✓ Making **Data Classification** Work for You

18 Things to Consider
When Choosing a Data
Classification Solution

Yuval Eldar, CTO

September 2014

secure islands

Why Data Classification?

If you’re reading this, there’s probably no need to explain the importance of data classification in your enterprise information security toolbox. The question is likely not “Does my organization need data classification?” but rather “Which data classification solution is right for us?”

Like any enterprise-level tool, data classification systems are complex and far-reaching. At the same time, ease of implementation is mission critical since the system needs by definition to interact with multiple other enterprise systems, and ease-of-use is even more important, since the solution is user-facing.

To help cut through the confusion, the information security experts at Secure Islands have put together the following list of tips and questions to ask when choosing a data classification solution.

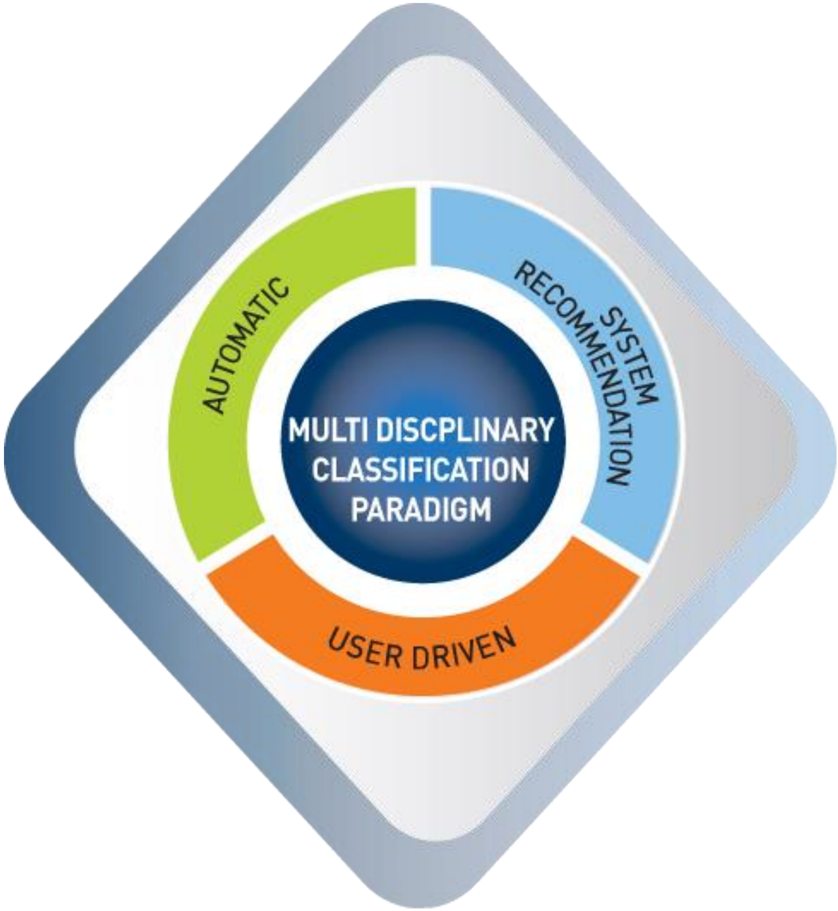
Tip 1: Choose a Hybrid

Much of your sensitive information can be deterministically classified with an intelligent, learning, automatic classification engine with minimal end user friction. At the same time, much will always need to be classified manually.

Make sure you choose a hybrid solution that offers:

- Automatic and transparent classification
- User-determined, manual classification
- A recommendation option which suggests classification options for the end user to confirm

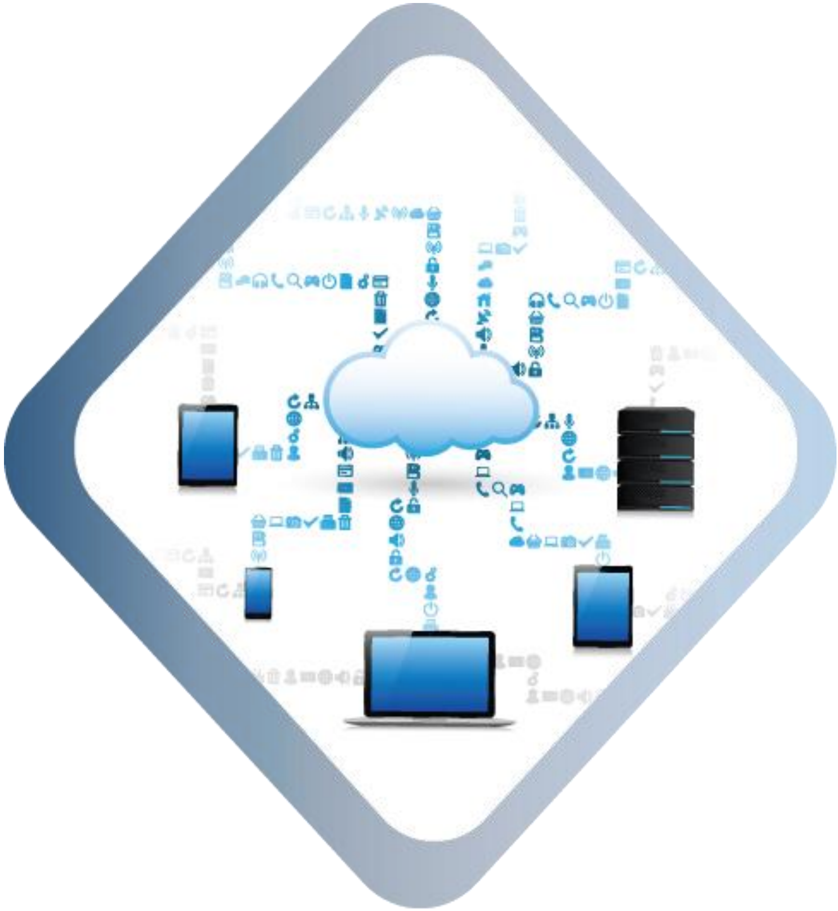
Moreover, selection of the classification methodology for each instance (automatic, manual, user prompt) should be itself automatic, based on data identification.



Tip 2: Policy-Driven Classification Analysis

When classification is automatic, it should be based on real-time analysis of content (phrases and patterns, thresholds, checksums, etc.), context (where is the information from, where is it going, who created it, what geolocation, etc.) and source.

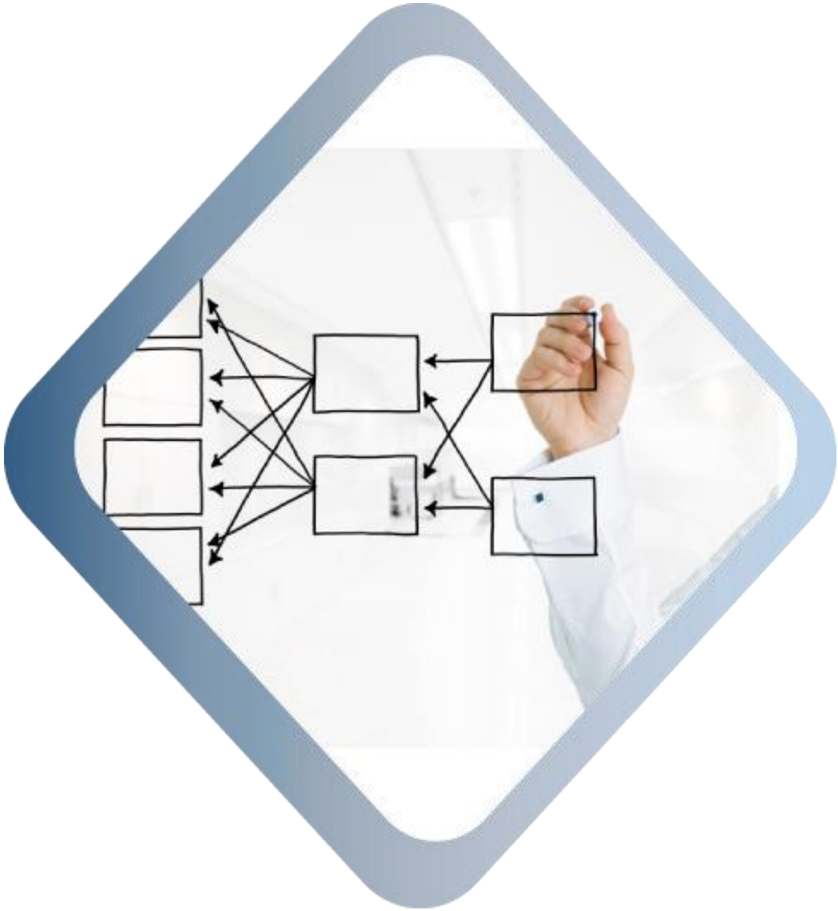
For each type of analysis parameter, your classification solution should allow highly-granular, policy-driven control.



Tip 3: Any Source

Sensitive information is everywhere in your organization, not just in commonly-protected applications.

Your data classification solution should intercept data and seamlessly classify content from many different sources, including cloud solutions, ECM (Enterprise Content Management) software like MS SharePoint, enterprise applications, storage networks, and all types of user-generated content.



Tip 4: Classification Triggers

To achieve the flexibility that complex business processes require, you need highly-granular control over the data interception events that trigger data classification.

For example, can your solution define where and when exactly classification occurs? On save? On upload to a specific location or service like Dropbox or SharePoint? On file open? On attachment to email via drag and drop? On copy between folders in Windows Explorer?

Make sure classification triggers are completely customizable, work in any application, and are policy-driven, enterprise-wide.

Tip 5: Beyond MS Office

Your organization runs on multiple applications from multiple vendors, not just on MS Office.

Make sure that the data classification solution you choose works smoothly and offers a seamless and uniform user experience in any application - from Adobe Acrobat, through CAD/CAM software, and everything in-between – not just MS Office utilities.

Tip 6: What about Pre-Existing Content?

There are millions of files in your repositories, many created long before you even thought of data classification.

Your data classification solution should be able to find and classify content generated in the past, as well newly-generated content.

More specifically, as part of the initial data classification implementation, your solution should scan your entire data repository to identify and classify valuable data - delivering immediate value to your enterprise.

✓ Tip 7: Classification Logic

Data classification does not exist in a vacuum. It is a critical part of your business processes, and is directly affected by evolving enterprise business strategy!

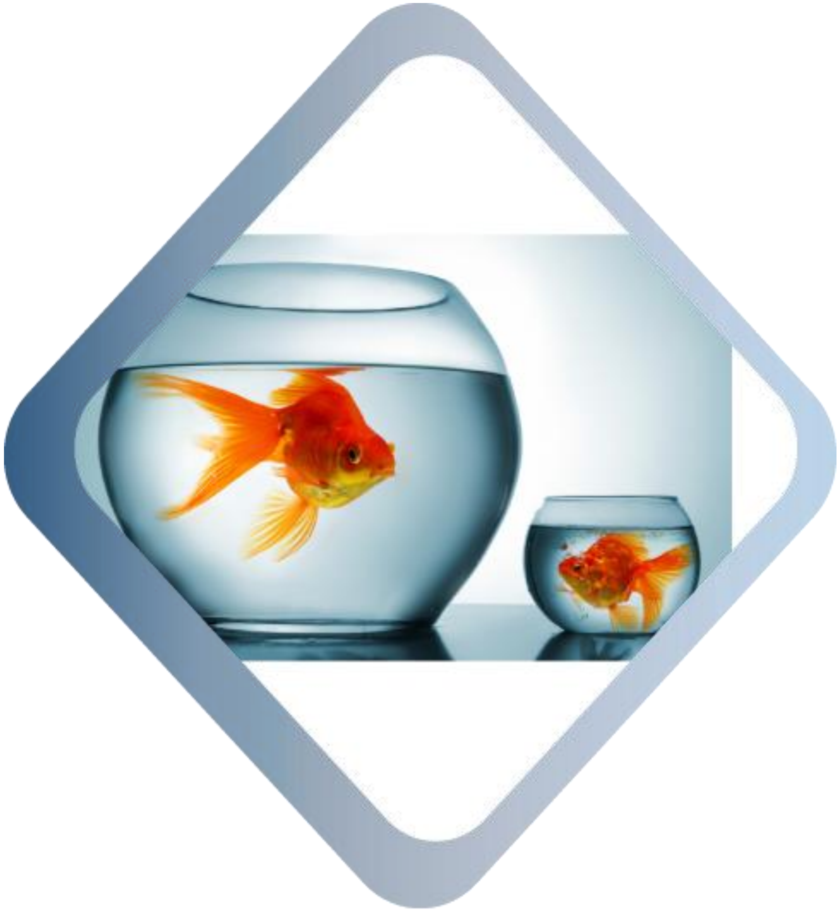
Make sure that data classification lifecycles and permissions are policy-driven, so they can remain in-line with changing business logic.

For example, can your classification policy specify who can increase or decrease the sensitivity of a given document? Who can declassify? Who can make classification mandatory or optional?

✓ Tip 8: Not One-Size-Fits-All

In large enterprises, different organizational units require different classification taxonomies.

Your data classification solution should enable business units, regional offices, and other semi-autonomous business entities to define their own classification policies.



Tip 9: Dynamic Classification Matrix

Data classification is a multi-layered, multifaceted art. Don't settle for a rigid solution that makes your organization adapt to preset classification attributes.

Make sure that you choose a solution that is flexible enough to adapt to your way of doing business. This can measurably impact both implementation and security.

Classification Levels
Lifespan
Jurisdiction
Projects
Hierarchy
Scope
Archive

Tip 10: Reporting and Analysis

Like any mission-critical security solutions, an enterprise-level data classification system must include extensive reporting, analysis, auditing, forensics, and risk assessment functionality.

For example, can your data classification solution identify with high granularity where exactly customer data is stored? Can it tell you where a given sensitive document was emailed most recently? How it was used before it was sent, and if it was reclassified?

secure islands

Tip 11: Leverage across Multiple Systems

To preserve investment in strategic enterprise tools, it’s a given that your data classification tool should integrate seamlessly with your DLP, archiving, e-discovery, and other enterprise solutions.

Moreover, make sure that these same enterprise systems can leverage data classification to extend their own native capabilities - enriching information management strategies, archiving and data retention, SharePoint categorization, search optimization, and more.

Tip 12: Flexible Enforcement

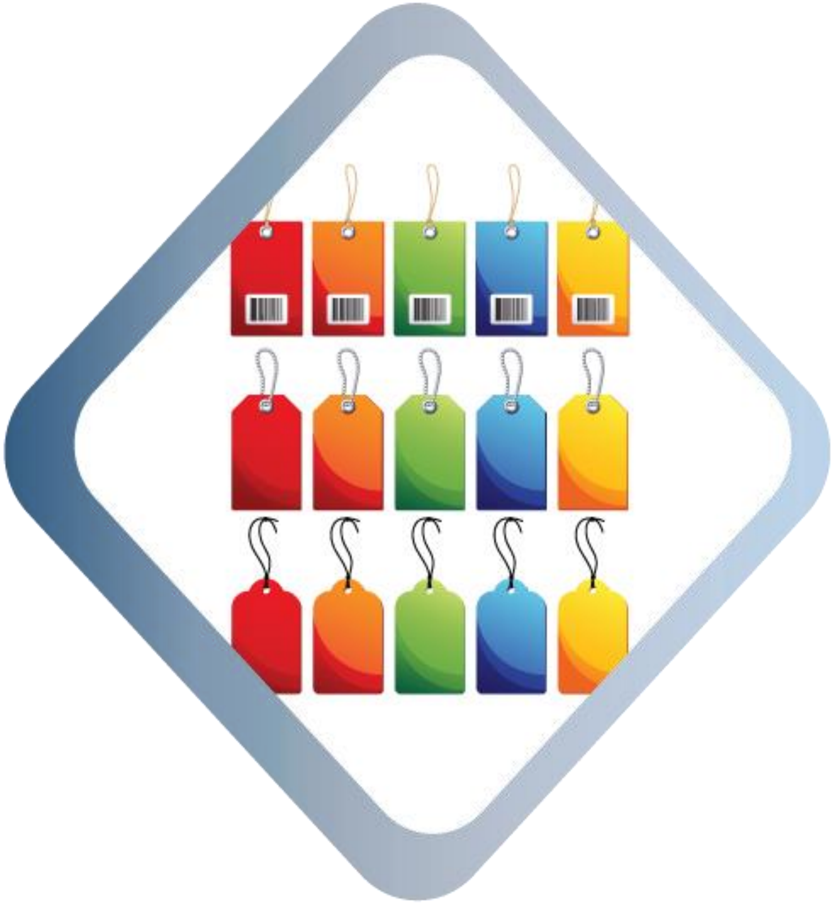
Your data classification solution should have built-in flexible and extendable enforcement capabilities, covering the entire sensitive information lifecycle.

For example, what happens exactly when information classified as sensitive is accessed or sent? Does your solution allow you to define whether requests should be blocked, allowed with automatic encryption or IRM protection, or just warned?

Tip 13: Persistent Tagging

Once classified, data needs to retain its classification no matter where it is in the data lifecycle – in use, in motion, in storage, anywhere.

For example, does cutting and pasting a file from a local drive to a USB drive remove classification tags from sensitive information? Does sending a classified PDF file via Outlook nullify classification? It shouldn't!



✓ Tip 14: Anti-Tampering

Although this seems like a given for any security solution, make sure that your data classification solution prevents users from maliciously removing or changing classification attributes without proper authorization.

Ensure that your data classification solution can provide alerts to a centralized auditing system, if such malicious activities are identified.

✓ Tip 15: Esperanto Not Spoken Here

A multinational organization needs a multilingual data classification solution. The solution you choose should not only classify multilingual data, but also have a multilingual user interface.

✓ Tip 16: Branding

Your brand is who you are, both to the outside world, and to your trusted internal users and partners. Like any end user-facing system, the user interface of your data classification system should be fully customizable to your brand's look and feel

✓ Tip 17: SIEM\SOC Compatibility

To avoid multiple points of control for key security systems, you have probably invested in a SIEM or SOC solution.

Treat your data classification solution just like any other mission-critical security system, and make sure it integrates seamlessly with your SIEM\SOC of choice.

✓ Tip 18: Truly Enterprise-Grade

Does your data classification solution offer a truly enterprise-grade feature set, including centralized classification policy management, seamless Active Directory integration with multi-forest capabilities, role-based administration, and health and operational monitoring components? Does it meet high-availability standards, offer load balancing, and support clustered deployment?

Conclusion

Reliable data classification is a key enforcement enabler for any enterprise information security policy. Advanced, enterprise-grade data classification packages provide an end-to-end solution that complements and extends existing security tools.

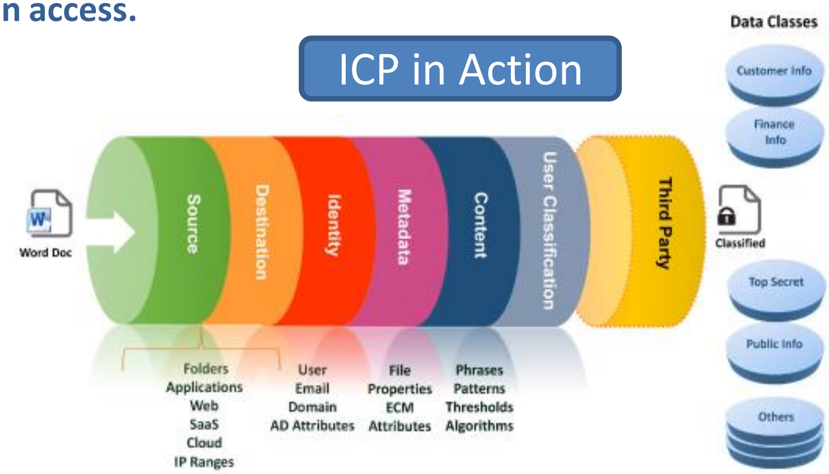
By choosing a data classification tool that works for you – and not the other way around – you can facilitate the secure, smooth flow of information within your borderless enterprise.

Secure Islands Information Classification

Based on unique and patent-pending Information Classification Prism™ (ICP) technology, and already implemented in large enterprises worldwide, Secure Islands automatically classifies sensitive content at creation or initial organization access.

Seeking out and classifying content from any source - endpoints, applications, server applications, mail systems, storage devices, the cloud, and more – Secure Islands then persistently tags sensitive data. This enables highly-flexible, policy-driven enforcement throughout the data life cycle – from data usage monitoring, user warning, blocking, and including persistent encryption via IRM and other enterprise solutions

ICP in Action



Secure Islands delivers:

- 100% classification accuracy powered by content and context analysis
- Multidisciplinary classification mechanisms – user, system recommendation, automatic
- Source-based classification between data source and destination
- Optimized classification cycle triggered by intercepted events (open, close, save, upload, download, copy, etc.)
- Full analytics of data usage events for all classification attributes - for enforcement, reporting, and audit