

National BILT Meeting Minutes
“Data Management” KSA Vote and Discussion

| | | |
|---|---|--|
| MEETING DATE: Friday, March 8, 2024 | MEETING TIME: 12noon-1:30pm Eastern | MEETING PLACE: Zoom |
| RECORDER: Mark Dempsey | RECORDING: Available upon request | PREVIOUS MEETING: First NITIC BILT meeting (Previous ITSS BILT “data management” KSA vote meeting, May 2023) |

MEMBERS PRESENT

| | | |
|--|--|---------------------------------|
| BILT: | Educators: | |
| 12 employer SMEs were present | Larisa Greenband, Maricopa County CC | Matt Rayman, Maricopa County CC |
| | Denise Kruckenberg, Maricopa County CC | Julie Stiak, Maricopa County CC |
| NITIC staff: Ann Beheler, Mark Dempsey, Christina Titus, Stephanie Schuler, Larry McWherter, Leah Palmer, Rajiv Malkan | | |
| ITSS staff: Terryll Bailey, Pete Maritato, Helen Sullivan, Gordon Snyder | | |

| Agenda items | Discussion |
|--------------------------------|---|
| Introductions and KSA overview | <p>Ann thanked the SMEs for attending and assured them we would end the meeting on time. She asked everyone to put their name and company in the chat box, but noted that the ITSS grant will preserve everyone’s anonymity.</p> <p>Ann showed SMEs the ITSS (IT Skill Standards 2020 and Beyond) grant’s working definition (see below) of “data management and engineering,” which was developed and verified by 100+ thought leaders convened by the ITSS grant to identify critical IT job clusters. “Data management and engineering” was one of those in-demand clusters.</p> <p>Database management and engineering typically utilizes specialized software to store, organize and secure data in both relational databases and in other non-relational formats. It includes capacity planning, installation, configuration, database design, migration, performance monitoring, security, troubleshooting, as well as backup and data recovery. This definition was adapted from Wikipedia.</p> <p>An SME noted that database management and data management are not the same thing. Ann agreed and summarized that this meeting will cover the “IT side of keeping and managing the data so that can be used and analyzed.”</p> <p>Ann explained that the KSA (knowledge, skills, and abilities) voting process is a version of PCAL (Performance Criteria Analysis) as developed by the US Air Force to determine needed job skills. SMEs will be asked to look at each job skill and prioritize them for hiring an entry level worker. Ann noted community colleges can prepare many students for entry-level data management jobs. SMEs are asked to consider the importance of the KSA item, the proficiency level of the item, the frequency the item is performed, and the difficulty to learn to learn the item. For example, a item with a high proficiency level but not done often and is hard to learn might be something better suited for on-the-job training rather than as part of college curriculum.</p> |

| | |
|---------------------|--|
| | <p>Ann explained the 1-4 voting scale. Each KSA item will receive its own separate vote.</p> <ul style="list-style-type: none"> * “4” means absolutely must be in curriculum * “3” means should be in curriculum * “2” means it’s a “nice to have” * “1” essentially means delete <p>The average of the SME’s KSA votes will form the basis of the discussion. Ann reminded the group to respect differing opinions. Today’s SMEs come from a variety of background. Ann then briefly explained the meeting roles (industry SMEs, educators who actively listen, and the facilitator).</p> <p>Using a QR code, the SMEs voted electronically. Ann reminded them to cast their vote based on what they want from an entry-level employee. We don’t have specific job titles in mind. She noted that two-year community college educate students for a wide variety of IT jobs.</p> |
| Discussion overview | <p>An SME asked about seeing specific job roles when evaluating the KSAs. While this would be ideal, Ann noted that community colleges often create more of a “generalist” with a wide spectrum of skills and knowledge to provide the best opportunity for employment. As a result, the KSA list has some limitations, but it helps educators decide what KSAs should and should not be included in curriculum.</p> <p>Ann explained that Column G on the spreadsheet is the vote average. Averages of 2.6 or less (which are pink on the spreadsheet) is likely to not be included in curriculum, so the first thing to do after a vote is to look for pink KSA items to discuss with the SMEs. After that, Ann often asks to discuss those KSA items with wide vote spreads – an item with lots of 4s and also lots of 1s. Aside from this, SMEs can ask to discuss any KSA item on the list.</p> |
| Knowledge | <p>Ann reminded the SMEs that knowledge is conceptual information. When faculty teaches knowledge, “we teach about more items we expect the student to be able to apply.” The Ks provide important background context.</p> <p><u><i>K23 - Knowledge of recent streaming data frameworks and protocols AMQP, (e.g., Kafka, RabbitMQ).</i></u></p> <p>An SME who voted 3 suggested that having knowledge of this is fine, though it depends on what kind of database you’re running. Ann reminded the group that Ks are “conceptual” – students aren’t necessarily doing anything with it. Another SME argued that it’s not essential for the classroom because this knowledge you’d be learning in the field. “It’s not the first thing they’d need to know” but it could be considered foundational knowledge. Other SMEs agreed that no entry-level worker would need to know this. Further, none of the SMEs wanted to argue for keeping it in curriculum.</p> <p><u><i>K27 - Knowledge of query languages such as SQL (structured query language).</i></u></p> <p>One SME said that this is important, but noted earlier KSAs items (T21) mentioned Python and R. Why are they separated like this? Another SME noted that that Python is often used by generalists, while R is favored by statisticians and data scientists. It’s one or the other. But SQL is a tool that everyone knows how to use. Should all three be together? Ann explained that colleges decide on their own which tool or language to teach. Another SME understood why they were separated – SQL is foundational, while Python and R are more about programming and scripting and may not be “must have” for everyone. Another SME agreed it’s okay to keep them separated. You can be a database engineer without structured query language, but you can be one without R and Python. The SMEs seemed okay leaving it as is for now.</p> |

| | |
|-----------------|--|
| | <p><u><i>K32 - Knowledge of the characteristics of physical and virtual data storage media.</i></u> One SME asked how you define physical modeling versus the characteristics of physical data storage media. This SME explained that physical modeling is how data exists in memory. What are the actual data structure bits and bytes that composes it? How is it actually stored on the physical media? Another SME agreed and gave an example of a PCI complaint related to a credit card transaction. Everywhere a credit card transaction is stored, it encapsulates the physical environment, which means it has to hit a higher standard based on the rules of law. This SME tried minimize where the impact points especially when using cloud based storage systems because then your physical network becomes convoluted to say the least. AWS and Azure move your data around based on expediency, so if you commit a violation and a state is suing you because of PCI, it's a "nightmare" to pull in all of that data. When this SME was thinking of "physical modeling," he was thinking that employees better know where the data is when it comes to the law and where it's stored. Where is it stored and in what state? What's the state in storage? What's the state in motion? When employees work with cloud-based data sets, they often don't understand where the physical data is or even how to back it up. Other SMEs agreed. One is dealing with lots of new hires that have no concept of physical hardware and how or where data is physically. This SME thinks this is a big hole in curriculum.</p> <p><u><i>K48 - Knowledge of understanding data ownership and data stakeholders.</i></u> One SME noted that the phrase "data stewardship" did not appear on the KSA list. This K item could be enhanced with that addition. Other SMEs agreed. ACTION: add the phrase "data stewardship."</p> |
| Skills | <p>Ann explained that skills refers to the capacity to do something. What will these entry-lever workers be doing on the job? She noted the SMEs seem to be in close agreement with the votes.</p> <p><u><i>S5 - Skill in generating queries and reports.</i></u> There was interest in talking to the SME who voted a 1, but that person may have left the meeting. One SME suggested adding the word "data" which would tie back to the earlier SQL discussion. ACTION: reword to "Skill in generating data queries and reports."</p> |
| What's missing? | <p>Ann asked the SMEs what important items were missing from the list.</p> <p>One SME suggested a stronger awareness of the different types of databases (S15). There are other options besides relational databases.</p> <p>Another SME likewise suggested there are other tools beyond PowerBI and Tableau. Those are popular, but there are others. Maybe there should be an awareness – or even a skill – related to these tools. One SME noted that Excel is a data tool as well. Often you start modeling data in Excel before moving it to more powerful tools.</p> <p>Another SME suggested that "data privacy" is becoming a more specialized discipline area. This is Different than "data security." Students needs an awareness of "data privacy." While there is no nationwide law covering this, but some states do have privacy laws. What are the impact on consumers? PII covers some of this, but not all of it.</p> |

| | |
|--|---|
| | <p>AI is a big topic as well. Generative AI attracts media attention. Large language models are prevalent, but small language models are coming soon. Generative AI brings with it lot of ethical and liabilities concerns.</p> <p>One SME suggested that data management is separate from data engineering, which are both probably separate from data AI. It could be three different KSA lists. Ann recapped the talk about acknowledging that students need an awareness of AI’s role in data and what AI tools and capabilities are most common. She’s not sure a two-year degree can help students understand everything about AI, but it’s good for them to be aware.</p> <p>Data management connects to AI. For AI to work, you need “curated data” to avoid bias and hallucinations. Curated data doesn’t guarantee removal of those problems, but good curated data from trusted data sources can go a long way to solving AI problems. Other SMEs agreed. AI still has to query to get data. Entry-level workers need to know that you have to isolate the queries – you cannot give AI permission to access everything. One SME thinks of AI as a person, not a database. This person may be unique in that they have more knowledge and faster access speeds, but they still need to be governed. Another SME wasn’t sure about considering AI like another employee. To that SME, AI is like a copilot modeler and designer to help cross the divide between data management and data engineering.</p> <p>From the SME Zoom chat box:</p> <ul style="list-style-type: none"> * Data encryption as well as data classification levels (unclassified, Secret, TS, FOUO, CUI) Awareness of what data privacy is (laws and legalities) * Memory Leaks and detecting them and rectifying them quickly. This often comes from poor application calls or query handling. * Query pulls on transactional databases, which is related to governance and policies. |
| <p>Next steps</p> | <p>Ann explained that this is the second KSA vote meeting for data management. The next step is to synthesize the feedback from that meeting and today’s meeting. The SMEs will be invited to attend a short follow-up meeting to discuss that synthesis and also to help develop key performance indicators (KPIs).</p> <p>Ann thanked the SMEs for their time and expertise.</p> |
| <p>Next Meeting: Trends Discussion - Tuesday August 27, 2024 (9:30am-10:30am Central/10:30am-11:30am Eastern)</p> | |