

## Forschungskolloquium Neuromorphe Elektronik

---

**Donnerstag 07. April 2022 -- 16:30Uhr**

---

### **CMP KNOWLEDGE FOUNDRY AND BIG DATA ANALYTICS**

**Ara Philipossian (1), Negar Monfared (1), Yasa Sampurno (1), and Jason J. Keleher (2)**

*(1) Araca, Inc. Tucson, AZ USA; (2) Lewis University, Romeoville, IL USA*

Data collection in a CMP process (or any other complex manufacturing process for that matter) is relatively simple. However, several problems surface when massive amounts data get collected day in and day out over a period of several years at multiple research sites throughout the world. First and foremost, one needs to solve the issue surrounding data storage in a centralized data bank that is retrievable based on certain key search criteria by all parties (i.e., research sites and scientists working therein). The other problem has to do with how to best analyze 100s of millions of simple and complex data that can add to the planarization knowledge base. Finally, process modeling and visualization capabilities need to be implemented that provide insight. The simple expression below sums up the hierarchy and importance of it all.

INSIGHT>>KNOWLEDGE>DATA

As such, any activity that does not use raw data to help build knowledge, and in turn, provide deep insight into a complex problem is an exercise in futility. To help address this fact, we at Araca, Inc. have started creating an experimental network of Copy EXACTLY! tools and methodologies with several exclusive long-term strategic partners with whom we are beginning to create a “Global CMP Knowledge Foundry”. In the meantime, we are looking to replicate the model with one or two research institutes in Europe who are working in the area of CMP. We well recognize that >99 percent of all data generated by our polishers and cleaners is “DARK MATTER”! As such, our Araca Insights® product is helping our strategic partners organize their data gathering processes, bring more data to light, put robust systems in place for centralized R&D data banking, management, analyses, modeling, and sharing.

On the data storage side of things, our software provides:

- ✓ Highly-organized central data center containing all of the data gathered from various tools owned by us, and our strategic partners in an easy-to-access format with a descriptive file nomenclature,
- ✓ Highly secure storage service with double authentication access for users,
- ✓ Automatic storage of newly collected data in our cloud servers,
- ✓ Automatic QC of newly collected data,
- ✓ Automatic notification to ALL users when data is uploaded to servers,
- ✓ Automatic parallel storage and backups for each single set of data,
- ✓ Parallel uploads from different centers, and,
- ✓ Elastic and scalable expansion of centers as the amount of metrology tools or knowledge foundry centers increase.

On the ANALYTICS side of things, our software provides:

- ✓ Recognition of data types, data cleaning, anomaly detection and QC of data,
- ✓ Automatic aggregation of required properties to be easily accessed,
- ✓ Ability to filter test properties to reach tests done in the same category, and,
- ✓ Interactive live-selection and visualizing of required property relations within desired ranges.

On the GUI side of things, we provide multi-variate filtering, relational plots, correlation analyses, ANOVA analyses and significance factor determinations, data animations, and many more features that are best demonstrated rather than described.

The BETA version will be available in April 2022. AI, ML and DIGITAL TWIN features are currently considered AS longer-term activities (2023 and 2024) depending on general sentiment, customer feedback, and available budget.

I trust you will all enjoy the presentation.

---

Meeting link: <https://tu-ilmenau.webex.com/tu-ilmenau-en/j.php?MTID=mce05275ba0921ab9731603824920eb42>

---