Data science, entrepreneurship, and economic inclusion

J.P. Allen
School of Management
University of San Francisco
Digitization creates (lots of) value -> wealth

My interests: More inclusion in digital wealth creation. Addressing the challenges of sharing digital value through labor markets and entrepreneurship.

‘Data science’ represents a new chapter in digitization.
Wealth shift to digital sectors (IT, Finance) from commodities (Energy, Materials)

Inside the black box of the new digital era

“On Technological Expectations”: The shaping of technological capabilities

“Learning by Using”: The mediating factors between capabilities and commercial success

(Rosenberg, 1982)
Inside the black dot.
Create your Lex bot

Amazon Lex enables any developer to build conversational chatbots quickly and easily. With Amazon Lex, no deep learning expertise is necessary—you just specify the basic conversational flow directly from the console, and then Amazon Lex manages the dialogue and dynamically adjusts the response. To get started, you can choose one of the sample bots provided below or build a new custom bot from scratch.

**OUR SHOP**

- Create your own: Custom app
- Try a sample: BookTrip, OrderFlowers, ScheduleAppointment

**Bot name**
(OrderFlowers)

**Intents**
- DeliveryOrder: A particular goal that the user wants to achieve

**Utterances**
- I'd like to order flowers.
- What kind?
- 1 dozen roses please.
- Where should we deliver?
- Can I go ahead with your order of a dozen roses to your home?
- Yes.
- Thank you. Your order went through successfully.

**Slots**
- Data the user must provide to fulfill the intent

**Prompts**
- Questions that ask the user to input data

**Fulfillment**
- The business logic required to fulfill the user's intent

**IAM role**
- Select an IAM role (if applicable)

[Create button]
## Pricing Example 2

A bot that sends 500 speech requests and 500 text requests in December. Monthly charges for December:

<table>
<thead>
<tr>
<th>Input requests to process</th>
<th>Cost per request</th>
<th>Number of requests</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 speech requests</td>
<td>$0.004</td>
<td>500 requests</td>
<td>$2.00</td>
</tr>
<tr>
<td>500 text requests</td>
<td>$0.00075</td>
<td>500 requests</td>
<td>$0.38</td>
</tr>
</tbody>
</table>

**Total Amazon Lex charges for 500 speech and 500 text requests**

$2.38
Alexa's persona

Alexa's personality is friendly, upbeat, and helpful. She can handle daily tasks with ease and accuracy. She's honest about anything blocking her way, but also fun and personable, able to make small talk without being obtrusive or inappropriate.

**Key elements across Alexa**

There are some guiding principles used across all Alexa skill experiences. These include:

- Speaking naturally and conversationally
- Having variation in the responses given
- Engaging the customer with questions
- Staying true to Alexa's personality

Learn more about Alexa's personality.

---

**Amazon Alexa**

Alexa's advice to 'kill your foster parents' fuels concern over Amazon Echo

Smart speaker’s remarks, apparently quoted from Reddit, come as Amazon tries to boost speaker’s conversational capacity

---

(Amazon, Alexa Design Guide)

(theguardian.com, 12/22/18)
Will I like this restaurant?

MLModelId: "<MODEL_ID>",
PredictEndpoint: "<ENDPOINT_URL>",
Records: {
    age: "20_34",
    budget: "20_to_30",
    cuisine_type: "Continental",
    price: "35_49",
    gender: "female"
}

Prediction
Prediction: {
    rating: "excellent"
}

Datasource (S3)
Another option is to use Redshift

Customer Reviews

Customer
<table>
<thead>
<tr>
<th>Category</th>
<th>Icon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytics</td>
<td><img src="image" alt="Analytics" /></td>
</tr>
<tr>
<td>Application Integration</td>
<td><img src="image" alt="Integration" /></td>
</tr>
<tr>
<td>AR &amp; VR</td>
<td><img src="image" alt="AR &amp; VR" /></td>
</tr>
<tr>
<td>AWS Cost Management</td>
<td><img src="image" alt="Cost Management" /></td>
</tr>
<tr>
<td>Blockchain</td>
<td><img src="image" alt="Blockchain" /></td>
</tr>
<tr>
<td>Business Applications</td>
<td><img src="image" alt="Business" /></td>
</tr>
<tr>
<td>Compute</td>
<td><img src="image" alt="Compute" /></td>
</tr>
<tr>
<td>Customer Engagement</td>
<td><img src="image" alt="Engagement" /></td>
</tr>
<tr>
<td>Database</td>
<td><img src="image" alt="Database" /></td>
</tr>
<tr>
<td>Desktop &amp; App Streaming</td>
<td><img src="image" alt="Desktop &amp; App" /></td>
</tr>
<tr>
<td>Developer Tools</td>
<td><img src="image" alt="Developer" /></td>
</tr>
<tr>
<td>Game Tech</td>
<td><img src="image" alt="Game" /></td>
</tr>
<tr>
<td>Internet of Things</td>
<td><img src="image" alt="Internet of Things" /></td>
</tr>
<tr>
<td>Machine Learning</td>
<td><img src="image" alt="Machine Learning" /></td>
</tr>
<tr>
<td>Management &amp; Governance</td>
<td><img src="image" alt="Management &amp; Governance" /></td>
</tr>
<tr>
<td>Media Services</td>
<td><img src="image" alt="Media" /></td>
</tr>
<tr>
<td>Migration &amp; Transfer</td>
<td><img src="image" alt="Migration &amp; Transfer" /></td>
</tr>
<tr>
<td>Mobile</td>
<td><img src="image" alt="Mobile" /></td>
</tr>
<tr>
<td>Networking &amp; Content Delivery</td>
<td><img src="image" alt="Networking &amp; Content" /></td>
</tr>
<tr>
<td>Robotics</td>
<td><img src="image" alt="Robotics" /></td>
</tr>
<tr>
<td>Satellite</td>
<td><img src="image" alt="Satellite" /></td>
</tr>
<tr>
<td>Security, Identity &amp; Compliance</td>
<td><img src="image" alt="Security, Identity &amp; Compliance" /></td>
</tr>
<tr>
<td>Storage</td>
<td><img src="image" alt="Storage" /></td>
</tr>
</tbody>
</table>
PERIMETER LAYER
AWS data center physical security begins at the Perimeter Layer. This layer includes a number of security features depending on the location, such as security guards, fencing, security feeds, intrusion detection technology, and other security measures.

EXPLORE »

INFRASTRUCTURE LAYER
The Infrastructure Layer is the data center building and the equipment and systems that keep it running. Components like backup power equipment, the HVAC system, and fire suppression equipment are all part of the Infrastructure Layer.

EXPLORE »

DATA LAYER
The Data Layer is the most critical point of protection because it is the only area that holds customer data. Protection begins by restricting access and maintaining a separation of privilege for each layer. In addition, we deploy threat detection devices and system protocols, further safeguarding this layer.

EXPLORE »

ENVIRONMENTAL LAYER
The Environmental Layer is dedicated to environmental considerations from site selection and construction to operations and sustainability. AWS carefully chooses our data center locations to mitigate environmental risk, such as flooding, extreme weather, and seismic activity.

EXPLORE »
1-16 of over 10,000 results for "my little pony"

Show results for

**Toys & Games**
- Action & Toy Figures
- Stuffed Animals & Plush Toys
- Dolls
- Kids' Party Supplies
- Action & Toy Figure Playsets
  - See more

**Girls' Fashion**
- Girls' Clothing
- Girls' Fashion Hoodies & Sweatshirts
- Girls' Sleepwear
- Girls' Dresses
- Girls' Tops & Tees

**Books**
- Children's Books
- Children's Beginner Readers
- Children's Horse Books
- Children's Action & Adventure Books
- Children's Humor

**Movies & TV**
- Kids & Family
- Prime Video
- DVD
- Movies
- TV
- Comedy

**Amazon's My Little Pony Page**
Discover books, read about the author, find related products, and more. [More about My Little Pony]

**My Little Pony Mermaid Blanket for Teen and Adult MLP Fans - Super Soft My Little Pony Fleece Blanket Featuring Three Favorite Ponies**
by Blankie Tails

$19.95 ✚ prime
Get it as soon as TODAY, Dec 30
FREE Shipping

**My Little Pony My Little Pony Necklace Activity Set**
by My Little Pony

$9.98 ✚ prime
Get it as soon as TODAY, Dec 30
FREE Shipping

More Buying Choices
$9.28 (20 used & new offers)

Amazon's Choice

Manufacturer recommended age: 3 - 10 Years

Product Description
... Design your very own My Little Pony necklaces with this fun easy to use ...
How is Amazon’s Choice selected?

Much like with search ranking, Amazon does not spell out exactly what its criteria are for Amazon’s Choice, and how they are weighted against each other. But they do provide some information to the shopper, which gives us an indication of what criteria is being taken into account.

- Highly rated from user reviews
- Competitively priced
- Popular with shoppers searching for that keyword
- Bestseller in its category
- Available to ship immediately
- Eligible for Prime delivery
- Lower return rate than similar products

We’ve also seen that Amazon’s Choice products consistently have Prime delivery, so marketplace sellers must be using FBA or Seller-Fulfilled Prime. Many items with the badge are branded items sold by Amazon themselves, but there are also plenty of Amazon’s Choice items from third-party sellers, including those with their own private label brands.

(webretailer.com, 2018)
Vocabulary for a new ‘data science’ black box

<table>
<thead>
<tr>
<th>Information</th>
<th>Datafication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automation</td>
<td>Algorithms</td>
</tr>
<tr>
<td>Communication</td>
<td>Platforms</td>
</tr>
</tbody>
</table>
Questions of entrepreneurship and economic inclusion in the new ‘data science’ world

Openness to new entrants - businesses and labor.

Openness to digital business models with different types of value sharing.

Entrepreneurship and skill-building in the context of giant platforms, operating through complex and opaque algorithms, and uniquely detailed data sets.
Three quick examples

- **Datafication**: Predictive policing in Netherlands
  Growth of intermediary role: Intelligence officers. Questions of reinforcement and bias.

- **Algorithm**: Predictim baby-sitter screening startup
  Data scientists meet the real world. How do sitter-entrepreneurs react?

- **Platform**: YouTube ‘Up Next’ video recommender
  Algorithm focused on engagement-maximizing business model. As many human content moderators as engineers.
Fig. 1. An example of a CAS heat map.

(Waardenberg et al., 2018)
<table>
<thead>
<tr>
<th>Type of data included</th>
<th>Input variables of the algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crime reports</td>
<td>· Number of times the crime occurred</td>
</tr>
<tr>
<td></td>
<td>· Specific location of the crime</td>
</tr>
<tr>
<td></td>
<td>· Specific time of the day</td>
</tr>
<tr>
<td></td>
<td>· Season</td>
</tr>
<tr>
<td>Socio-demographics and location-specific data</td>
<td>· Number and kind of companies present in the neighborhood (e.g., bars, coffee shops, banks)</td>
</tr>
<tr>
<td></td>
<td>· Distance to closest known offender (e.g., mugger, robber, burglar)</td>
</tr>
<tr>
<td></td>
<td>· Mean distance to 10 closest known offenders</td>
</tr>
<tr>
<td></td>
<td>· Distance to nearest highway exit</td>
</tr>
<tr>
<td></td>
<td>· Number of burglaries, robberies, etcetera, in several different time periods (relative to the reference moment)</td>
</tr>
</tbody>
</table>
Algorithms arrive from CS with a set of practices and understandings that open up relatively inexpensive value creation opportunities, as well as resources for creating legitimacy.
We trained our AI algorithms using billions of social media data points to increase the accuracy of our reports about important personality traits. Our proprietary technology can identify social media content that is aggressive, abusive, explicit, or offensive with a very high accuracy rate, and it continues to improve.

Parts of a Predictim Report

Report Summary

<table>
<thead>
<tr>
<th>Feature</th>
<th>Risk Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullying / Harassment</td>
<td>Very High Risk</td>
</tr>
<tr>
<td>Disrespectful Attitude</td>
<td>Very High Risk</td>
</tr>
<tr>
<td>Explicit Content</td>
<td>Very Low Risk</td>
</tr>
<tr>
<td>Drug Abuse</td>
<td>Very High Risk</td>
</tr>
</tbody>
</table>

Personality Features

Each Predictim Report contains an assessment of four different personality features: propensity towards Bullying / Harrassment, Disrespectfulness / Bad Attitude, Explicit Content, and Drug Abuse.
As algorithms take over, YouTube's recommendations highlight a human problem

A supercomputer playing chess against your mind to get you to keep watching.

Within a few clicks, a search for "Saturn" and other science topics on YouTube lead to recommendations for clusters of conspiracy and propaganda videos. 

(nbcnews.com, 2018)
Responses to technology-related inequality in developed economies

- Education and skills
- Regulations (taxation, IP, antitrust)
- Gender and ethnic inclusion efforts
  - *Alternative business models and platforms*

(Allen, 2017)
Empowering non-Computer/Data Science people in this new digital era

- Quickly prototype
- Create or access unique data sets
- Experiment and learn faster in ‘use’
- Shift between multiple business models

(Allen, 2019)
Data science, entrepreneurship, and economic inclusion

J.P. Allen
School of Management
University of San Francisco

Thanks for your time!
jpallen@usfca.edu