The KIT

Knowledge & Information Technology

cébé
IT & Knowledge Management

No. 281 - 1 February 2021

In this Issue

Our Consulting Services

IT Strategy and Roadmap
Enterprise Architecture
Business Process Modeling
Enterprise Software Selection
IT Innovation Briefings
IT Due Diligence for M&A
Cloud Computing Adoption
Enterprise Security Maturity
Knowledge Management
Communities of Practice
Taxonomy Development
Enterprise Social Media Adoption

Digital Rights and the Cloud

The Minnesota chapter of the Cloud Security Alliance held an excellent webinar on January 28, entitled **"Digital Rights and the Cloud,"** with guest speaker Brittany Kaiser, an advocate for privacy rights and co-founder of the Own Your Data Foundation. You can watch the replay here.

Ms. Kaiser was one of the whistleblowers who revealed how the political consulting firm Cambridge Analytica had harvested the data of 87 million Facebook users, without their consent, before the 2016 U.S. elections. While you might expect from her an activist, anti-corporate stance, she is very pragmatic in her recommendations. She advocates a scheme where people's ownership of their data is recognized, protected, and the owners can make fully conscious decisions to sell specific data, earning a "digital dividend." For example, you could allow the city where you live to record your location in order to monitor traffic speeds, or allow researchers to remotely collect vital signs in order to participate in paid medicine trials. People may earn a couple dollars a day, which is negligible in our "first world" societies, but could make a difference for

people in other parts of the world.

She is also taking a nuanced view of the need for laws and regulations, wanting them to be narrowly focused on privacy violations that are deliberate or the result of grossly negligent.

The Q&A session inevitably (given recent events) touched on the right to anonymity on social media. Kaiser thinks that anonymity can be useful and acceptable, but protection should end when someone violates the law.

→ Edge Computing (or "Caution: Overused Buzzword Ahead")

The Edge Computing Technologies in Oil & Gas conference, held virtually on January 27, offered a chance to see if that industry is catching up with others in the adoption of Industrial Internet of Things (IIoT) technology. There were some hopeful signs: the speakers did not all stick to generalities, but mentioned several concrete examples; they discussed the accelerating effect of the pandemic on remote operations; and a security discussion focused on concrete measures rather than vague fears.

COVID-19 is making it harder to send personnel on site, resulting in an increase in the use of remotely operated equipment and virtual or augmented reality (AR/VR), a change that will persist post-COVID. Yet this is not about edge computing *per se* (placing compute resources near IoT devices, whatever the word "near" is means within a certain connectivity context), but more about IoT in general. In fact, the speakers and panelists seemed to frequently meander between digital transformation, IoT, and edge computing, which are related but distinct scopes.

There were a few specific use cases of edge computing proper, such as real-time detection of leaks in pipelines or sand intrusion in production wells (which requires the local processing of data from downhole acoustic sensors). But there were also some puzzling statements, such as "you can't have edge computing without cloud computing" (from a representative of an investment advisory firm).

Michael Lewis of Chevron demonstrated his clear understanding of IoT security challenges -- none of which, incidentally, are unique to the oil & gas industry. Regarding software vulnerabilities, he is tracking the NTIA initiative to create a "software bill of materials" (SBOM) standard. He advocated the adoption of a "zero trust" approach. in the long term, he said to watch for the threat posed to data encryption algorithms by quantum computing. And in the short term, he noted that the pandemic makes it harder to deliver physical authentication tokens (e.g., smart cards) to company personnel.

→ The Cost of Poor Software Quality

A trillion here, a trillion there, and soon you're talking about real money. No, this is not about a stimulus program during the pandemic -- it is about the cost of poor software quality. A research report for the Consortium for Information and Software Quality (CISQ), written by Herb Krasner, retired Professor of Software Engineering at the University of Texas at Austin, "concludes that poor software quality cost the U.S. upwards of $2.08 trillion dollars in 2020," adding up losses from operational software failures (75%) and poor quality legacy systems (25%). This represents close to 10% of the country's GDP, and an increase of 9% over the 2018 estimate of $1.91 trillion. In addition, unsuccessful projects are estimated to have cost $260 billion.

Bite-Sized Taxonomy Boot Camp

The annual London Taxonomy Boot Camp, an annual event produced by Information Today, is adapting to COVID, like all other conferences, by changing its format. Instead of just going virtual with a similar agenda to the last in-person event (October 2019) they are splitting it up into a series of "bite-sized" installments lasting just 3 hours, making them watchable at a manageable time from almost anywhere in the world. The first one is on March 2, from 2 to 5 pm UK time (GMT). There will be three talks, two of which relate to the use of taxonomy in healthcare; the third talk will be on how to handle vagueness in ontologies and taxonomies. Registration costs £79 (about $100 or €90).

Knowledge/Property Graph Drawing Made Simple

The emergence of knowledge graphs and property graphs as key form of knowledge representation and as sources of data for AI software has been fairly rapid. The concept is not new -- the rate of adoption is.

Three issues seem to have stood in the way:
- One is the "religious war" between those who believe that only RDF (Resource Description Framework) is the correct representation of a knowledge graph, and those, mostly centered around the company Neo4j, who prefer "property graphs," in which nodes can be labeled and edges can have attributes. We've heard the arguments from both sides, and the jury is still out. Both forms may have advantages depending on the use case.
- A second issue is that exploiting graphs requires new and rather non-intuitive query languages, such as SPARQL (for RDF graphs) or Cypher (for Neo4j property graphs).
- Finally, data scientists tend to manipulate very large graphs (from thousands to millions of nodes and edges), therefore visual display and editing tools have not been their priority. But what about the rest of us, who would like to create and draw much smaller graphs, including for simulation, experimentation or education?

On that last point, here comes some help in the form of arrows.app, a cloud-based graph editor with a very intuitive interface, created by Alistair Jones of Neo4J and Irfan Karaca of Kale Yazılım. Graphs created with this application can be exported as images or to the other Neo4j tools.

Seen Recently...

"The open.engineer site […] managed to automate the previously incredibly labor-intensive process of looking like stupid a**holes."

-- Corey Quinn, cloud analyst and commentator, talking about a site that scrapes data from LinkedIn as well as (apparently) GitHub in order to provide a very primitive rating of software engineers

"According to [Jakob] Nielsen, 1% of content generators in social media contribute 90% of publications, 9% contribute 10%, and 90% of users produce no content at all."

-- Carlos Viniegra Beltrán, Mexican economist, consultant, and former government official, comparing various forms of inequality in a journal article on income distribution. Nielsen's work is from 2006.