



ANALYTICS CONFERENCE:
TECHNOLOGY, NEW DATA STREAMS, AND MARKETING STRATEGY
PHILADELPHIA, PA, MAY 4 – 5, 2023
HOSTED AT WHARTON AI & ANALYTICS FOR BUSINESS

MSI 2023 Analytics Conference: Technology, New Data Streams, & Marketing Strategy - Session Summaries

Contents:

Day 1 Presentations:

- [Livestream Shopping and Dynamic Customer Interactions](#)
- [Applying “Explainable” AI: Using Theory to Understand AI Emotion Models](#)
- [MSI’s Marketing Mix Modeling Initiative Progress](#)
- [“Responsible AI” and Data Ethics](#)
- [Customer Interest Graph for Dynamic RealTime Hyper-Personalization](#)

Day 2 Presentations:

- [Using Simulated Data to Validate Marketing Mix Models](#)
- [Mega or Micro? Influencer Selection Using Follower Elasticity](#)
- [AI at the Front End of Innovation](#)
- [InnoVAE: Generative AI for Patents, Innovations, and Firms](#)
- [How Behavioral Science Can Help Data Analytics](#)

MSI Research Priorities 2022-2024

Day 1 – May 4, 2023 | 10:30 am – 7:00 pm EDT

Livestream Shopping and Dynamic Customer Interactions

Speaker:

Xiao Liu – *Assistant Professor of Marketing, New York University, Stern School of Business*

Overview:

In this session, Xiao Liu (New York University) takes a deep dive into the practice of live-stream shopping. In the opening of her presentation, Liu described Livestream shopping as the new format of e-commerce, connecting brands and influencers with consumers to sell

products and services in real-time. While giving a background on the livestream shopping industry, she pointed to Taobao Live, the largest livestream platform, which boasts over 500 billion sales transactions, 2 million influencers and sells products/services from a variety of industries. In addition, she mentioned American livestream shopping platforms, including Amazon Live, Facebook Live, TikTok, etc. Liu noted that consumers can be engaged in live-stream shopping by leveraging coupons. In her presentation, Liu examined a large-scale field experiment conducted in partnership with the live-stream platform Taobao. Lui indicated at the time of the study coupons were pushed out randomly to engage consumers. The high frequency of delivering these coupons created a "dynamic targeting problem." In the future, the platform wanted to understand who they were reaching, how much the discount should be and the optimal sequence to deliver these coupons to consumers. The study aimed to develop a theoretical framework that incorporated the intertemporal tradeoffs in dynamic personalized pricing, along with designing a policy maximizing revenue. Additionally, they wanted to empirically evaluate the performance and understand the gains and the mechanism that can explain the gains. A personalized pricing strategy (dynamic coupon targeting strategy) was leveraged using an artificial intelligence algorithm and deep reinforcement learning, outperforming all study benchmarks.

Takeaways:

Industry Background on Livestream Shopping

- Taobao Live, is the largest livestream platform, boasting over 500 billion sales transactions, 2 million influencers and sells products/services from a variety of industries.
 - American livestream platforms include TikTok, Amazon Live, Wayfair, Facebook Live and Shoploop (Google).
- Coupons are often leveraged in live-stream as a powerful tool to engage consumers and promote sales.
 - These coupons are pushed via the device screen creating many opportunities for platforms and influencers to engage with consumers.

The Experiment

- **In conjunction with the livestream shopping platform Taobao, a large-scale field experiment was conducted** to address the high frequency and random nature of delivering coupons to consumers, creating a "dynamic targeting problem."
 - The platform wanted to understand who they were reaching, how much the discount should be, and the optimal sequence to deliver these coupons to consumers.
- The study aimed to **develop a theoretical framework that incorporated the intertemporal tradeoffs in dynamic personalized pricing, along with designing a policy maximizing revenue. In addition, they wanted to empirically evaluate the performance** and understand the gains and the mechanism that can explain the gains.

- **The study presented a personalized pricing strategy** (dynamic coupon targeting strategy) using an **artificial intelligence algorithm and deep reinforcement learning**.

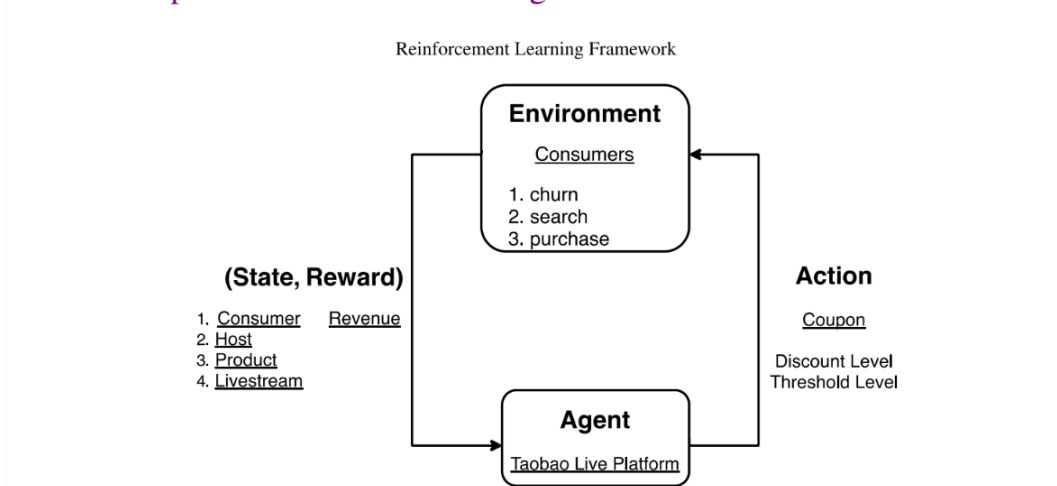
Research Questions

- How can we develop a theoretical framework that incorporates the intertemporal tradeoffs in dynamic personalized pricing and designs a policy that maximizes revenue?
- How can we empirically evaluate the performance?
- What are the gains and what mechanism can explain the gains?

Takeaways:

- Batch Deep Reinforcement Learning (BDRL) increases GMV by 63% in a **field experiment** compared to the status quo (random allocation)
- BDRL is 34% more effective than static targeting policies, because it incorporates intertemporal tradeoffs, especially the **reference price** effect
- BDRL is 20% more effective than an econometrics model, because it mitigates **model bias**
- Recommend a small discount for more attractive hosts and **price skimming** over time

Model Setup: Reinforcement Learning Framework



- **The results from the study demonstrated that the proposed dynamic personalized pricing strategy outperforms all the benchmarks.**
 - **Policy 1: When to Target?**
 - The dynamic policy gradually increases the high discount level coupons to avoid the reference price effect.
 - **Policy 2: Who to Target?**
 - Better deals for less attractive hosts.
 - **Policy 3: When + Who to Target?**
 - Low spenders: More price sensitive, bigger discounts and faster increase.

- High spenders: Less price sensitive smaller discounts and slower increase.
- This **generalizable framework** can be applied to dynamic, high-frequency interventions and customer lifetime value optimization.

Applying “Explainable” AI: Using Theory to Understand AI Emotion Models

Speakers:

Hortense Fong – *Assistant Professor of Marketing, Columbia University, Columbia Business School*

Overview:

Speaker Hortense Fong (Columbia University) discussed how to create AI models that are more "explainable." Specifically, Fong explored how to use theory from the Natural Sciences, Social Sciences or managerial expertise to design AI models that are more predictively accurate but also achieve explainability. Using GPT-4 as an example, Fong demonstrated that AI has been successful in the completion of exams (e.g. Uniform Bar Exam, LSAT, SAT, GRE, etc.) and in creative areas (e.g. photography, text, audio), but she acknowledged that this improved performance of AI comes at the cost of explainability. In her discussion, Fong indicated that explainability increases generalizability, leading to a model that is more useable and transferable to other subjects. Fong pointed to the study she conducted with Vineet Kumar and K. Sudhir, [A Theory-Based Explainable Deep Learning Architecture for Music Emotion](#), which examined the link of music to emotion, a key driver of emotion in the consumer journey. She provided examples that demonstrated the use of music and how the same commercial with different music could inspire different emotions (e.g., exuberance, sadness). She then discussed the application of a deep learning model that can predict emotion from music for optimal ad insertion of YouTube mid-roll ads to achieve brand recall and avoid ad skipping.

Takeaways:

- The increased use and visibility of AI has been achieved through its ability to make successful predictions and generation of creative assets.
 - AI has performed well on standardized tests and the generation of creative assets such as texts, images and audio.

CNBC


Google CEO: A.I. is more important than fire or electricity

Sundar Pichai says it is artificial intelligence.

Simulated exams


| | |
|--|------------------|
| Uniform Bar Exam (MBE+MEE+MPT) ¹ | 298/400 ~90th |
| LSAT | 163 ~88th |
| SAT Evidence-Based Reading & Writing | 710/800 ~93rd |
| SAT Math | 700/800 ~89th |
| Graduate Record Examination (GRE) Quantitative | 163/170 ~80th |
| Graduate Record Examinat | 160/170 |

GPT-4
estimated percentile



AI art piece wins international photography competition

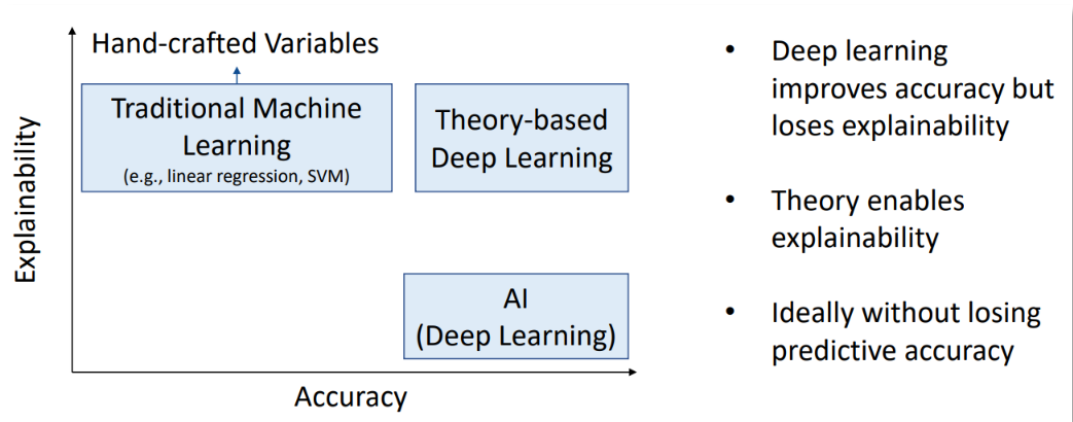
A German artist has turned down an international photography award as he revealed the photograph was AI-generated.



- **The improved performance of AI by leveraging AI (Deep Learning) for improved accuracy has come at the cost of explainability**, which is vital for areas such as managerial trust in predictions, generalizability, creating improvements and ethical concerns.
 - Explainability increases generalizability leading to a model that is transferable to other subjects.

Incorporate Theory to Gain Explainability

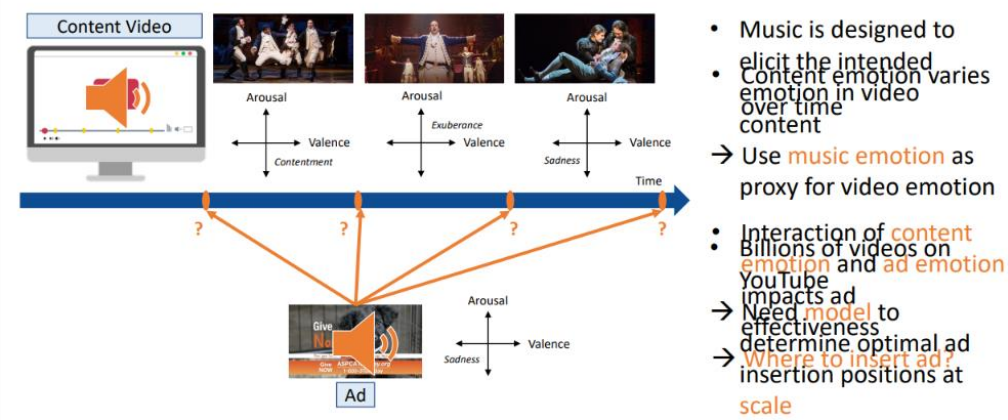
- Incorporating theory into the design of a deep learning model can lead to explainability.



- Indicated by the Behavioral Sciences, **emotion impacts behavior all throughout the consumer journey, with music being one of the key drivers** of emotion.
- **Emotion induced by content impacts ad effectiveness when an ad is placed in an optimized position within the content**, as demonstrated using YouTube mid-roll ads during the viewing of the musical Hamilton.

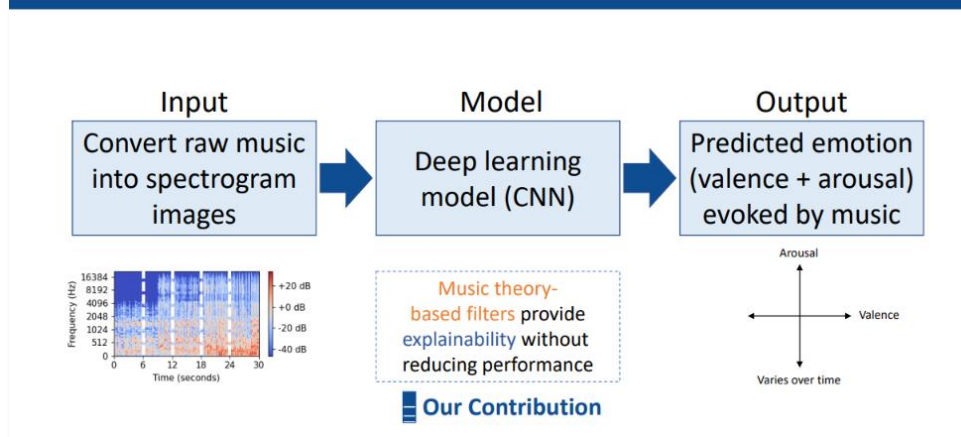
Emotion induced by content impacts ad effectiveness

Where to insert an ad based on content emotion?



- The model proposed by Fong, Kumar and Sudhir takes the **raw input of music**, **converts it into a spectrogram image** (visualization of music) and **passes this visualization through a deep learning model (CNN) to predict emotion** (valence and arousal evoked by music).

Predicting Emotion from Music



MSI's Marketing Mix Modeling Initiative Progress

Speakers:

Elea McDonnell – Associate Professor of Marketing and Associate Dean of Research at Drexel University, LeBow College of Business

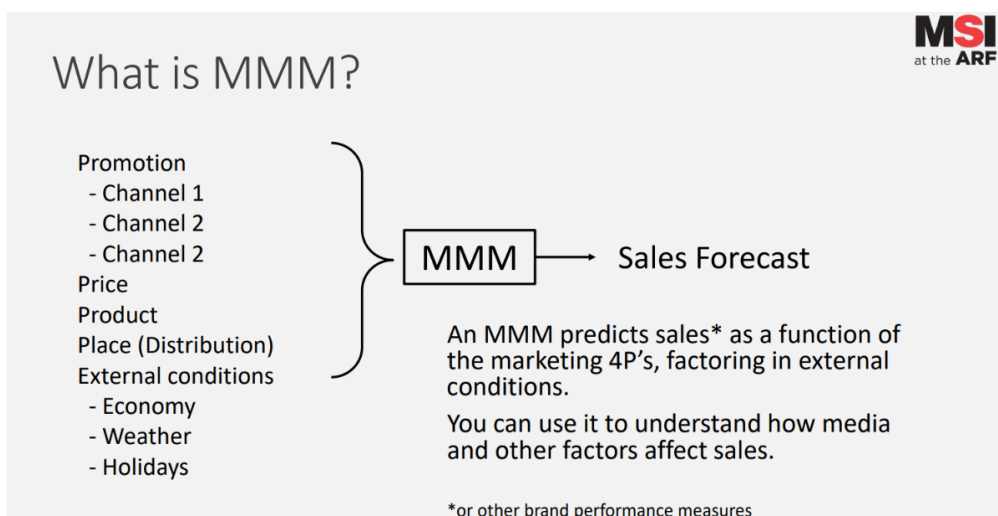
Alice Li – Associate Professor of Marketing, Ohio State University, Fisher College of Business

Overview:


Speakers Elea McDonnel (Drexel University) and Alice Li (Ohio State University) discussed the progress of the Marketing Mix Modeling Initiative at MSI. In the presentation, the speakers provided background and history on marketing mix models (MMM), starting with the introduction of MMM, along with the 4P's by Professor Jerome McCarthy in 1950. They then provided details on the function of MMM, which predicts sales (or other brand performance measures) "as a function of the marketing 4P's," factoring in external conditions. McDonnel and Li provided a distinction between marketing mix models vs. media mix models. Furthermore, the speakers examined the resurfacing need for MMM, brought about due to a variety of reasons, most importantly rising privacy regulations and the deprecation of cookies. Additionally, they recapped the important goals of this initiative, the academics involved in the process and the details behind the phases that guide it. Stakeholders were cited as data providers, designers, users and investors. Finally, they discussed important data and optimization issues surrounding this initiative.

Takeaways:

- Marketing mix modeling (MMM) determines how much impact is driven by each of the 4Ps and forecasts the future impact of altering or optimizing the marketing mix.
- **An MMM predicts sales or other brand performance** measures as a function of the marketing 4P's, factoring in external conditions.



- MMM is making a resurgence in importance as privacy regulations, loss of granular information at the cookie level, and walled gardens restrict data access to marketers.
- **The goal of this MSI Marketing Mix Model Initiative is to develop and disseminate "best practices for the design, validation and use of MMM to increase trust and ensure marketing spend is as effective as possible."**

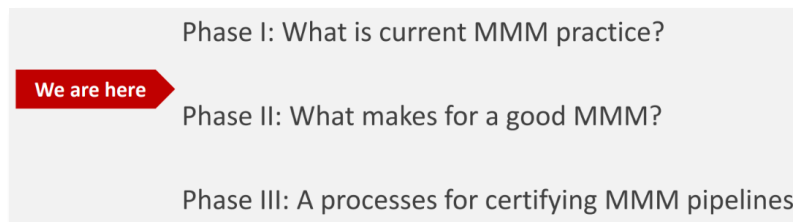


What issues are important today?

Data

- Accuracy
- Frequency, latency & agility
- Granularity, across geography, time and channel **Panel Interest**
- Adjusting to changes in ad tech
- Lack of variation, e.g. a new channel with no history
- New data: sentiment, creative quality, media attention
- Optimizing marketing spend using MMMs **Panel Interest**
- Marginal versus total ROI
- Efficiency versus effectiveness
- Training investors to use MMM

- There are **three phases in the MSI Marketing Mix Modeling Initiative:**



- **Phase I:** Scoping current MMM practice through academic and industry panels, developing objectives for phases II and III, collecting user cases, etc.
- **Phase II:** Establish industry standards
- **Phase III:** Creating a process to certify MMM pipelines.

“Responsible AI” and Data Ethics

Speaker:

Mary Purk – *Managing Director of Wharton AI and Analytics Research Centers, University of Pennsylvania, Wharton School of Business*

Miriam Vogel – *President and CEO, EqualAI*

JoAnn C. Stonier – *Chief Data Officer, MasterCard*

Overview:

In this session, Mary Purk (Wharton School, U. Pennsylvania) led an illuminating panel discussion on responsible AI usage and data ethics in marketing. Along with Miriam Vogel (EqualAI) and JoAnn C. Stonier (MasterCard), the group discussion examined the opportunities and ethical challenges created by the use of AI in marketing. In opening the

conversation, Vogel and Stonier discussed definition(s) of responsible AI and ethics in data governance. The two provided faceted feedback regarding areas such as trust, conditions under which the data was obtained found, and the avoidance of bias in AI. In terms of considerations for academics, Vogel reiterated her primary concern as trust in AI and Stonier noted that in terms of guidance and regulation, this is a "community challenge." In addition, context was brought up as a concern with generative AI, in particular, to avoid biases (e.g. metrics for "attractiveness" or "effectiveness"). Both Vogel and Stonier agreed that generative AI is a significant breakthrough, with Vogel suggesting that it is the "next phase in an industrial revolution involving knowledge." The discussion extended to topics such as governmental influence and legislation, the effects AI can have on the end user and employees, and the accelerated pace of AI's impact in many areas (e.g. workplace, healthcare, education, training, society, etc.).

Takeaways:

Definitions of Responsible AI and Ethics in Data Governance

- "What it comes down to, is **can we trust the technology we're using which is obviously so interrelated with data**, given that AI has two main ingredients: algorithms and data." We need to understand that algorithms and data are supporting us and beneficial for us and safe for us. – Miriam Vogel
- "We're really talking about the combination of safety, security and concerns. Sometimes we get so fascinated by the technology that we forget that the **technology was put in place to help all of us and we're at a really interesting juncture now where the technology is learning independently from humans.**" We're now at a juncture where the whole notion of responsible AI and data ethics is coming into play, due to the concern that we still need to be able to interact with AI and ensure that it is still working for society. – JoAnn C. Stonier
 - Responsible AI needs to ensure the reduction of bias, better results, better data, improved efficiencies and outcomes.
- Responsible AI and data ethics is a global societal issue. – JoAnn C. Stonier
- When companies begin to leverage AI, regardless of their status as a product or service company, they become AI companies. - Miriam Vogel
- Being "stewards of powerful information" places importance on the methodologies in place. **When you look at AI, it is crucial to understand "What conditions can be found in the data? What does that tell us?"** – JoAnn C. Stonier

Topics of Concern for Academics to Consider in Responsible AI

- **The primary concern with AI is trust** relating to AI systems. Consider, biases, risks, transparency, safety and the expedited environment we are currently in. – Miriam Vogel
- In terms of **generative AI**, we are in a moment where engineers and data scientists are looking to regulators and lawyers for guidance. In turn, regulators and lawyers

are looking to scientists and engineers for guidance, unveiling **this as a community challenge** – JoAnn C. Stonier

- Solving and addressing issues in AI (generative) needs to **consider the context of the problem** (e.g. metrics for attractiveness or effectiveness) – JoAnn C. Stonier

Is Generative AI a Significant Breakthrough?

- **Generative AI is the "next phase in an industrial revolution involving knowledge."** Each facet of our lives will be transformed by how AI is impacting our world. Consider the impact AI will have on employees depending on whether it is implemented responsibly or not. In addition, consider the effects of AI on the end user (e.g. culture, language, age group, ADA compliance, etc.). – Miriam Vogel
- The pace of change is accelerating due to the impact AI is having. We need to be in a position to handle this rapid change in a variety of outlets (e.g. workplace, healthcare, education, training, society, etc.). – JoAnn C. Stonier

Privacy Balance and Data Access

- There are new technologies emerging that can do a better job at data protection by leveraging proxies that can avoid the use of personal data (design thinking). – JoAnn C. Stonier
- AI is constantly learning. The scrutiny and testing of AI are vital to avoid the effects of AI-learned biases. There is a constant struggle concerning data and AI between privacy, transparency and accuracy. – Miriam Vogel

Customer Interest Graph for Dynamic Real-Time Hyper-Personalization

Speaker:

Nitzan Mekel-Bobrov – *Chief AI officer, eBay*

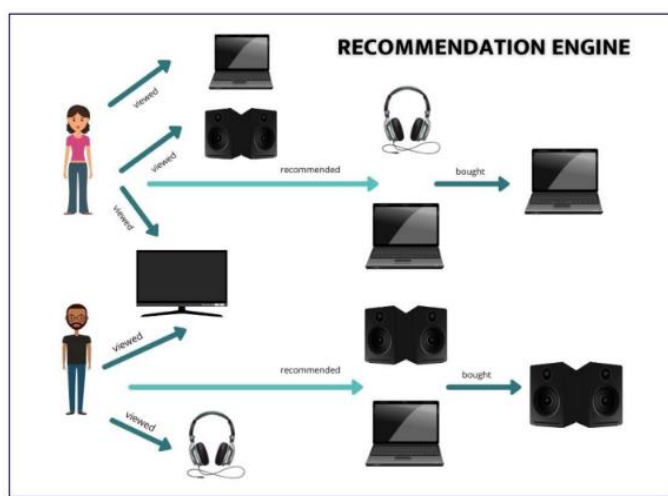
Overview:

In his presentation, Nitzan Mekel-Bobrov (eBay) discussed issues and practices of personalization using generative AI and the value of customer interest graphs to enhance the online shopping experience. He noted that personalization is now an expectation from online consumers who responded positively to this enhanced experience. Mekel-Bobrov examined attributes of the typically used "recommender system" model, which matches consumers with products of interest based on prior actions. He noted that this system has been upgraded through the addition of a deep learning model, which greatly improved personalized recommendations by capturing more facets of a product, but essentially still used the same paradigm. Mekel-Bobrov then demonstrated the more dynamic "sequence model" and the adoption of interest graphs, which not only accounted for products viewed but the time between the items as consumers viewed them, providing more clues on the strength of the consumer's interest in a product.

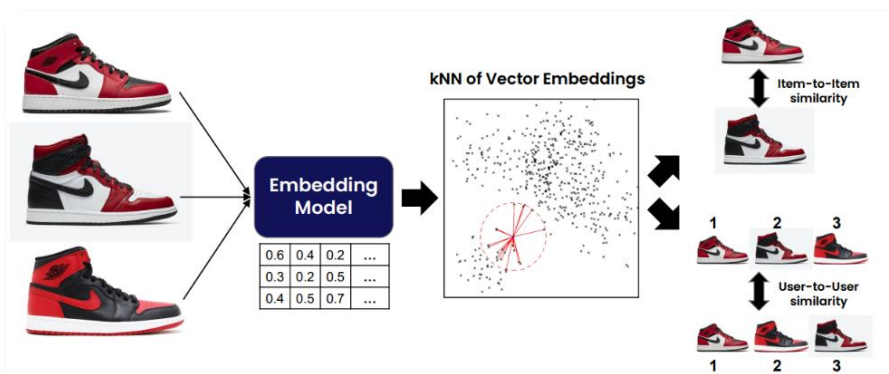
Takeaways:

Problem Definition

- Personalization has become a **"baseline expectation"** for online shoppers.
 - 56% of shoppers expect to find what they need in 3 or fewer clicks, 71% of customers become frustrated when the online shopping experience is impersonal, and there is a 28% increase in brand loyalty among Millennials when communication is personalized.
- Currently, **personalization in e-commerce typically follows** an "old-fashioned paradigm" known as **the recommender system model**, which matches consumers to related items based on past behavior or "a sequence of products they click on."



- **Deep learning can greatly improve personalized recommendations to consumers**, which allows the capture of so many different dimensions of a product or sequence of products without creating any features manually but **still using the same paradigm**.



- **This paradigm is problematic because it is created from the POV of the company**, not the consumer's, especially at the discovery stage.



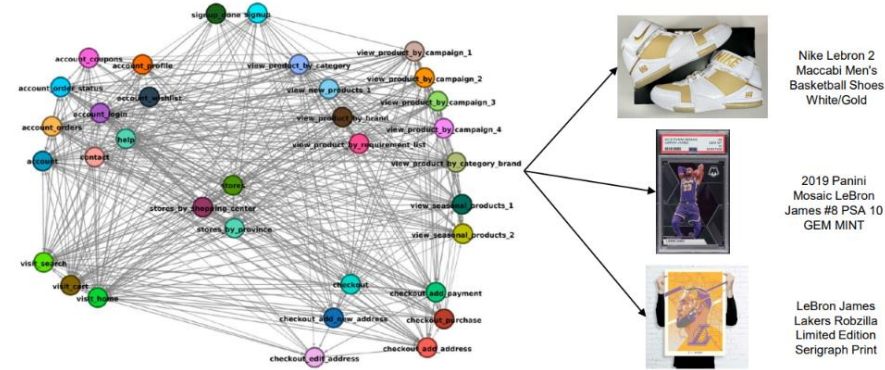
Customer problem statement

- “Let me be **inspired** with products I never thought exist”.
- “Let me shop my **values** and **interest**”.
- “Show me **diverse, fun** recommendations”.
- 70% of Gen Z’s - **discovery** is the best part of shopping.

Interest Graphs

- Building an **interest graph** can create a **multitude of actions** to dynamically find in real-time customers’ interests as they emerge.

We built an Interest Graph of 63 actions to dynamically find in real-time customer’s interests as they emerge



16

ebay

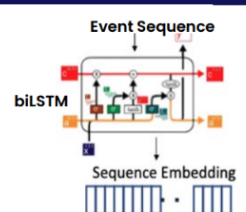
- **Creating a sequence model** accounts for not only the products viewed but the time between them as they are viewed, **providing clues on the strength of the consumer's interest in a product**.

Time Attention Based Behavior Sequence Embedding

| | | | | | | | | | | | | | |
|---------------------------|----|------------------------------------|----|------------------------------------|----|-------------------------------------|----|---------------------------------|----|----------------------------------|----|-------------------------------------|-----|
| Landed on Page [Homepage] | 1s | Viewed Product [Camera X, Image 1] | 3s | Viewed Product [Camera Y, Image 1] | 2s | Clicked Product [Camera Y, Image 1] | 1s | Viewed Product [Toy X, Image 2] | 5s | Clicked Product [Toy X, Image 2] | 7s | Viewed Product [Toy X, Description] | 30s |
|---------------------------|----|------------------------------------|----|------------------------------------|----|-------------------------------------|----|---------------------------------|----|----------------------------------|----|-------------------------------------|-----|

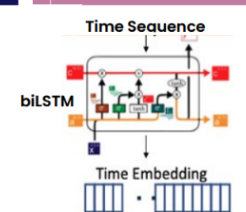
| | | | | | | | | | | | | | |
|---------------------------|------------------------------------|------------------------------------|-------------------------------------|---------------------------------|----------------------------------|-------------------------------------|----|----|----|----|----|----|-----|
| Landed on Page [Homepage] | Viewed Product [Camera X, Image 1] | Viewed Product [Camera Y, Image 1] | Clicked Product [Camera Y, Image 1] | Viewed Product [Toy X, Image 2] | Clicked Product [Toy X, Image 2] | Viewed Product [Toy X, Description] | 1s | 3s | 2s | 1s | 5s | 7s | 30s |
|---------------------------|------------------------------------|------------------------------------|-------------------------------------|---------------------------------|----------------------------------|-------------------------------------|----|----|----|----|----|----|-----|

Event Sequence



Sequence Embedding

Time Sequence



Time Embedding

+

↓

Click Stream Embedding

Customer Interest

18

ebay

Day 2 – May 5, 2023 | 8:00 am – 12:10 pm EDT

Using Simulated Data to Validate Marketing Mix Models

Speaker:

Jessica Nguyen – *Quantitative Researcher, Meta*

Overview:

In this session Jessica Nguyen (Meta) examined the value and challenges of Marketing Mix Models (MMMs), which are seeing a major resurgence due to increased privacy concerns and regulations, the lack of causal outcomes from some attribution methods, and the increased ability and efficiency in building these models. In her discussion, Nguyen noted that though MMMs provide many qualities, they do come with some setbacks. Specifically, she indicated that MMMs are difficult to validate due to the lack of ground truth data and time series data. Additionally, Nguyen recognized that the success of MMMs requires years of data that newer advertisers do not have access to. In overcoming these obstacles, she provided examples of methods to achieve validation such as backtesting, calibrating with experiments and testing on simulated data. Additionally, Nguyen provided more insight on the use of simulated data to validate MMM through the use of SiMMMulator, an open-source R package allowing users to simulate data from scratch.

Takeaways:

Resurgence in Marketing Mix Models

- Marketing Mix Models (MMMs) have seen a resurgence stemming from **increased privacy concerns and regulations** making access to granular data difficult, concerns over the **lack of causality from attribution methods**, and the ability to **create the models faster, incorporating machine learning**.

Validating Marketing Mix Models and its Challenges

- Validating MMMs can be challenging due to the **lack of ground truth data**, the **lack of time series data** and the **lack of sufficient data points** available to new advertisers to make an accurate model.

MMMs are commonly used, but validating them is difficult because ...



Lack of ground truth data

Advertisers may have limited data with ROIs of various channels that predictions can be compared to



Time series data

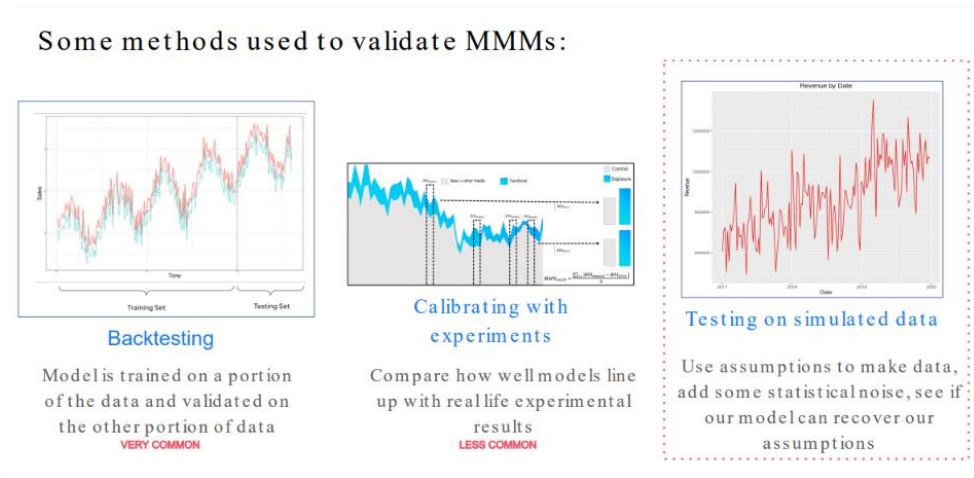
Since past data influences subsequent data, it's hard to pick the right time period for a holdout



Requires many years of data

Newer advertisers may not have sufficient data points to make an accurate model or do model validation

- Some **methods used to validate MMMs** include: **backtesting**, **calibrating with experiments (RCTs)** and **testing on simulated data**.



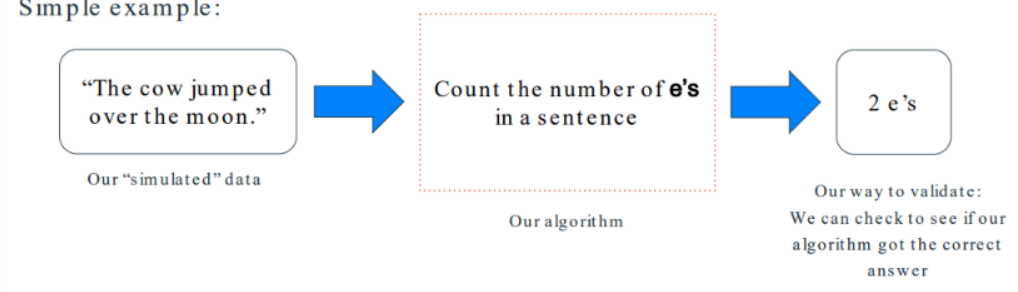
Using Simulated Data

- Simulated data is "data that we make up."** This can be done by taking "existing data and adding a lot of statistical noise to it," or starting with some "basic assumptions" and adding statistical noise.
- This method is useful because **"since we created the data, we know what a model or code is supposed to output."**

Why is simulated data useful?

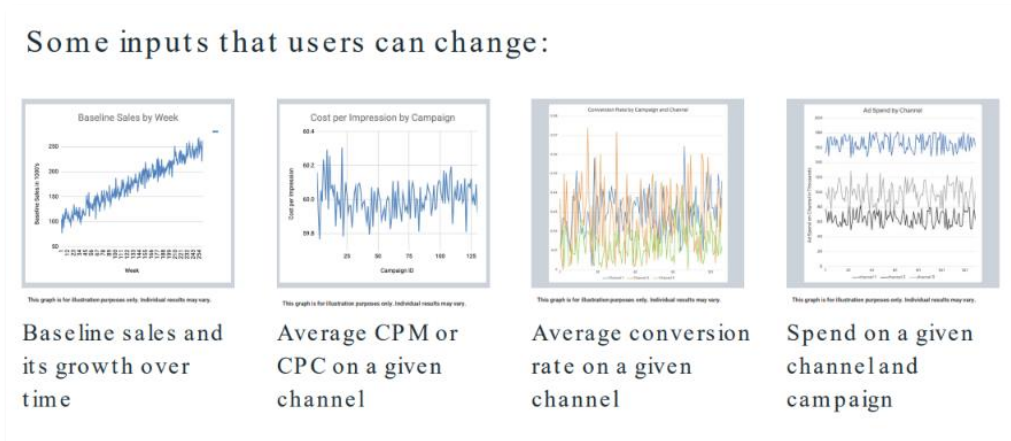
Since we created the data, we know what a model or code is supposed to output

Simple example:



- Simulated data be used to validate Marketing Mix Models by recreating the advertising process**, and plugging them into MMMs to see how close your MMM can get to the assumptions you put in.
 - Simulate various scenarios: see how your MMM responds, and what kind of environment your MMMs are sensitive to.
 - Quantify the value of innovations.

- **Leveraging siMMulator**, an open-source R package, **allows users to simulate data from scratch**.
 - By using siMMulator different data sets can be generated by putting in different inputs.
 - Simulate various scenarios to see how your MMM responds, and what kind of environment your MMMs are sensitive to.



Mega or Micro? Influencer Selection Using Follower Elasticity

Speaker:

Ryan Dew – *Assistant Professor of Marketing, University of Pennsylvania, The Wharton School of Business*

Overview:

In this session, Ryan Dew (Wharton School, U. Pennsylvania) examined the expanding area of influencer marketing. In his discussion, Dew noted the criterion organizations use when selecting an influencer, and the benefits and shortcomings of each type of influencer (e.g. Mega influencers, micro-influencers). Dew discussed the ambiguous nature of influencer marketing by acknowledging that more followers don't always yield the most impressions. Considering popularity versus impressions, Dew then [examined research](#) done in partnership with Zijun Tian and Raghu Lyengar, both also of the Wharton School of Business. Through this study, Dew examined a framework he and his coauthors developed to determine "follower elasticity" to help select the best influencer for an organization's purposes. Their methodology included a causal analysis between follower size and video impressions on TikTok, addressing the nonlinear treatment effect, the key confound of context, the heterogeneity of the relationship, and unobserved confounders.

Takeaways:

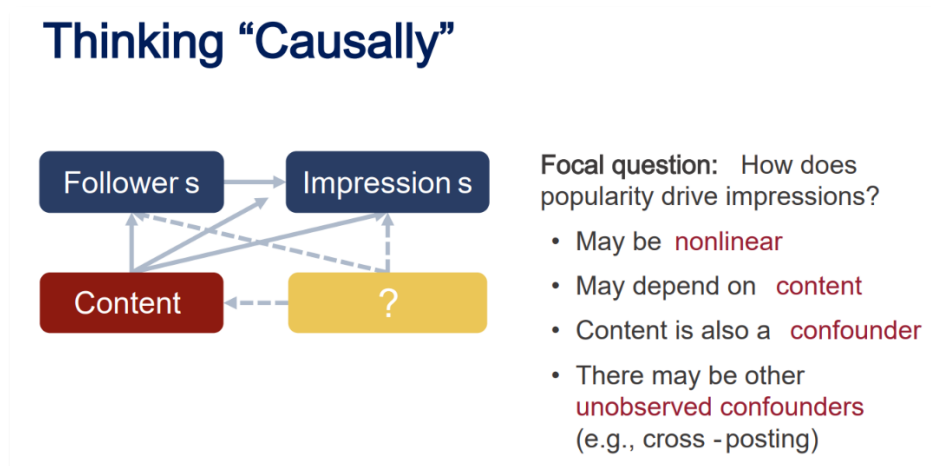
Who should you partner with?

- In terms of mega influencers and micro influencers, **the popularity of the influencer doesn't necessarily equal more impressions.**



The Study

- Collecting data from the Discover page of TikTok, the study accounted for all hashtags from October 2020 to April 2021, all videos posted under each hashtag, and all tracked from their introduction to maturity.
 - The focal question: How does popularity drive impressions?

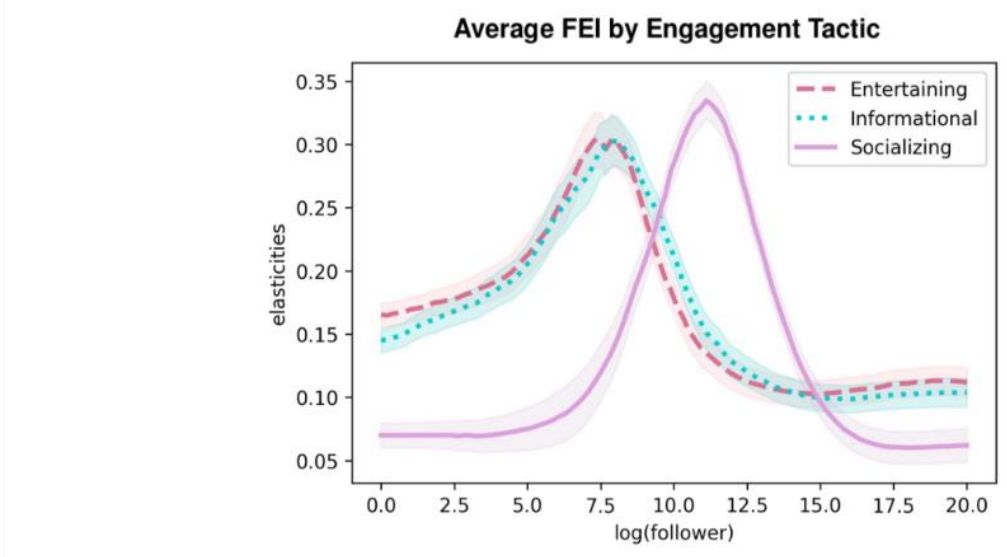


- The answer may be nonlinear and ultimately depend on the content and the effects of confounder(s)** (e.g. the influencer posts a lot of popular content that is irrelevant to your mission, cross posting, etc.).

A New KPI: Follower Elasticity

- Representation learning** can provide a **statistical representation** of content.
 - Through **content representations combined with a machine learning model** the follower elasticity of impressions (**FEI**) can help observe **engagement tactics** such as **entertainment, informational and socializing**. This is indicated through elasticity curves.

Key Findings



- **Results of the study showed** that the **Mid-tiered influencer is more engaging** and creates an atmosphere that tends to be viewed as more attainable to consumers.
 - Findings suggested **the content of a campaign plays a larger role** in how popular of an influencer should be sponsored.

AI at the Front End of Innovation

Speaker:

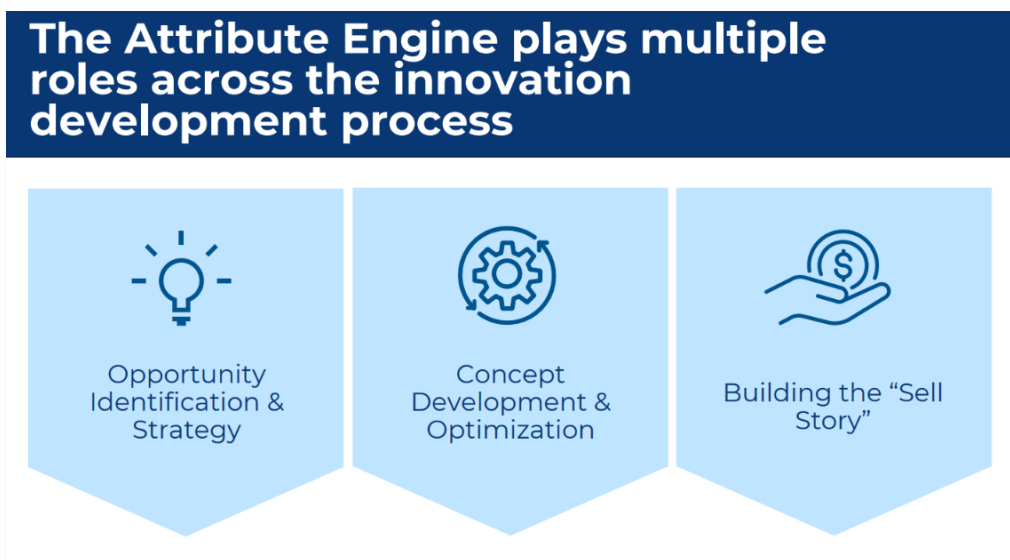
Jessica Yankell – *Director, PBNA Innovation Insights & Strategy, PepsiCo*

Overview:

Jessica Yankell (PepsiCo) examined the important role of AI in product innovation. In her discussion, she indicated that in terms of product innovation and the expansion of new offerings, the challenge for PepsiCo is that they “compete across a wide range of evolving categories” and place great importance on understanding “which consumer trends to act on and when.” Yankell acknowledged that PepsiCo “had to rethink how we innovate, finding new more agile ways to bring the consumer in” besides the traditional focus group. She then demonstrated ways they leverage “AI-fueled insights for innovation” through a tool known as the Attribute Engine, which can aggregate deep insights and conversations of beverage-related feedback across multiple platforms to inform new products.

Takeaways:

- Old methods such as qualitative focus groups are still relevant in new product development but still have deficiencies, particularly when **there is a need for more efficient, comprehensive and agile feedback.**
 - By leveraging their Attribute Engine, **PepsiCo can look across "all the spheres that influence beverage trends"** such as TikTok, Pinterest, YouTube, Reddit, Instagram, etc.
 - The Attribution Engine takes all the information derived from all the platforms, and this **information is parsed out into a very beverage industry-specific taxonomy.**
 - The information in the Attribute Engine can be used to **"create the universe of conversations that are happening around that consumption occasion to help us build the concepts"** to test with consumers.



- **AI enables real-time consumer-centricity at scale** in a rapidly evolving landscape.
- **"Quality data, relevant taxonomies and integrated frameworks"** are vital to success.
- There is still a **balance between the role AI and human** actionability of outcomes.

InnoVAE: Generative AI for Patents, Innovations, and Firms

Speaker:

Dokyun Lee – *Kelli Questrom Associate Professor of Management, Boston University, Questrom School of Business*

Overview:

Dokyn Lee (Boston University - Questrom School of Business) examined [research](#) he did in conjunction with Zhaoqi Cheng (Boston University - Questrom School of Business) and Prasanna Tambe (Wharton School, U. Pennsylvania). In the opening, Lee provided a brief overview of Generative AI and its powerful capabilities such as the ability to learn to generate complex objects (images, documents, patents, portfolios, etc.). Additionally, Lee examined the shortcomings of Generative AI such as only holding the ability being able to learn from data that has been digitized. In his discussion, Lee explored the development of a "generative deep learning model based on a Variational AutoEncoder ("InnoVAE") that converts unstructured patent text into an interpretable, spatial representation of innovation ("Innovation Space")." This model serves as a tool that enables the interpretation and grouping of patents and/or firms which can provide a variety of deep and actionable insights such as potential outcomes when combining patents, ranking companies based on a technological factor, firm movement in an innovation space over time, etc.

Takeaways:

Generative AI Overview – Strengths and Weaknesses

- Generative AI is a model data generative process. In short, Generative AI "generates things."
 - **Generative AI can learn to generate any complex object** (e.g., images, documents, patents, jobs, portfolios, digital twin, etc.). In doing so this technology learns object space and compositions in scale.
 - Subsequently, a trained model can map out the object space and **"provide deeper insights** (compare and contrast)."
 - The trained model can also **"augment the purposeful synthesis of a new and creative object."**
- **Generative AI can only learn from data that has been digitized** (e.g. Web or available data). Generative AI is "lacking data on senses, muscle memory, high resolution of feelings and emotions."

Can AI Represent Innovation? How is this Useful?

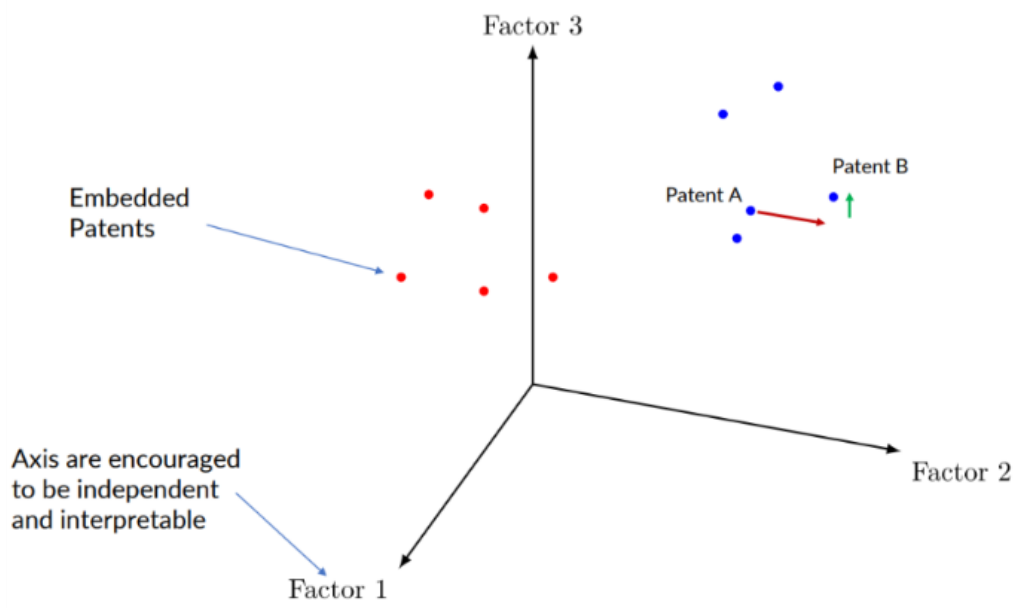
Research Question: Can AI **Represent** Innovation? And how is this useful?

1. **InnoVAE** (variational autoencoder) to estimate **disentangled representations** of **patents** using multimodal data
 - map **real-world objects** → low dimensional vector space
 - (disentangling) each dimensions extracted such that:
 - Statistically more independent (semantically more orthogonal)
 - Movement within the space rendered meaningful/understandable
2. Patents in an **interpretable vector space** of **factors of innovation**.
 - e.g., Computing patents may reside in dimensions like “security”, “human-computer interaction”

Results & Contributions Overview

- **Innovation Space (IS)** - can enable explorations into patents, innovation and firms by providing distance and movement measures, which can be used on any business object such as jobs, firms and products.
- **The image below is a 3D visualization of Innovation space from InnoVAE.** The Images shows that the factors are distinct and data-driven and that similar patents are near each other.
 - The image indicates that **Patent B has increased factors 2 and 3 compared to Patent A** (i.e., B is more exceptional).

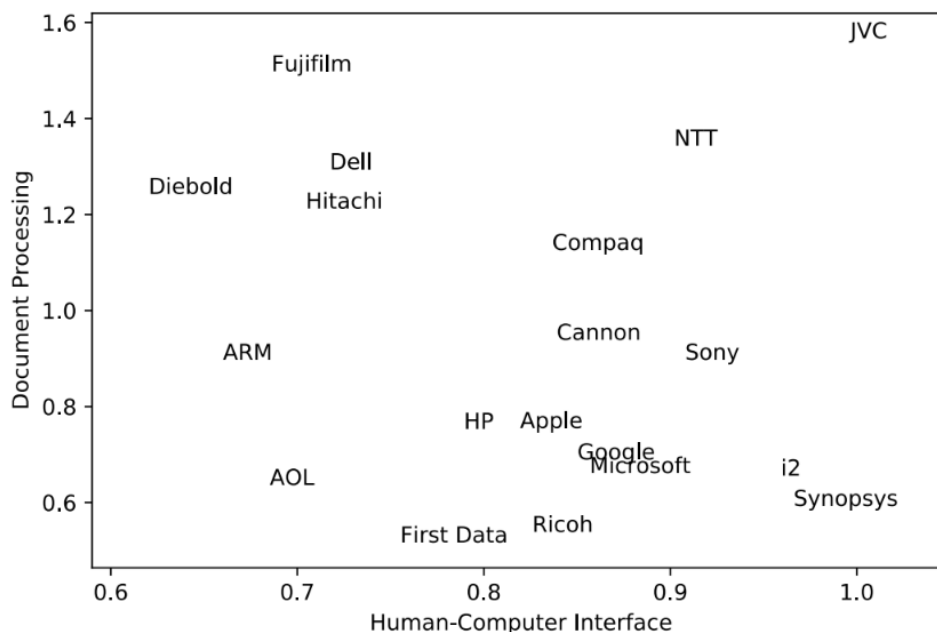
Visualization of Innovation Space from InnoVAE



- This technology can also **represent firms in the Innovation Space**, as shown in the following graphic.

- This can assist in the **comparison and contrast** of tech companies based on document processing and human-computer interface to be useful in a business strategy.

Situating Firms in the Innovation Space (2D Example)



- **This Innovation Space allows the exploration of outcomes when combining** patents (e.g. automate combinational creativity), how exceptional a particular patent is with respect to a specific technological factor (e.g., user interface), ranking companies based on a technological factor, how firms move in Innovation Space over time, innovation activity after merger, etc.

How Behavioral Science Can Help Data Analytics

Speaker:

Stefano Puntoni – *Professor of Marketing, University of Pennsylvania, Wharton School of Business*

Shiri Melumad – *Assistant Professor of Marketing, University of Pennsylvania, Wharton School of Business*

Patrick Moriarty – *SVP of Analytics Practice in North America, Kantar*

Overview:

In this panel discussion Stefano Puntoni (Wharton School, U. Pennsylvania), Shiri Melumad (Wharton School, U. Pennsylvania) and Patrick Moriarty (Kantar) explore the merger of the Behavioral Sciences with AI and data analytics to improve the reception and application of the technology on a human and societal level. The panel examined the research on human factors, user experience (UX), and the adoption of these technologies in improving customer experience. In opening the conversation, Patrick Moriarty discussed the highly relevant but under-evaluated impact on humans and the opportunities and challenges created by AI. He also noted the delicate and strategic process of implementing these technologies. Shiri Melumad echoed a common theme of the conference, acknowledging that "context really matters when you're trying to interpret data." For example, research shows that responses can change based on the device subjects are using (i.e. PC, handheld device, voice technologies, etc.). For his part, Stefano Puntoni noted the importance of keeping the consumer and the human element at the center of implementing and making decisions based on new technology.

Takeaways:

- As we think about AI, the notion of **"the human impact is highly relevant but often under-evaluated."** – Patrick Moriarty
 - In addition to being a proving ground for programmers, it is important to understand how leveraging AI can improve workplace productivity and how people interact with these technologies as workflows change.
- **"Context really matters when you're trying to interpret data.** Where consumer data is coming from can really alter the nature of that data." – Shiri Melumad
 - Specifically, Melumad noted that different technological devices can alter the nature of consumer data. The generation of content on a PC for example can change the nature of responses. She pointed to her research that indicated that people tend to be more personal in their responses via their handheld devices than responses given through a PC.
- "It is easy for companies to be caught up in technical problems" because it's such a large and new problem. This in turn can make companies lose track of why they are implementing these technologies in the first place—which is the decisions and the people. – Stefano Puntoni
 - **Advertising is often obsessed with predicting behavior** (e.g. who is going to click) but the true problem of advertising is a "problem of incrementality" and understanding for "whom a particular message is going to have the biggest effect."

MSI Research Priorities 2022-2024

1. Data challenges from business disruption and missing information.

- 1.1. Effects of privacy regulation on customer value creation
- 1.2. Effects of ability to generate value from advertising
- 1.3. Analytics challenges following changes in firm strategy

2. Measurement and analytics

- 2.1. Measuring returns to analytics with greater ability to support causal claims
- 2.2. Analytics for short-term versus long-term effects
- 2.3. Brand measurement
- 2.4. Attention, engagement and customer experience

3. Long-term changes in how customers and firms interact.

- 3.1. Effects of changed patterns of living and working on customer demand
- 3.2. Effects on intra-firm processes

4. Inflation and supply chain disruption.

- 4.1. Effects on reconfigured supply chains
- 4.2. Coping with inflation

5. Corporate mission shifts from shareholder value to stakeholder value.

- 5.1. Healthcare pricing for access by lower-income consumers
- 5.2. Tele-medicine
- 5.3. Responsible production and consumption and role of new food technologies
- 5.4. ESG influence on marketing budget allocation
- 5.5. Brand purpose, political ideology and consumer behavior
- 5.6. Firm externally – focused activism versus internally – focused action
- 5.7. Diversity and inclusion

6. Regulatory and public policy issues affecting marketing.

- 6.1. Effects of privacy policies on competition
- 6.2. What aspects of privacy matter to consumers?
- 6.3. Ethical forms of exchange for consumers to opt in to share data
- 6.4. Regulatory barriers to innovation to improve health and well-being
- 6.5 Achieving profitability in ways consistent with United Nations SDGs

7. The influence of marketing in the firm

- 7.1. Organizational structure and the influence of marketing
- 7.2. Customer value versus brand value