**First Interview:**

**Codes**

* **The Scientific Data Officer**
* **The importance and need of the role, and current state of their companies**
* **Challenges they are facing**
* **The benefits of utilizing dark data**
* **Changes to the current structure in the company**
* **The role of the SDO**
* **Selling the value of the SDO**
* **APARAVI**
* **New Classification System**

**Themes**

* **First Theme:**
  + **The Scientific Data Officer**
  + **The importance and need of the role, and current state of their companies**
  + **Challenges they are facing**
  + **The benefits of utilizing dark data**
* **Second Theme:**
  + **The Scientific Data Officer**
  + **Changes to the current structure in the company**
  + **The role of the SDO**
  + **Selling the value of the SDO**
  + **Challenges they are facing**
* **Third Theme**
  + **APARAVI**
  + **New Classification System**

**Results:**

He thinks that the SDO is a very new role and believes the topics of ROT and Dark Data are very important. They are currently exploring these topics because when it comes to new deep learning techniques, they rely on data that they usually don't capture or have, such as unstructured data, especially like LLMs or new multimodal models, as they work with unstructured data.

Recently, they have been shifting their perspective on data to see what other data they haven't captured before that they can now capture and utilize.

They didn't capture it because they might have thought it was not important, they didn't have the capabilities to process it at the time, or the data was highly unstructured (and they didn't have time to make it structured), or because of the difficulties in capturing this data.

Nowadays, they have the capabilities to work with unstructured data.

Therefore, he thinks that the role of the SDO might be very helpful for organizations, as he can meet with every department, identify their projects, and then identify which pieces of data need to be captured. Therefore, he can be the point of contact for all departments that they can refer to when they need specific pieces of data.

He thinks that there might be one person in the organization that is an SDO, or the additional roles of the SDO can be assigned to a data scientist working in the company.

He can see the value of this role and thinks that this role is definitely needed in organizations. He believes that this role will exist in most, if not all, large organizations in the near future as the amount of unstructured data that organizations are capturing is increasing exponentially, and therefore, it is posing more challenges.

He also thinks that the idea of the research is doable and very relatable for them.

Also, he says that they are currently at the stage where they need a position like the SDO to manage all their data.

They have a CDO.

He also thinks that we need to question how the introduction of this role might change the pipelines on how they capture and utilize their data because sometimes there might be some specific process in which they fail to capture its actual data, and the data is just being said not recorded, or it is being written but is not comprehensive enough, and they are facing this a lot.

He gave us an example of a simple task of summarizing a conversation between you and a user. It is not being summarized properly, or it might just say "problem solved" without providing additional details. However, we are more interested in knowing how the problem was solved, what the original problem was, what the solution was, what the resolution time was, etc. All of this, if properly captured, can help the organization in identifying many other things that they did not anticipate previously.

He suggests that in addition to the roles of the SDO, we need to add the task of identifying what other types of data need to be captured, how we should capture it, what process should we change in the pipeline, and in what format to store it.

He hadn't heard about the role previously, but he kind of knew that a similar role was coming as there are a lot of gaps that needed to be covered by a role that doesn't exist yet, which the SDO covers. This is also due to the fact that a lot of new roles are emerging recently, such as the role of prompt engineering, AI engineer, generative AI, etc., and he thinks that all of these new roles are emerging because of all the new technologies and what they require.

And he thinks that the SDO's main objective will be to handle all the dark data.

He thinks that the SDO will help in data-driven decision making, as we are building the baseline of the new technologies that will solve more difficult and complex problems. It might help you identify some problems in the organization that will lead you to make some data-driven decisions.

It is not always easy to manage the metadata as there could be some difficulties in formatting it correctly, or simply time constraints of not having enough time to properly write the metadata, as there is no way to capture everything.

He thinks that it is the responsibility of the SDO to change the process on how we formulate our metadata so that what is captured can be comprehended and used in a more timely manner.

He thinks that in order to convince the departments or management to capture specific data, we have to give them the use case of the data and what value it brings, by giving them a proof of concept (such as a simple dashboard) so that they can buy into the idea and understand the value of doing this (you have to show them value then they can work).

He gave us an example that previously they didn't have much work as not much of the management bought into the department, but after they showed them the value that they can add, they became overwhelmed with work.

He was captivated by the introduction of APARAVI’s Intelligent Management of Unstructured Data, seeing it as a tool to enhance the SDO’s toolkit. He recognized APARAVI’s potential to revolutionize managing the bank's burgeoning unstructured data. He viewed APARAVI as a powerful ally in navigating and curating vast amounts of unstructured data, reducing dark data prevalence.

He believed in APARAVI’s ability to streamline data lifecycle processes in accordance with FAIR principles, aligning with visions for robust, ethical, and efficient data management strategies. He envisioned the SDO using APARAVI to capture and maximize the utility of information, transforming data asset leveraging. APARAVI classifies the data by using over 200 classifiers that are specialized in specific topics in unstructured data such as email labeling, CV labeling, etc.

He was equally intrigued by the concept of New Classification System Development for strategic enhancement of the SDO’s data management capabilities. He saw the potential of the system to enable proactive categorization of new data by employees, distinguishing between business-critical and potentially dark data.

He viewed the system’s ability to assign retention periods to ROT data as an effective strategy against dark data challenges. He appreciated the system empowering the SDO to make informed decisions on data archiving or deletion, optimizing storage, enhancing security, and ensuring compliance

**Second Interview:**

**Codes**

* **The Scientific Data Officer**
* **The importance and need of the role, and current state of their companies**
* **Challenges they are facing**
* **The benefits of utilizing dark data**
* **Changes to the current structure in the company**
* **The role of the SDO**
* **Selling the value of the SDO**
* **APARAVI**
* **New Classification System**

**Themes**

* **First Theme:**
  + **The Scientific Data Officer**
  + **Challenges they are facing**
  + **The benefits of utilizing dark data**
  + **The role of the SDO**
* **Second Theme:**
  + **The importance and need of the role, and current state of their companies**
  + **Selling the value of the SDO**
  + **SDO’s skill set**
* **Third Theme**
  + **APARAVI**
  + **New Classification System**

**Results:**

He stated that even structured data contains a lot of dark data (data that is stored but not used) and that it doesn't only relate to unstructured data.

They discovered that when utilizing this data, they found out that it has great potential and that it helped the organization increase its profits by making data-driven decisions by using the data that they used to store and not utilize. They started using this data to elevate their marketing strategy, increase profits, reach new market segments, and gain new customers.

He thinks that the role of the SDO is a new challenging role that is needed and good.

He also thinks that the roles of the SDO are not only his efforts, rather it requires the collaboration of all the different departments, all the different employees, to be able to really take the full potential of this new role.

He thinks that based on his experience, dark data is an actual problem and a big problem as in their company as they faced many instances where they have been asked to capture various pieces of data that they have stored but never used till this day.

Their company isn't a big data company, but all their customers that they deal with are big data companies, as most of their clients are big data companies, as they work with multiple telecom operators, big-name insurance companies, and other big data companies.

They newly started forming their AI team about a year ago, but unfortunately, they still haven't reached a point in which they have a chief data officer.

They do not have a CDO.

He thinks that the SDO will become a standard role in IT companies if this role proved its true value in the industry. He mentioned an example of AI, as previously not much wanted to invest in AI and the roles of AI personnel in their companies, but after AI showed its true value and potential in the industry it has become a must-have department in almost all big companies, and he thinks that the SDO will go through a similar evolution. He says that without this value no one will invest in this role.

The SDO has to have a good knowledge of the data in the company and should have good judgment on whether the data is important or not, whether it has value or not, whether it could be utilized or not, if it is dark or not. He also believes that the SDO should have a specific set of skills that will help him in doing these roles.

He also needs to have strong communication skills in order to be able to effectively communicate with the other departments. When trying to deliver a point on whether the company should store, remove, or utilize a specific piece of data, he has to also be able to fully and properly justify each decision that he proposes along with presenting the value that making this decision will add to the company.

He thinks that some of the SDO's skill sets must include: data management, critical thinking, exploration for data using different approaches, ability to see the big picture (to understand all the data that the company stores), he should take the initiative to understand what is going on and what we are storing and why, presentation and convincing skills, strong communication skills, willingness to improve oneself, ability to structure unstructured data, along with other skills that are required by all roles in any company (e.g., time management, organization, social skills, etc.).

He was intrigued by APARAVI, a platform that could empower the Scientific Data Officer (SDO) in their organization by transforming the management and utilization of unstructured and dark data.

He saw APARAVI as an invaluable tool for the SDO, with its potential for dark data identification, classification, and practical handling methods.

He envisioned APARAVI driving the organization towards improved data transparency, operational efficiency, and strategic depth.

Despite his enthusiasm for APARAVI, he advocated for a comprehensive evaluation to ensure its compatibility with the organization’s data environment and objectives.

He was interested in the impact of New Classification System Development on enhancing the SDO’s data management strategy by proactively classifying new data and assigning retention periods.

He saw the system as a way to alleviate the burden of managing increasing data volumes, optimizing storage, mitigating risks, and unlocking hidden data value.

Acknowledging challenges, including the need for meticulous data assessment and overcoming potential cultural resistance, he was eager to explore how the system could enhance the SDO’s data management framework.