PA2578 – ASSIGNMENT 3

Asset Management Strategies and Open Source Strategies

PART I – ASSET MANAGEMENT AND STRATEGIES

I work within a central unit that provides IT & analytics support to other units in the telecommunication giant. As a part of the unit we own multiple analytical products and platforms that are used both within the unit as well as by other units within the company.

To manage recurring and new user access requests within the company, the IT team has developed an access management portal that enables each user to view, request and manage their access requests to different internal and external products. This access management portal also enables review of requests to ensure approval before gaining access to sensitive products and services. This helps provide a consolidated view of products and services at both user as well as organizational level. To foster rapid response, the access management tool is tightly integrated with another in-house self-service tool to not only download and install softwares but also activate and manage licensing for each user. The self-service tool also enables the central IT team to centrally manage version control and patch updates in a secure environment without user intervention. The tool also serves as a place to centrally manage migrations and disposal of software products.

Since the company is diverse with many teams using different tools to manage specific needs and productivity gains within individual units , the IT team has also developed an in-house portal that catalogs all available products and services. This enables teams to quickly discover products relevant to their use-cases and get access without needing to draw-up individual licensing contracts, centralizing software procurement. However, to manage the growing needs for new products, a separate procurement portal is in-place that ensures new vendor and software requests are managed via established policies and approval processes.

While the requests are managed in JIRA to maintain information flow across the hierarchical levels and wide-spread teams, to manage deployment and development of in-house tools and services, the unit has established a well outlined continuous integration and continuous delivery (CICD) via Github to maintain version control and source code. Github is connected to run on both on-prem as well as on-cloud hardware depending on individual requirements via Github actions that allows teams to automate deployment tasks and establish pre-defined environment parameters for different use-cases, products and services. Eg. A github action flow to run machine learning algorithms via docker container managed by kubernetes, the artifacts stored within an central artifact store (artifactory) and Min-IO buckets to manage storing of any output generated during the process. A separate Github action exists to industrialize data pipelines connecting with Airflow and internal data warehouses. Such github action flows not only allow streamlined and standardized deployment practices but also improve developer productivity through centrally managed production procedures.

Knowledge management and decimation is another area where the organization has made considerable developments and standardized processes to ensure efficient knowledge as an asset management. The

organization-wide knowledge repository Confluent is used to store key documents capturing strategy, standardized processes, domain specific information, unit specific playbooks, technical documentation as well as product and service documentation. The in-build user management within Confluent allows selective dissemination of sensitive information and caters to a wide audience. The platform serves as a key repository for all knowledge artifacts (slightly critical, vigilant, handy & strategic). While the product documentation is captured as a part of every sprint cycle associated with a product release or major product change, it often becomes obsolete especially in products with high velocity or those with limited usage.

As a part of Externalization strategy, each capability team (eg. data science, data engineering, data analyst, tester etc) maintains their respective crowd-sourced playbook that captures not only the major processes of daily work but also parts of the tacit knowledge gained through experience. These playbooks hosted on Confluent forms a key source of information dissemination within the teams among experienced as well as new joiners. To further promote peer dissemination of tacit knowledge, multiple discussion forums exist to socialize and contribute towards development of a wider ecosystem. Eg. Data Science group that connects data scientists across teams to meet every 2 weeks and share learnings as well as respective work, thereby also contributing towards discovery of collaboration opportunities.

The above also illustrates how knowledge management and dissemination is closely integrated into product strategy. To improve product adoption as well as gather feedback the socialization forums such as product management forum, network management, core network management, Sweden infrastructure management, data science etc forums provide an efficient platform and is often well integrated as a part of product strategy. The product artifacts such as technical documentation, architecture, as well as product user guide are also developed and shared via Confluent and the development is often included as a part of product development sprints and strategy. Multiple product strategies and how they work together to help the wider unit strategy and vision are also illustrated within Confluent to help provide overview to people within and outside of the unit.

PART II – PRINCIPLES OF INDUSTRIAL OPEN SOURCE

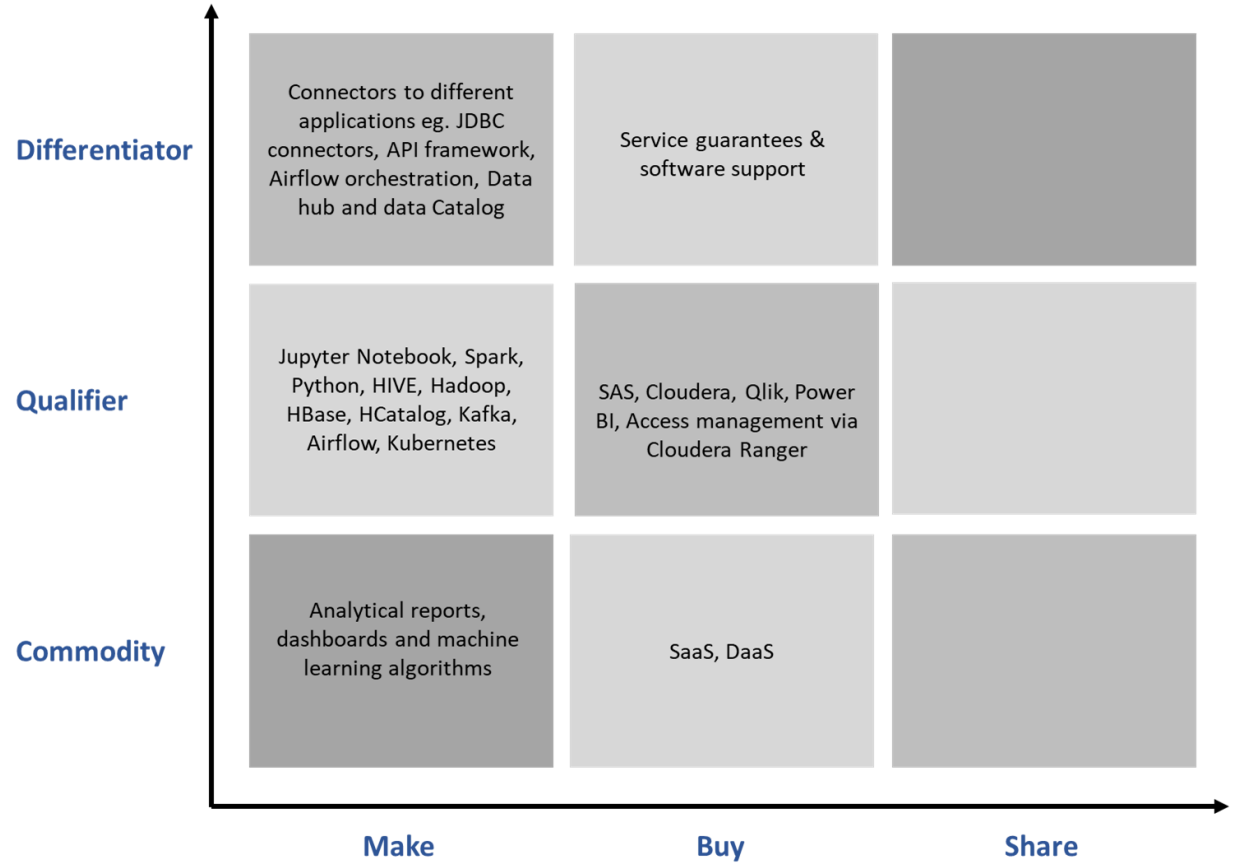
Within our unit we have established products and services using a combination of open source software solutions, 3rd party software products as well as in-house developed components. I have come to observe different parts of the unit belonging to different phases of their respective Open Source software journey. The differentiation exists owing to highly sensitive data and organization’s inclination towards maintaining a high level of security and privacy compliance.

The data science teams often rely on algorithms developed via open source communities and policies are in-place to ensure security and compliance when installing a new package. The production environment encapsulates the widely used modules from a trusted list of repositories to ensure a stable and secure system. However, the policies and adoption of open source softwares is relatively restricted among the data analyst community that develops analytical visualizations and dashboards on sensitive data fetched from internal data warehouses. The adoption is limited to 3rd party procured softwares that ensure strict privacy and legal compliance. Being end-user facing, these softwares are often bought with service agreements, something that remains limited within the open source community of similar solutions.

A good mix of open source components and in-house components can be observed when looking at the data science workbench or analytical platform as a whole. The platform team over the years has successfully established policies to evaluate and use open source components, integrating them with in-house developed

components to provide multiple customized solutions that cater to the diverse needs within the company. As a policy the team first evaluates the need followed by security and additional effort needed for the open source software to be integrated within the ecosystem. Open source solutions are screened through a thorough process evaluating multiple aspects such as number of contributors, adoption among the community, available production support as well as extent to which the solution may be future proof. While the unit lacks a dedicated open source officer, the role is fulfilled via a core group of Staff data architects, Staff machine learning engineers, Staff data engineers and product managers. The core group evaluates the need, available solutions and closely works with the Data, Privacy and Legal unit to ensure strict compliance. Once security and legal compliance is in place, the team forks the repository into internal Github repositories and starts working on a POC (proof of concept) within the development environment. This provides a safe and secure environment to not only test functionality but also evaluate the additional effort needed to build, maintain and put into production. During this phase we also evaluate if maintenance can be out-sourced and if customized components needed for integration can also be sourced via open source community or other paid solutions.

The following picture illustrates make-buy-share analysis for the analytical platform.



Although the organization not only lacks intent but also infrastructure to share the in-house developments within the wider open source community, I have observed an increasing awareness and shift in mindset to start contributing and exploring open source models as additional revenue streams. Hopefully in the near future, the unit would invest in developing this further.

PART III – RETURNING TO YOUR BUSINESS MODEL CANVAS FROM ASSIGNMENT 1

As a part of the 1st assignment, I described the AI Traffic forecast that I am building. The solution consists of an ML algorithm, data preprocessing and postprocessing and a visualizing dashboard. The solution is built on an in-house analytics workbench that is stitched together using open source and 3rd party components. The production code however remains independent of the tech stack and can potentially run via docker containers. This makes the Indirect business model difficult to implement and commercialize.

Looking closely we can find the solution itself compiles multiple open source algorithms customized only to match the underlying data structure and patterns within the data. This layered structure makes it difficult to open source the solution across a wide community. An extended business model would mean developing very specific customisations that may be difficult to scale.

However, there may be niche users especially within telecommunication and large retail where this may provide significant cost savings and benefit from a quick to implement solution. Providing the integrated solution coupled with a visualization layer (dashboard) as a free service in exchange for data may also pose challenges given data security and privacy concerns prevalent within the industry. This makes the Asymmetric Business Model difficult to implement and commercialize.

However, with management approval and limited budget constraints, we can provide a low-cost SaaS solution in exchange of model parameters (as opposed to data). This would help build a more generic AI model, capturing reality and a wider set of real-world use cases. A AI model that can be commercialized to produce highly accurate forecasts in future once trained on enough data. A monetisation model that may be classified as Asymmetric Business Model and that is similar to that adopted by companies offering GPT models eg. Google, AWS, Azure, Open AI etc.