de:hub
KARLSRUHE
Artificial Intelligence

HUB
FRANCE
IA

Summary

Keynote I: AI in the Aisle: Re-imagining Grocery Shopping by ERAN KRAVITZ CTO & Co-Founder at Shopic

A few years ago, grocery shopping took a new turn thanks to the introduction of stores without a checkout concept. It was introduced mostly by Amazon, but other companies have been working on this concept for a while.

Shopic, a company founded in 2015, works with top tiers retailers to digitize supermarkets. This company of 40 people based in Tel Aviv is specialized in AI and computer vision. Thanks to the device they developed, a clip-on on carts with cameras distributed all over the store looking in different directions, they can determine what goes in and out of each cart smoothly.

Thanks to the data collected by the cameras, they put the effort into making this experience as frictionless for customers as possible. Shoppers seem to enjoy this shopping experience as the company observed a very high rate of return to the stores after a first experience.

Session 1 – Data spaces in AI

Speaker 1 : Anne-Sophie Taillandier from IMT & Data Hub

Data Hub is a nonprofit data and AI platform to experiment with architecture and algorithms on real data. To do so, DataHub developed a business model to cover the cost of the platform and offer a trusting, neutral, and sovereign environment in which they act as a data playground box.

The main question in each project developed thanks to the DataHub platform, is about interoperability of data. The second main question is about trust and sovereignty.

On the topic of interoperability, the challenge resides in interoperability between the different cloud service providers to avoid vendor locking. To create dataspaces based on European values, we need trust and sovereignty. For data to be interoperable, it's required to know where the data is, how it is processed, who can access the data, and for which use case.

To solve these problems, DataHub put user requirements as the core of the development of Gaia-X. To be able to connect the data, they built an architecture based on a common underlying framework for all domains. This should allow to scale to interoperability and help to collaborate to get the emergence of new business models and digital infrastructures.

Speaker 2 : Sébastien Picardat from Agdatahub

Agdatahub's database is a European platform made for data exchange and dedicated to the farming industry. This database is key for the agri-food industry as Data and AI are the two main pillars to develop digital services for farmers.

The main issue with data in agriculture in Europe is that farmers are surrounded by partners' data from the 10 million farms in Europe but without interconnection between the different sources of data. Data are also collected thanks to the Internet of Things through various machines and information systems that issue the data. Although, without interoperability, most farmers can't analyze nor make it available for the AI algorithm.

To fix this issue, AgDataHub works at the European level with DG Connect and working groups to develop interoperability with some partners' platforms. Together, they aim to develop standards of interoperability of data and implement technical standards about AI while acknowledging farmers' concern about consent in data-sharing and respecting it.

Speaker 3: Sascha Rank from Karlsruhe Institute of Technology

What is federated machine learning? It is a way to train models across private data. Each client trains its model with its local data. Then, this model is centralized with the other models trained by the other clients. Finally, the serve aggregates all the updated models and generates the new updated model. Federate learning allows to create a machine learning model using the data of each client, but the local data never leaves the clients, only the models are exchanged.

Federated learning could be a simple solution to share data between companies to increase the available amount and overcome the common issue of lack of data. The solution could be to share the machine learning model and parameters between companies instead of sharing the data.

However, federated learning also comes with limitations. First, it has a relatively low failure safety as one central server is responsible for aggregating all the models. If it fails, the entire system stops working. Second, it has low transparency. The clients can't check how the models are aggregated. Finally, there could be legal implications as it's unclear who the owner of the final updated model is.

Speaker 4: Dr. Emmanuel Bacry from the Health Data Hub

Health databases are full of information. Although, they are still largely underexploited, and this is mainly because of three reasons. The governance of each base (hospital, cohorts, Assurance Maladie) is intricated, data is scattered among multiple databases and finally, they lack interoperability. These limitations mean that it is difficult to identify that you have the same patient between the different databases.

Health Data Hub created a data catalogue for researchers from public institutions or private companies to be able to operate public interest research. Thanks to the Assurance Maladie, the French reimbursement database, information about the healthcare of everybody in France is available. Of course, due to the sensible data, access to this platform must be controlled and authorization must be granted.

The creation of this Health data space is game-changing for research: it will include about 20 databases to get contextual information about the patients and allow to offer patients the best long-term treatments, save time, and reduce delays in diagnosis.

Keynote II: How reinforcement learning can be applied to IoT by Cameron Schuler CCO and VP, Industry Innovation at Vector Institute (Canada)

Learning is valuable to help people go where they want to go, even if they don't know the way. Ultimately, having adaptability in systems is incredibly important. We need systems that learn from the experience and adapt to the environment. Imagine a computer that will cause an action (virtual or physical) in its environment. As it learns on the way through it needs the ability to go back and have that continuous feedback loop.

Reinforcement learning came with Udacity's discovery. Sebastian Thrun created Udacity, a massive online learning company and they found that some salespeople were twice as productive as other ones. Rather than firing their unproductive salespeople, they collected all the data to understand what highly efficient salespeople do. This information becomes training data to help less productive salespeople become more productive.

Another interesting example is Amazon. In the UK, people are buying particular chemicals. Amazon recommended to people who bought this chemical, also to buy those chemicals. And those are the foundations for homemade bomb-making supplies. So that brings us back to when you create models you need to be quite thoughtful about your outcome. They can have unintended consequences.

Reinforcement learning is based on reward systems. Computer maximization is rewarded, when you think about how we learn as humans, we get penalized for bad behavior, and we get rewarded for good behavior. It just really is discovering and trying things to understand what the best result is. Reinforcement learning allows the system to learn from the inputs and outputs to be able to be more effective. It is very ideally suited for control problems. Thinking about IoT, lots of those don't control problems. This is not to say that reinforcement learning is easy to use, but there's a lot of promise in that area.

Thanks to simulation we also can see incredible outcomes. For instance, a company said that for some of their robotics, they're 99% real-time in terms of things they do in their virtual environment, what they get, deploy them and get that feedback and the accuracy level. Some pretty incredible changes will appear over time, based on things like digital twins.

Session 2 – Cybersecurity for AI and with AI

Speaker 1: Ingmar Baumgart from FZI Research Center for Information Technology

Today, the Internet of things is everywhere around us. It can be through sensors using microphones and/or cameras or through embedded systems. Sensors can sense very private information and need security systems to protect the privacy of people that are interacting with these systems. Embedded systems connected via the Internet increase the attack surface enormously. Due to the Internet, attacks can appear from everywhere around the world.

These attacks are very scalable as they do not need many resources to attack a large number of devices simultaneously as soon as you detect a security vulnerability in one of these devices. In addition, there is a very low budget for implementing IoT security measures. Especially for consumer products as they have to be affordable and consumers don’t want to pay extra for security and privacy. Moreover, it is very expensive to keep these products safe as you have to provide security updates during the whole lifetime of the product.

How can AI be a saviour for IoT Security? AI is useful for intrusion detection, malware detection, and the area of vulnerability discovery. Thanks to machine learning analyzing the logs and executables in

the object, AI can be a very promising approach to go toward more security and privacy in IoT. By automatizing these processes, it's possible to reduce the costs for testing these devices' security issues.

Speaker 2: Victor Vuillard CTO at Parrot

Due to the increased connectivity and use of drones, cybersecurity is becoming a primary matter in creating trusted drones. Nowadays, drones can be flown with a 4G connection which increases the attack surface as it is no longer controlled through a direct private radio link between the remote controller and the drone. The drone will benefit from AI only if it can be trusted.

To have better cybersecurity and privacy, Parrot implemented a full data governance and allows users to keep control over data. The user can stop sharing but also decide to delete every data shared previously with Parrot. Alongside, encryption and authentication were implemented.

Finally, when it comes to cybersecurity in IoT, transparency is a key element. An open-source policy lets people see how secure the product is. Going through audits and publishing the results also allows building trust in your products.

Speaker 3: Dr. Carmen Kempka from WIBU Systems

The topic is how to protect AI and especially machine learning against attackers. An attacker could try to manipulate the training data, tamper with the training algorithm to get a falsely trained model or tamper with the model afterwards to achieve fraud or steal the model.

To protect machine learning in AI, various tools are available. First, it is possible to protect the program used for machine learning through tools like AX protector which protects python code. This will protect the program from reverse engineering. It's also possible to put licenses on the software or to transfer some part of sensitive code into secure hardware.

What's important when it comes to protecting the machine learning lifecycle is focusing on data, training process and data collection.

Session 3 - Edge AI

Speaker 1 : Dr. Julius Pfrommer from Fraunhofer IOSB

Edge AI is deployed in various areas from autonomous cars to industrial productions. It requires three broad categories. First, runtime environments where the Edge AI nodes can live. Secondly, there is also a need for specialized hardware, especially AI accelerators, FPGAs and GPUs. Finally, management platforms must be able to manage the fleet of devices version management, model updates, etc.

To function properly, Edge AI also requires cloud solutions and data. The cloud solution must be available for the long term and be under control as the field is developing really fast. Data has to be available in real-time so minimum latency is necessary along with it being available in a single location.

To develop Edge AI, the goal is to develop a systematic engineering practice for developing AI-based applications. The overall goal of AI system engineering is to enable the use of AI systematically to develop tools, methods, and processes for developing these types of solutions that also help us to predict the quality of an AI-based solution before it is deployed.

Speaker 2 : Christian Verbrugge from GrAI Matter Labs

The goal of GrAI Matter Labs is to provide brain-inspired technology in objects to enable machines to be naturally more efficient and help people in their everyday job. To do so, Edge AI is the key in order to be more accurate because the Data will be collected directly on the field.

Edge AI is crucial to bring Industry back to Europe. It's a key element to gather data and find easy to use solutions. GrAI Matters Labs' technology is based on fast processing thanks to parallelization in the core and high-speed memory to gather data and compare data to previously collected ones. This allows their chip to have 20 times better efficiency compared to Google Edge CPU.

To go toward better AI, Edge AI is key even if it's difficult to train and use AI in the real world compared to a data-centric approach. That's why we need to change the way young engineers or students are looking at AI.

Speakers 3: Dr. Robert Pesch & Fabian Meyer from inovex GmbH

Small and medium-sized companies must keep up with digitalization to remain competitive. By using edge components and natural language processing in the end, a solution was created that aims to address the entire service life cycle from corrective identification of incidents wire and predictive maintenance. All thanks to data collected at every step in the industrial IoT.

However, the major issue is to collect data to develop machine learning models. Most of the time, companies don't want to share their data because they're sensible. Experts must explain why it's crucial to share it, to make more benefits.

Speaker 4: Joel Rubino from STMicroelectronics

Nowadays, we are building smart technology, but the intelligence is in the cloud, not in the devices. There are several issues with having intelligence in the cloud instead of in the devices. That's why it's needed to make the products smart. Indeed, clouds can lead to privacy issues and it feeds data to the big companies owning the cloud servers.

ST Microelectronics provide a solution running on computers to create a machine learning library that will allow predictions, classification etc.