alteryx

Humanizing Big Data

A White Paper from CITO Research

The Mission to Humanize Big Data

In many businesses, the wrong conversation is taking place around Big Data. Businesses recognize that the data being generated by connected devices and consumer activity holds potential, but most conversations are driven by technology platforms that emphasize volume, variety, and velocity, leaving out any discussion of value. To get value from Big Data, you must add contextual information and place analytical capability in the hands of those who need it. In other words, Big Data needs to be "humanized": taken from the world of bits and bytes and converted into real insight for real businesspeople. Big Data needs to be brought down to earth where people who know business can use it to help drive decisions and unlock its value.

Alteryx is on a mission to humanize Big Data, to take it from "isolation among experts" and make it accessible and useful and help draw out its story. We take inspiration from Jer Thorp of the *New York Times*, who has shown in his graphics and in his November 2011 TED presentation, "Make data more human," how to take complex statistics and help them tell stories we can easily understand.

Humanizing Big Data is dependent on two critical elements:

1. **Making Big Data easy to access:** The ability to access, integrate, and analyze Big Data should be available to the data and business analysts who drive strategic decision making across the organization.
2. **Helping Big Data tell its story:** Big Data can provide full stories that drive business value only if it is enriched by the full context of all data available and if advanced analytical capabilities can be applied without the need for data science or statistical expertise.

The Big Data Workflow

Organizations are sitting on a mountain of data that they could use to make decisions. Answers to business questions lie in Big Data: a vast array of sources, from traditional data warehouses, to unstructured, machine-generated data and free-form text. But the answers are difficult to reach. First of all, most BI tools create backward-looking reports and dashboards, based on structured data. This is often only internal data—no market insight, competitive intelligence, or location data—that tells only part of the story.

The current Big Data Workflow has many constituent parts. Data must be acquired from myriad sources and cleansed. It must be sorted and joined so that queries can be made against it. It needs to be stored in a file system that will accept unstructured formats. Analysts and programmers must then work together in a statistical environment such as R, SAS, or SPSS to query the data. Then the data must be visualized in some format—a static report, or perhaps in a 2D or 3D visualization tool.

The problem is that all of this work with data is not handled by a business user working closely with a data analyst. It is in large measure done by a team of specialists behind the scenes in IT, and every step in this process requires getting someone else involved, who already has a substantial backlog of work. The Big Data Workflow involves a number of handoffs, resulting in delays based on high demand for specialized data and analytic skills. The person closest to the business user, the data analyst or business analyst, can't do most of the work, and so the time from question to insight involves numerous delays. In fact, it is often the case that decisions are made based on limited information long before the answers come back from the Big Data Workflow. In effect, the Big Data Workflow—as it stands today—doesn't flow. It's more reminiscent of a logjam than a flowing stream (Figure 1).

The Big Data Workflow involves handoffs and delays based on high demand for data and analytic skills.

Business User

Business Question
"How do I achieve hyper-local marketing success across digital channels?"

Business Decision
"Establish channel-specific customer segment targeting"

Data Analyst

Identify Data Sources
Hadoop, CRM, Excel, Data Warehouse

Present and Adjust
Build application or report

Analysis
Explore output and iterate

Data Scientist IT Expert

Unify and Assemble Data
Access and combine multiple sources

Cleanse and Enhance
ETL, data cleansing, etc.

Append and Add Context
Location, demographics, and segmentation

Build Analytic Workflow
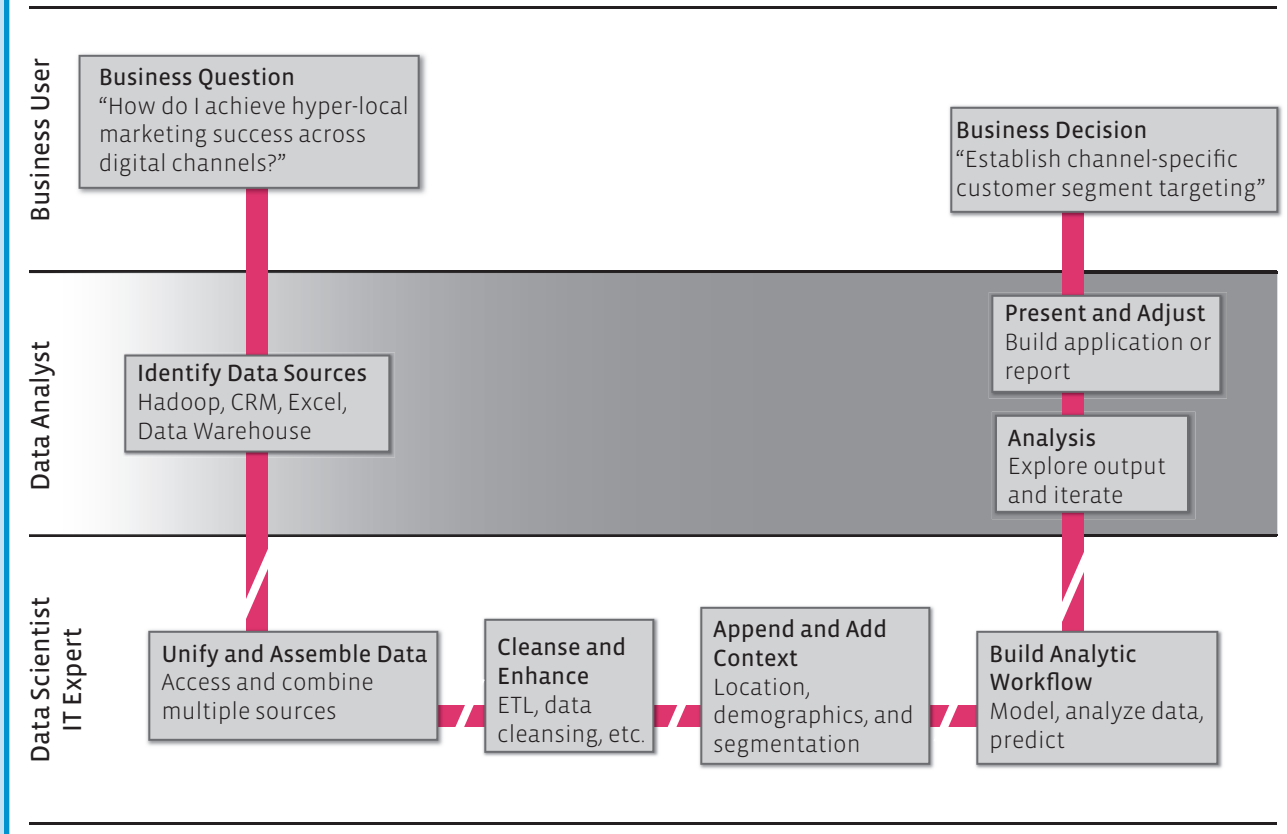Model, analyze data, predict

Figure 1. Big Data Workflow

Humanizing Big Data

Humanizing Big Data makes it accessible for analysts who operate in today's enterprise business units, giving them capabilities usually available only to IT. It's rendering data into information that is easily accessible and highly relevant. It's making analysis based on Big Data effortless and natural. Instead of relying on specialized skills in programming and statistics, data can be humanized by adding appropriate context and offering straightforward tools for building analytical applications. Humanizing Big Data means working with the data directly so that it tells its story. Having the full story leads to business insight. It also means a new ability for data analysts to hone their craft and expand their ability to do analytics independently. They become, in effect, data artisans.

The Rise of the Data Artisan

In the past, data analysts have had strong skills in statistics or in business. By providing analytical tools and contextual data that have traditionally been accessible only by data scientists or IT experts, data analysts can become data artisans.

An artisan is a person who takes raw materials (in this case data) and uses skill, knowledge, and vision to craft it into something of unique value.

Data artisans have a deep understanding not only of the data that they are looking at but also of the business and its problems. Because data artisans understand both the data and the business, given the right tools, they can really hone their craft and apply it to business problems in a repeatable way.

Data Artisan

Today's leading-edge data analysts are more artisans than reporters, applying creativity and insight to their role in strategic analytics. Data artisans apply skill to their craft, infusing their deeply embedded knowledge of strategic analytics into the analytic process and associated applications. They not only understand their organization's business drivers and problems, but also where to find the right data for strategic decisions. Often, however, data artisans are not skilled in pulling data for analysis or consolidating enterprise data structures. That's where tools can help them hone their craft.

The work of the data artisan is at the core of humanizing Big Data. The data artisan is like the chief humanizing officer when it comes to Big Data. The data artisan creates data and analysis workflows that provide the foundation for helping the data tell its story, whether delivering an answer to a specific question or creating a new application for business users at the end of the process.

Design Principles for Humanizing Big Data

Humanizing Big Data involves several key design principles when it comes to creating solutions that deliver real insight:

- **Ingest and integrate data from anywhere:** Systems of record, social media, and sensor data are all fair game, as is information from the data warehouse.

- **Seek patterns:** Patterns hold the key to predicting future outcomes. Don't look for a pixel-perfect report that's accurate to the penny when searching unstructured data. Fuse the qualitative structured data with the important context provided by unstructured data.

- **Make insight available at the point of decision:** Insights are best when widely available. A store manager knows his market; he sees it every day. With powerful analytical tools that were formerly centralized, he can make effective and informed decisions.

- **Reuse analytical IP:** A data artisan can create a data object and share it with a full range of decision makers, who can adapt and build on it. Each time the story extends, it also gains focus.

The goal of the humanized Big Data approach is to get these capabilities into the hands of analysts in business units, allowing them to create analytic reusable workflows.

Alteryx: Humanizing Big Data

Alteryx is a strategic analytics platform for business analysts that can close the gap between the Big Data you have and the value you need, humanizing Big Data along the way. Alteryx is for people who can't wait for long, drawn-out IT processes or statistician-driven projects to get access to the information needed to drive decisions.

Making Big Data Easy to Access

With Alteryx, complex business questions can be answered by line-of-business analysts and business decision makers in a tenth of the time taken by other solutions. Using Alteryx, data artisans can evolve their skills to include areas such as predictive analytics, Big Data, and information management, without the need for a multi-year training course or a Ph.D. program.

Alteryx provides a complete platform for integrating, cleansing, enhancing, and analyzing data, and sharing the insights with others by allowing data artisans to quickly build applications and visualizations. All the steps in the Big Data Workflow—data integration, cleansing, creating analytic apps, and defining the output—can now be in the hands of the people who most need to extract that value. While there is incredible value in each specialized area of the Big Data Workflow, the biggest value comes when each area is integrated with the rest of the workflow that actually drives business decisions.

Alteryx helps data analysts and data artisans by providing them with a workbench that allows them to conduct the workflow steps that formerly were handled by IT and data scientists, as outlined in Figure 2.

An analytic workflow is an important piece of intellectual property that gains value the more it is shared, reused, and appended.
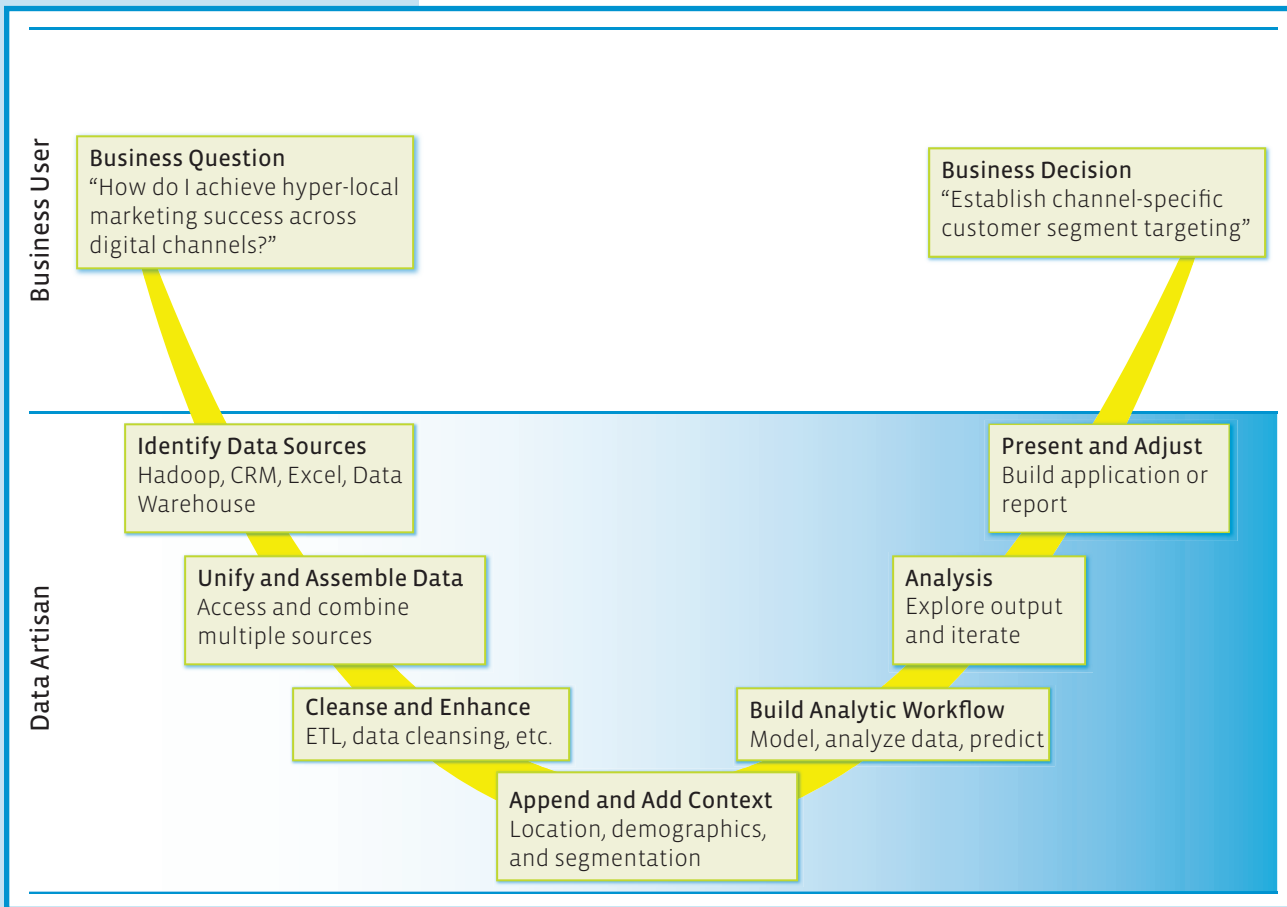
Business User

Business Question
"How do I achieve hyper-local marketing success across digital channels?"

Business Decision
"Establish channel-specific customer segment targeting"

Data Artisan

Identify Data Sources
Hadoop, CRM, Excel, Data Warehouse

Present and Adjust
Build application or report

Unify and Assemble Data
Access and combine multiple sources

Analysis
Explore output and iterate

Cleanse and Enhance
ETL, data cleansing, etc.

Build Analytic Workflow
Model, analyze data, predict

Append and Add Context
Location, demographics, and segmentation

Figure 2. Humanizing the Big Data Workflow

Alteryx takes advantage of technologies like multicore processing, solid-state drives, and in-memory computing to bring powerful analytical tools to the desktop.

Alteryx helps data artisans deliver sophisticated analytics to answer complex business questions much more quickly than is possible with solutions such as SAS or SPSS. Because of its drag-and-drop, workflow-based user interface, which contains all of the data, all of the predictive capabilities, and all of the analytical capabilities the data artisan will need, it's much easier and faster to use than a platform in which most of the work has to be done by coding.

Alteryx delivers out-of-the-box data that enriches and adds customer and location context to any data, bringing context and reality to the story the data is telling. Contextual data comes from key data providers such as Experian, Tom Tom, and Dun & Bradstreet. By building this contextual information right into the platform, data artisans are empowered in new ways.

Contextual Data (Traditional)	Contextual Data (Alteryx)
Purchase data from multiple third-party data providers Data experts join the external data with internal data	Data from Experian, Tom Tom, Dun & Bradstreet, the US Census, and more is built in to the platform

Alteryx supports a vast range of inputs and file formats, across structured and unstructured data sources, including delimited text files, Hadoop, MongoDB, Google Earth (KMI), Microsoft Access and Excel, ODBC drivers, Oracle OCI, XML, social media (Twitter, Facebook, Yelp and Foursquare APIs), CRM, and enterprise cloud data sources such as SalesForce.com.

Helping Big Data Tell Its Story

With Alteryx, data artisans can quickly organize ideas, then convey them seamlessly by publishing the data, and the applications surrounding that data, to the cloud. Think of Alteryx as a personal ETL and predictive analytic workbench that enables data artisans to publish mashups and applications to a cloud service, which can then be put to broad use across the organization.

With an Alteryx-generated analytic application, the business user can do his own slicing and dicing and get answers right away, seeing the story the data is really telling and acting on that story. This is the essence of making data-driven decisions.

For enhanced predictive capabilities using Big Data, Alteryx has fully integrated the R statistical language directly into the end-to-end workflow. Or, if the analyst wants to use a statistical function but is not an R expert, one of more than 25 prepackaged statistical tools built using R can be integrated into any analytic workflow.

Helping Big Data Flow

Alteryx performs the entire Big Data Workflow in one platform, where other analytical solutions require separate ETL, integration, packaged data, and analytical add-ons. A data artisan can use Alteryx to answer sophisticated business questions and perform predictive analysis on Big Data, without highly skilled programmers. He or she can build applications that fuse capabilities rarely provided to business users, such as providing spatial analytics to those responsible for product merchandising mix.
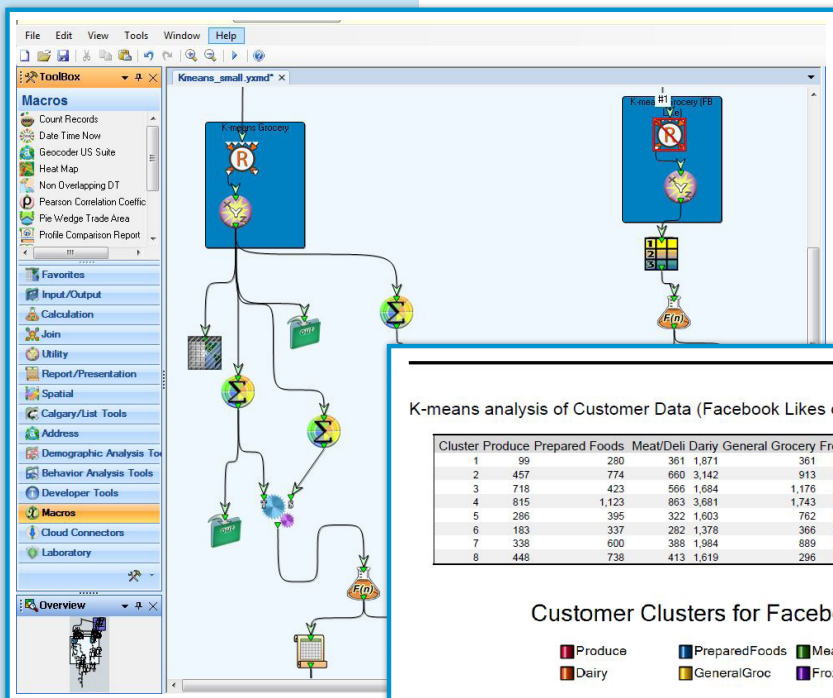
An example Alteryx workflow Integrating Big Data and driving predictive analytics
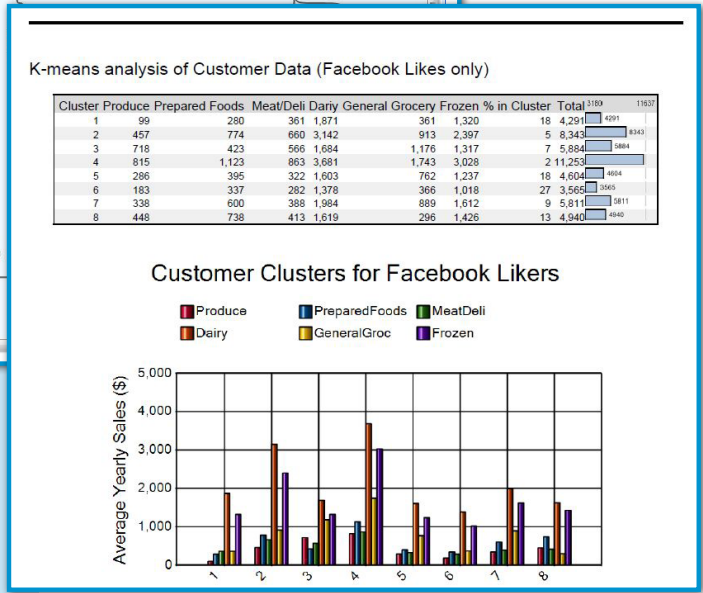
An example of Big Data delivering business value – created by Alteryx partner Farsite

Alteryx in Action

Companies large and small use Alteryx to gain new insights from humanized Big Data. With new datasoures and business models have come a need to answer new, contemporary business questions that don't just look at past performance. Instead, the need is to drive strategic decision making using strategic analytics. This is what Alteryx customers such as Walmart, McDonalds, and AT&T do every day. Some examples of Alteryx customer use cases that specifically take advantage of Big Data insight are outlined below.

Analyzing Set-Top Box Data to Drive Revenue Growth: A major US cable company needed to differentiate itself in a highly competitive advertising market and drive revenue growth. Through their set-top boxes, the company had access to very large amounts of semi-structured data that is generated constantly from millions of homes across America.

By integrating this non-traditional data with pre-packaged customer and market data, such as demographics, along with their core customer data, this Alteryx customer is now able to deliver deeper and more accurate insight into audience viewing habits than was possible using traditional sources. Instead of sampling to guess at what people are watching, the set-top box provides actual (though voluminous) data. The ability to quickly and efficiently analyze this data in turn enables the company to increase its overall advertising revenue through more targeted ad purchases.

Proving the Value of Social Media Marketing to Clients: Proving real value from investments in social media based marketing is always a challenge. For one online contact management agency, the process of driving and proving value, through actual customer purchases, was costly and time consuming due to manual integration and cleansing of unstructured social media data with their clients' point-of-sale data. This was compounded by the need to use scarce highly educated statistical experts for the analysis of the data.

Alteryx allowed this customer to humanize their Big Data by making it easy for a single data artisan to access, integrate, and analyze social media, structured POS data, and loyalty card data in a single workflow. They are able to show the true impact of social media investments by their clients and help them to increase their ROI by having the right analytics in the hands of key marketing decision makers.

Conclusion

The value of humanizing Big Data is that, with the right tools in place, your organization can develop tremendous business value without adding a new layer of skills to your personnel base. Tools such as predictive and spatial analytics have historically been isolated in the hands of the very few. But if organizations are truly to unlock the value of Big Data, they must place powerful tools in the hands of those who will actually implement the insights they generate. This accelerates innovation and value creation.

Through Alteryx, Big Data is "humanized" when sophisticated concepts around data can be delivered to the business user in an understandable format, without sacrificing any of the computational abilities in the background. Alteryx empowers the data artisans in your company to assemble, overlay, and analyze any combination of your enterprise data, market insight, and spatial analytics into a single picture, so you can take action immediately.

CITO Research

CITO Research is a source of news, analysis, research, and knowledge for CIOs, CTOs, and other IT and business professionals. CITO Research engages in a dialogue with its audience to capture technology trends that are harvested, analyzed, and communicated in a sophisticated way to help practitioners solve difficult business problems.

alteryx