IMPACT ANALYSIS

™

SCIANTA ANALYTICS
DEEP INSIGHT™

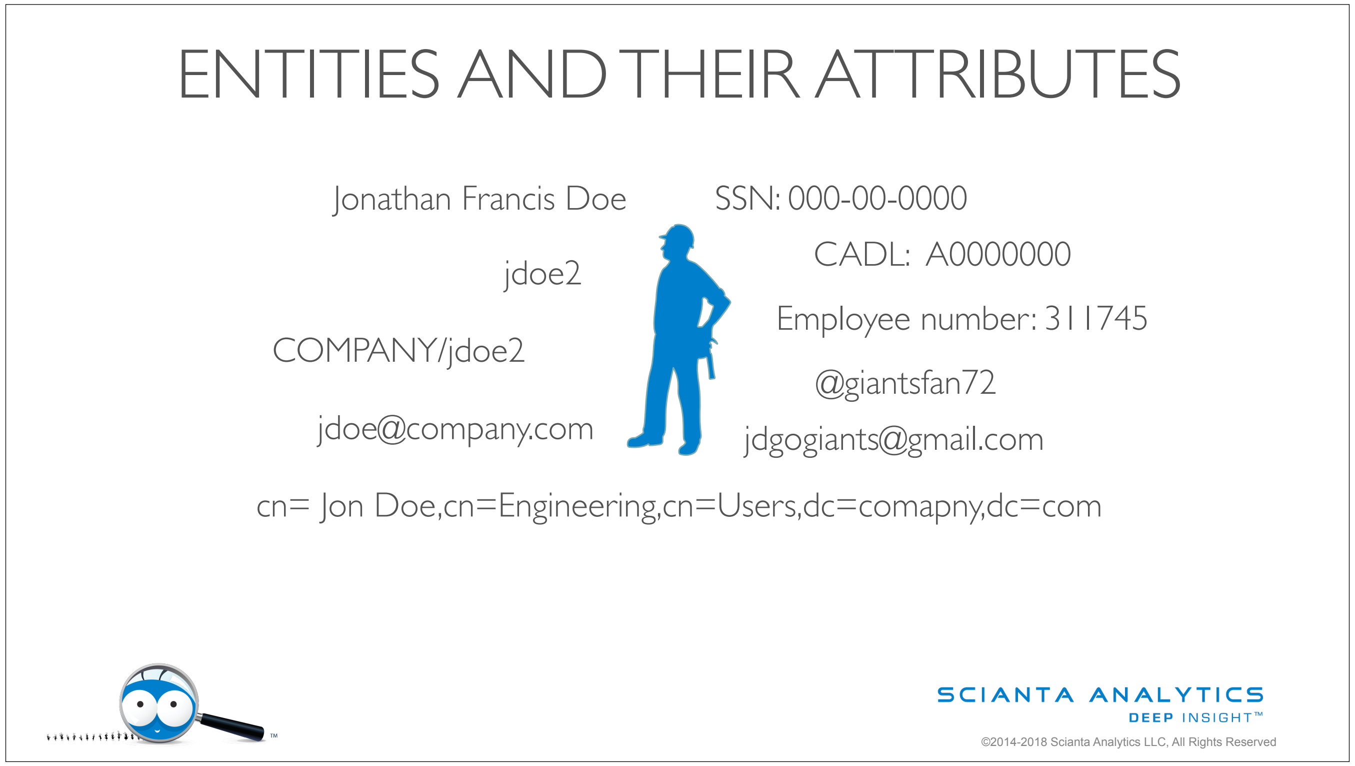
“The natural evolution of machine learning, Cognitive Computing attempts to imbue, in computer systems, the same insight and understanding we see in humans.”

Earl Cox
Chief Scientist, Scianta Analytics
Splunk .Conf 2013

AGENDA

Introduction to Machine Intelligence	Data Handling 1	Data Handling 2	Anomaly Detection	Transactional Behavior	Impact Analysis
Academic Concepts	Collection	Retention	Anomaly Definition	Defining Transactions	Organizational Visibility
Data Systems	Storage	Format	Measuring Normality	Transaction Relationships	Types of Impact
Maturity Curve	Security	Labeling		Probability Measurement	Responsiveness

SCIANTA ANALYTICS
DEEP INSIGHT™



ENTITIES AND THEIR ATTRIBUTES

Jonathan Francis Doe

jdoe2

COMPANY/jdoe2

jdoe@company.com

SSN: 000-00-0000

CADL: A0000000

Employee number: 311745

@giantsfan72

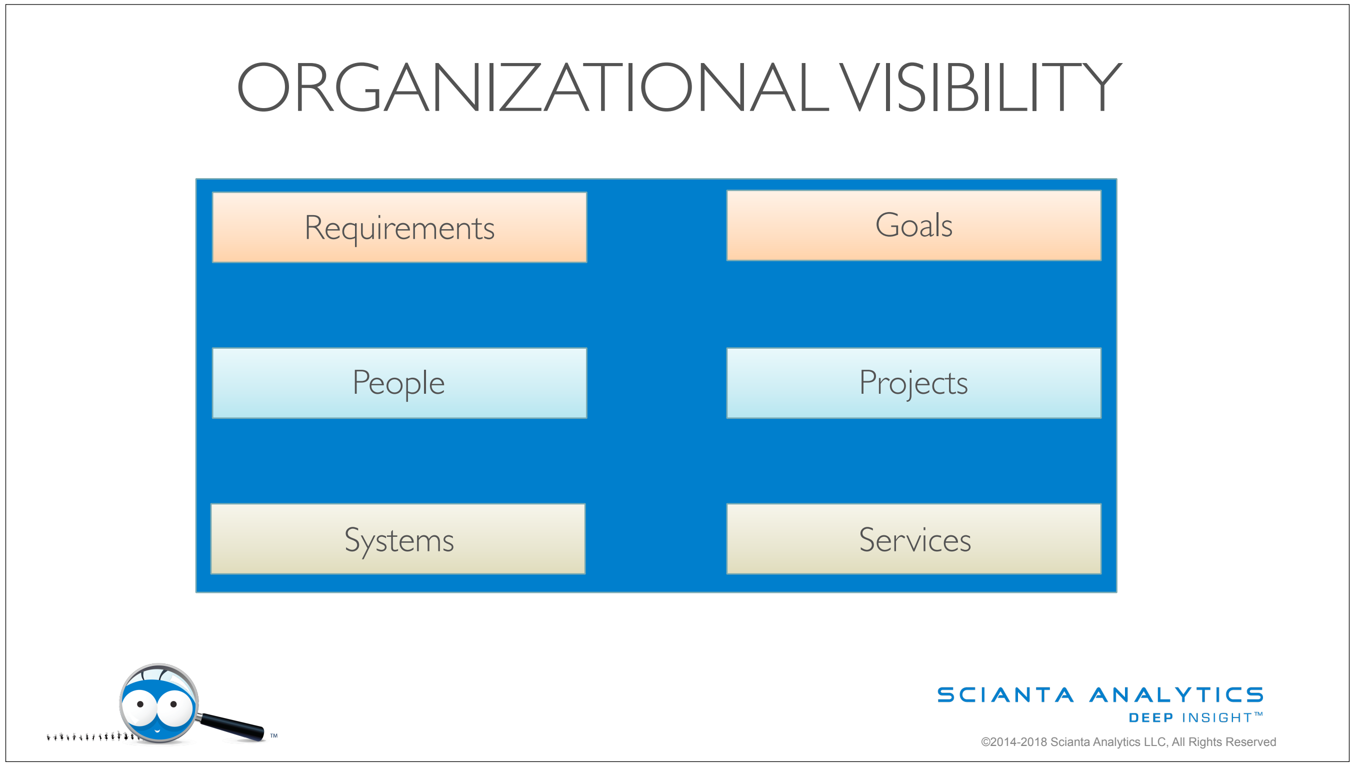
jdgogiants@gmail.com

cn= Jon Doe,cn=Engineering,cn=Users,dc=comapny,dc=com

SCIANTA ANALYTICS
DEEP INSIGHT™

We've talked about actors and assets, and there's an interesting problem in that. The thing is, we don't really care about what an email account or an IP address did, we care about what a person did or whether a service is working right. Our goal is to provide visibility to the organization, and that's about people and things, not addresses.

The problem is that lots of data sources only record an attribute, like "jdoe2" or an employee number. Finding a complete picture of what a person is doing takes a lot more data, you need to know all of their attributes. Dig deep enough and you might be linking Jon to his laptop and iPad, and they've each got lots of attributes too. This kind of mapping is too expensive to solve completely and continuously, but it is valuable enough to link common attributes when they're available.



ORGANIZATIONAL VISIBILITY

Requirements

Goals

People

Projects

Systems

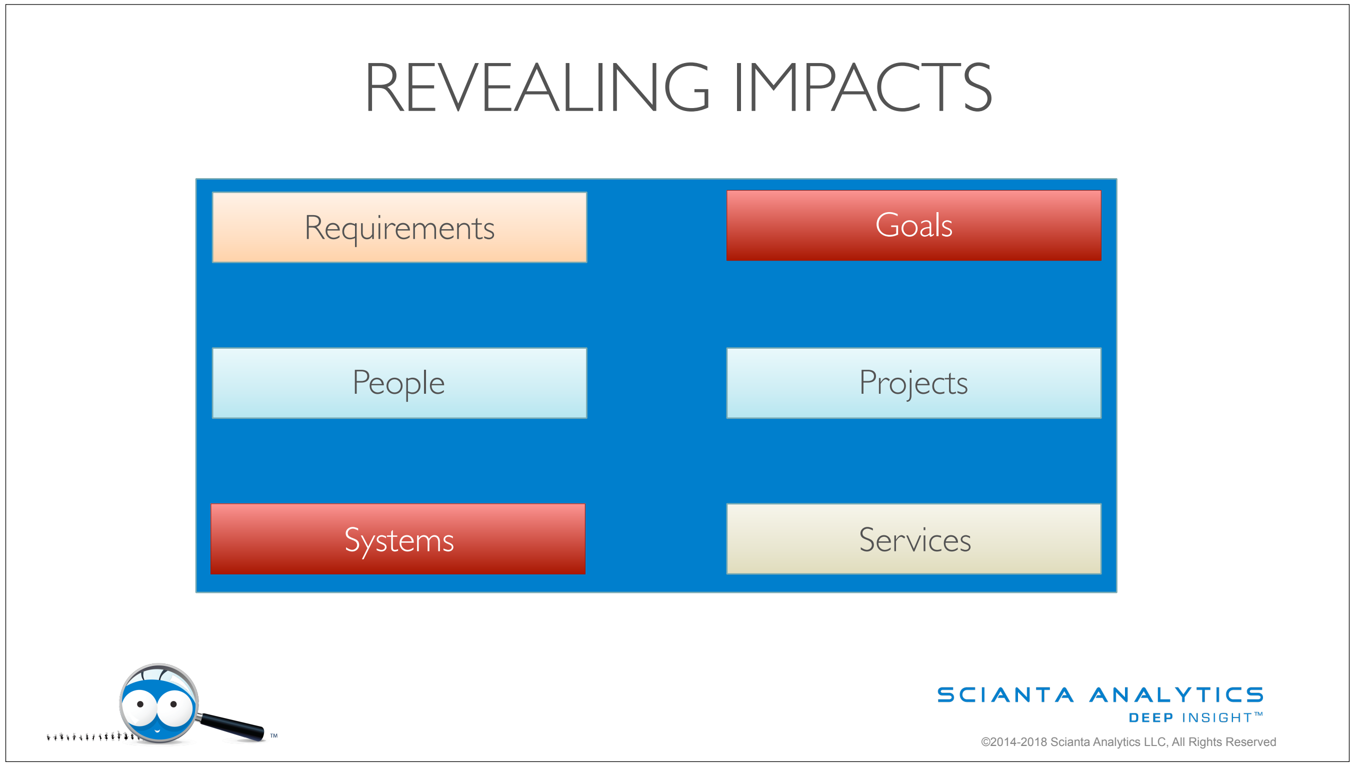
Services

SCIANTA ANALYTICS
DEEP INSIGHT™

At the risk of being obvious, every organization needs to meet requirements and achieve goals. People work together in teams to complete projects, and they depend on systems and services. So far so good, but there's a challenge: there's a lot more data at the bottom of the stack than the top, and it's not always clear how to draw a line from the bottom to the top.

There's two valuable tools for drawing those lines, from bottoms up and from tops down. The first is mapping the attributes of observed systems to entities that we know are important. We may not know the purpose of every fiber channel number in a SAN, but if we can determine that a set of them are allocated to the Hail Mary project that meets our organization's most important goal, we can pay extra attention to alerts concerning those.

Second, we can use behavioral awareness to understand what resources are used by a given team and project. Looking at that same Hail Mary project, we can assume that systems those team members interact with are important, and increase the attention that they get.



REVEALING IMPACTS

Requirements

Goals

People

Projects

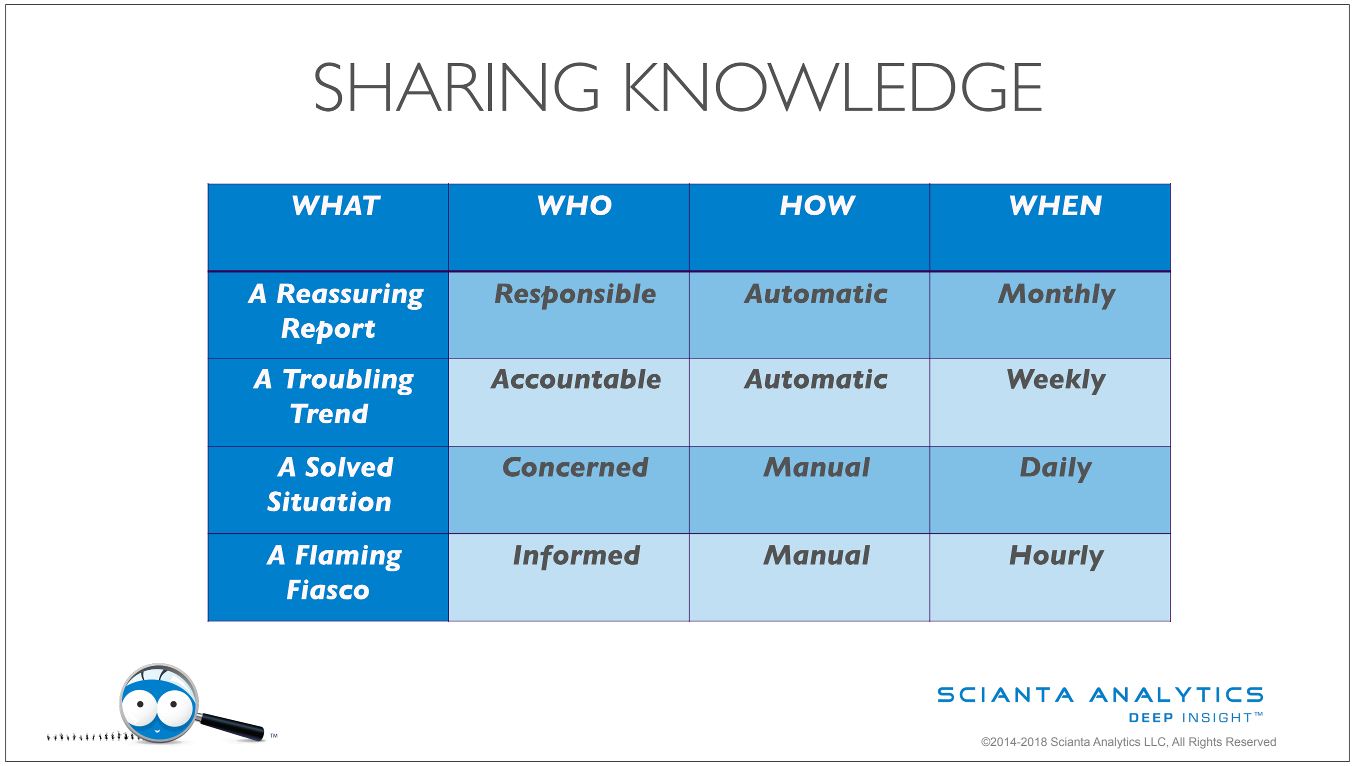
Systems

Services

SCIANTA ANALYTICS
DEEP INSIGHT™

It may seem like a lot of extra effort, but it's effort that pays off. Everyone can assume that a given running system needs to stay that way… but if you can show that a broken cable directly stops work on a project that decides the organization's success, that's better than an assumption.

To do this, we need to use dependency and adjacency between transactions. To continue our example, perhaps we have a transaction that describes data flowing to and from the SAN, a transaction that describes its processing, and a transaction that describes the presentation of results. It's not likely that we can model transactions like "sponsor approval", "board presentation", and "analysts pitch day", but if we can get enough systems visibility humans can use their contextual knowledge to complete the picture.



SHARING KNOWLEDGE

WHAT	WHO	HOW	WHEN
A Reassuring Report	*Responsible*	*Automatic*	*Monthly*
A Troubling Trend	*Accountable*	*Automatic*	*Weekly*
A Solved Situation	*Concerned*	*Manual*	*Daily*
A Flaming Fiasco	*Informed*	*Manual*	*Hourly*

SCIANTA ANALYTICS
DEEP INSIGHT™

You've found something important that you want to pass on, and now you've got a whole new problem. Who needs to know this? How does it need to be positioned for them to accept it? When do you give them the information? This is potentially a deep problem, but here's a simple framework you can use for making initial decisions. This is a matrix that classifies the type of knowledge that we want to pass along, and the people who might want to know it. We're using the standard RACI model here, where each lower level includes the higher levels. We can also assume that this matrix is for people outside of your immediate team! Whoever is involved with daily efforts probably wants to know what is going on first.

Is this just a regular ping that lets people know that things are okay? Then the Responsible folks want to hear. Think of the managers and architects that have committed to your team's success; they're the targets for automatic, not too frequent reports.

But what if the report is showing a problem? A growing issue that can still be corrected needs more focus, and so we expand the audience to Accountable people. This might be the directors or VPs who your team's managers report to, for instance; the people who will need to allocate resources or provide air cover if it goes wrong. They're called "Accountable" because they're on the hook even if they didn't do anything, so keep them advised.

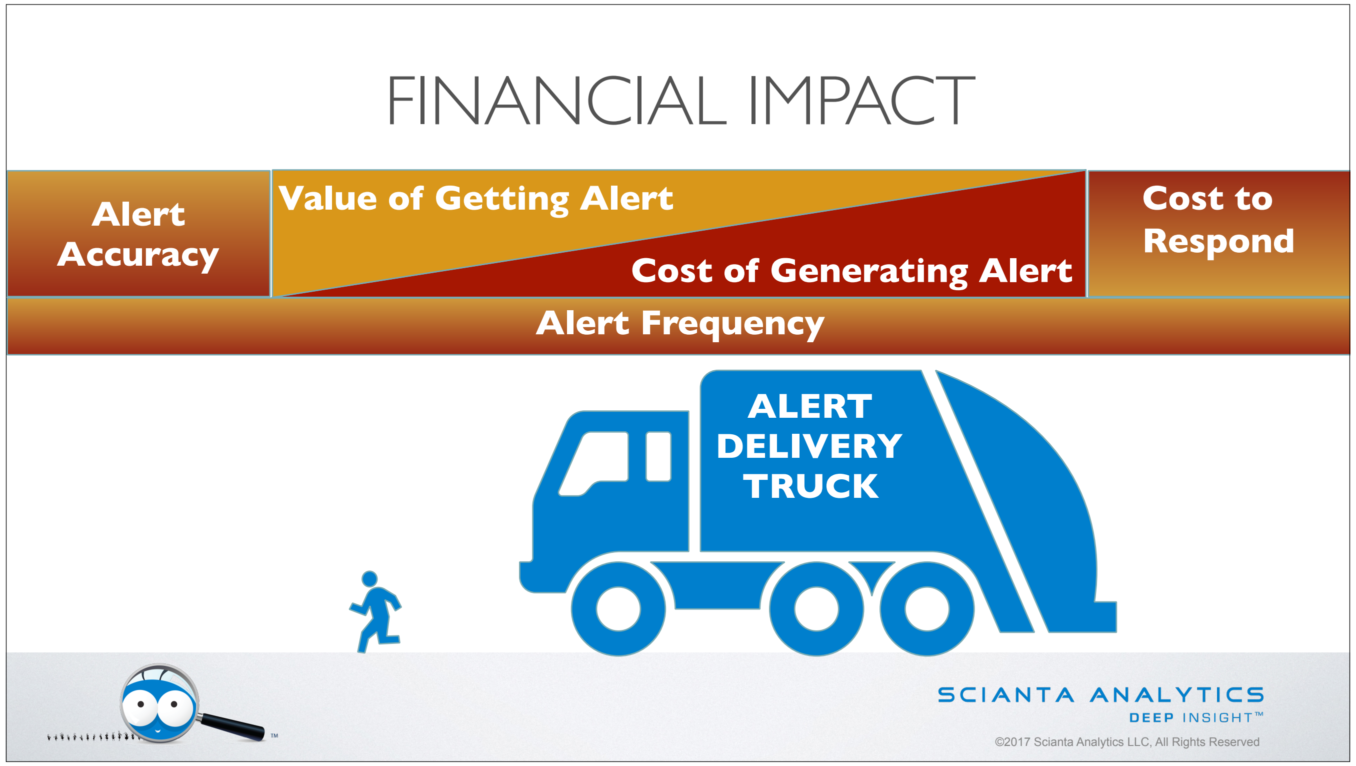
What about problems that you've got a handle on? You've solved it, you know how to solve it, it's an example of what you know how to do. These are an opportunity to demonstrate capability; expanding the audience to your Concerned folks is wise. This is the managers and leads who depend on you to make their projects successful. They're internal customers, either your best advocates or the end of your road.

Lastly… it's all gone wrong, no one knows why, and it might be getting worse. Your instincts say "hide"! This is the point where you need to broaden your audience, and provide calm, clear, accurate, and frequent reporting to the Informed set. They're all worried, they're all involved, and they all want to know

AGENDA

Introduction to Machine Intelligence	Data Handling 1	Data Handling 2	Anomaly Detection	Transactional Behavior	Impact Analysis
Academic Concepts	Collection	Retention	Anomaly Definition	Defining Transactions	Organizational Visibility
Data Systems	Storage	Format	Measuring Normality	Transaction Relationships	Types of Impact
Maturity Curve	Security	Labeling		Probability Measurement	Responsiveness

SCIANTA ANALYTICS
DEEP INSIGHT™



FINANCIAL IMPACT

Alert Accuracy

Value of Getting Alert

Cost of Generating Alert

Cost to Respond

Alert Frequency

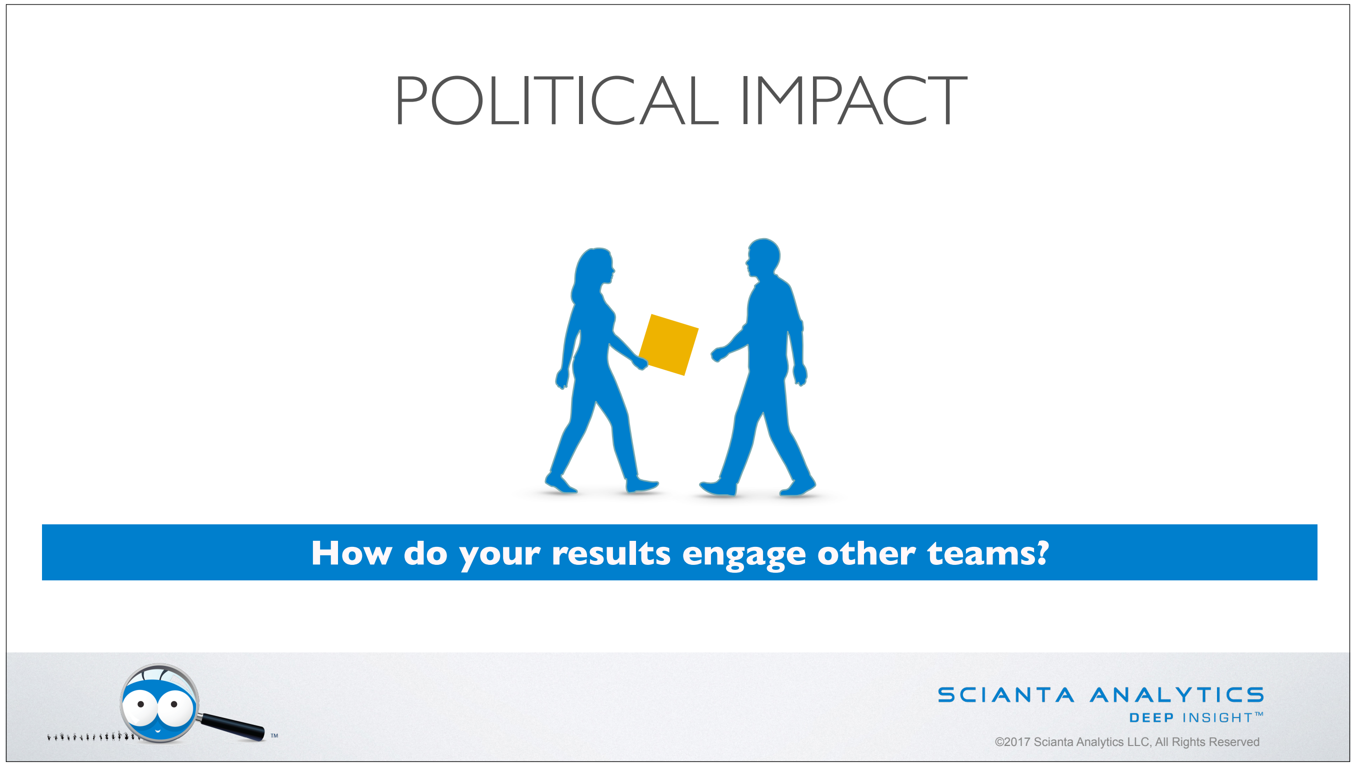
SCIANTA ANALYTICS
DEEP INSIGHT™

Figuring out how to generate an alert is not the entire picture, unfortunately. What is the value of the alert to your organization? If it saves the organization from certain doom, it’s a pretty good alert! Unfortunately a lot of alerts just remind you of what you already knew.

How much does it cost to generate that alert? Or respond to it? Worse yet, what if the alert isn’t perfectly accurate?

If the alert saves you five bucks, costs a dollar to generate, costs a dollar to respond to, and is 80% accurate, then you’ll find its net value grows slower than your gross costs. Keep adding that kind of alert, and soon your costs will exceed value.

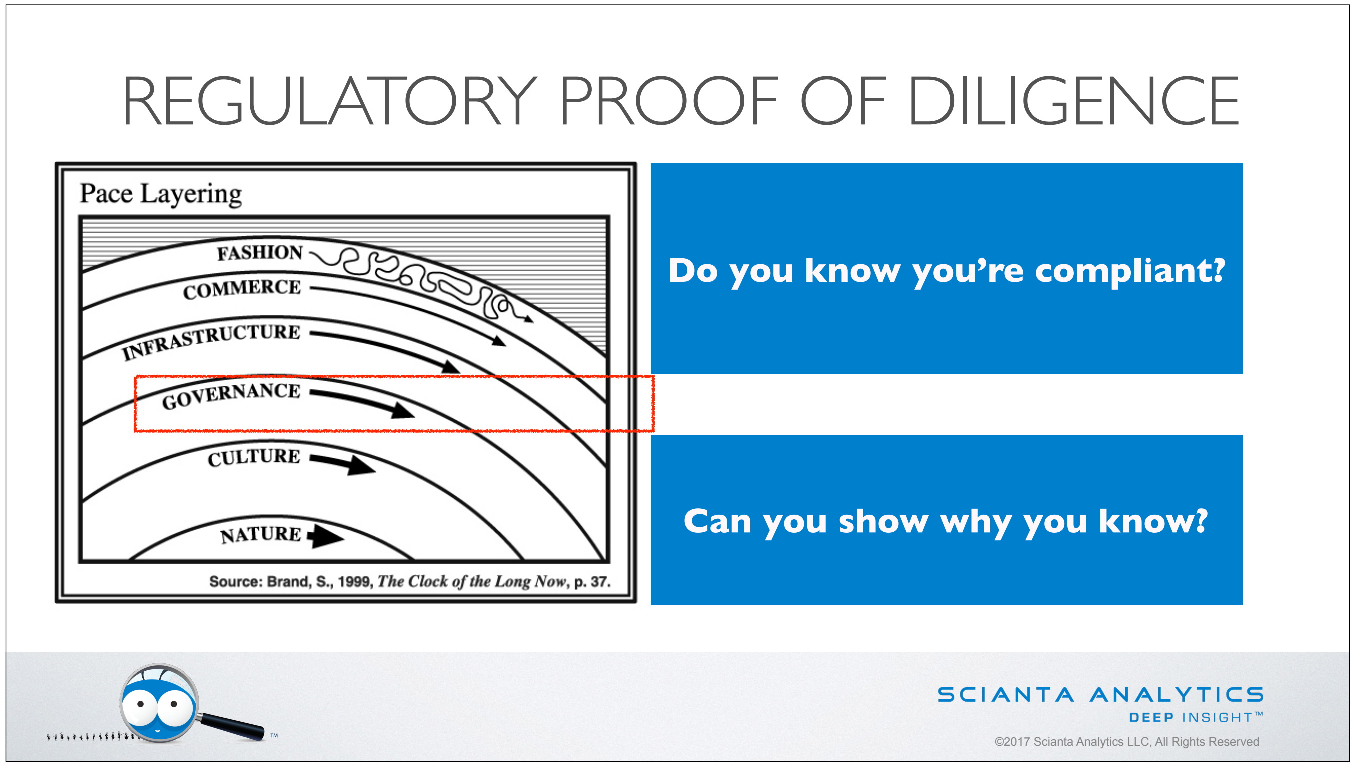
Every team has a certain amount of bandwidth available for responding to alerts. If that bandwidth is exceeded, then the alerts are no longer helping the team.



POLITICAL IMPACT

How do your results engage other teams?

Another angle to be aware of when generating alerts is what impact they have on other teams. Whether you're merely notifying or expecting action, the alerts that leave your team reflect on your team. You want them to be high value and high accuracy. But most of all, you want them to be easily actionable. A cognitive computing system is a relatively smart and expensive system compared to a simpler tool like a firewall or a solenoid. That means the ideal response to an alert from the cognitive computing system probably lies with another team that will execute on advice produced by your system. That other team won't thank you for false alarms and bad advice! It's critical to continually evaluate the strength of your cognitive computing system's recommendations, as well as their eventual correctness when measured against reality.



REGULATORY PROOF OF DILIGENCE

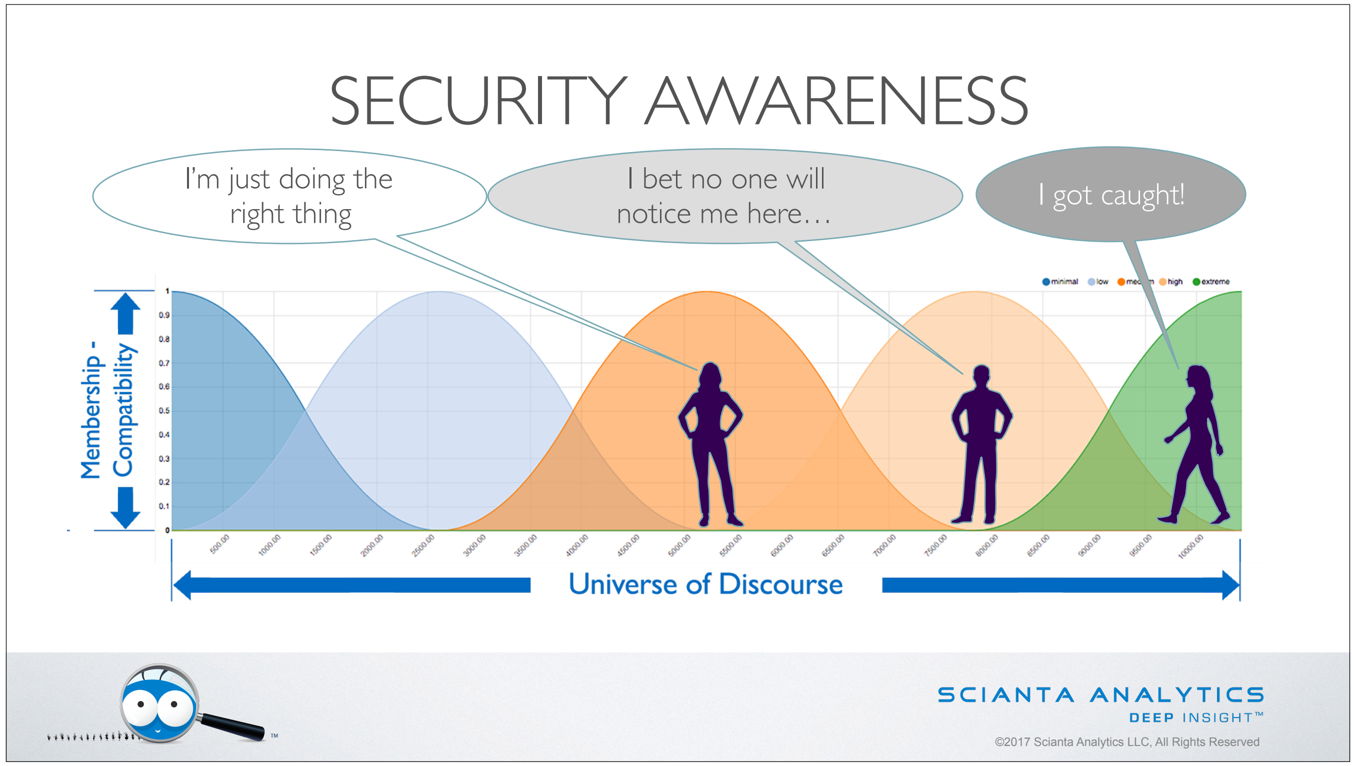
Source: Brand, S., 1999, *The Clock of the Long Now*, p. 37.

Do you know you're compliant?

Can you show why you know?

The ultimate "other team" is the regulatory compliance regime that affects your organization. Every organization operates in a forest of overlapping requirements, and often there's an auditing requirement as well. When the time comes, you'll want to be able to show that your cognitive computing system provides good guidance and doesn't break the rules.

The principle that allows this is "discoverability" — you want to be able to discover why the cognitive computing system has decided to issue an alert, so you should be recording every signal with a key that allows it to be traced in relation to other signals. Since alerts from a cognitive computing system are often driven by many different tests working in concert, it's critical to be able to start at an alert and discover each of the elements that went into that alert.



SECURITY AWARENESS

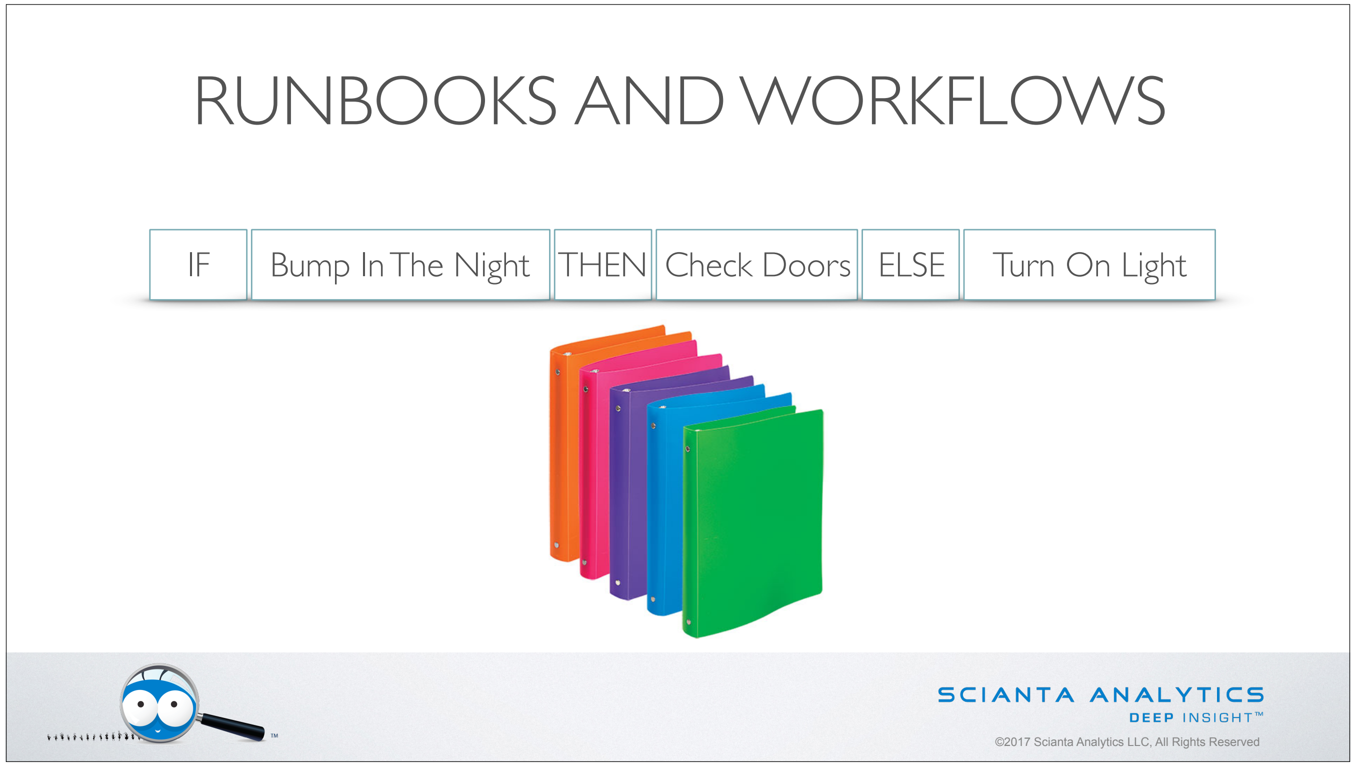
SCIANTA ANALYTICS
DEEP INSIGHT™

The biggest challenge facing a cognitive computing system probably won't surprise you. It's human intelligence. A human attacker, whether inside or outside of your environment, can be a very tough adversary indeed. There is no silver bullet answer, because every organization has its own, ever-changing threat landscape and attackers. But, a qualitative expression system can help quite a bit by letting you write flexible rules that respond to observed data, uncover outliers, and measure many risk factors in correlation to each other. A good cognitive computing system enables forensic teams and threat hunters, who in turn enable front line security analysts.

AGENDA

Introduction to Machine Intelligence	***Data Handling 1***	***Data Handling 2***	***Anomaly Detection***	***Transactional Behavior***	***Impact Analysis***
Academic Concepts	***Collection***	***Retention***	***Anomaly Definition***	***Defining Transactions***	***Organizational Visibility***
Data Systems	***Storage***	***Format***	***Measuring Normality***	***Transaction Relationships***	***Types of Impact***
Maturity Curve	***Security***	***Labeling***		***Probability Measurement***	***Responsiveness***

SCIANTA ANALYTICS
DEEP INSIGHT™



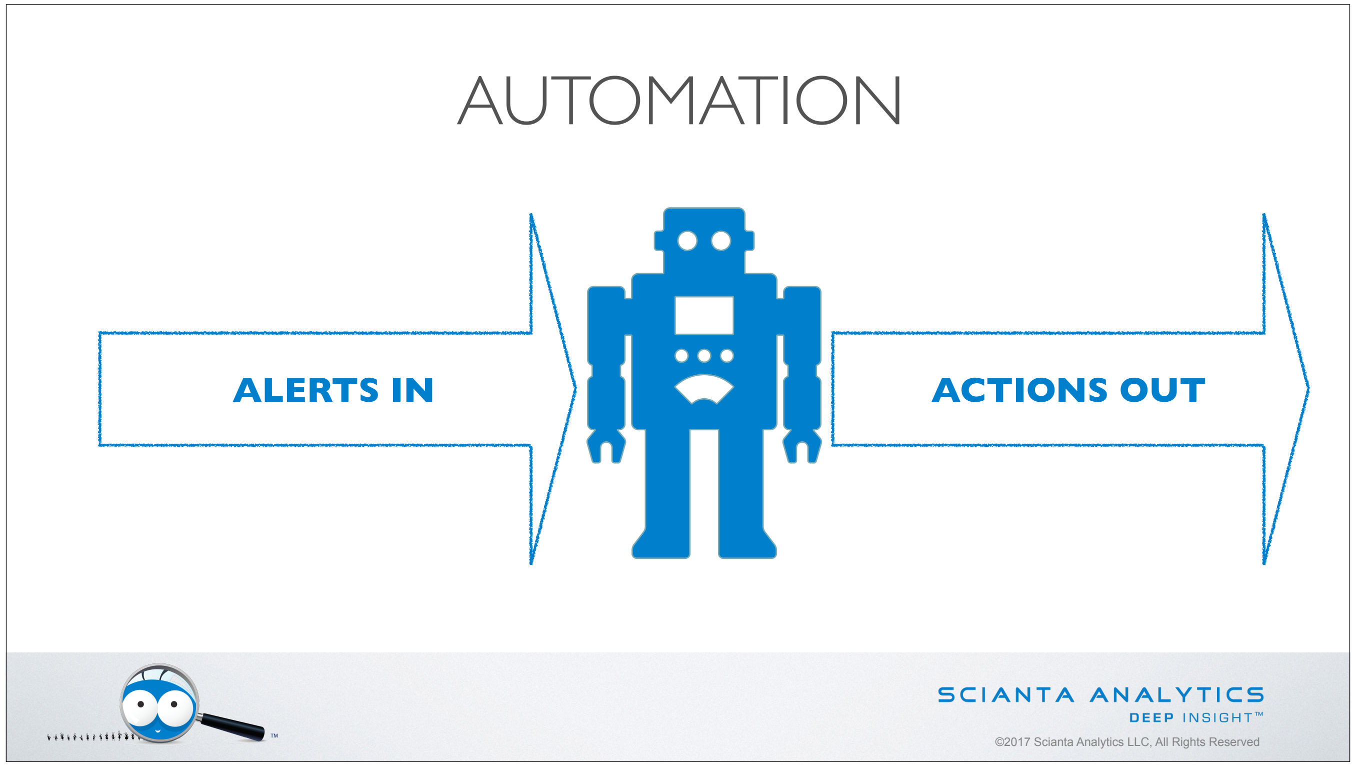
RUNBOOKS AND WORKFLOWS

IF | Bump In The Night | THEN | Check Doors | ELSE | Turn On Light

SCIANTA ANALYTICS
DEEP INSIGHT™

We've touched on a number of different results that could occur from an alert, but this is potentially an entire area of its own. While there are certainly many areas that need human intuition and experience, sometimes there are opportunities to give that human guidance. Many organizations provide runbooks to support portions of their processes, or in some cases allow automated workflows to support processes from software.

These automation or partial-automation systems may impose additional requirements on your cognitive computing system, ranging from formatting results to allowing script executions from within the system. Investigate these requirements in light of the financial and political impacts discussed in this session.



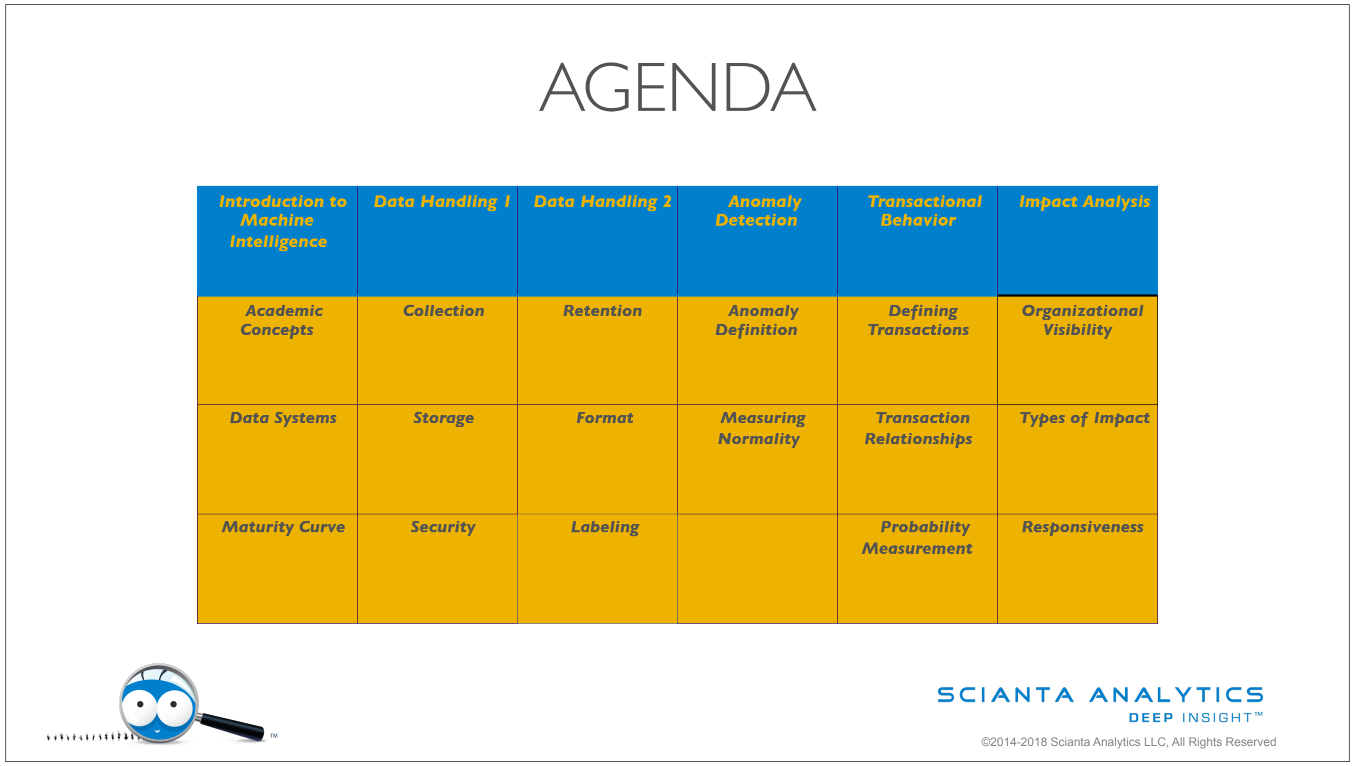
AUTOMATION

ALERTS IN

ACTIONS OUT

SCIANTA ANALYTICS
DEEP INSIGHT™

Once we've begun to evaluate runbooks and workflows, it's a slippery slope to scripted automation. This is a beneficial approach for many processes, and automating simple activities saves resources for more complex work. It is nevertheless very important to consider automation carefully when applying it to cognitive computing systems. These systems are strong at behavioral analytics, anomaly detection, finding correlations, and balancing risk. Those are highly valuable processes that can be prone to confirmation bias and classification errors. It is highly advisable to retain human oversight when automating responses. If you're not 100% sure that you understand every possible input and output of a system, then it only makes sense that you will treat that system's output as recommendations, not commands.



AGENDA

Introduction to Machine Intelligence	Data Handling 1	Data Handling 2	Anomaly Detection	Transactional Behavior	Impact Analysis
Academic Concepts	Collection	Retention	Anomaly Definition	Defining Transactions	Organizational Visibility
Data Systems	Storage	Format	Measuring Normality	Transaction Relationships	Types of Impact
Maturity Curve	Security	Labeling		Probability Measurement	Responsiveness

SCIANTA ANALYTICS
DEEP INSIGHT™

We’re all done with our overview! Make sure to look into our advanced trainings and seminars at https://www.scianta.com/learn

Thank you!