



Using AI to Revolutionize Global Payments

Dr Hui Wang, Sr. Director Global Risk and Data Science, PayPal

Oct. 17th, 2017



极客时间

重拾极客精神·提升技术认知

每天10分钟,邀请顶级技术专家,为你传道授业解惑。



扫一扫,试读专栏

AiCon

全球人工智能技术大会 2018

助力人工智能落地

2018.1.13 - 1.14 北京国际会议中心



扫描关注大会官网

主办方 **Geekbang** & **InfoQ**
极客邦科技

ArchSummit

全球架构师峰会 2017

12月8-9日 北京·国际会议中心



APSEC 2017



APSEC 2017

24th Asia-Pacific Software Engineering Conference
4-8 December 2017, Nanjing, Jiangsu, China

12月4-8日

中国南京



了解详情

- What is PayPal
- How do we use AI to revolutionize payments
- World class data science team

Global Scale

193M*

Consumer Accounts

17M*

Merchant Accounts

We are a trusted part of people's financial lives and a partner to merchants in 200+ markets around the world.

PayPal's global customer service team provide support in more than 20 languages.

Our customers can accept payments in 100+ currencies, withdraw funds to their bank accounts in 56 currencies, and hold balances in their PayPal accounts in 25 currencies.



2016 Full-Year Results

\$10.84B

REVENUE¹

\$354B

TOTAL PAYMENT
VOLUME²

6.1B

TRANSACTIONS³

\$102B

MOBILE PAYMENT
VOLUME

2.0B

MOBILE PAYMENT
TRANSACTIONS

¹ **Revenue** is presented on an Non GAAP FX-neutral basis, and includes the impact of pro forma adjustments directly attributable to the separation of the company from eBay Inc. on July 17, 2015 had they existed historically.

² **Total Payment Volume** is the value of payments, net of payment reversals, successfully completed through our Payments Platform, excluding transactions processed through our gateway and Paydiant products.

³ **Payment Transactions** is the total number of payments, net of payment reversals, successfully completed through our Payments Platform, excluding transactions processed through our gateway and Paydiant products.

The power of our platform

PayPal operates one of the
largest

PRIVATE CLOUDS

in the world

6.1
BILLION
payments in 2016

~600
payments/
second at peak

+2.5
MILLION
developers

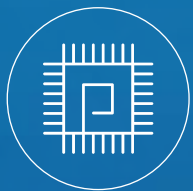
Dedicated to provide the best
customer experience through a
strong, highly scalable,
continuously available platform

4.5T
database calls/
quarter

197M
active customer
accounts

42
petabytes
of data*

PLATFORM.

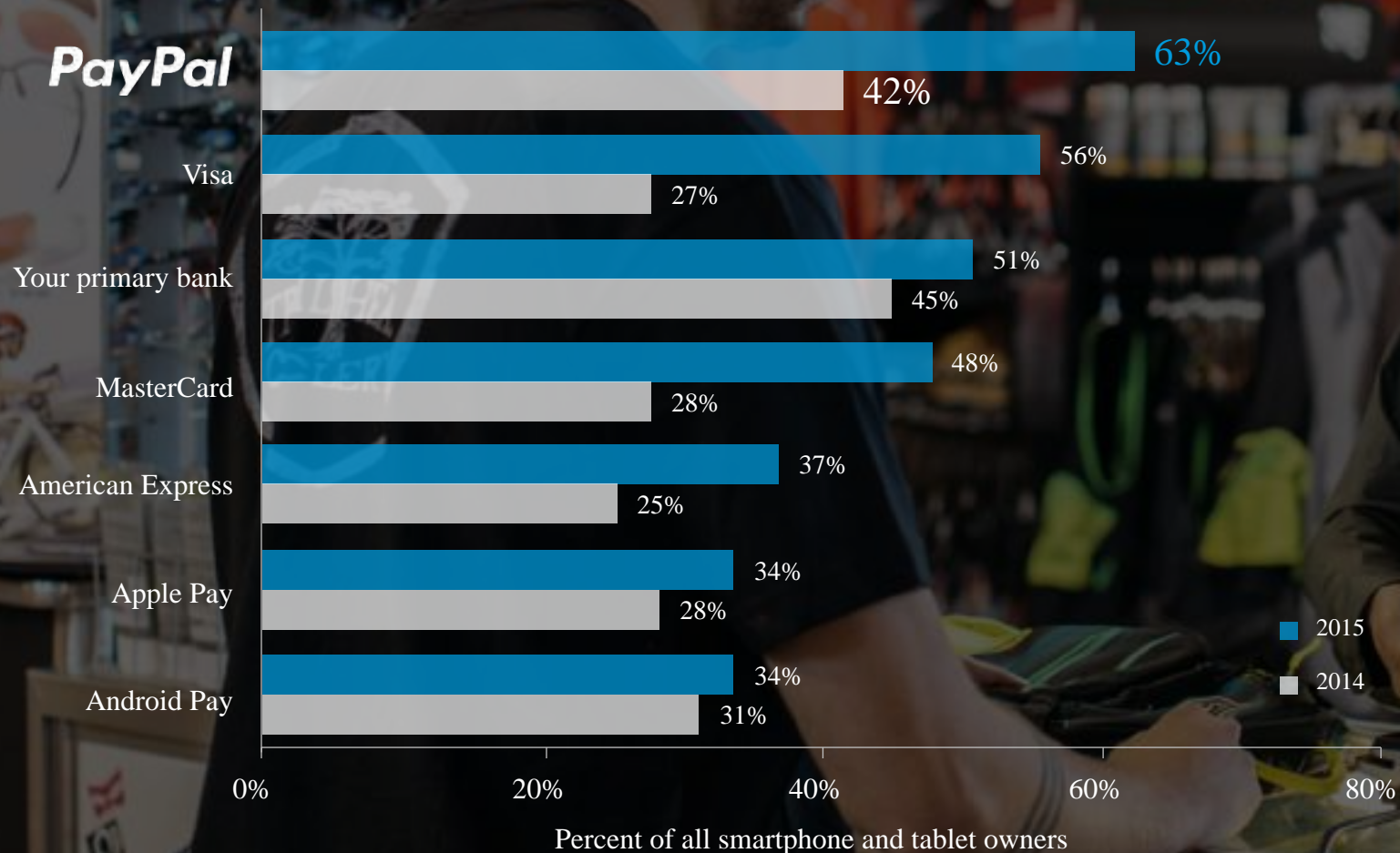


Our technology robust Platform
enables us to:

- Process payments at tremendous scale
- Accelerate the innovation of new products
- Engage world-class developers & technologists

Consumers choose PayPal for trust and protection

According to a Javelin study, 63% of consumers said they preferred PayPal



PayPal

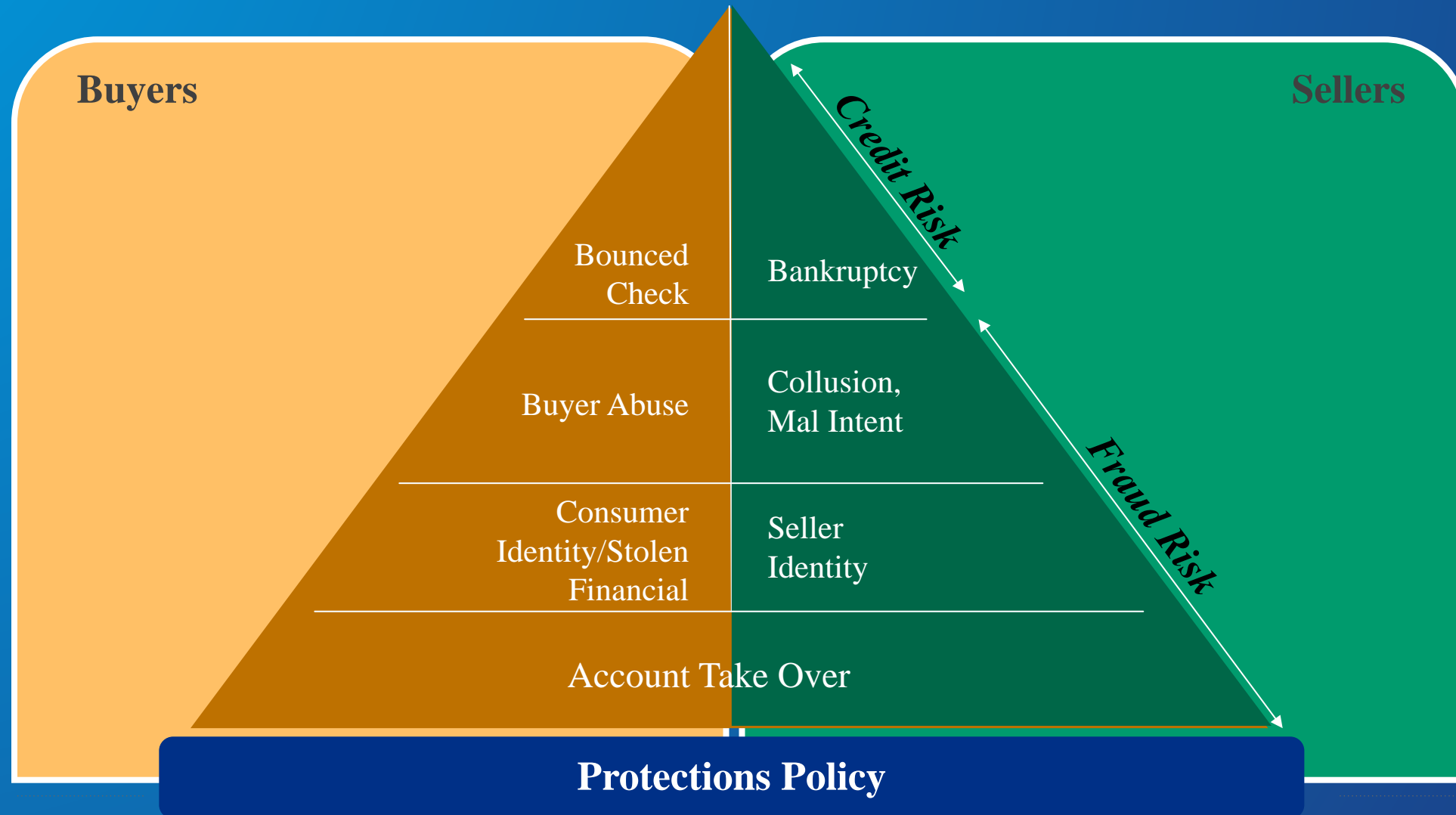
Market Advantages



RISK MANAGEMENT USING AI - BRIDGING TECHNOLOGY AND BUSINESS

Financial Risk Management Expertise

By connecting buyers and sellers, PayPal manages a broad range of risks



Fraud Risk - Identity Theft

Credentials Available Online for a Price

| Credit Cards | Price in 2013 | Price in 2014 | Recent Prices |
|--|---------------|---------------|---------------|
| Visa and MasterCard (U.S.) | \$4 | \$4 | \$7 |
| Visa Classic and MasterCard (U.S.) with Track 1 and Track 2 Data | \$12 | \$12 | \$15 |
| Visa Classic and MasterCard (Canada, Australia, and New Zealand) with Track 1 and Track 2 Data | \$19 – \$20 | \$19 – \$20 | \$25 |
| Visa Classic and MasterCard Standard (EU) with Track 1 and 2 Data | \$28 | \$28 | \$40 |

Hacking Email and Social Media Accounts

| | Recent Prices |
|---|-------------------|
| Popular U.S. Email Accounts (Gmail, Hotmail, Yahoo) | \$129 |
| Popular Russian Email Accounts (Mail.ru, Yandex.ru, and Rambler.ru) | \$65 – \$103 |
| Popular Ukrainian Email Accounts (Ukr.net) | \$129 |
| Popular U.S. Social Media Accounts | \$129 |
| Popular Russian Social Media Accounts (VK.ru and Ok.ru) | \$194 |
| Corporate Email Accounts | \$500 per mailbox |
| IP address of Computer User | \$90 |

Our “Secret Sauce”



**Human approach
to Big Data**



**Award winning home grown
Enterprise Fraud
Management System**



Brilliant brains

More than 200MM accounts

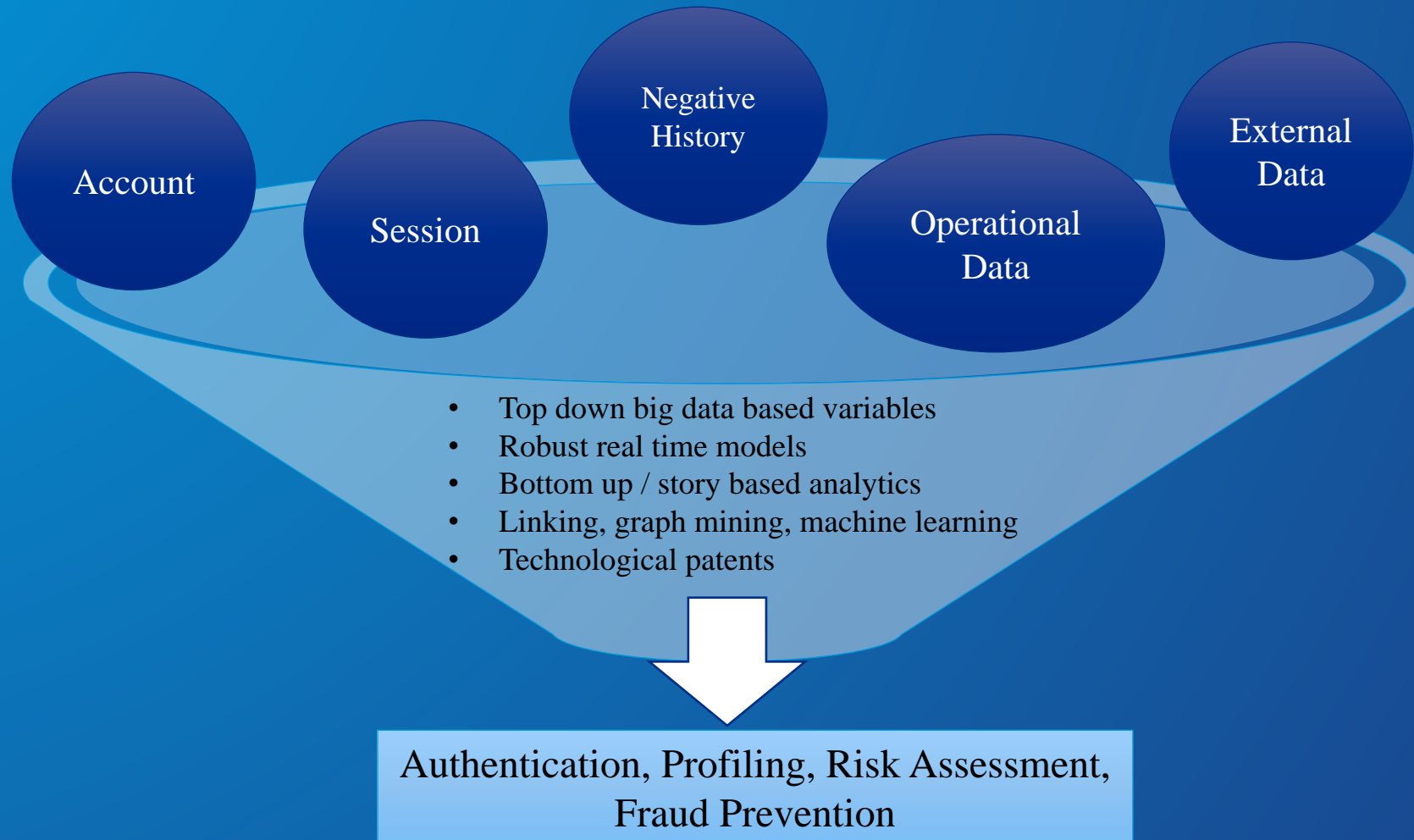
More than 6 Billion
transactions per year

Hundreds of Billion Events per
year

More than 5 Billion unique
entities

Thousands of aggregated
features per entity

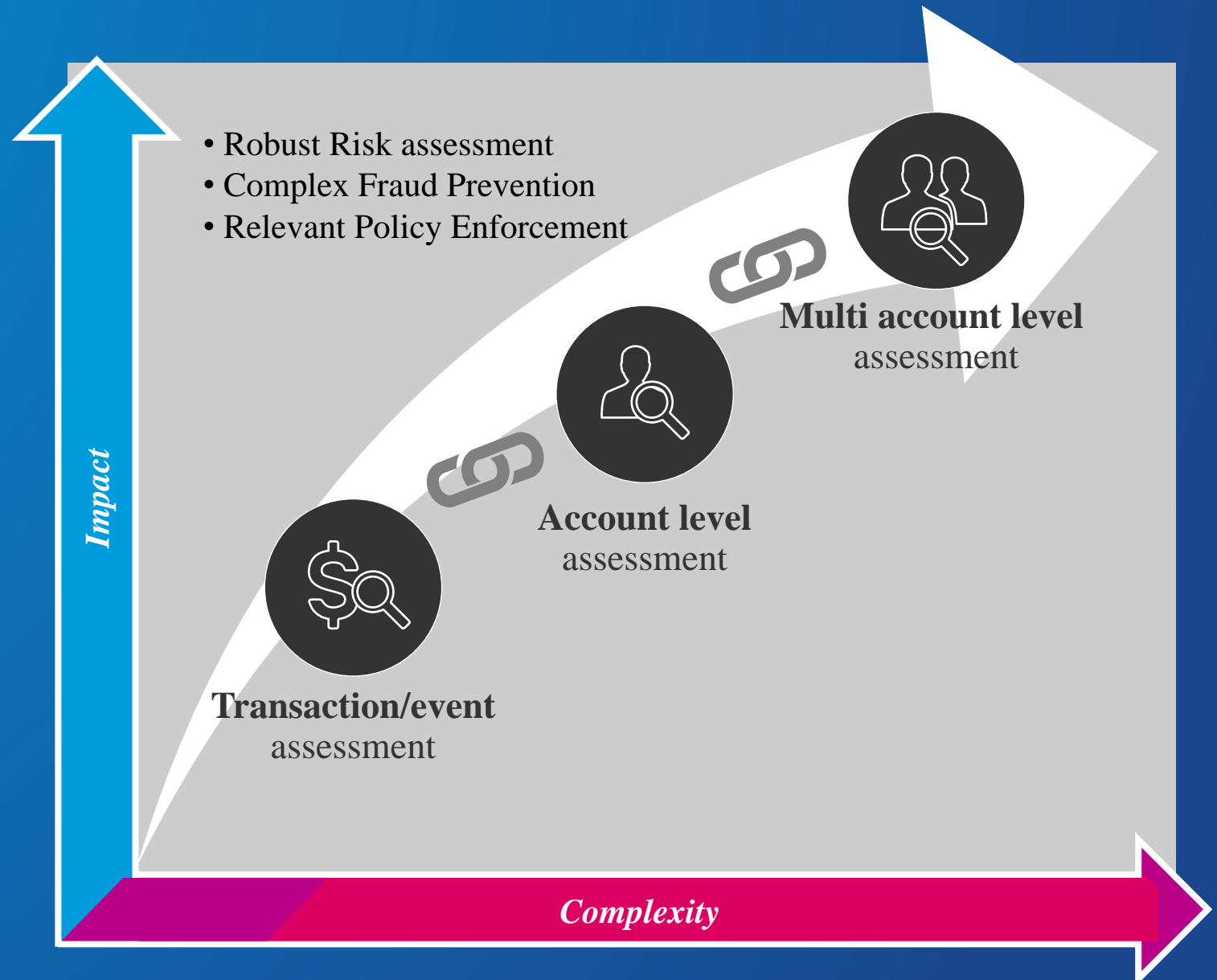
PayPal leverages internal/external data, real-time models, advanced analytics and tech patents for our risk assessment



Home Grown Innovation Helps PayPal Overcome Complexity of Risk Management

We are evaluating multiple “linked” accounts when making our real-time assessments:

So how do we do this?



Risk Management is Complex

Matching users poses a series of analytical and technical challenges

10^{17}

Potential account pairs to compare

Алекс =
Sasha ?

Aliases, Fuzzy matching, Languages,
spelling errors

$A \leftrightarrow B \leftrightarrow C$

Transitive relationship

Story-Based Approach Decreases False Positives

Example – NOTE: Sam is a not a real customer

Large purchase is made from Sam Loe's account. Sam is a U.S. resident living in NY but the transaction attempt originated from an IP address located in Thailand and at 2:00 am NY time.

What should we do?

Decline & Lock Sam's Account?

Common Bad Story

- Statistically most common scenario is that Sam's account has been taken over by a fraudster based in Thailand.

OR

Approve?

Possible Good Stories

- Sam's on vacation at a Thai resort
- He's on a business trip in Thailand
- His parents are from Thailand and he's currently visiting
- Sam serves in the U.S. army and is stationed in Thailand

Story-Based Approach Reduces False Positives

How does PayPal accurately separate good and bad transactions?

1) Ask the right questions:

- Is the transaction attempt consistent with Sam's account history?
- Can the attempt be associated with a trend?
- What do we already know about the IP and device used for this transaction?

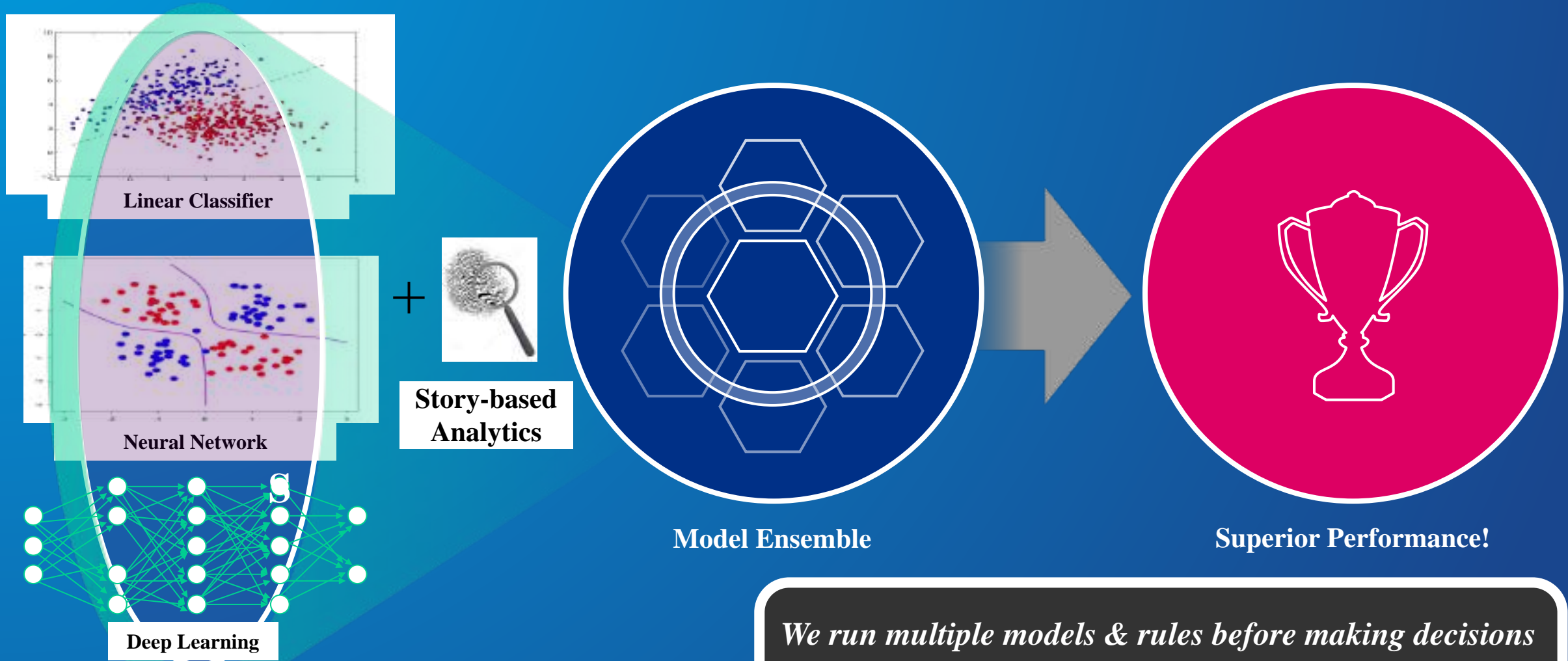
2) Build risk models to ask the right questions:

- Attributes – gather required data top-level Risk models to determine likelihood of someone traveling
- IP Connection Type tagging – story driven tags for IPs
- Seller industry tagging – check for consistency with past transactions
- And focus on understanding and validating important stories

Storybook Ending

In this case, PayPal's risk algorithms determined that Sam was on vacation in Thailand at a resort. Within a few seconds of clicking the payment button, his transaction was approved. Sam is a happy PayPal customer.

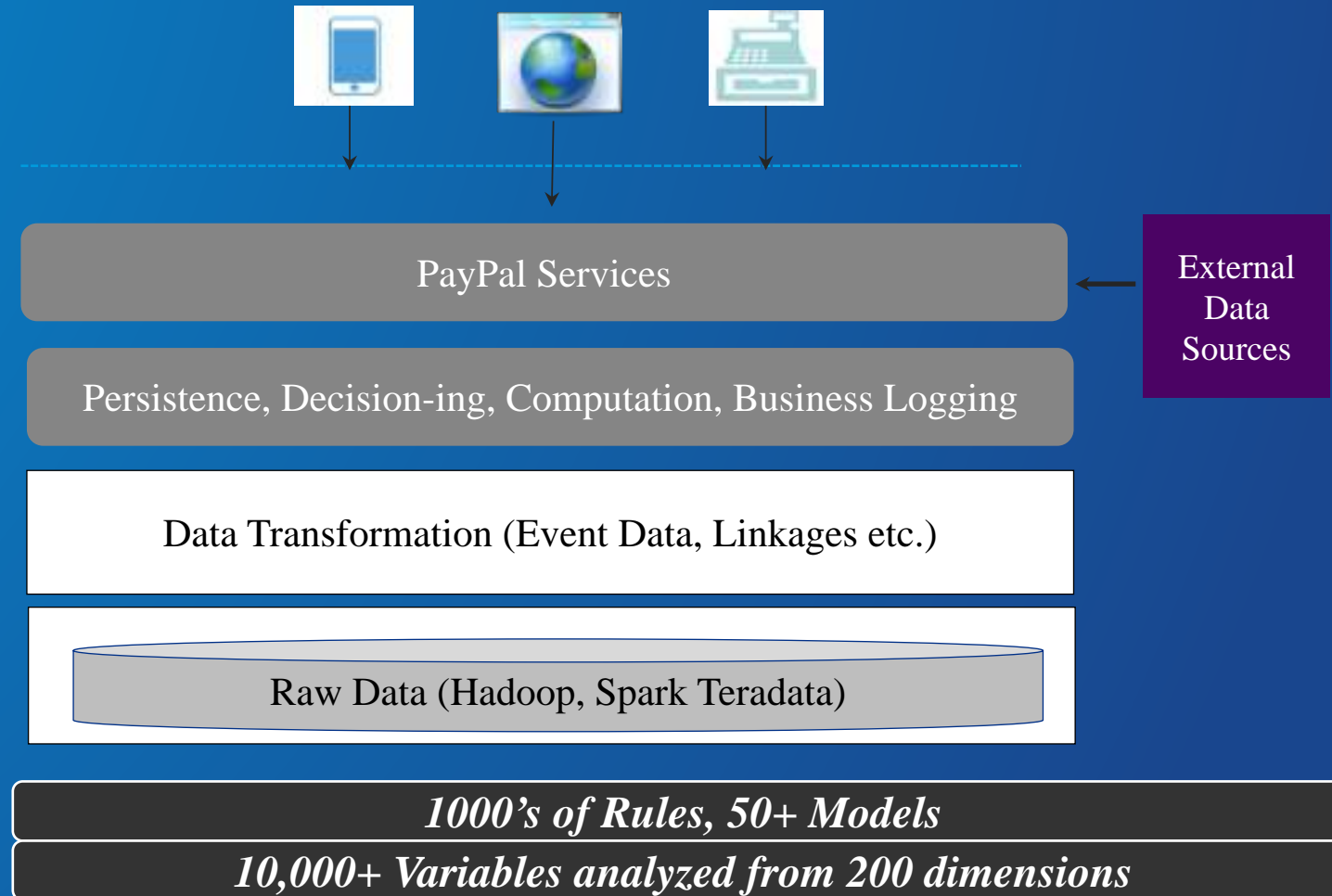
Advanced Machine Learning combined with PayPal Proprietary Story-based Analytics Approach drives Better Risk Decisions



We run multiple models & rules before making decisions on transactions – all within milliseconds (SLA improving constantly)

Risk Decisions in Milliseconds built on Fast, Scalable and Flexible Infrastructure

- Massive Data
 - 1.5 petabyte & growing data
 - >10 years of event histories, profile, third party data
- Daily New Events
 - 1+ billion new events added daily
 - Near real-time streaming to data warehouses



PayPal Security

Consistently ranked among the top in consumer trust & security

Strong Front Door

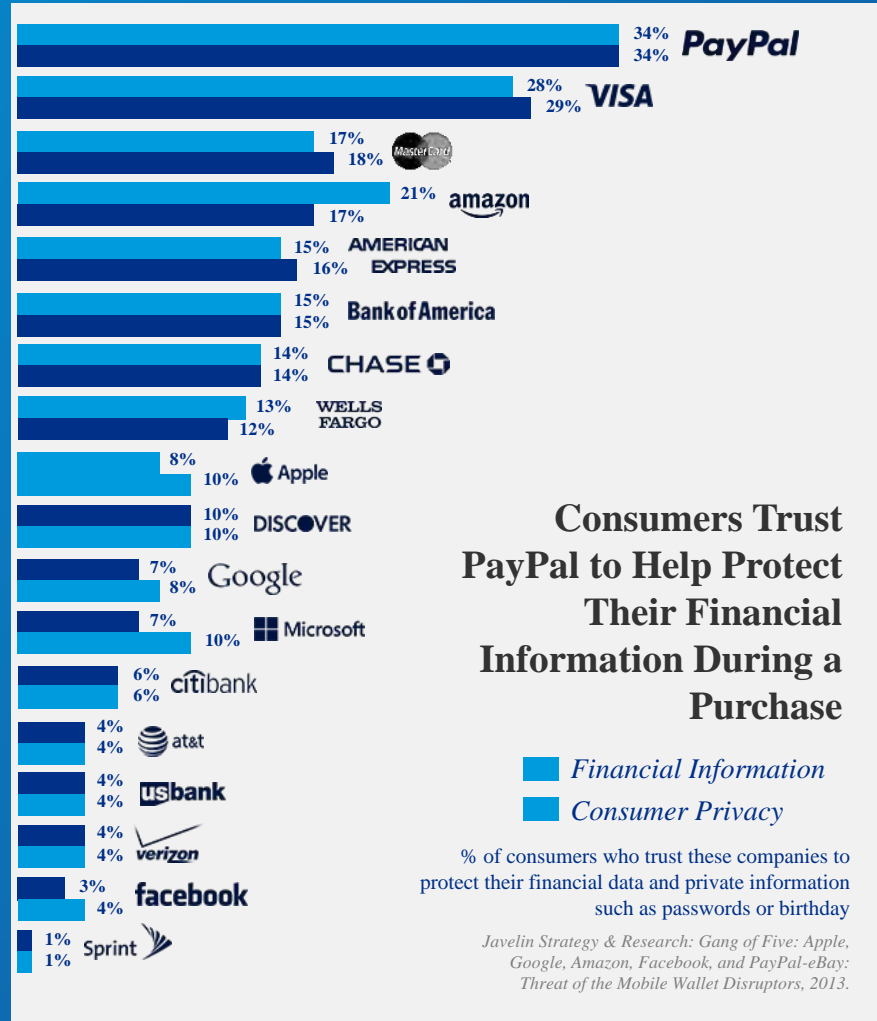
- 0100 11001 Encryption on every PayPal transaction
- XXXXX PayPal does not share financial information with merchants
- PayPal verifies a person's identity for payments

16.7 MILLION PAYMENTS

processed daily by PayPal

Strong Foundation

- 24/7 data analytics combined with human oversight to quickly spot suspicious activity
- Constant innovation to advance our risk management techniques
- Seller and buyer protection offered for eligible transactions

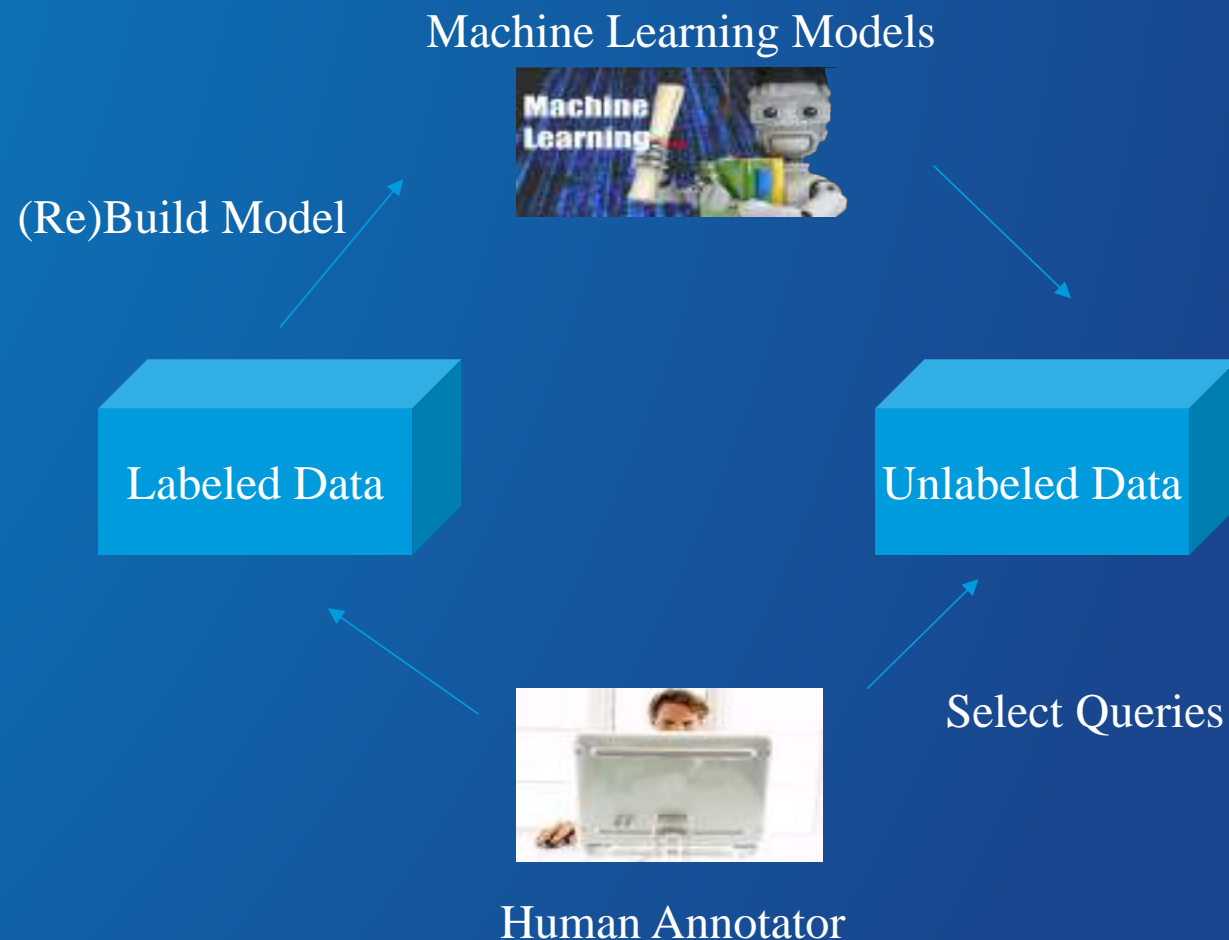


Industry Engagement

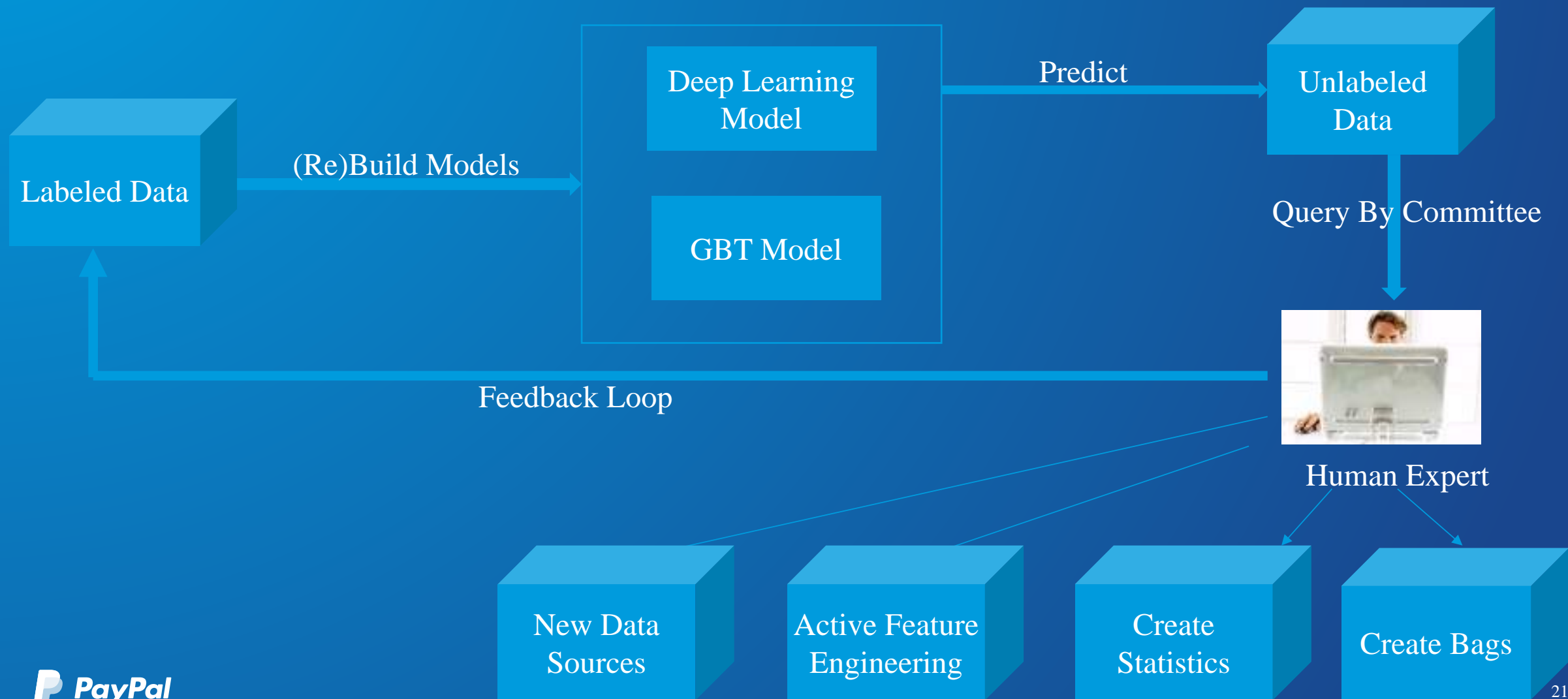
- Founding member of the FIDO alliance
- PayPal chairs the DMARC initiative to reduce phishing attacks against all Internet users
- PayPal has been doing tokenization for 15+ years, securely storing customers' financial information in the cloud.

AI in Action – Active Learning

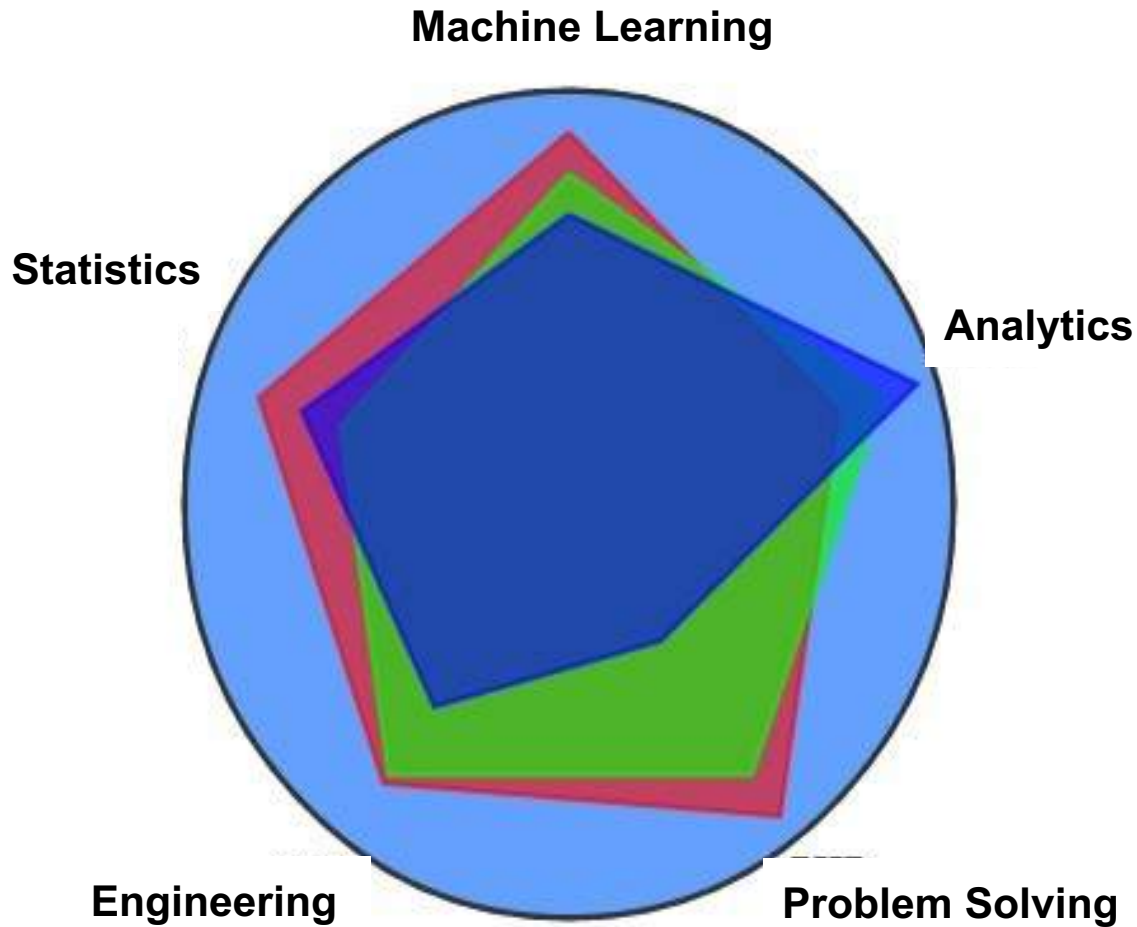
- Supervised learning algorithms require data to be labeled
- Labelling is difficult, time-consuming and expensive : Active Learning to the rescue
- Idea – ML Algorithm can achieve better accuracy if it is allowed to “choose the data” from which it learns*
- Overcome labelling bottleneck by asking queries (unlabeled data) to be labeled by human



AI in Action – Active Learning



World Class Data Science Team



Healthy Team Characteristics:

- Each member has experience in most of the areas, but likely master only one or two domains.
- Mixture of Background:
 - Different industries and disciplines
- Mixture of Tenure:
 - Complement industry experience with the latest from academia
- Work side-by-side to achieve short cycles and facilitate knowledge sharing

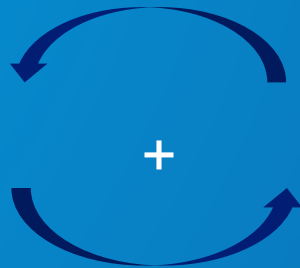
Common Mistakes:

- *Overly invest in a single area*
- *Expect super-heros – everyone can master everything*
- *Team members with orthogonal skills*

... Empowering Effective AI Solution

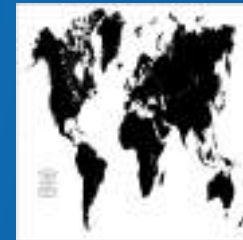


Discover *insights*



Refine till *accurate*

+



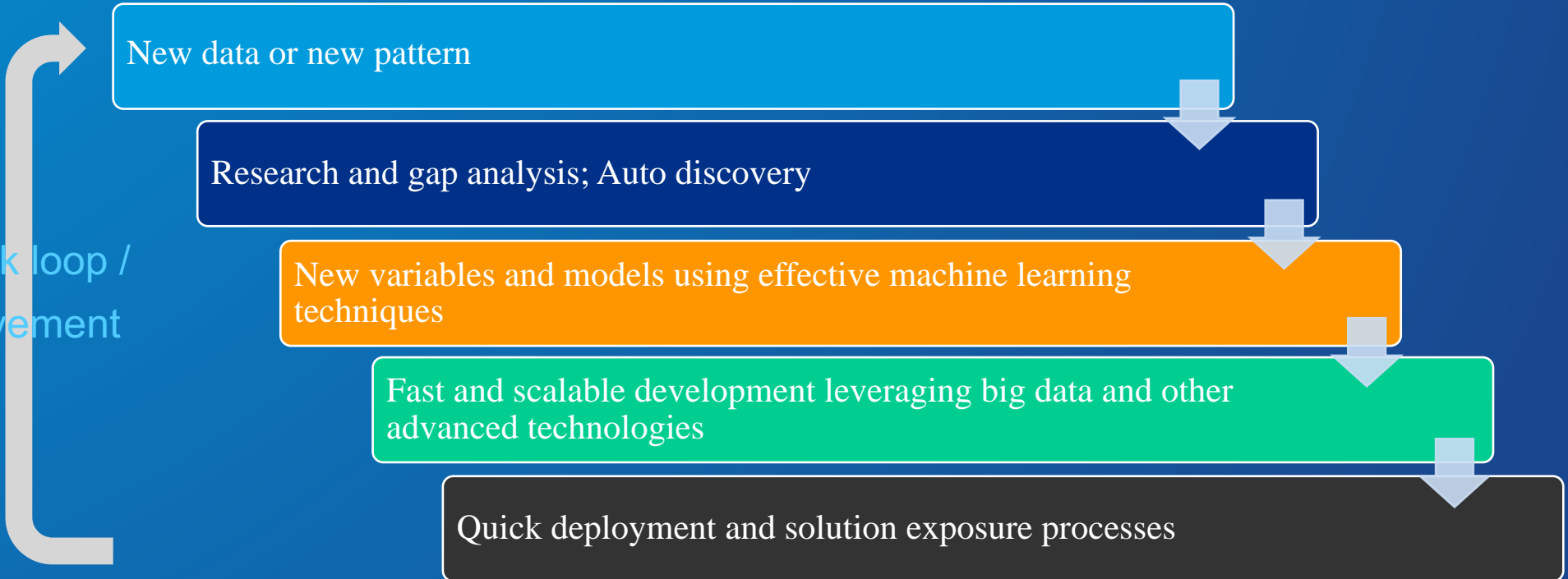
Apply at *scale*

=



Effective *Solution*

Constant Feedback loop /
continuous improvement



THANK YOU



Hui Wang
Sr. Director, Global Risk and Data Science
PayPal
huiwang@paypal.com