

Current solutions compromise signal for privacy



USING SYNTHETIC DATA

Rarely finds anything beyond the rules used for the synthesis



USING AGGREGATED DATA

Severe loss of data richness and granularity

Signal



Privacy

- Aggregated data
- Synthetic data

...or privacy for signal



REMOVAL OF PERSONAL IDENTIFIABLE INFORMATION

PII becomes de-anonymized in the process



HOMOMORPHIC ENCRYPTION

Unable to scale, and at risk to side-channel attacks

Signal

● Anonymized data

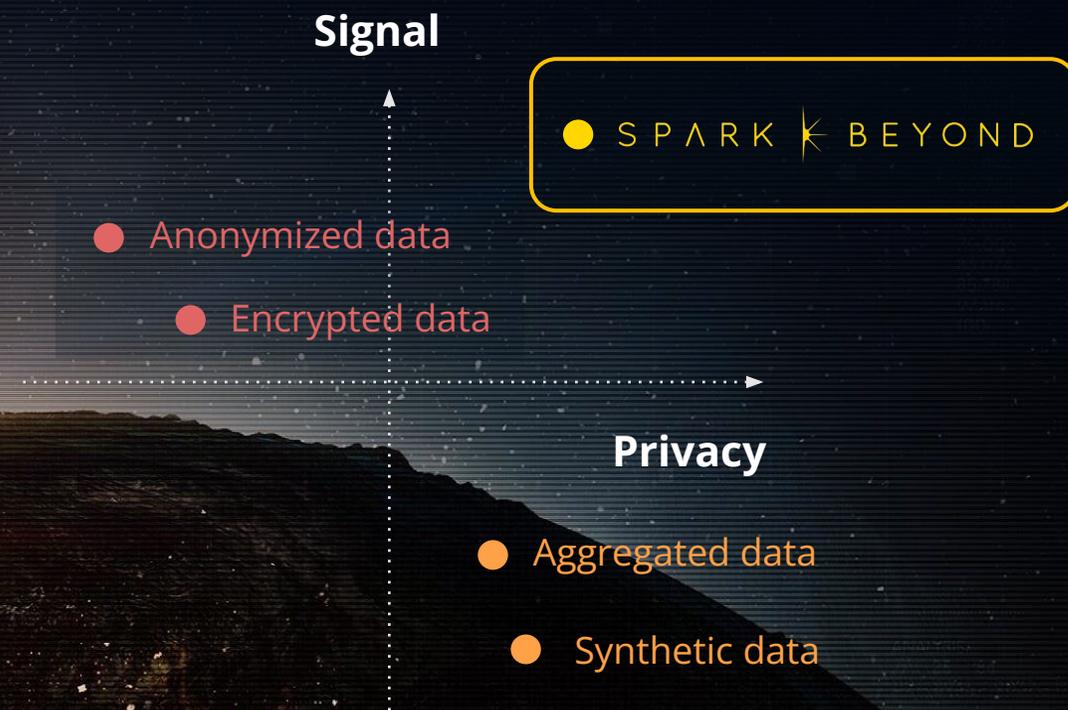
● Encrypted data

Privacy

● Aggregated data

● Synthetic data

SparkBeyond liberates sensitive data, allowing deep analysis and collaboration **without compromising security and privacy.**



HOW IT WORKS

1

Leverage AI to make
pattern & insight
discovery autonomous

The engine autonomously analyzes data to generate insights and composite features, discover patterns, or hook datasets together.

Export | Showing 300 from 400 | Sorted by Lift & RIG

HistoricalAverage of auto_renewal < 0.5
Add description...

| Class | Lift | Support | RIG |
|-------|--------|------------|-------|
| 1 | x2.960 | 17% 14,578 | 0.139 |

Source Group
Payment_Info In the last 10 weeks before prediction_date by subscriber_id

Stddev of days_since_prior_order ≥ 10.277
Add description...

| Class | Lift | Support | RIG |
|-------|--------|------------|-------|
| 1 | x1.715 | 32% 27,744 | 0.050 |

Source Group
Order_Histo... In the last 1 year before prediction_date by subscriber_id

HOW IT WORKS

2

Circumvent exposure of humans to sensitive information

Row-level information can be hidden from certain datasets, while still providing view-only access to the data schema.



HOW IT WORKS

3

Analyze sensitive data without diluting its value

Work with unaggregated and unprocessed data as AI platforms sit between data providers and users, surfacing patterns from merged, "unseen" data sets.

| | | |
|---|------------------|---|
| historicalAverage(auto_renewal of subscriber_id over all weeks before prediction_date) | x2.96 for churn: | 1 |
| stddev(days_since_prior_order of subscriber_id in the last 1 year before prediction_date) | x1.72 for churn: | 1 |
| Females-30-34 of zipcode of subscriber_id ≥ 893 | x1.57 for churn: | 1 |
| sumOfValues(days_since_first_order of subscriber_id in the last 5 weeks before prediction_date) Many events of days_since_first_order in the last 5 weeks before prediction_date | x1.10 for churn: | 0 |
| sumOfValues(order_number of subscriber_id in the last 5 weeks before prediction_date) Few events of order_number in the last 5 weeks before prediction_date | x1.57 for churn: | 1 |
| minSlope(order_number of subscriber_id in the last 1 year before prediction_date) An increase in order_number in the last 1 year before prediction_date | x1.11 for churn: | 0 |
| average(lookupFromTimeSeries(lookupFromTimeSeries(order_id of subscriber_id in the last 1 year before prediction_date))) | x2.37 for churn: | 1 |

EXAMPLE // Unlocking new value through shared data



CASE STUDY // Data partnership unlocks millions in sales

THE CHALLENGE

Top global CPG wanted to prioritize limited sales capacity, empowering their reps with dynamic trade insights to share with retailers.

A data partnership with retailers would help them discover brand-level consumer micro-segments.

THE APPROACH

SparkBeyond's Blindfolded Analytics analyzed both the CPG's and retailer's confidential and unshareable data to reveal granular drivers & micro-segments without exposing analysts to raw data.

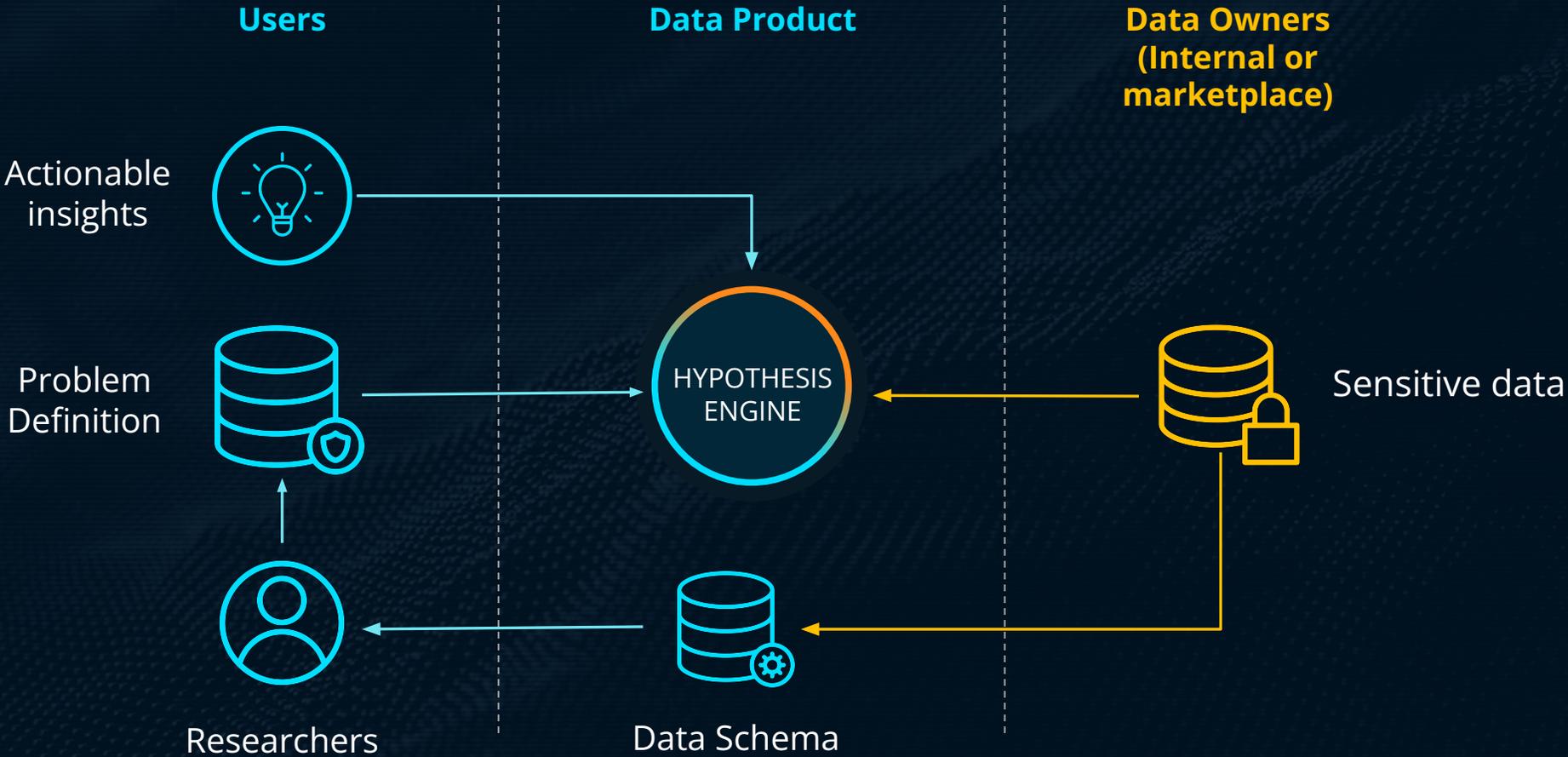
RESULTS

Renewed partnership between global CPG and EPOS data vendor by facilitating analytics on daily-level category & brand sales data.

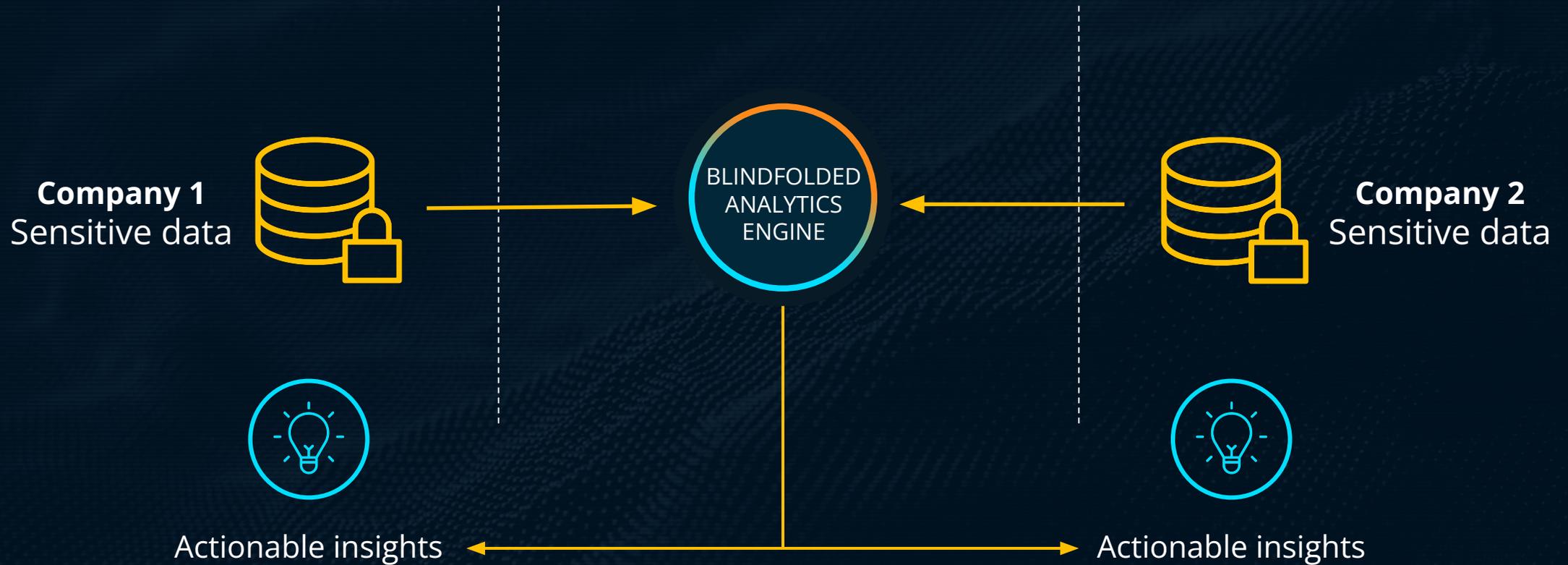
SparkBeyond uncovered brand and geo-specific drivers to inform GTM (patterns are dynamically updated at the pace of change), unlocking over \$5M in sales in 5 months.

- ✓ **Facilitated analytics on daily-level category & brand sales data**
- ✓ **Uncovered brand and geo-specific drivers to inform GTM**
- ✓ **Unlocked >\$5M in sales in 5 months**

ARCHITECTURE // Single-company scenario



ARCHITECTURE // Multi-company collaborative partnership





Under the Hood

THE SPARKBEYOND HYPOTHESIS ENGINE

Creative AI designed for impact

Insight Discovery at Enterprise Scale

Connects the dots between disparate events and data sources, discovering patterns and correlations in complex data

Made for a Dynamic World

Adapts insights and models to ever changing market conditions to ensure lasting predictive signal and accuracy

Expanded Data Universe

Augments internal data with millions of external data inputs to uncover 'unknown unknowns'

Human-Centric and Explainable

Embeds trust, transparency and ownership in analytics and AI outputs with end-to-end visibility



Unique solutions to ubiquitous challenges



AI-GENERATED INSIGHTS

Continuously explore millions of insight hypotheses on complex data ecosystems, surfacing high impact ideation



DEEP SEGMENTATION

Build hyper-granular audience and SKU segments, powering nuanced interventions and personalization across all channels



BLINDFOLDED ANALYTICS

Unlock analysis and collaboration opportunities for sensitive, confidential or un-shareable data



FEATURE DISCOVERY

Automatically build, test, and select composite features for Auto ML that drive business outcomes



SCORING & PRIORITIZATION

Automated data sharding enables large batch scoring at scale (billions of records via hundreds of models) and fast validation



OPERATIONALIZATION

Low-code deployment environment with no CI/CD required between research environment and production environment

