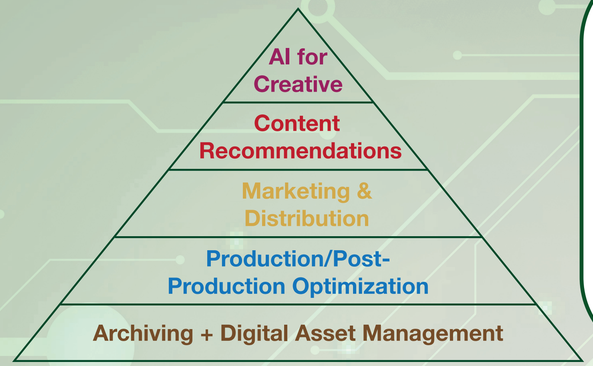
AI and Machine Learning

Explanation of wallchart purpose and what it illustrates

After 50+ years of false starts and marketing hype, artificial intelligence is starting to get deployed throughout the media industry. But as it rises as a real tool, it is cursed with a triple disease: (1) it's tremendously important to our future (2), it's a conceptual nightmare, and (3) it's a great story, hence it easily lends itself to fantasy. This wallchart attempts to clear out some misconceptions and create clear guidelines for what it is, what it means, and where to apply it first.

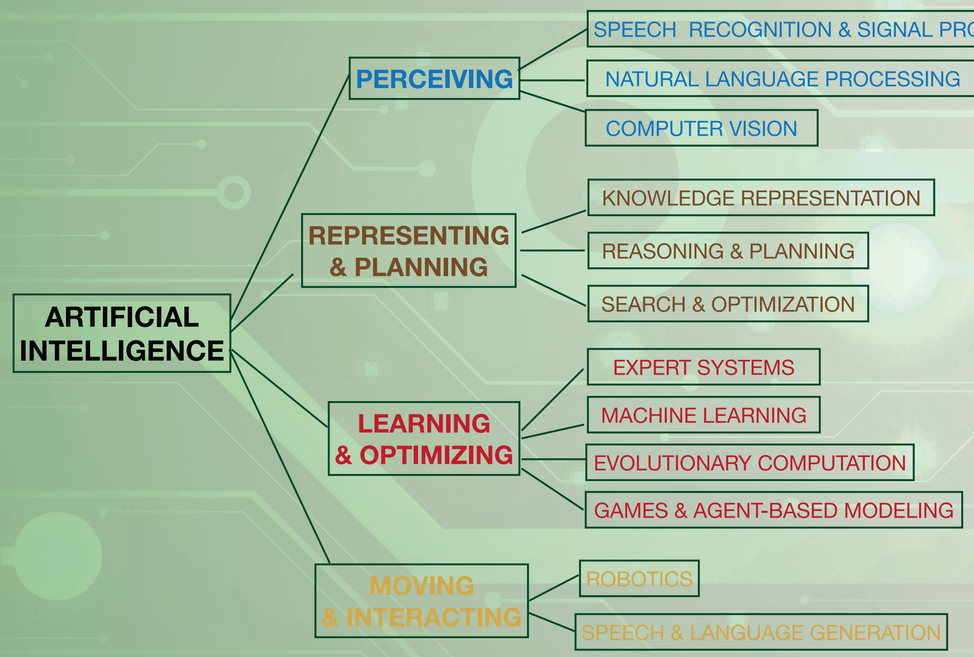
M&E'S AI Hierarchy of Needs

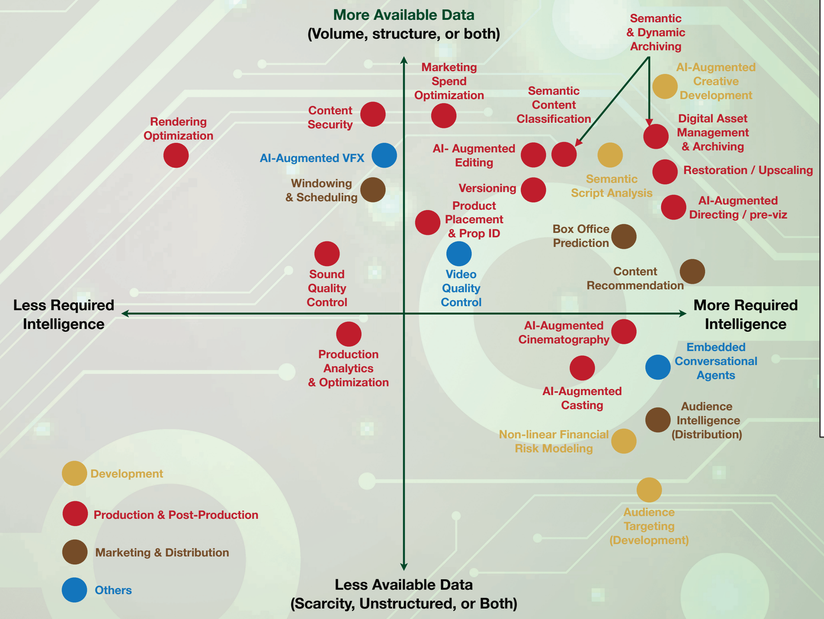


Hierarchy Explanation

The media industry is a technology industry. It is a computational industry that ingests, processes, and produces data. Media is one of the most data-rich industries by way of its products (pixels) and how much its customers talk about them (conversations). As such, its needs for AI are as diverse as they are critical. At the most fundamental level, the industry needs to break down its unstructured products (text, video, audio) into structured symbolic features (colors, character arcs, scenes, shot types, narrative structures, objects) for the purposes of archiving and digital asset management. Once this is done, these features can be fed throughout production and post-production processes. Text is another prized yet highly unstructured dataset: audiences are talking about what they like and want through billions of threads. Once again, extracting symbolic features (topics, ideas, sentiment, affinities) from this highly unstructured dataset (words) informs marketing and distribution strategies. Finally, understanding product features in the context of audience features – and vice-versa- is the Holy Grail of content recommendations, which increasingly drive creative decisions.

The Media Industry's "Hierarchy of Needs" in AI/ML





Artificial Intelligence

Artificial Intelligence is a domain of science at the intersection of many disciplines, including computer science, mathematics, philosophy, psychology, neuroscience, electrical engineering, linguistics, and information theory. Too often machine learning is painted as AI, but intelligence is broader and more complex than statistical learning. The latter is also too dependent on large, hand-labeled training sets, and ungodly amounts of computation, to qualify as intelligence. Of AI's 4 major components, none is perhaps more essential, difficult, or misunderstood than knowledge representation and reasoning. Because to qualify as AI, an application needs agency. A classifier trained to recognize dogs from cats from a labeled training set is a large-scale computational puppet. A self-driving car capable of making decisions in the real world is an agent that reasons and understand causality. It's no surprise, then, that agency is quoted in most of the academic definitions of AI, none better than Shane Legg's "optimized behavior of agents in unknown computable environments." But to have agency, an AI application must "know" the world we live in and be able to reason about it (specifically in causal terms). Reality as we perceive it is a Russian doll-like structure of lower-level symbols clustered together into higher-level symbols. Sub-atomic particles cluster into atoms, which cluster into cells, which cluster into organs, which cluster into organisms, etc. Being able to understand this structure, being able to abstract each cluster of symbols into another, more abstract, higher-order symbol, and reason about how they all relate to one another, requires intelligence. This is what knowledge representation does: building machine-based logic to abstract low-level data (numbers, a set of pixels, a set of words) into symbols (objects, emotions, concepts, shots, scenes, stories) that can then be articulated to form knowledge. This is why knowledge representation is the essence of AI.

Hard Quadrant

AI goals spread across a vast landscape of data availability, computational requirements, and model complexity. And for those eager to get on the road to building intelligent computational solutions in the media industry, nothing is more important than identifying low hanging fruits and showing early - and preferably cheap- "wins." This "hard quadrant" attempts to identify them: areas where there's substantial data available and making sense of it doesn't require too much computation or experimental modeling. Rendering, for example, produces a vast amount of log data: it's possible to see how fairly simple clustering models could be built to predict which rendering jobs are bound to fail before they do, and thus optimize cycle time and compute costs. On the other end of the spectrum, audience targeting at the development stage, or even baseline embedded computational agents (virtual characters), require levels of intelligence and training datasets that are much more difficult to muster. Large language models like GPT-3, for example, require petabytes of training data and millions of dollars of compute to produce extremely brittle conversations, far below the minimal standards of a virtual character.

Special thanks to Yves Bergquist for his contribution to this wallchart. Bergquist is program director in the AI & Neuroscience in Media Dept. at the Entertainment Technology Center @ University of Sourthern California.