DARK MATTER:
SHARED DRIVE FILE ANALYSIS APPLIED TO UPSTREAM OIL & GAS

24th September 2015

Matthew Gordon
Lockheed Martin (Information Systems & Global Solutions)

Tel +44(0)1224 611368
Mob: +44(0)7720 535516
E: matthew.Gordon@civil.lmco.com

LOCKHEED MARTIN

100 YEARS OF ACCELERATING TOMORROW

WHAT IS FILE ANALYSIS?

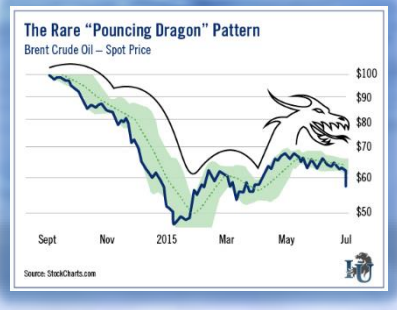
> File analysis tools analyse, index, search, track and report on file metadata and, in some cases, file content, enabling organisations to take action on files according to what was identified.
>
> *Gartner, Inc. | G00262949*

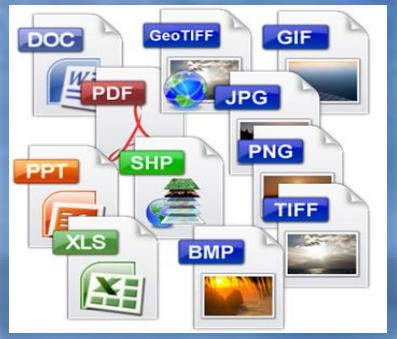
Main Reasons for usage:

1. Increase Operational Efficiency
2. Lower Cost
3. Mitigate Risk

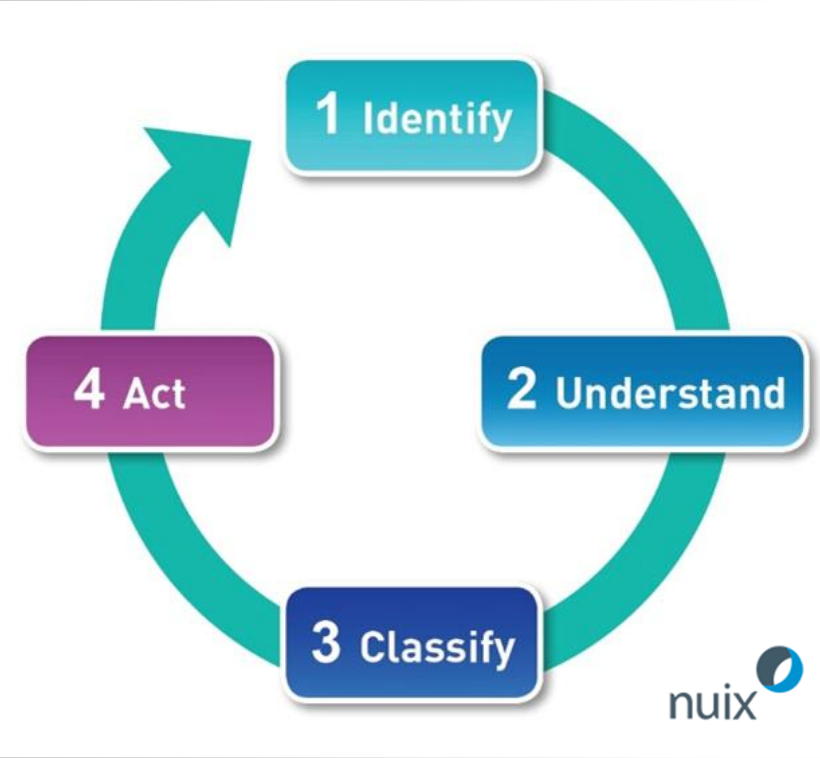
LOCKHEED MARTIN PROPRIETARY INFORMATION

THE PERFECT STORM





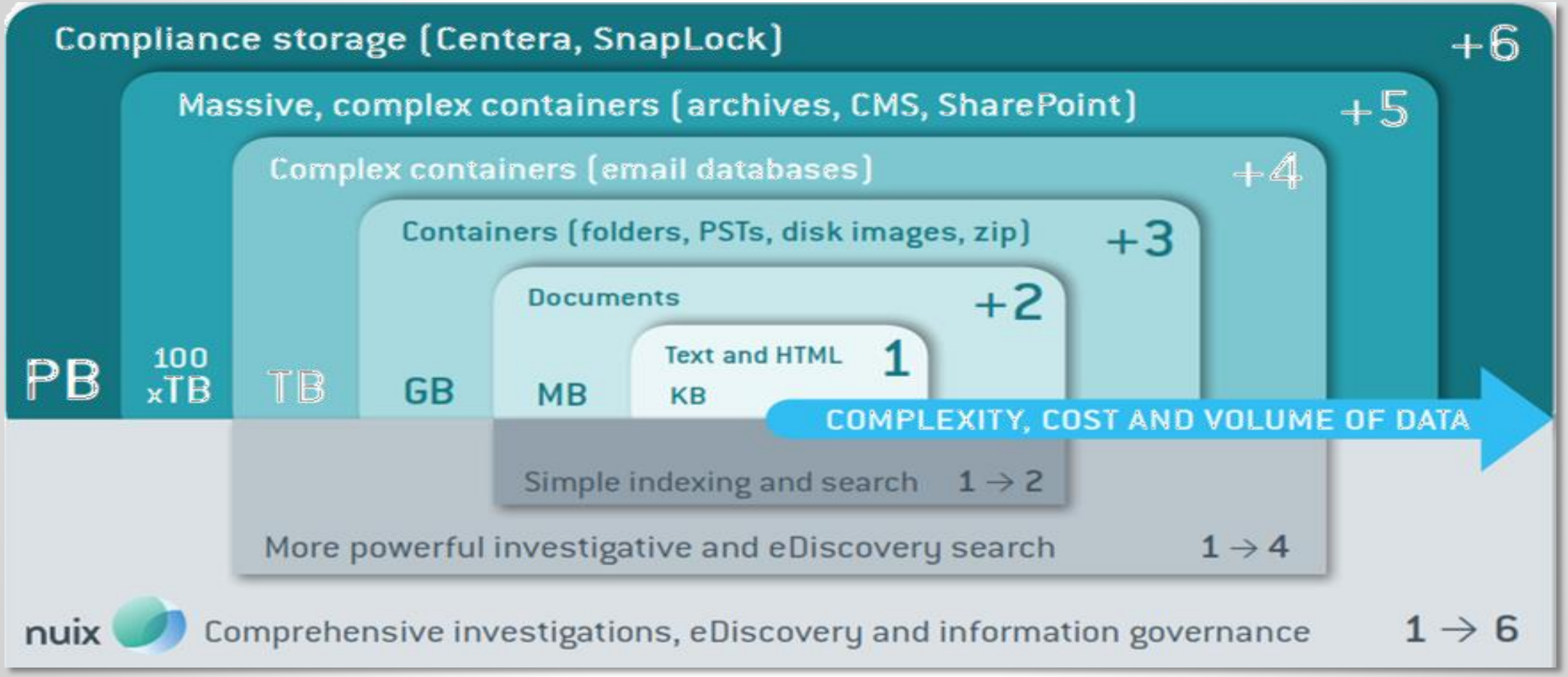
FOUR STAGES OF FILE ANALYSIS



1. **Identify** the content and context of unstructured information scattered across myriad locations
2. **Understand** the age, ownership, format and content of each item
3. **Classify** categorise the content based on informed governance decisions (discipline, purpose, value, versioning, criticality, sensitivity, etc.)
4. **Act** on those decisions by optimising storage systems; enriching, migrating and protecting data; deleting defensibly or making data more readily available for disclosure or information sharing requests

LOCKHEED MARTIN PROPRIETARY INFORMATION

SIX LAYERS OF UNSTRUCTURED DATA



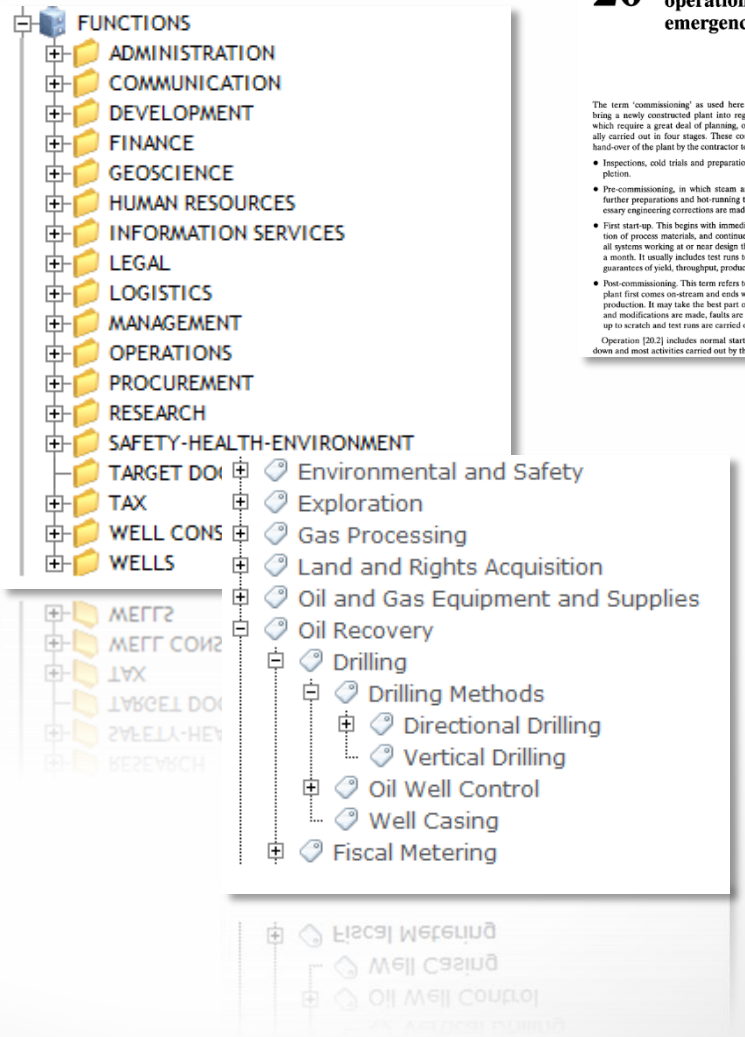
LM FILE ANALYSIS SOLUTION – WHAT IS IT?

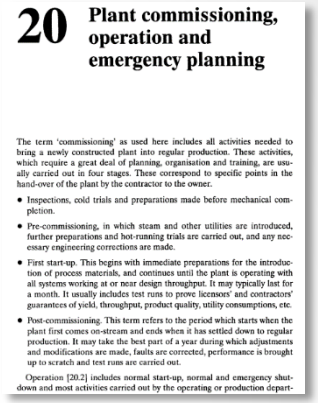
Customer & project specifics

nuix ENGINE

Predefined O&G workbench, Taxonomies, Searches & Actions

Upstream IM Expertise & know-how



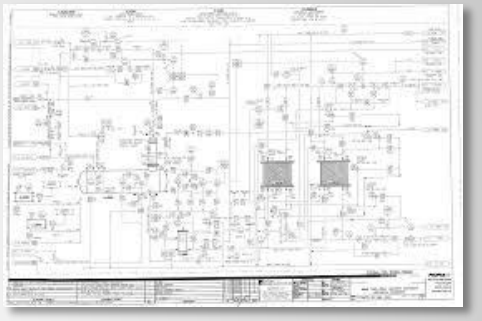


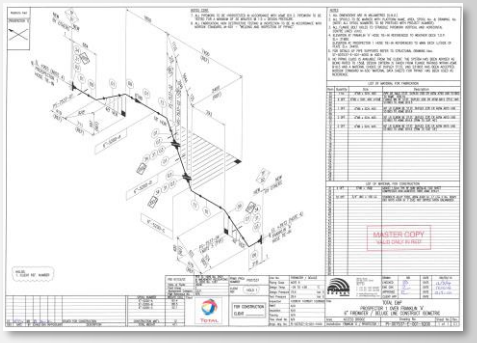
20 Plant commissioning, operation and emergency planning

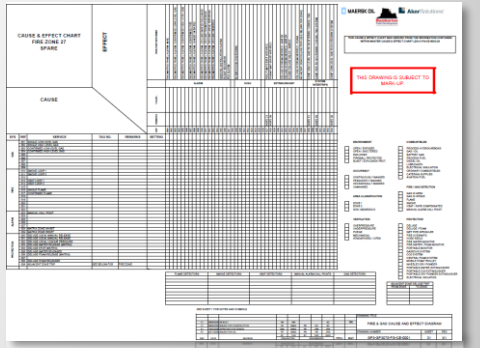
The term 'commissioning' as used here includes all activities needed to bring a newly constructed plant into regular production. These activities, which require a great deal of planning, organisation and training, are usually carried out in four stages. These correspond to specific points in the hand-over of the plant by the contractor to the owner.

- Inspections, cold trials and preparations made before mechanical completion.
- Pre-commissioning, in which steam and other utilities are introduced, further preparations and hot-running trials are carried out, and any necessary engineering corrections are made.
- First start-up. This begins with immediate preparations for the introduction of process materials, and continues until the plant is operating with all systems working at or near design throughput. It may typically last for a month. It usually includes test runs to prove licensors' and contractors' guarantees of yield, throughput, product quality, utility consumptions, etc.
- Post-commissioning. This term refers to the period which starts when the plant first comes on-stream and ends when it has settled down to regular production. It may take the best part of a year during which adjustments and modifications are made, faults are corrected, performance is brought up to scratch and test runs are carried out.

Operation [20.2] includes normal start-up, normal and emergency shut-down and most activities carried out by the operating or production depart-







BACKGROUND

Timescales:
5 man-days conducted across 3 weeks (April/May 2015)

Scope:
Analyse ~1TB of Upstream Shared Drive Information

Business Objectives:
- Reduce Costs associated with Shared Drives
- Find misfiled information of corporate value (potential records)

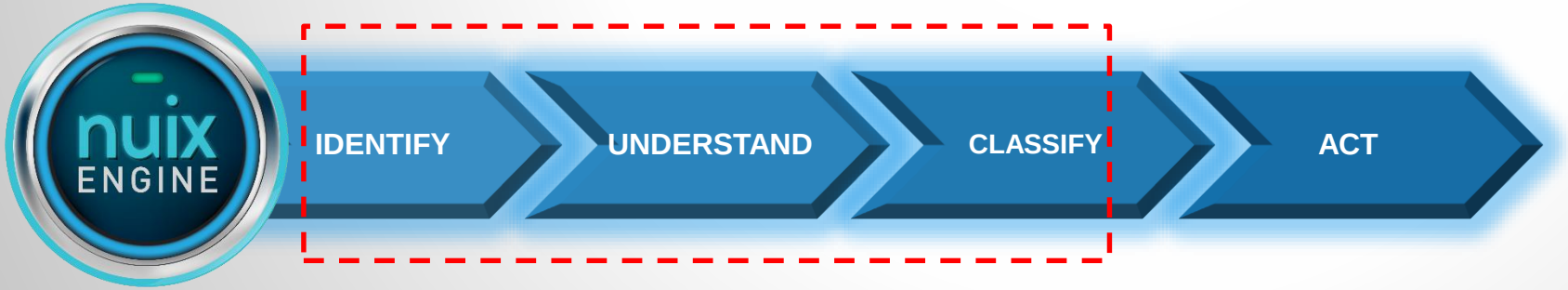
Success Criteria:
- Identify 40%+ ROT / Duplicates
- Use Taxonomy to Auto Classify

LOCKHEED MARTIN PROPRIETARY INFORMATION

METHOD

- Customer-provided laptop, LM Nuix Solution installed
- Initial scan: 700GB ingested and analysed on site
- OCR performed
- Taxonomy applied
- Company-specific pattern matching
- Review and refine classification
- Present findings

5 man-days effort



LOCKHEED MARTIN PROPRIETARY INFORMATION

SUMMARY OF FINDINGS

Analysis of 700GB (Upstream Shared Drive Data)

- 5 man-days (conducted across 3 weeks, Q1 2015)
- ~50% ROT and Duplicate Information identified
- ~7% potential records identified

Business Case:

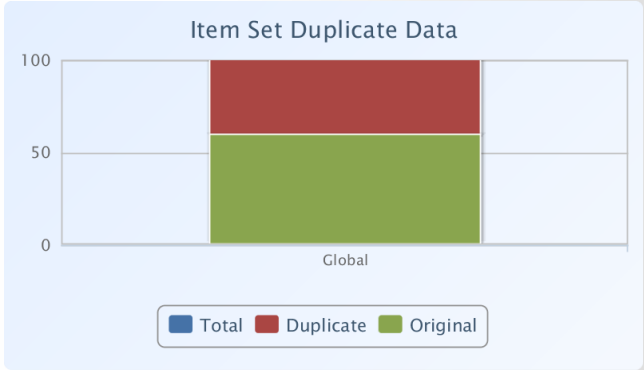
- 3 Year ROI: 290%
- Payback: 10-14 months

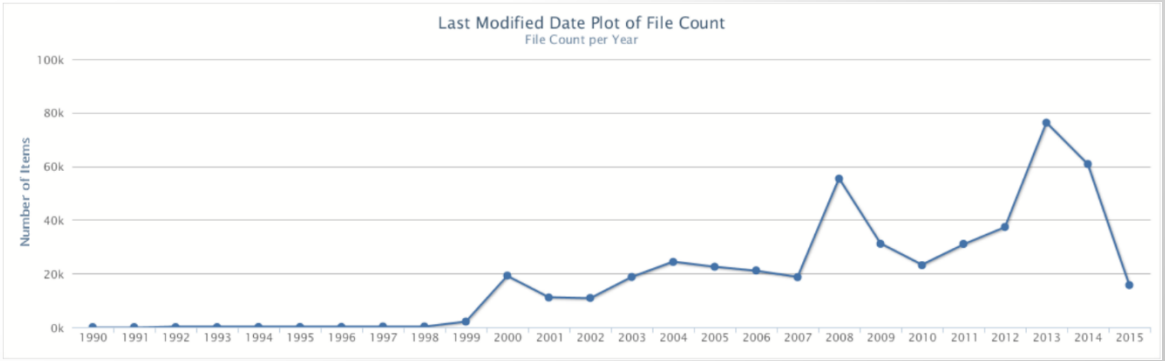
Risks

- Non-searchable PDFs: 43,000 files
- Corrupted: 1,400 files
- Encrypted: 500 files
- Credit Card Nos.: 1,160 files
- Aged documents:

Duplicates & ROT

- 41% Item Set Duplication
- 7% Additional ROT





SUMMARY OF FINDINGS

CLASSIFICATION:

Lockheed Martin Upstream O&G Taxonomy

Identified potential records:

- By Asset
- By Discipline

Prioritise IM areas of interest:

- Asset 1
- Asset 2
- Safety
- Electrical
- Schematics / Layouts

	Electrical->Control Schematic	Electrical->Protection Schematic	Electrical->Switchboard / Distribution Board GA, Layouts & Schedules	Instruments->Cause & Effect Diagram	Structural->Platform & Site GA / Plot Plan / Elevation	HSEQ->Fire & Gas Layout	Electrical->Cable Schedule	Subsea & Pipelines->Field / Subsea Layout & Approaches	Electrical->Cable Layout/Routing Drawing	Electrical->Equipment Interconnect and Termination	Process->Process Data Sheet	Technical Safety->Safety Case Documentation
[redacted] Platform	1,798	1,622	1,348	1,204	829	828	778	766	764	743	599	529
[redacted] Platform	1,354	1,117	1,037	701	828	818	780	303	891	702	605	394
[redacted] Platform	1,311	1,189	638	754	855	472	726	314	586	722	683	288
[redacted] FPSO [redacted]	1,274	1,003	1,261	896	284	359	541	633	424	502	390	273
[redacted] FPSO [redacted]	1,158	967	723	575	445	266	401	559	352	528	353	223
[redacted] Platform	873	721	598	484	488	254	339	298	283	552	386	285
[redacted] Platform	928	593	413	430	461	478	363	86	361	523	319	253
[redacted] Platform	921	858	517	419	661	164	471	270	381	465	371	228
[redacted] FPSO	596	484	485	313	195	162	227	237	161	246	164	136
[redacted] Platform	483	520	239	240	283	116	278	150	225	212	230	143
[redacted] FPSO [redacted]	490	432	362	204	219	116	223	221	227	239	153	72
[redacted] Platform	346	335	207	233	201	90	211	136	153	156	106	120

SUMMARY OF FINDINGS

PATTERN MATCHING:

Identifying Records via company specific document numbering standard

- Matched 488 of 8,000 documents scanned

Name	File Type	Document Number(s)
A11-30-SH-0008 A11 Ops Guide A01 Approved.docx	Microsoft 2007 Word Document	56962-AE-P-CB-6010
July 1st 2011 A11-30-SH-0008 R02.docx	Microsoft 2007 Word Document	56962-AE-P-CB-6010
TN Pierce-Brynhild RBM FIV Erosion Flow Restriction MPFM Piping r0 BOP-25-SH-004.docx	Microsoft 2007 Word Document	21480-[redacted]-RA-0009
Nemo Report Liam input.docx	Microsoft 2007 Word Document	56962-AE-P-RB-0018
ORP04210 Krabbe Concept Study Report 25April.docx	Microsoft 2007 Word Document	56962-AE-P-RB-0018
Comments_Bryn Manifold P&ID.docx	Microsoft 2007 Word Document	21480-LUNAS-O-FD-0001

- Enables enhanced classification & metadata enrichment
- Easily extensible:
  - Apply other document, tag and well numbering standards
  - Industry standards + company specific

OTHER OPPORTUNITIES & USE CASES

- Apply toolset to:
  - SharePoint
  - EDMS
  - Email
  - Archives

- Validation and enrichment of metadata in SharePoint
- Auto-classification (contribute to ever-greening)
- Scalability at speed: GB – TB – PB (local, regional, divisional, group)
- Records Management (identify records = reduce risk, improve classification, improve search)
- Divestments & Acquisitions
- Portable storage devices / archives
- eDiscovery

LOCKHEED MARTIN PROPRIETARY INFORMATION

SUMMARY

1. We’re in the midst of a crisis: the time is right for this to make a real impact to your business
2. Don’t aim for perfection – look for the biggest, fastest wins and ensure they deliver (then repeat)
3. Connect the business with IM SME’s to succeed