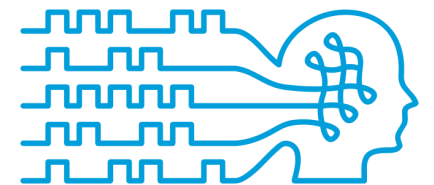


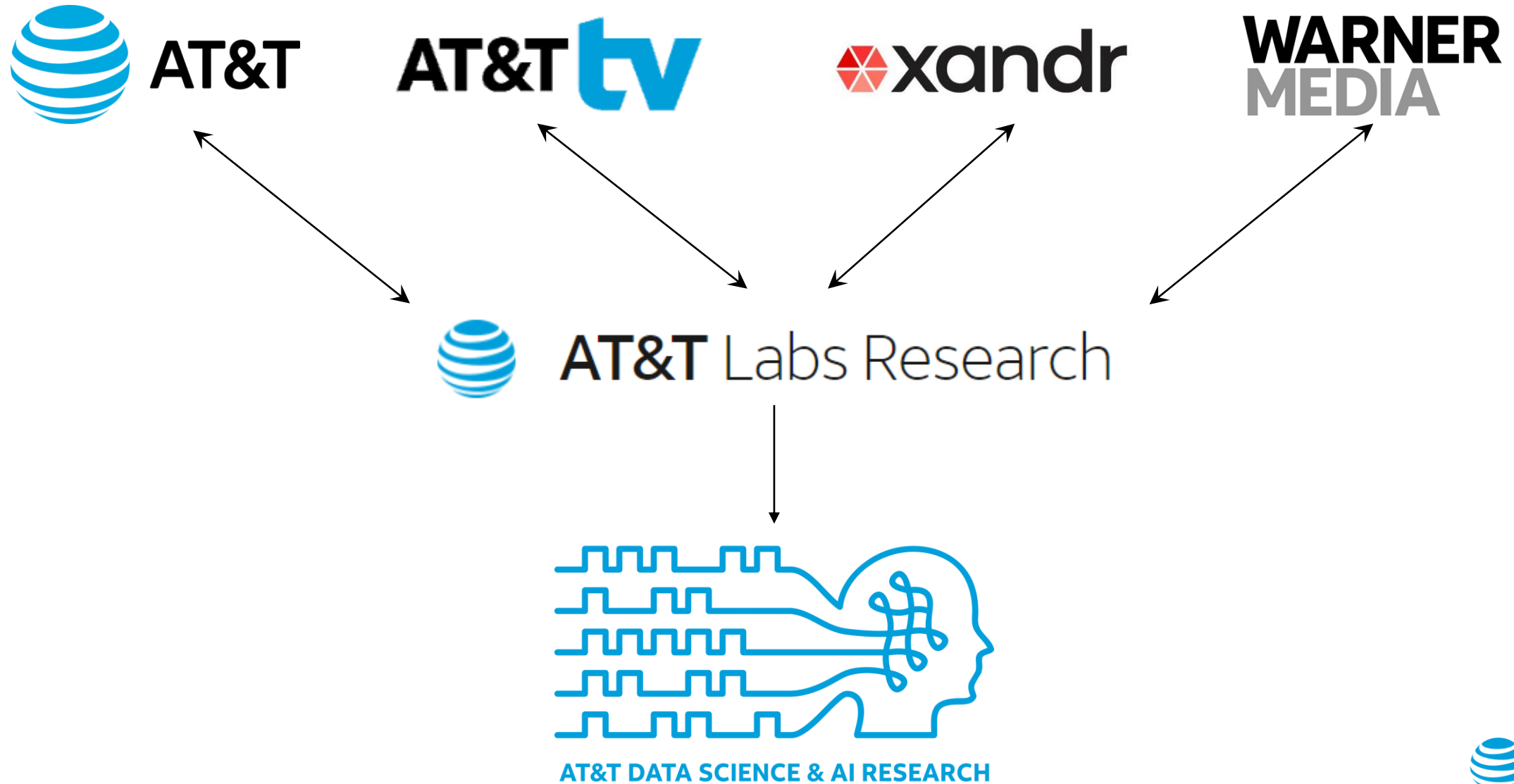


Data Science in AT&T: a Labs perspective

Subho Majumdar
Data Science and AI Research, AT&T Labs
Wednesday, February 19, 2020



Who we are

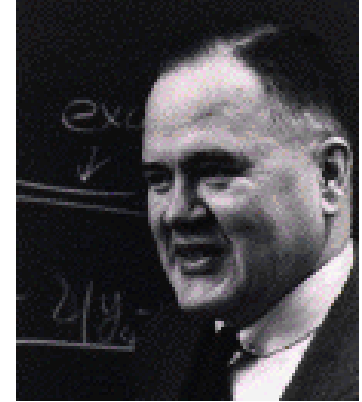


History

- Erstwhile AT&T Bell Labs.
- Long history in Statistics: Shewhart's control charts (1924) and Dodge and Romig's double-sampling inspection (1930's), Wilk, Gnanadesikan, Mallows.
- Current:
Birthplace of the R language (originally known as S).
Our team won the Netflix prize!



Walter A. Shewhart



John Tukey



Chris Volinsky

Current work

- Customer experience modeling
- Machine Learning-based media planning
- AI-powered video metadata and curation
- AKG: Automated Keyword Generation for metadata enhancement
- Data quality and anomaly detection in data streams
- Real-time visualization of large spatiotemporal datasets using Nanocubes[®]
- Machine Learning for 5G
- Fairness, transparency and data privacy research
- Open source involvements: R, RCloud, Nanocubes, Acumos

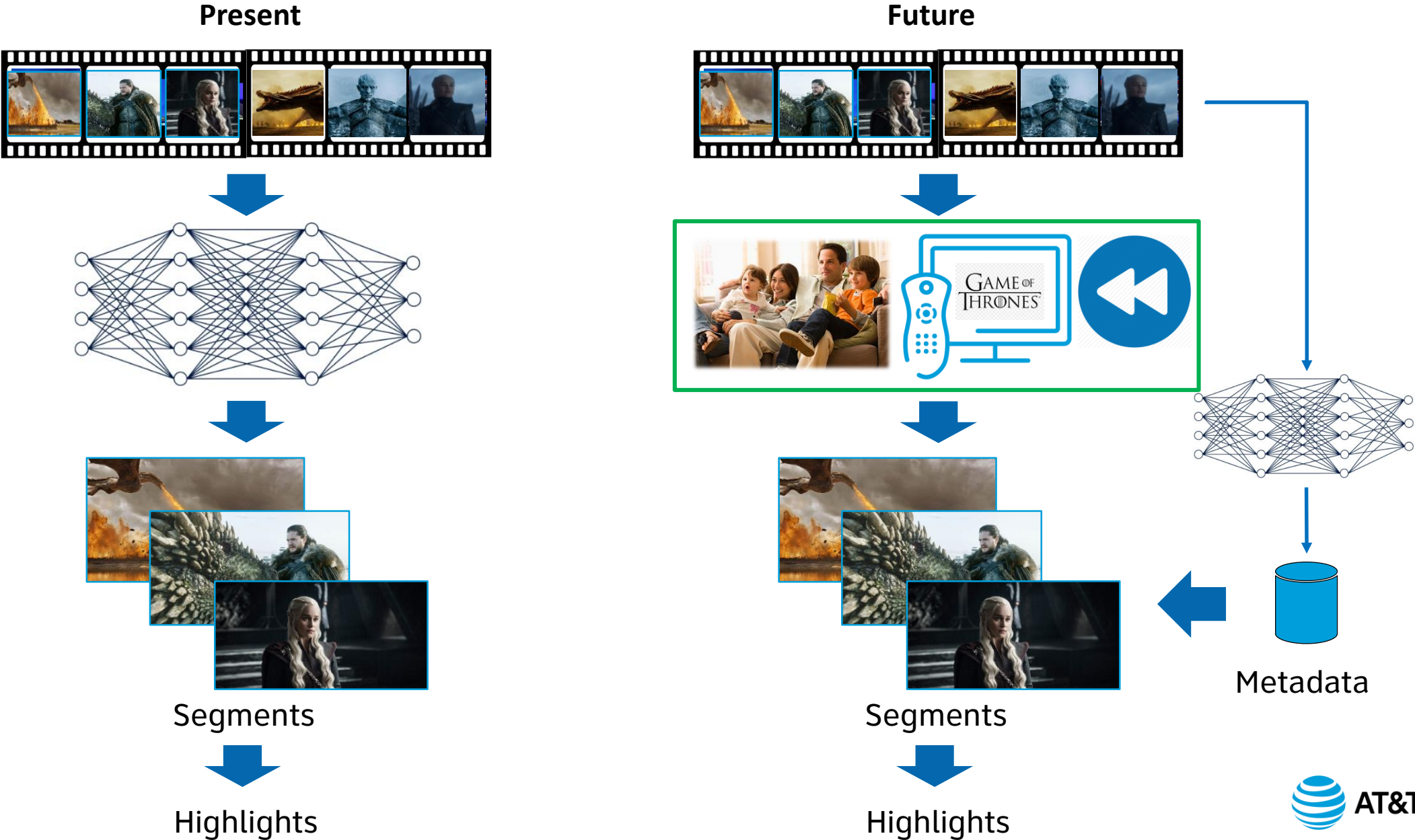


Simon Urbanek

The good, the bad and the ugly

- Data, data and more data!
- Iterative process of cleaning, feature generation and modelling
- Any modelling has to be done in a scalable way
- With smart feature engineering, simple models often perform as good as complicated models
- Not only the methods are important, but how the outcomes are presented to stakeholders

Example: Viewership-centric Content curation



Thanks!

