

**ISKO Singapore – Irena Zadonsky on Data Governance
June 17, 2022**

Edited chat transcript notes

The large scale move to the cloud, especially in the wake of the Covid pandemic, has created new governance challenges - more distributed data, more access points and pathways, less clarity over who is responsible for what.

Data volumes are increasing, manual governance is no longer feasible. Automated governance is increasing, which means being able to define rules around data management and use. Rules around data use can come from different sources, including data vendors.

How do we accelerate the move to the cloud while retaining the guard rails?

Metadata is key to all of this: every use of data creates metadata, so it is important to be able to track this.

Access: how does access governance change when you aggregate data from different sources, and when they may have different rules and licenses?

Transparency and Fairness: this is about eliminating bias from the way that data is used. How do we ensure there is no bias built into automation, use of AI and machine learning?

Automation allows us to be metadata driven.

Gie: Could the pandemic have affected how we have shifted our views on what's critical?

Sridhar: Increasingly, we're also seeing data ethics becoming critical in how roles and responsibilities are being defined across orgs.

Data governance was previously responsibility of data managers, now the focus is pointing to business leaders and business owners, appropriate sourcing and use.

We are seeing a convergence of data governance, IM, KM, RM governance.

Data governance should cover data in other forms, not just structured databases e.g. data contained in or extracted from documents. To the data folks, unstructured data is unfamiliar territory. But it's becoming more important.

Initial drivers for data governance were defensive - oriented to risk and compliance, but increasingly governance driven by the cloud/AI/ML is offensive - this is about creating new value from data, especially when aggregated and in large volumes.

Governance cannot just limit its scope to operations, decisions take place at all levels in the organisation, from strategic to operational, and so the governance framework needs to cover all levels.

Eugene: A lot of parallels in my view between information security and data governance, also the Software Development Lifecycle (SDLC) so one can borrow a few ideas.

Sridhar: @Eugene- exactly, Governance is about having checks and balances all along the lifecycle.

Bill: The open data movement has asked or required that data made available cannot be restricted. It is a high level statement which in practical terms results in many questions, e.g. can we make money from open data, where does privacy start and openness end, just thoughts

Patrick: @Bill, yes and countries such as the USA are now requiring that data generated within publicly funded research should be made open.

In the context of governance, "change management" planning is now being reframed as "engagement management" - it is about the appropriate engagement strategies to use at different points of growing governance. We need to do a lot of sensemaking up front, to figure out the core principles that will apply to us, and the areas of support we will need. We need to gather use cases.

Patrick: Participatory, Adaptive, Iterative - sounds like KM!

On the convergence of Data Management and Knowledge Management, this is becoming more important. We need to look at what is in common across the structured and unstructured data sets.

Eugene: I think Knowledge Organization and creating faceted taxonomies of what data is about can help sidestep some of the arguments about data models (when building data catalogues). There is a convergence here between data management and knowledge management.

Meghan: When we talk about being metadata driven, knowledge management is key. Lots of threads between KM and metadata storage/access.

In finding the right governance model for themselves, organisations need to "feel their way" into the right mix for them, e.g. where they sit in the spectrum between centralised and distributed. They might swing between different models before settling on the right model for them. The operating model may be emergent out of competing interests and forces, and it has to adapt to current needs - goals, environment. How to set boundaries between what's centrally governed vs distributed? There are some core elements and some that clearly need to be distributed. The guiding focus should always be on the value to the customer. Data strategy should not be separate from business strategy.

Meghan: Also, when you distribute decision making it can slow down processes. Having a central authority model is sometimes more efficient.

How do we know if our governance model is working or not working? Monitor for culture rejection signs. Different interest groups may have opposing interests - e.g.

control and compliance vs opportunity. The governance organisation should be facilitating the necessary dialogue rather than parties retreating to their silos. You know you are doing well when people come to you to help you solve their issues, or to ask for help in getting their data in order. Quick wins: it is very important to maintain energy and commitment - use cases gathered up front help us define the challenges and to define what success looks like.

In a sense, data governance acts as an interpreter between the business and IT.

Is there a risk that we choose quick wins simply because they are easy to achieve but not scalable? Deciding on that can be based on prioritisation of use cases: they will give us criteria for business value, scalability.

Key enablers for data governance: data literacy, enterprise data catalogue (lots of tools available now to compile and manage data catalogues). Work with ambassadors to promote adoption and communication. Communities of practice also helpful - sharing knowledge, building alliances, spreading good practice and good stories of results.

For people who don't want to be data stewards, make it easy for them. Build trust and credibility, do things together. Set goals based on where you are now.

Computational governance = metadata driven governance... then the question arises, what happens to the data governance team when it's all automated? Automation is only the execution, the definition and adaptation continues to be needed, so the data governance team is still giving high level direction and monitoring effectiveness.

“The role of governance will be defining good enough” and steering the application of governance to data in an evolving environment.

Edited by Patrick Lambe