An Oracle White Paper
March 2014

The 5 Rules of the Road for Enterprise Data Discovery
Introduction ...................................................................................................................... 3
Pain Points ....................................................................................................................... 4
Solving the Problem .......................................................................................................... 5
The 5 Rules of Enterprise Data Discovery ................................................................. 5
  1. Govern Self-Service for Results Without Risk .................................................. 5
  2. Blend Diverse Data for Deeper Insights ............................................................... 6
  3. Equip Yourself with Integrated Search, Navigation, and Analytics .................. 6
  4. Always Have a Dialogue with Your Data ............................................................... 6
  5. Enrich Diverse Data and Keep On Discovering ................................................... 7
Oracle Endeca Information Discovery ........................................................................... 7
  Enterprise Self-Service Discovery ............................................................................... 8
Realizing the Value ........................................................................................................... 8
  Benefits for the Business ............................................................................................. 8
  Benefits for IT .............................................................................................................. 8
Conclusion ....................................................................................................................... 9
Introduction

The fundamental premise of analytics is to use information to gain insight, and to have that insight drive action, which in turn leads to greater success: better competitive results, greater mind and market share, increased customer loyalty, higher profits, lower costs. The challenge companies face today is that insights remain elusive, even though data and the tools to handle it have proliferated. Organizations have dedicated thousands of hours to collecting, cleaning, rationalizing, wrangling data—and to crafting solutions around it, starting from the early days of online analytical processing (OLAP) and evolving through dimensional analysis, cubes, and universes to the current complex business intelligence landscape. These solutions share common traits:

1. They are predicated on a known set of data, which is contained within an enterprise, typically in the enterprise data warehouse.
2. They address a known set of questions, which tend to be operational in nature (e.g., annual sales, broken down by region by quarter).
3. They are built for predictability, reliably delivering regular reports. As highly-curated systems, they are usually neither quick to implement nor easy to change.

The legacy of the information superhighway, these are the well-traveled roads that companies use to manage much of their day-to-day operations.

The path to higher performance, however, is not always so clearly marked. Achieving insight is a process that requires combining intuition with data in an agile cycle of discovery, with no set route and no preordained destination. In other words, sometimes achieving insight requires going off-road, and when it does, it’s helpful to have a GPS. Interactive, intelligent guidance—that high-speed dialog with the data which fuels innovation—is especially necessary in this age of big data, where such varied information is produced so rapidly that it threatens to engulf those who try to navigate it.

This paper defines a solution and the rules of the road, so companies can understand how they can achieve the insights they desire, from data big or small, without being overwhelmed in the process.

Two roads diverged in a wood, and I—
I took the one less traveled by,
And that has made all the difference.

Robert Frost
Pain Points

Organizations that use analytics to drive decision-making are more successful than their less data-driven competitors. Why, then, aren't more companies analytics overachievers?

Simply put, because it's hard. Despite the wealth of information, abundant tools, and increasing computing power, numerous obstacles remain:

- Access to much of the information within the enterprise is still somewhat limited and slow.
- Massively varied data and sources make it very difficult to acquire, understand, and use non-traditional information effectively.
- Traditional platforms often are built around complex formal models, which means they take too much time to change and make IT a bottleneck.
- Because of the complexity, users spend too much time waiting for their requests to make it through the IT backlog, and too little time using their expertise to innovate and move the business forward.
- Many of the alternatives to existing systems are siloed tools—they seem to offer a quick fix initially, but do little to address the enterprise need for better, faster, deeper analytics across the organization.

This list leaves everyone frustrated and confused: the business is under growing pressure to deliver results in an increasingly competitive market but can't get to the raw materials; IT is struggling to reinvent itself as a center of innovation and value while maintaining corporate data standards and quality; and executives are wondering how to help the two groups work together more effectively to get even greater value from existing analytics.

The result is a shift in control, with lines of business more frequently owning analytics budgets and setting departmental strategy. This includes using new tools and experimenting with new data, which leaves IT in the unenviable position of trying to get value from current systems while maintaining governance over a growing analytics footprint in the face of a decreasing budget.

The question, then, is how to move beyond the current stalemate. How can companies be more agile and be highly confident in the data and results of their analysis? Must they abandon their current platforms and infrastructure, or is there a way to extend the value of their existing environment (that doesn't require almost completely rebuilding it)? And is there any way to get business users what they want and make them more independent, without compromising IT’s charter to provide the quality and governance that are the cornerstones of corporate accountability?
Solving the Problem

The theme that emerges is the need for fast, easy, intuitive access to any data, for both the business and IT. That’s a tall order, and for many organizations it’s hard to achieve. After all, how can companies justify the significant expense and effort of expanding current systems, or of buying/building a new system, to meet all those requirements when it’s not clear what the value of all that data might be?

Fast, intuitive access to any data, no matter how diverse, is the very problem that enterprise data discovery is built to solve. Enterprise data discovery simplifies access to wildly diverse information, making it available for instant exploration, regardless of its state—structured or unstructured, dirty or clean, rich or sparse, jagged or uniform, internal or external—giving both the business and IT unprecedented visibility into its value to the organization. Instant exploration also requires simplified access, in the form of a highly intuitive and dynamic interface that is friendly, familiar, and fast, so users can spend time interacting with the data and uncovering insights, instead of learning the software. Easy access and instant exploration are enhanced by natural language processing, which uncovers sentiment and enriches analytics by deriving structured attributes from text. The speed, simplicity, and enriched experience are all made possible through a highly flexible architecture and significant in-memory processing power, which provides extreme performance at scale.

Enterprise data discovery is a natural complement to existing business intelligence practices and platforms, offering analytic agility without compromising quality. It’s the GPS that enables businesses to be more confident in the decisions they make, even as they make them more quickly.

The real innovation here is that we can ask questions and get the answer back before we have forgotten why we asked the question in the first place.

– Hilary Mason, Chief Scientist Bit.ly

The 5 Rules of Enterprise Data Discovery

With the above in mind, then, following are the five rules for enterprise data discovery solutions. To be considered enterprise data discovery and deliver the value of instant, intuitive exploration, solutions must meet all 5 rules in a single platform, without requiring additional third-party products, custom development, or up-front data modeling.

1. Govern Self-Service for Results Without Risk

Imagine this: you’ve seen a spike in sales in one of your regions and want your analysts to find out why, so you can replicate it nationally. Your analysts have the sales reports that show the spike, but nothing to show why. They have a hunch that a recent media campaign could be playing a part, but need more information to be sure. That would require combining sales data—not just the report, but the actual raw numbers—with results from the media campaign. But how to quickly get sales data and combine it with
other sources? Even more importantly, how can you be confident in their conclusions? You've been in the room before when off-road analysis lead to a dead-end instead of new insight.

Governed self-service discovery is the answer. It provides quick, easy access to both gold-standard data from IT and to other diverse data, within the organization and outside it. It's made possible by an enterprise discovery platform, giving IT full visibility to the data and analytic applications at the same time it meets the business's need for self-service. Business users are empowered with the independence they want, to upload and combine data on their own, with the confidence they'll get good results. IT benefits from managing discovery based on a world-class platform that offers the scalability, security, and reliability they expect, and the transparency organizations require for strong data stewardship. Everybody wins.

2. Blend Diverse Data for Deeper Insights

Suppose the combination of sales data and media campaign results still didn't quite get you there. What next? Your business users suggest using social feedback to get more color on the campaign. While that would normally make IT cringe, in your case, they're happy because they know a governed self-service discovery platform makes it easy to access any type of information, no matter how diverse the source. Providing easy access to a wide array of data allows organizations to make the most of existing assets, including unstructured and text-heavy information that is often not included in today's analytic environments. It also means business users can quickly blend current data with new information of any type to find out what's driving the spike in sales. All that data business users always want but is not in traditional analytics—blog posts, customer reviews, survey text, Tweets, call logs—is now available to give you a whole new view of what's driving your business.

3. Equip Yourself with Integrated Search, Navigation, and Analytics

When you're looking for that deeper insight to enhance your competitive edge, there's always the risk of going down a blind alley. That's why you want to have everything you need for exploration, to help you stay on course. Full-featured search, guided navigation, and powerful visual analytics built together from the ground up are your critical toolkit for discovery. On your quest for the cause of that spike in sales, your business users can search all that diverse data just as they would on the Web, without needing any training. Intelligent navigation suggestions driven by the data help steer users away from dead-ends without constraining the questions they can ask. Dynamic visualizations maintain the context from search and navigation, helping users get answers quickly. The ongoing interaction among search, navigation, and visual analytics guides users through the data—leave out any one of the three, and it's easy to get lost.

4. Always Have a Dialogue with Your Data

Traditional analytics take you on a well-traveled path through your operations, based on established questions. With enterprise discovery, the data is in the driver’s seat, giving you the agility to follow the road less traveled. Interesting—that spike in sales occurred in the middle of the media campaign. What might cause that? Your business users can ask the data to find out, and discover that during the time of the spike there was a major local sporting event. There was no way to plan for that question, or the others that will follow from it. This new clue prompts a new question: what could a sporting event have to do with the spike? Again, the data reveals its value by providing a new answer—one of the ads from the campaign got additional play at the event. Now you have something solid to go on. And you have the speed and flexibility you need to continue the dialogue, because enterprise discovery features
in-memory performance combined with an innovative, data-driven architecture that means information is indexed as-is rather than having to conform to a model. The result is a unique experience that empowers users to get the answers they need, when they need them.

5. Enrich Diverse Data and Keep On Discovering

You're near the end of the sales spike journey. You know the ad got more play, but was that really the source of the spike in revenue? Traditional analytics don’t typically include qualitative information—the hidden meaning and critical connections that reveal what people really think and what's really driving their behavior. Yet that's often the key to root cause, and explains your sales spike—posts in social media show that the crowd at the sporting event loved the ad, which drove a viral reaction across the region. This kind of insight is priceless, and that's why enterprise data discovery gives business users the power to get the most out of raw text. Enriching text takes a couple of clicks and makes it possible for users to unlock the hidden insights in unstructured content. Combining text enrichment with enterprise-class sentiment analysis means you can see the numbers side-by-side with key themes and sentiment, giving you a complete view of your business and enabling you to make truly informed decisions.

Oracle Endeca Information Discovery

Oracle Endeca Information Discovery is built upon the 5 rules and is a complete solution for agile data discovery across the enterprise, offering fast, intuitive access to both traditional analytic data, leveraging existing enterprise investments, and non-traditional data, including external and unstructured information. This allows organizations unprecedented visibility into all relevant information, to drive growth while saving time and reducing cost.

Only Oracle delivers powerful self-service discovery as part of a complete enterprise platform that offers a framework of governance and standards, and which is integrated with Oracle Business Intelligence Enterprise Edition, enabling faster and more confident decisions, reducing the IT backlog, and increasing innovation.
Enterprise Self-Service Discovery

Oracle Endeca Information Discovery (OEID) also offers organizations a distinct advantage when it comes to empowering business users to do their own analysis. Unlike desktop discovery tools that allow users to upload data to their own computers, OEID shuns silos. Users enjoy not only tremendous ease-of-use, but also leverage the full strength, security, and scalability of enterprise-class infrastructure. In practice, this means that users can easily upload data into a centrally-managed platform and then perform analysis from their computers—and the experience doesn't stop there. Because the platform is shared across the enterprise, users can collaborate quickly and effectively, with other users in their department or with other groups. Importantly, 'other groups' includes IT. Since the platform is centrally managed, IT not only benefits from the ease-of-use and agility offered by discovery, but also has total visibility to data and definitions. OEID means that business users can help themselves in the context of a secure world-class platform, in balance with IT governance and strong data stewardship, enabling governed self-service discovery for the enterprise. This fundamentally changes the dynamic between lines of business and IT, driving rapid iteration and collaboration, and ending the analytics stalemate.

Realizing the Value

Organizations that leverage enterprise discovery as part of their analytics strategy realize significant benefits, achieving greater success through better competitive results, greater mind and market share, increased customer loyalty, higher profits, and lower costs. These are the direct result of discovery's core value proposition:

- **Agility.** Putting the data first creates a dynamic relationship between business and IT, driving iterative prototyping, collaborative decision making, and faster results.

- **Innovation.** Combining existing data with non-traditional content surfaces new meaning, fostering the broader perspectives and deeper insights that reveal new opportunities.

- **Confidence.** Empowering the business to self-serve while maintaining IT stewardship ensures responsible decisions the enterprise can trust.

Benefits for the Business

Today, the business is faced with overwhelmingly rich information that is all potentially relevant to decision making. Enterprise data discovery enables agility and drives innovation through easy access to any type of data, empowering business users to visualize new insights and get fast answers with confidence.

Benefits for IT

For IT teams under pressure to manage costs while still being responsive to the business, enterprise data discovery provides an enterprise-class platform that balances empowering business users with IT governance. Leveraging self-service capabilities on a foundation of enterprise standards allows users to load diverse data sources, configure easy-to-use applications, and iteratively expand them in a fraction of the time typically required to deploy analytics applications.
Conclusion

Are you able to offer the best of both worlds to IT and the business, promoting greater analytic agility without compromising quality or governance? That question is no longer hard to answer, because now organizations can test drive their solutions using the 5 rules of enterprise data discovery. In addition to helping measure the value that an analytics environment can provide, the rules also function as a GPS for implementation, helping organizations plan their approach to incorporating enterprise discovery into their infrastructure and business practice. Oracle Endeca Information Discovery is a complete solution that embodies all five enterprise discovery rules, allowing companies to have it all, analytically speaking: the agility to harness any data from any source; the confidence that comes from leveraging gold-standard data alongside new sources; and the deep insight that drives innovation.