# Not All Data Is Created Equal

The Science of Ensuring Quality and Transparency
in Consumer Behavior Data

Executive Summary

The burgeoning consumer behavior data intelligence market holds great promise for decision-makers and technical teams looking for solutions to help them make more accurate decisions. This is especially true in high-stakes areas such as selecting new store locations, betting marketing budgets on consumer segments, or investing in new commercial endeavors where minimizing risk is just as important as maximizing profit. This data market is relatively new; data sourcing and preparation is complex, and most vendors are a black box in terms of their processes – therefore, customers are left on their own to figure out who has the best mousetrap.
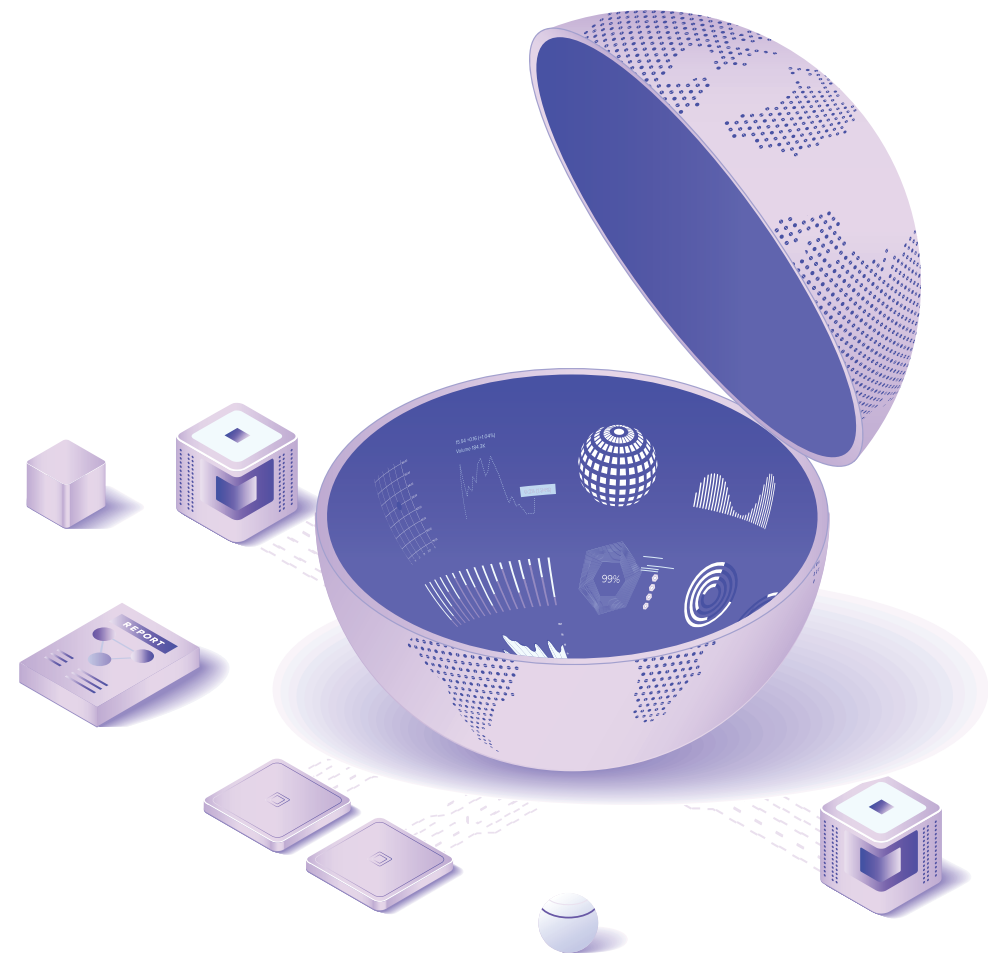
Many data leaders and decision-makers are lured by promises of the largest data sets and start working with new data types they don't really understand, only to discover that the black box they were sold doesn't meet their needs. They find out the data isn't reliable, it isn't delivered in the most usable formats, and the correlations between raw data and insights are difficult to make. Eventually, they lose trust. And since the ROI hasn't materialized, many enterprises grow skeptical of working with or onboarding large, new datasets such as consumer behavior data considering the time they have already invested in similar efforts with less-than-optimal results. This frustration isn't surprising; in a recent survey, Gartner revealed that almost 50% of digital workers struggle to find information needed to perform their jobs effectively[1]. This problem will only compound as in 2023, we will generate three times the volume of data generated in 2019.[2]

What you're not told when you start talking to vendors about consumer behavior data is that not all data is created equally. More specifically, few providers bring the necessary experience, business rules, and advanced data science needed to process, screen, model, aggregate, score, and interpret the data so it reliably delivers real business value.

That is why Azira's Consumer Behavior Data Intelligence platform was built on three pillars:

Global, Quality Data:
Ensure the highest quality data is used to fuel and train the models that convert raw data into insights.

Privacy and Transparency:
Develop actionable insights at any level of granularity without compromising consumer privacy.

Data Science:
Use advanced techniques to detect anomalies and uncover hidden trends that can fuel analyses and sound business decisions.

For business leaders, these pillars deliver confidence in their decisions, including identifying the optimal location for a new store, understanding the right tenant mix for properties and knowing who and when to target for a new marketing campaign. This confidence is achieved through thorough understanding: linking profiles across physical and digital spaces to provide decision-makers with deep insights about the places their customers frequent, as well as their brand and competitor preferences.

Overview: Consumer Behavior Data Defined

Consumer behavior data delivers a powerful competitive advantage to businesses looking to understand patterns through real-world signals such as foot traffic, traffic hours, demographics, online behaviors, lifestyle, and brand affinity. Companies can use this information to segment, target, and better engage consumers across many categories. They can develop a deep understanding of how customers move through the physical world to shop, dine, travel, or simply use the services and amenities within their communities.

Informing high-cost, high-risk business decisions with quality, up-to-date data is critical. At the same time, privacy concerns are increasing worldwide, as are regulations that restrict the collection and use of data. And although these two requirements — access to more granular, high-volume data while adhering to tight privacy standards — seem like they're mutually exclusive and cannot coexist, there is a way. It is possible to have data that respects privacy, anonymizes individuals, adheres to and surpasses government regulations, and yet still delivers the business insights needed to answer pressing business questions such as:

Technical and Operational:

- Can Azira's consumer insights help with staffing, inventory, or merchandising decisions?

- How does my small team begin sifting through all this aggregated data to extract insights?

- How have trade areas, visitor/resident travel behaviors and consumer shopping patterns changed?

Marketing and Audience:

- How do I discover audiences that are in-market for products and services like mine?

- Where do people shop today, what stores do they visit, and what are they buying?

- How do I build better campaigns that drive real-world results, like store visits?

Executive Leadership:

- How are my competitors doing?

- How can I leverage what I know about consumers to increase investor confidence?

- How can we deliver the best experience to our customers and tenants?

Market Context and Data Challenges

According to Statistica, there will be 15 billion connected devices by the end of 2023[3], and IDC predicts that this number will skyrocket to 55.9 billion connected devices by 2025 (150 billion if you count RFID tags)[4]. Given those statistics, it's no surprise that 97% of executives in Fortune 1000 organizations[5] say they are investing heavily in data initiatives[5] – but the path to becoming a data-driven company is far from straightforward.

Transparency and Compliance Pressures Are On the Rise

Although most leaders widely recognize data's importance in driving key business decisions, many struggle with internal priorities and challenges to effectively transform their operations. In a recent 2023 survey by Azira and Hanover Research of 590 global data and operational leaders working in the retail, restaurant, commercial real estate, and travel & hospitality industries, findings revealed three main themes that underscore Azira's mission around Data Quality, Privacy, and Accuracy. Of respondents:

>90%

Said ensuring data accuracy was a top priority for their organization.

>83%

Said they were worried about poor or unreliable data quality.

>77%

Said compliance with privacy regulations (GDPR, CCPA, Australian Privacy Principles) was a top priority.

While respondents clearly valued data and its importance to the organization, many are still falling short when it comes to applying it:

59% A majority (59%+) said changes in consumer behavior have caused their organization to rethink business strategy.

40% 40% said their teams have missed opportunities due to missing data.

With digital products and connected devices becoming almost ubiquitous, data collection on users is surging, causing consumer concerns about their information and trust issues in the data industry, especially in light of a handful of bad actors.

Data is Challenging Data Teams

To understand consumers today, data teams need to have visibility into who they are, where they're going, and the products and brands they're interested in. Most companies have limited access to such extensive data and struggle with one or more of the following challenges:

Third-party data may vary in quality:

External data providers often collect incomplete or incorrect data. Purchased lists may be broad and contain errors, while devices and surveys can provide misleading information. In addition, many consumers remain "invisible" due to privacy preferences.

Data is not privacy-focused:

Collecting data without transparency raises ethical issues and harms businesses with inaccurate data. Consumers often turn off location settings and use ad blockers, excluding vital demographics. Relying on platform holders' changing privacy practices, like Apple's IDFA, leads to outdated information. The best third-party data partners utilize explicit opt-ins, create consistent identities, and respect privacy and privacy-conscious consumers.

Data is accurate but incomplete:

Understanding today's complex marketplace requires data that represents individuals' behaviors, not just their characteristics. While the US Census offers demographic insights, it doesn't reveal consumer brand interests or location viability for new businesses. To obtain these answers, Census data is necessary, but additional information is also needed.

First-party data has its limits:

It only captures the touchpoints between a business and its clients' physical or digital boundaries. It does not contain details about competitors or other customer affinities and preferences.

Identities are fractured:

Consumers have multiple devices and/or profiles, with data spread across various data sets. This proliferation leads to difficulties reconciling digital identities for analysis or targeting.

Lack of Data Usability Hinders Productivity

Data complexity isn't the only challenge for data teams. A recent MIT Management Sloan School study discovered that business leaders feel like they are effectively using data to run their businesses because they've hired data scientists. The truth is these internal data teams are bogged down by tasks that don't add value to the business and come with many direct and indirect costs.[7] They spend:

be significant in terms of time, accuracy, and business performance. A company may waste its marketing budget on irrelevant audiences and spend additional resources on rectifying issues. They may also forfeit revenue that could have been generated by making better-informed decisions.

90%
of their time working on data discovery, integration, and cleaning.

90%
(of the 10% that is left) to fix data-cleaning errors.

1%
of their time on actual analysis and data science.

Ultimately, the opportunity cost for a company selecting an inadequate consumer behavior data intelligence platform can

The Azira Approach

Bad data costs U.S. businesses $3 trillion. Not only that, but researchers predict that up to 40% of business objectives fail due to inaccurate data leading to more lost revenue [8]. In fact, according to Gartner, the average financial impact of poor data quality on organizations is $9.7 - $15 million per year[9], and this only includes costs that can be measured like time spent on data quality, resource re-allocation etc. They don't include the unknown opportunity costs that are hard or virtually impossible to measure.

A Robust Consumer Behavior Data Intelligence Platform

Since 2016, Azira has built a rich history of working on data privacy and compliance. It has taken rigorous steps to align with global regulatory frameworks. All products are built with privacy-led design principles, and Azira focuses on unleashing the full potential of consumer behavioral insights through its commitment to innovation, use of AI and advanced data science techniques, and unwavering focus on privacy - all of which help its users drive better decisions and outcomes.
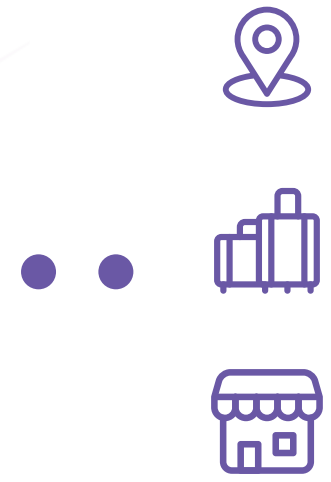
Multiple Sources of Raw Data
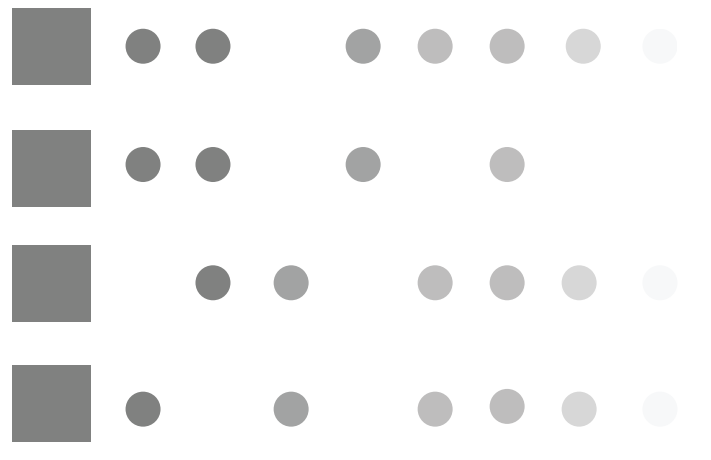
Processing and Contextualization
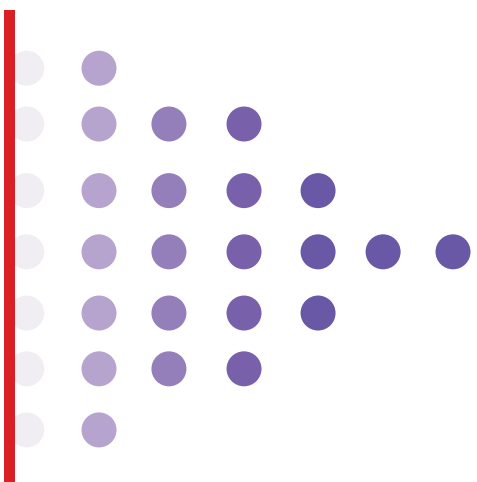
Refined Data Supply

Consumer Behavior Data

The Azira Platform never stores or deals with Personally Identifiable Information (PII). All data leveraged by Azira comes with consumer consent. The Azira platform has built-in processes to forget and purge user data upon request.

Consumer Behavior Data Platform Checklist

Evaluation Criteria

With respect to data, look for partners who:

☐ Focus on data quality assurance achieved by data preparation, validation, cleansing, bias analysis, quality and monitoring.

☐ Ensure data sourcing is both global and compliant, and has the market reach and richness needed to deliver results.

☐ Have access to robust global and regional datasets as your markets grow and needs expand.

☐ Provide insights on consumer behavior that are actionable for both operations and marketing.

☐ Deliver consumer insights using data enrichment, stitching and identity resolution.

☐ Can deliver data in various formats accompanied by quality monitoring.

☐ Are equipped to handle the scale needed to process geospatial data.

☐ Source data across multiple industries and markets.

Ensure privacy and transparency is not an oversight. Work with partners who:

- [ ] Make global privacy and compliance adherence to new regulations a top priority.

- [ ] Allow users to have control over their information regarding their data and privacy.

- [ ] Adhere to international data standards when data is transported internationally.

- [ ] Make dedicated investments in machine learning and artificial intelligence to help safeguard user privacy.

- [ ] Are transparent about where the data comes from, how it has been processed and how it is used.

Leverage advanced data science to filter out bad data and deliver deeper insights. Look for partners w

- [ ] Have a clear strategy around the central role of AI and ML and consumer behavior insights.

- [ ] Use proprietary algorithms to generate advanced insights based on existing data sets.

- [ ] Use natural language processing technology to access and interpret audiences, segments and markets.

- [ ] Implement automated anomaly detection and other advanced methods to ensure the accuracy and reliability of the underlying data.

Azira: A Foundation Built on Global Data, Privacy and Data Science

Let's take a closer look at the cornerstones of Azira's Consumer Behavior Data Platform, specifically the three foundational pillars (Global, Quality Data, Privacy and Data Science) which drive Azira's vision.
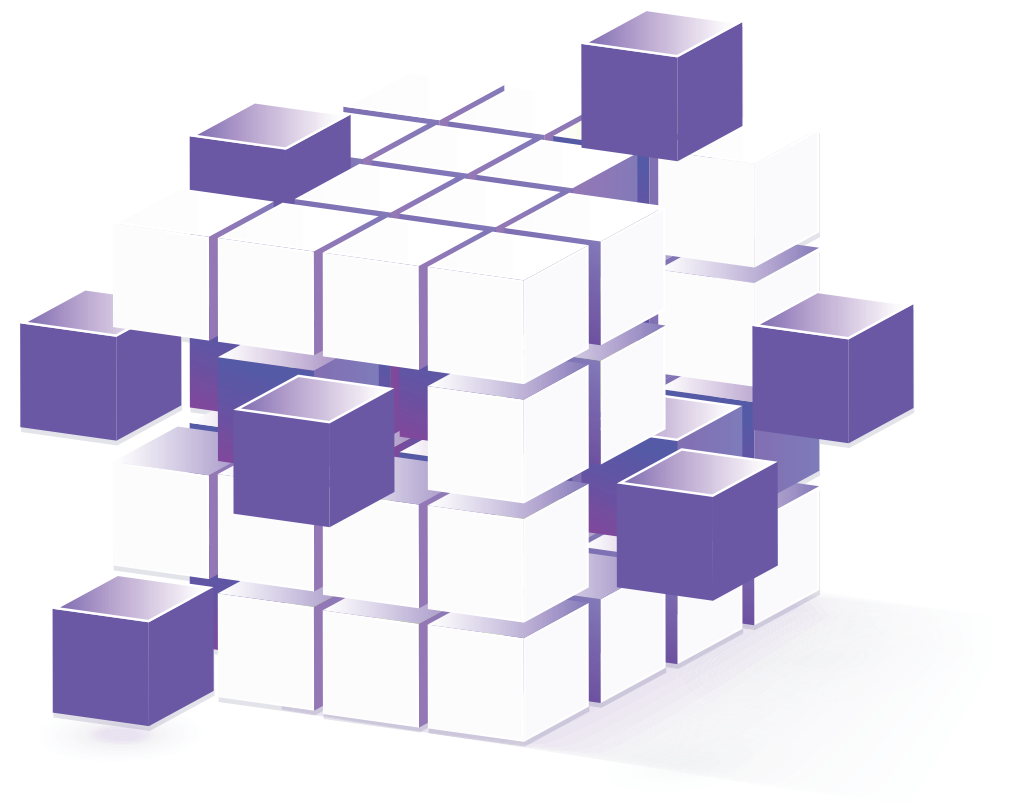
Foundational Pillar #1:
Global, Quality Data

Azira specializes in providing high-quality insights about people and places for numerous industries and applications. Like crude oil, most raw data requires a tremendous amount of refinement to become useful. Azira has built complex processes to convert the raw data inputs to jet fuel that generates actionable insights for customers.

Data Delivery

Transparency:

Many data partners act as a "black box": they deliver finished dashboards but do not give the user access to view, slice, dice and perform their own analysis on the underlying data. Azira's approach to data delivery is fundamentally different:

- Empowering clients to work in the analytics environment of their choice

- Allowing users of all skill levels to unlock powerful insights from the data, from line-of-business users to technical analysts and data scientists

Identity Resolution:

Consumer interaction with the real and digital world can create multiple identities. Using AI, Azira's proprietary, persistent identity solution integrates user identities into a single identifier (U.S. Patent No. 10,979,848). With it, businesses can obtain a single view of customers, remove the redundancy in the data collected, and unlock the value of other third-party data sets.

Data Formats:

To accommodate a large range of client requirements, Azira can ingest and output location data in various formats, such as GeoJSON, KML, and Esri shapefiles, in addition to providing flexible delivery options, which include Azira's own intuitive UI (with maps, charts and graphs), direct API access, pushed data feeds, and bulk downloads. All event data is available in common tabular formats (CSV, TSV, etc.).

Data Enrichment

Azira enriches its mobile device data with location contextualization, demographics and audience affinity. This additional color enables customers to create custom audience segments for marketing, understand propensities around shopping behavior, and conduct many other types of industry-specific analyses.

Data Scale

Global Mobile Data Collection:

As an advertising demand-side platform (DSP), Azira participates in ad exchanges, allowing it to collect location data in the process of displaying mobile ads. Additionally, Azira has partnered with other location-collecting SDK providers in order to increase its reach of mobile device users. Combined, Azira collects information on 1.6B people worldwide from over 100,000 apps. Unlike most location data providers, Azira's datasets reflect global coverage of mobile devices, representing sizeable daily activity from over 44 countries and growing.

Highly Scalable Infrastructure:

The Azira platform employs a distributed computing framework in AWS to manage geospatial data at scale. Proprietary compression and indexing technology allows Azira to process more than 10TB of data daily and work with millions of geofences. Azira also leverages computing technologies like Spark to bring speed and efficiency into analytics and data science workloads.

Data Pipeline

Data Cleansing and Validation:

Mobile device data is notoriously "dirty" in its raw form. All data ingested by Azira undergoes a thorough, automated cleansing process to filter out duplicate data points, fraudulent data, anomalies, and other irregular data. As this step is crucial to maintaining data quality and building a reliable foundation for deriving insights, Azira regularly evaluates and updates its filters to prevent new irregularities from being introduced to the data.

Data Integrity and Monitoring:

Azira employs a customized data monitoring platform to automatically identify any changes in data volume and other attributes that may impact data quality. In addition, there are various measures to address any ongoing challenges in the mobile data ecosystem, which include conducting regular audits and reviewing data quality reports with partners.
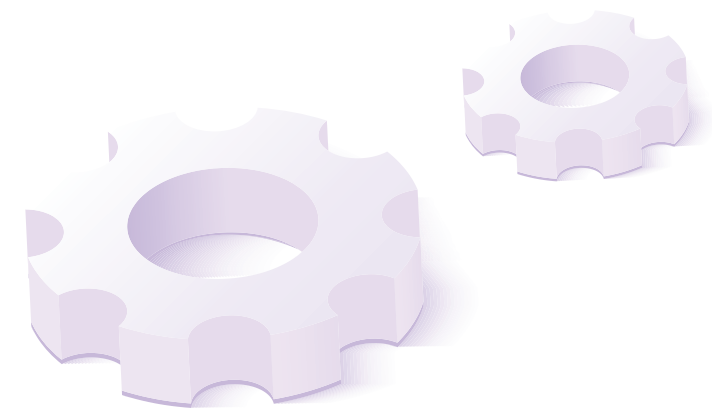
Foundational Pillar #2:
Privacy and Transparency

At Azira, data privacy and transparency are not just buzzwords – they are deeply ingrained in the DNA of the company. The Azira platform is built on the principle of privacy-by-design, ensuring that data privacy considerations are integrated into every aspect of our products, processes and practices.

Compliant with Global Privacy Regulations:

Azira monitors emerging privacy requirements in all markets to stay compliant with the most stringent global privacy regulations. Additionally, Azira is transparent about transferring personal information across national borders and to other countries where Azira and its partners operate.

Customer Data Use:

Every use case undergoes rigorous vetting before a customer receives access to Azira data. Azira will terminate a customer upon discovery of any misuse, dishonesty or illegal practices.

Focus on Privacy by Design:

Azira focuses on the following key privacy-by-design principles to ensure value and trust, including:

- Adherence to international data protection standards.

- Hashing mobile device identifiers (MAIDs).

- Managing end-user preferences to facilitate consent, data sharing, and data usage permissions, including the deleting or exclusion of data associated with their identifiers.

- Securely managing sensitive personal data so that no re-identification of individuals is possible. This includes the use of its patented AziraID for audience targeting, strict data retention policies, and selective storage of device attributes to ensure that no re-identification is possible when records are joined with publicly available data.

- Deploying encryption systems that meet current standards (such as FIPS) to minimize data loss in the case of a cyberattack. Data is secured at rest and in motion to avoid tampering and misuse.

- Obfuscation of home and work locations to further protect the privacy of individuals. Azira adds jitter to the GPS coordinates. Additionally, this information is removed for devices from any country covered by GDPR.

Data Security Policies and Procedures:

Azira's secure infrastructure and tight access controls provide an added layer of defense for its systems. Azira also conducts periodic penetration and breach tests and engages white hat hackers to keep its systems robust. Software-based data security techniques coupled with regular training and audits to enforce policies and procedures ensure a safe and secure data environment.

Empowering Users with Transparency and Control:

Azira provides users with full control over their information, outlining their choices and rights so they can make informed decisions about their privacy. The Azira privacy policy provides detailed information about how the company collects, uses, shares, and protects personal data.

Data Ethics and Responsible Use:

Azira is deeply invested in the ethical implications of its data insights. Customers are prohibited from using location data to target vulnerable or sensitive communities. Azira does not acquire or offer data on certain sensitive locations, and all customer geofencing requests are vetted to ensure that they do not violate Azira's data policy or any local laws. At every step, Azira actively manages the potential risks of sharing its data with bad actors.

Trust, Verification and Privacy in Partnerships:

Azira ensures its data vendors get explicit permission from people whose data is collected and shared. Azira only uses data for which informed consent has been obtained, and the individual has full control to request deletion or deny permission to use their information under certain contexts.
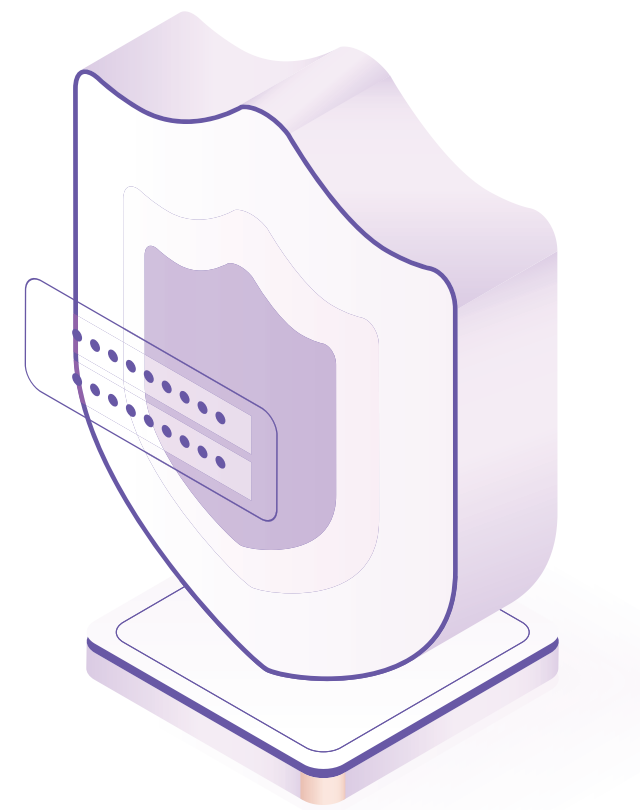
Ask your data provider:

Do they partner with the right companies?

Consumers concerned with protecting their privacy do not differentiate between partners, vendors, or clients. Everyone in the ecosystem is responsible for data stewardship, and every company needs to be a responsible data partner. This means partnerships need to be well-vetted and built on trust.

Is compliance a top priority?

Many companies are not meeting basic compliance requirements, which could have legal implications for their customers. Compliance with constantly changing privacy laws requires a framework. Vendors should be proactively reviewing all internal data use and taking a closer look at vendor and partner data privacy and compliance practices.

What security and internal data handling processes are in place?

Companies must ensure secure and fool-proof processes not only for internal data handling but also to vet the data and related compliance and privacy processes from any third parties.

What investments are they making in privacy beyond compliance?

Compliance with privacy laws is table stakes. Consumers today demand safety and transparency. They want to know how their information is being used and have the right to request to delete their data. The process should be easy and straightforward. What steps are they taking toward anomaly and fraud detection? Are they investing in the latest AI technologies to constantly improve their models and accuracy?
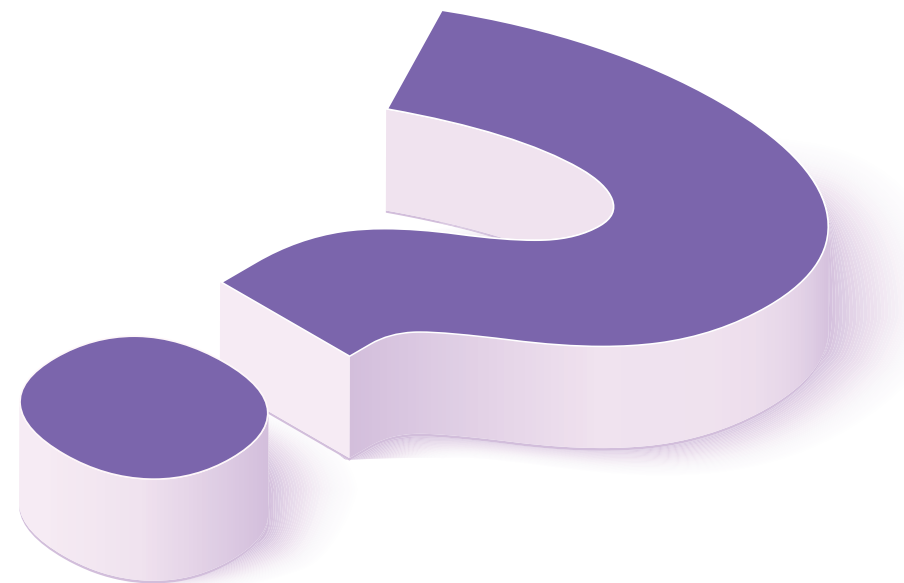
Foundational Pillar #3:
Data Science

Successful organizations depend on accurate and specific insights to inform their decisions. However, the value of these insights is only as good as the data they are based upon. Too often, lower-quality data providers and mobile apps drown data feeds with large amounts of inaccurate, fraudulent or anomalous data. The business consequences of using unexamined data in decision-making can be profound. Thus, Azira continues to make significant investments in research and development to incorporate data science technology into all aspects of its data intelligence platform:

High-Quality Business Logic Models:

Azira's models convert massive quantities of simple data into complex insights using three key data points: location, timestamp, and anonymous/hashed user ID to produce answers to such business questions as, "Where do consumers live and work" (common evening/daytime location?), "How do they get to my store and where do they go afterward?"(pathing), and "What other locations do they visit and for how long?" (brand affinities, dwell time).

Traffic Estimation:

Mobile location data can only provide a limited sample of customer behavior. By combining a deep understanding of industry-specific foot traffic trends with normalization and machine learning, Azira has developed models for estimating the true number of visitors for a specific time and location.

Anomaly Detection:

The mobile ecosystem continues to expand, and with it with an increasing amount of anomalous data that can distort or misrepresent behaviors in location data. Azira uses pattern recognition techniques to identify complex and/or hidden anomalies in consumer behavior data and deploys a variety of filtering methodologies for their detection and removal.

Audience Segment Creation:

Azira uses Natural Language Processing to allow customers to query its vast repository of consumer data. Unlike other vendors who reverse engineer generic open-source NLP models, Azira NLP is a proprietary generative AI technology trained specifically on consumer insights. For instance, users can create highly bespoke audience segments just by entering open-ended text descriptions such as "Women over 40 who went to a Starbucks during the past month," while most audience segmentation tools offer limited dropdown select menus.

Conclusion
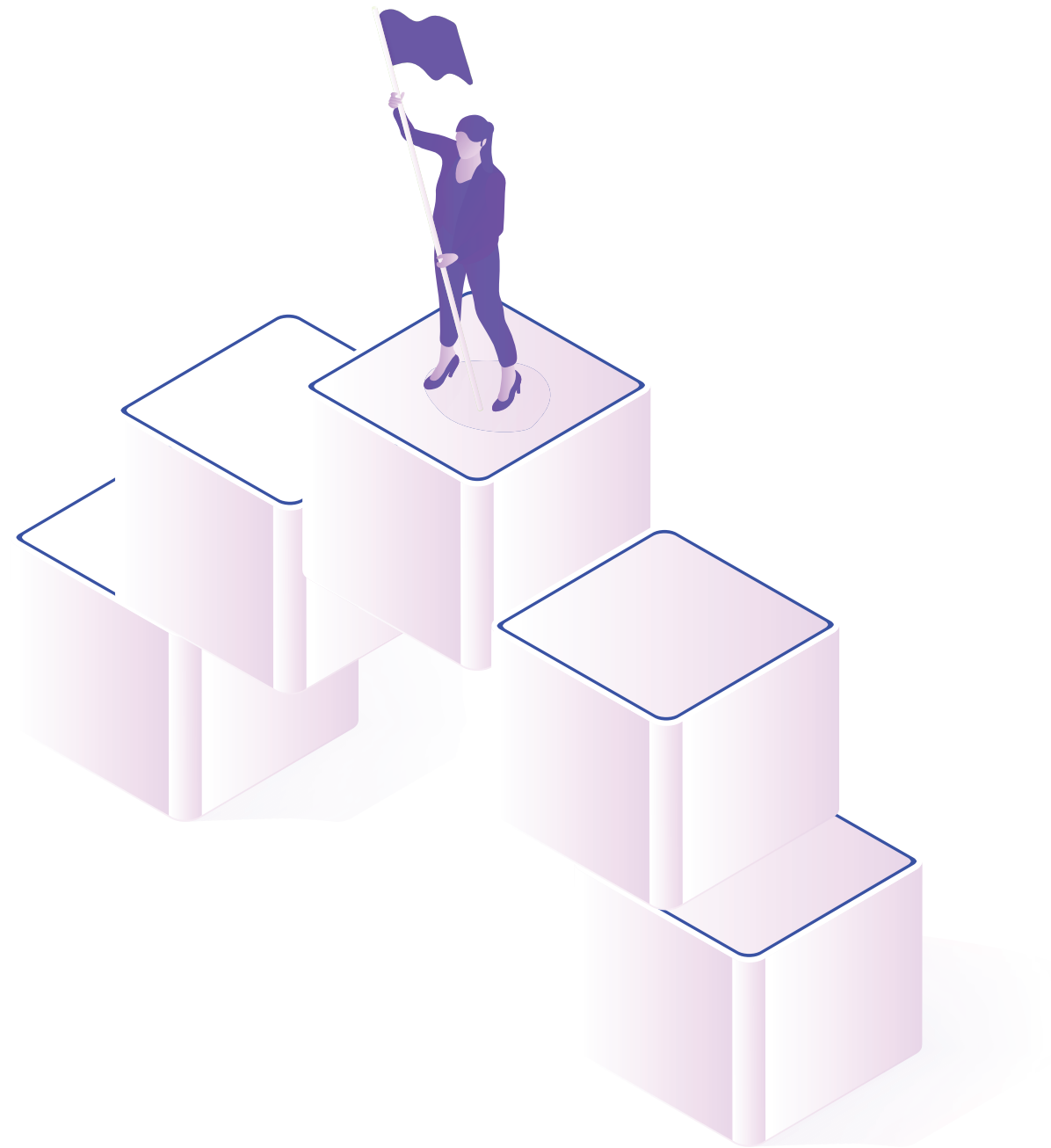
Consumer Behavior Insights versus Commodity Data

By leveraging advanced algorithms, machine learning models, and expert data scientists, Azira deciphers the subtle and not-so-subtle nuances that commodity data alone cannot reveal. Azira delivers comprehensive, refined, and accurate consumer insights that empower businesses to discover intelligence about people and places and excel in their competitive landscape.

Deep Machine Learning and AI Underpinnings

The Azira Platform harnesses the power of self-learning artificial intelligence, enabling it to analyze massive amounts of data from various sources, such as location, transactional, and demographic data, to provide highly accurate and actionable consumer insights, which are continuously refined as consumer data changes. By employing sophisticated algorithms and NLP models, Azira can uncover hidden patterns and trends, allowing businesses to make well-informed decisions based on the ever-evolving consumer landscape.

Balancing Data Utility with Privacy

A balanced approach between utility and privacy is key to unlocking data's commercial and societal potential. One of Azira's key strengths is in harnessing data from interactions in the real world to build context and business intelligence. Locking away all the data weakens the insights that empower better decision-making while making individual data available at too granular a level poses a privacy risk. Azira focuses on striking this balance while arming customers with the necessary guardrails to do the same.

Azira goes to great measures to ensure all data is anonymized and aggregated, eliminating any risk of individual identification. Azira's privacy-first approach and robust consent management framework empower users to control their data usage. As a tenured, responsible and transparent player in the data analytics industry, Azira fosters trust and confidence in users, partners, and clients.

Data Insights, Any Format

Azira believes in providing users with data-driven insights that can be easily understood and put to work, using its visualization techniques to transform raw data into information everyone can understand. Azira provides user-friendly data visualization tools and flexible access options so businesses can work with Azira data in accordance with their in-house skillsets and their preferred analytics environments.
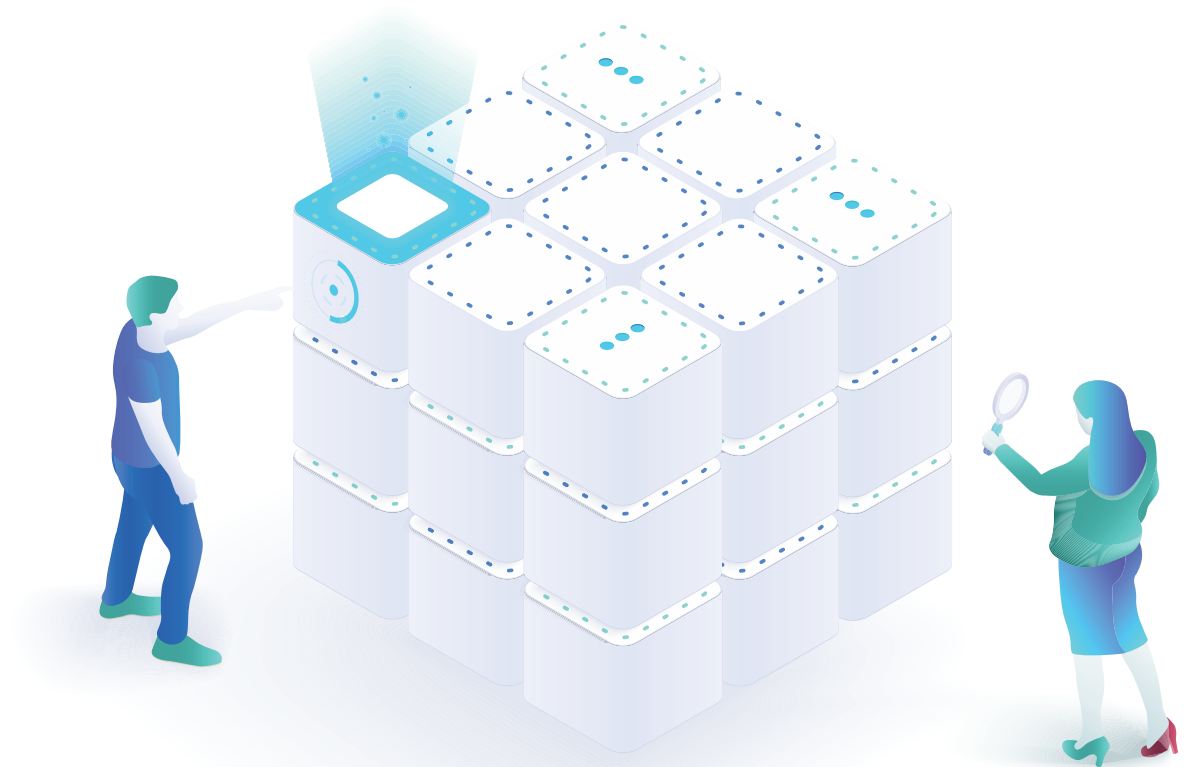
Actionable Guidance

Partner with a trusted team to handle compliance

Azira's global investments around privacy and compliance include numerous resources and experts who understand the implications of non-compliance and who continuously ensure Azira - and by extent, its customers - are at the forefront of data protection.

Make accuracy a priority

Accuracy is the ultimate measure of trust and confidence with regard to insights about people and places. Differentiate datasets with broad strokes around consumer foot traffic from those that deliver precise insights about market areas and how people move inside of them, as this distinction is what will ultimately drive success and minimize risk.

Maintain effective internal practices around data

Bad data costs businesses trillions annually, highlighting the need for an effective data governance and management framework in every organization. Azira's data undergoes multiple filtration processes and rigorous validation so it can be ingested into any new data ecosystem without introducing additional risk.

Invest in the right consumer data insights provider

The adage "you get what you pay for" is especially true when discussing consumer behavior data. Azira cuts through the complexity of today's vast data marketplace with a keen business focus, delivering measurable benefits in three key areas: accuracy, reduced resource costs, and enhanced insights for better results.

Azira™

Sources

Gartner Digital Worker Survey

IoT Connected Devices, Statistica

Today's Growing Datasphere, IDC

2023 survey by Azira and Hanover Research

Statistica 2022 Big Data, Digital and Trends Report

MIT Management Sloan School – 10 Big Data Blunders

Business Wire – Quest to Achieve Data-Driven Leadership

National Law Review – How Much Does Bad Data Actually Cost?

Gartner Research – How to Create the Business Cas for Data Quality Improvement

Azira LLC, a global Consumer Insights platform, helps marketing and operational leaders improve their effectiveness with actionable intelligence to drive business results. Its mission is to create a more relevant world where brands are empowered to reach and build relationships with their consumers. With a profound commitment to partnership, trust and transparency, combined with decades of expertise in consumer behavioral analytics, Azira delivers innovative marketing solutions to curate audiences, activate omnichannel campaigns, and understand footfall attribution. It also provides operational insights for use cases such as site selection, trade area analysis, competitive intelligence and more. Azira serves enterprises in retail, hospitality, travel, real estate, financial services and media. A global company, Azira is headquartered in Los Angeles with offices in Paris, Bangalore, Singapore, Sydney, and Tokyo. To learn more, please visit https://azira.com.