2020-04-17 Call Note - Data Integration Critical Capabilities

Data Integration – Critical Capabilities

Date: 4/17/2020 2pm EST

Presenter – Jim Phelps (University of Washington)

Scribe: Darrell Cook (Yale University)

Materials / Deck: Data Platform Critical Capabilities 2020-03-03

Meeting Notes

- Began with just quick Introductions of name and school / university (39 Participants on call)
- Why write this and goals of deck
  - Need to be posting architecture foundations (and what should we be calling that) that can be published and put out. Foundation for how we are thinking about why we are doing it and shortcutting around conversations that are published and known to help others understand the reasoning why we are doing things the way we are doing them and that are architecturally important.
  - Builds on foundation of System Information Systems and their role in the future
  - Look to possibly put together some type of article, or set of articles or video's on the UW website.
  - Jim looked to build a story along with case studies so people can help put these things into context and what others have done before them.
- Digital Transformation is coming and not something we can put off.   Requires a really rich environment to act and run.   Initiated out of customer experiences.
- Pace Layers and Systems of Record: Rates of change in organizations.   Focus on maintaining integrity of data thru consistent business rules and business processes.   Systems should not change very often. Do not want to be doing heavy customization. Two case studies included here.
  - Pace Layers from the Itana wiki. https://spaces.at.internet2.edu/display/itana/Pace+Layers
- System of Differentiation also knowns as Systems of uniqueness: Keeping system of record as vanilla as possible and doing unique items above that.
- Review of case studies from the Student Information Systems work discussed last meeting and also Sun Microsystems.
- Connective Tissue Layer
  - One place that is often missed.   The area of connective tissue layer also allows for governance.   How often they are doing it, security and SLA is correct, etc.
  - Louis K – commented that 1 complexity of connective issue layer in example given is how transactional and realtime your organization is. Moving into more time based user experience on events and activities then those realtime considerations become greater and questions whether connective tissues layer affords us protection of systems of record.
  - John M. commented: When do you go to mediated API's vs cached layer in the connective tissue (Realtime event driven vs protecting the SOR vs transactional integrations)?   If ERP goes down for maintenance then users don't experience any issues because you have this near realtime changes.
  - Question: Anyone implemented a caching layer?   John M. did indicate they are doing this for subset of workday ERP HR data and looking to do the same for Finance.
  - Is Pub / Hub part of this layer? Sure.   Probably could be done in a queue.   definitely think pub/sub would be included here.
- MASA Model (Mesh Application Service Architecture) – layer of API's that are mediating systems behind them.   Mediations could be built fit for purpose (layer of customization).   Allows you to customize experience of end user and data they see and that is highly performant for them (getting to that hyper personalized experience).
  - JJ Commented that *"Governance becomes a bigger questions here when we start introducing mediation layers or identifying business services are needed and how are they instantiated."* Maybe another piece that needs to be included here (role of new governance structures)
  - Question from Jeff K - Is this like Mulesoft's three-layer approach to APIs with System Layer APIs talking system language, fronted by Process Layer APIs and topped with Experience Layer APIs?"   If your exposing data via API's then those layers will need isolation so Mulesoft has very similar structure where lowest levels typically use same data structures, semantics, etc.

@James yes, though it's excellent that both emphasis the crucial importance of providing optimized experiences at the API layer for builders calling those APIs (for others: https://blogs.mulesoft.com/dev/api-dev/what-is-api-led-connectivity/.   Mulesoft's model I think insinuates that there are governance layers between System (Central IT), Process (LOB), and User Experience (Devs).

  - Louis K commented on Pace layers: As you move up stack you in some ways democratize the data / governance layers that need to be in place.   As we move up to Innovation and uniqueness we may need to consider new governance types.

- Question from Jeff K - *how are folk shipping the data access rules, data semantics, and rules for things like "how to calculate academic standing" from the source systems to the Data-Rich Activity Hubs in ways that are faithful to the source, safe and usable, and not requiring separate maintenance of all that stuff in the Activity Hub?*   Don't have answer from UCSD.   Follow-up with them on this and maybe get them onto a call.   Seeing a goal there for UCSD was feeding data to a community of skilled and trusted _analysts_ might partly answer my question. We've also found this Gartner piece helpful in positioning descriptions about Data Hubs in general: Freidman, T. & White, A. (2019) _Data Hubs: Understanding the Types, Characteristics and Use Cases, Gartner for IT Leaders, Article ID #G00348875, available at https://www.gartner.com/document/3880121
- Other thoughts
  - Louis K. - Question in regards to what sits in each pace layer.   There are enterprise like systems that are implemented at higher level than enterprise systems (specific domains).
  - Jim mentioned this was done to thing about how do we educate roles, systems are, how to rethink what they are thinking.
  - Yolanda commented: *"What kind of education/training can be provided to them at a high level to understand this. As someone in a functional unit, I try to think of it as an institutional data repository but others don't. They don't want anyone to touch their data."*
  - Would like to continue to work and flush this information out and why we are doing what we are doing.

Additional Items

- Beth: Women in IT: Next week (Thursday) meeting for Women in EA (topic of unconscious bias) and look forward to everyone joining. It is on the Itana calendar
- Jacob: Meeting next Wednesday 2pm and doing WIKI refresh kickoff meeting. Active meeting and there is a pre-read and looking for others to join.   Take it from user centered design (kinda of things to solve and gaps to fill) and how we refactor the Itana wiki to make it as valuable resource to all.
- Dana: Business Architecture: Looking at starting back up on May 8th
- API: Monthly meeting. See calendar for details
- Approved for content at Educause meeting in October.   May do virtual Face2Face.   In Boston possibly
- Spring Call Program is up.   If you have other decks or information to share feel free to get input by other EA.   Reach out to Jim or Itana committee