Crazy Data Train

Data is the strategic asset – and it needs to be managed as such, by deeply understanding its sources and destinations, and mastering all ways of the enterprise to activate at the very heart of the business

Enterprises jump on the bandwagon of data as the key to better engage their customers, to intelligently automate their core business processes and to design and manufacture smart products. For that, data needs to be activated at the center of every crucial business activity. It's crucial to understand how to get it; not only from internal, but also from external sources. It requires a sharp 'procurement' market eye to get the right data. It needs an R&D-like vision to design how it will produce value. And it taps into the mindset of marketing to envision how to monetize data, in- and outside the enterprise. If data can be put on the corporate balance sheet, it will trigger even the most executive perspectives. Above all, it's a matter of deploying organizational data enablers and behaviors - leaving nobody behind for the ride. All aboard!

in **Anne-Laure Thieullent**
Expert in Residence

WHAT

- Any organization looking to activate data in their business first needs to create an inventory of data assets across the company, as well as data assets leveraged from outside.
- Data from external marketplaces, including brokers of open data and industry consortia – combined with own data – can create even more valuable insights.
- Suppliers of (labelled) training data provides crucial input for machine learning and analytical purposes, especially when the organization does not hold such data itself.
- New ways of accessing training data outside the organization – such as Federated Learning and Data Collaboration - open up previously unexplored data sources.
- Simulated environments and reinforcement learning can provide crucial (synthetic) training data in cases where real-life data is insufficient or unavailable.
- Data becomes a corporate asset when it provides tangible economic benefits and adds to the corporate purpose; its economic value should be on the corporate balance sheet..

USE

- The MELLODDY consortium aims to enhance analytics on the data of 10 pharmaceutical companies, without exposing proprietary information.
- Skywise is the first-of-its-kind open data platform, developed by Airbus for collaboration with the aviation industry ecosystem.
- Continental developed an AI-based virtual simulation program, which generates 5,000 miles of vehicle test data per hour, which otherwise would have taken over 20 days.
- 890 by Capgemini offers public, private and community marketplaces for curated datasets, providing companies 'as-a-service' access to data from external and internal sources.
- The EC Support Center for Data Sharing enables member states to share data securely with other other governments and organizations, whilst respecting intellectual property and privacy.

IMPACT

- Activating data at the heart of the business to serve key corporate objectives and purposes.
- Improving effectiveness and value-creation of existing business intelligence and analytics by adding external data.
- Shortened time to value for new insights, analytics or AI models by leveraging more diversely sourced training data as well as pre-trained solutions.
- Monetization of own data and augmented data with partners, suppliers, clients, potentially even creating new business models.

TECH

- Federated learning tools: TensorFlow Federated (TFF), Xaynet, Owkin, OpenMined,
- Data exploration: Informatica, Alation Data Catalog, Apache Atlas, Waterline, Microsoft, Collibra
- Data creation: Mighty.ai, Appen, Scale AI labeled, Lionbridge, Mostly.ai, Sogeti ADA
- Data monetization: Dell Technologies, Intersec, Inzata, Sisense, Oracle

read the full report here »