories we need to realise with the KNIME setup.

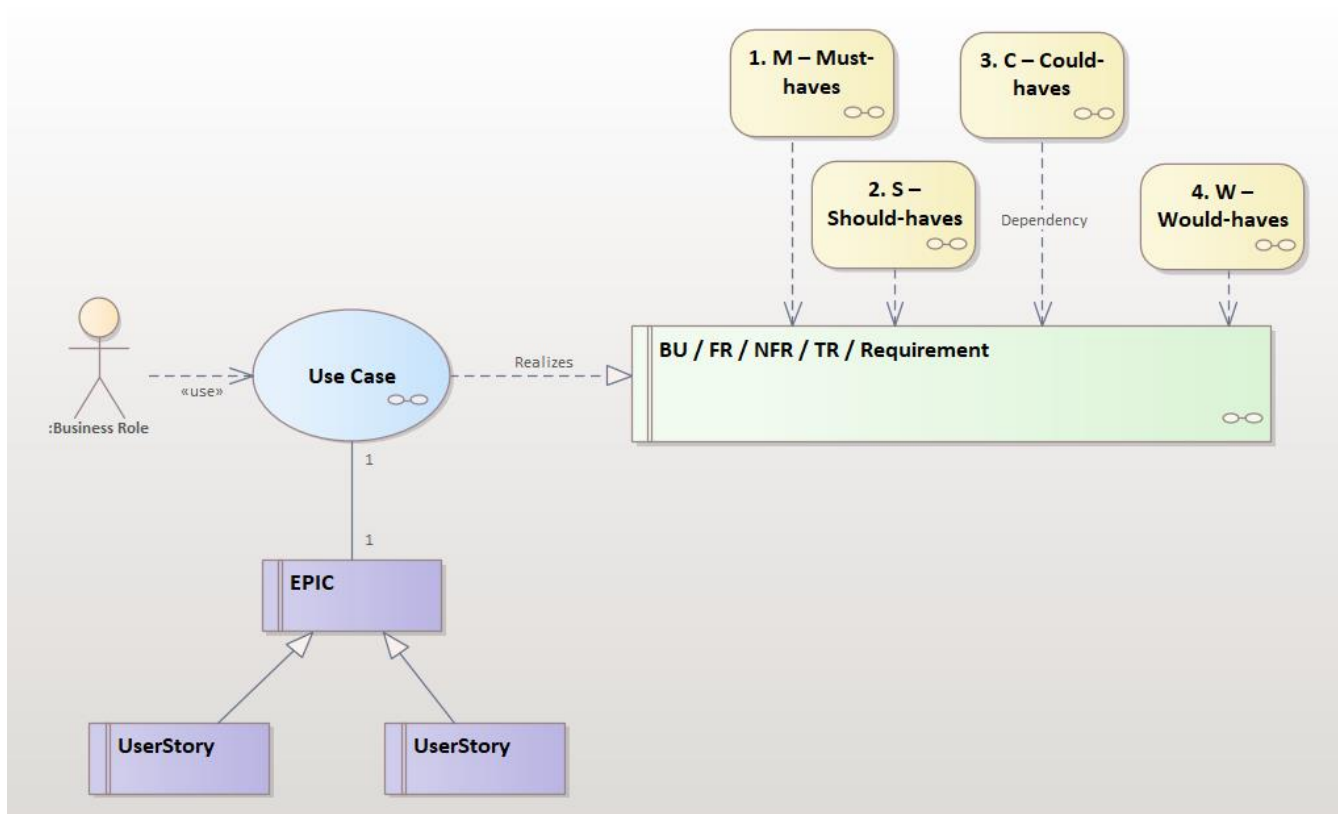
Deliverables:

- Background;
- Project Definition, including the major use cases;
- Outline Business Case;
- Project tolerances;
- Customer's quality expectations;
- Acceptance criteria;
- Project dependencies;
- Risks.

The estimate can be done after closing the first phase and will depend on the information that is already available in the IT department as well as in the R&D department.

Phase 3: In-depth analysis

This is where the rubber meets the road: based on the previously identified use cases we will formulate the user requirements to make sure the KNIME implementation meets their needs. Below we sum up a few of the considerations as an example.

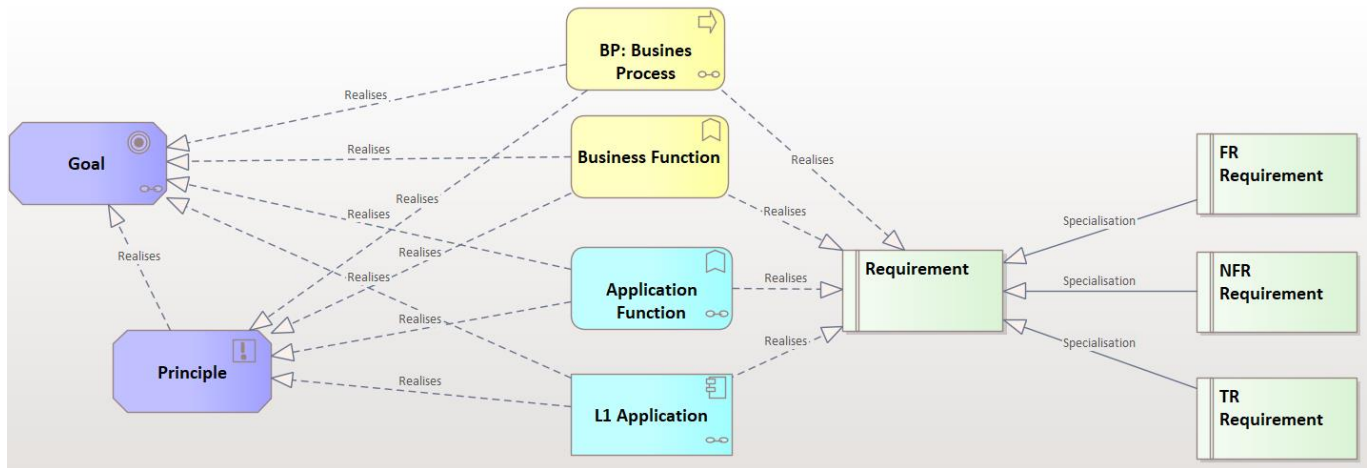


The analysis should provide us with answers to questions like:

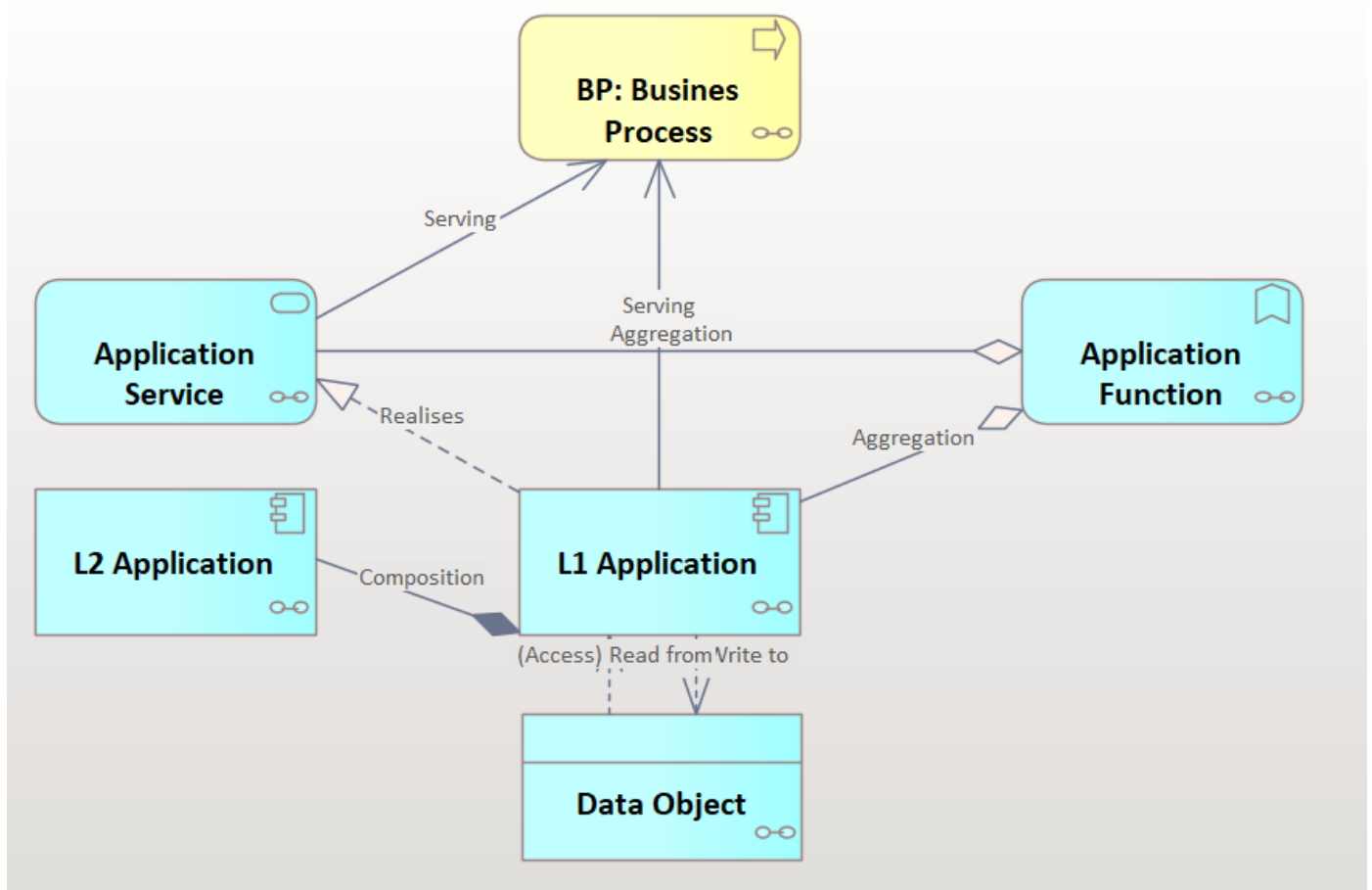
- Will there be a shared execution context or separate contexts , e.g. based on research domain, research methodology, algorithms used, etc...?
- Will the data be containerised per user type or shared throughout the department?
- Is there a need for extra Docker execution contexts?
- What level of monitoring and governance is acceptable?
- What is the process for job and workflow optimisation?
- Will there be a need for snapshot backups?
- What is the timeline of introduction and roll out per target group?
- Can users be members of multiple teams?

Phase 4: KNIME Business Hub design

The Business Hub design document with its design specs serves as a blueprint for the configuration phase. We will assess if there is a need for prototypes or mock-ups. Sometimes (depending on the risk aversity of the target groups) these can be important for getting early feedback from stakeholders, ensuring that the final product meets user expectations.



The data architecture will also have an impact on the Business Hub setup. Therefore, we will provide PHARMCO with an architectural design for the data sets used by the R&D community. The level of detail is to be determined by both users and IT department. If this is done in an agile way, we need to maintain a data lakehouse, capable of instantiating new data sets if the requirements change.



Phase 5: KNIME Business Hub configuration and testing

Based on the input from the previous phase, the already installed Business Hub is configured:

- Setting up contexts, teams and their team administrator as well as the team members;
- Setting up the first tailor made nodes, workflows and jobs together with the R&D users;
- Setting up the monitoring and governance processes and measurements;
- Adding data connections where necessary;
- Providing access to “read only” users;
- Etc...

Phase 6: Business Hub Deployment

As soon as the select group of convinced power users has given a “go” we can start rolling out the Business Hub to the rest of the team. The speed of roll out will be determined by the speed with which the users adopt the new way of working and the tool as well as the availability of resources in the IT and data science department.

Phase 7: Documentation and training development and delivery

Since we use an EA tool to document the entire process and the configuration, this phase should require minimal throughput time. Yet training the users will probably be a matter of exploiting the R&D’s available time to the maximum. To be analysed further.

Phase 8: Go-Live and Go-Live support

The first power users need to be on board with the previous phases. From the configuration and testing phase we should be aware of the potential problems users may face and have a response plan ready.

Phase 9: Post implementation review

Using the four-eyes principle, we guess this phase should be done by a PHARMCO authority, to be determined earlier during the project. The evaluation of the roll out process, the problems encountered and the way they were resolved should provide input for the further roll out to the rest of the R&D community.

Phase 10: Continuous improvement, maintenance and user support

The next batch(es) of users, teams and admins are added to the Business Hub as well as new workflows, jobs, node collections etc...

More information?

Contact Bert Brijs: 0475 81 80 30 or bert@linguafrancaconsulting.eu