The Preservation of your Data

SQA Conference March 2017

Keith Williams

Formpipe.

Keith Williams
Sales Director, Formpipe (Life Science)

An entrepreneur, director, and business manager with UK, European and US experience. He brings over 25 years of Life Sciences experience, particularly in a pragmatic approach to getting computerised systems compliant and keeping them compliant. He has worked in a manufacturing, laboratory and clinical environment and has more recently focused this experience to build compliant set of configurable SharePoint products for content and document management. He has a BSc in Microbiology and an MSc (Eng) in Biochemical Engineering.

www.formpipe.com/lifescience

Today

- Definition of Electronic Preservation

- What are the Issues?

- Paper Records

- Electronic Records

- Digital Preservation

- Some Cost-effective Solutions

Formpipe.

Definition of Electronic Preservation

Ensuring that electronic records and electronic signatures remain accessible through appropriate preservation policies and processes and timelines. In the context of electronic archiving, this means maintenance of the authenticity and integrity of electronic records and electronic signatures.

Formpipe.

What are the issues?

- Multiple sources of records

- Paper Records - Can we, and if so, how do we get them into electronic format?

- Electronic Records - System Lifecycle vs Record Lifecycle

- Digital Preservation - multiple file types, different record lifecycles

- The hidden cost – and risks – to the business of the preservation activity

- Electronic signatures - what do we do with these? (Not covered today)

www.formpipe.com/lifescience

Multiple Sources and Variety of Records

Multiple Electronic System Repositories

All have different regulatory and legal retention requirements

What to Do with Paper Records?

1. **Paper records without signature**

   - Paper records may be destroyed and a certified electronic copy may be retained in their place.

   - Paper records that originated in an electronic format may be destroyed and the original electronic record may be retained in their place.

   - Paper records which are printed from specialized instruments, and where the results are recorded electronically, may be destroyed.

2. **Paper records with signature**

   - A risk assessment must be conducted of the applicable legal and regulatory framework related to the retention of originals and certified copies. If the retention of original records is mandated by regulations or legislation, any originals bearing wet-ink signatures, seals, watermarks, or other physical markings such as hand-written notes should be retained.

Checking the New Electronic record after the Scanning Stage

A quality check with sign off should be carried out to confirm that all content on the original record has been transferred to the copy.

This will include checks for:

- inclusion of headers, where present
- inclusion of footers, where present
- inclusion of page numbers, where present
- correct rotation of pages
- absence of skewing
- all pages are present
- both sides of any double-sided page have been scanned

Electronic Record
System Lifecycle vs Record Lifecycle

www.formpipe.com/lifescience
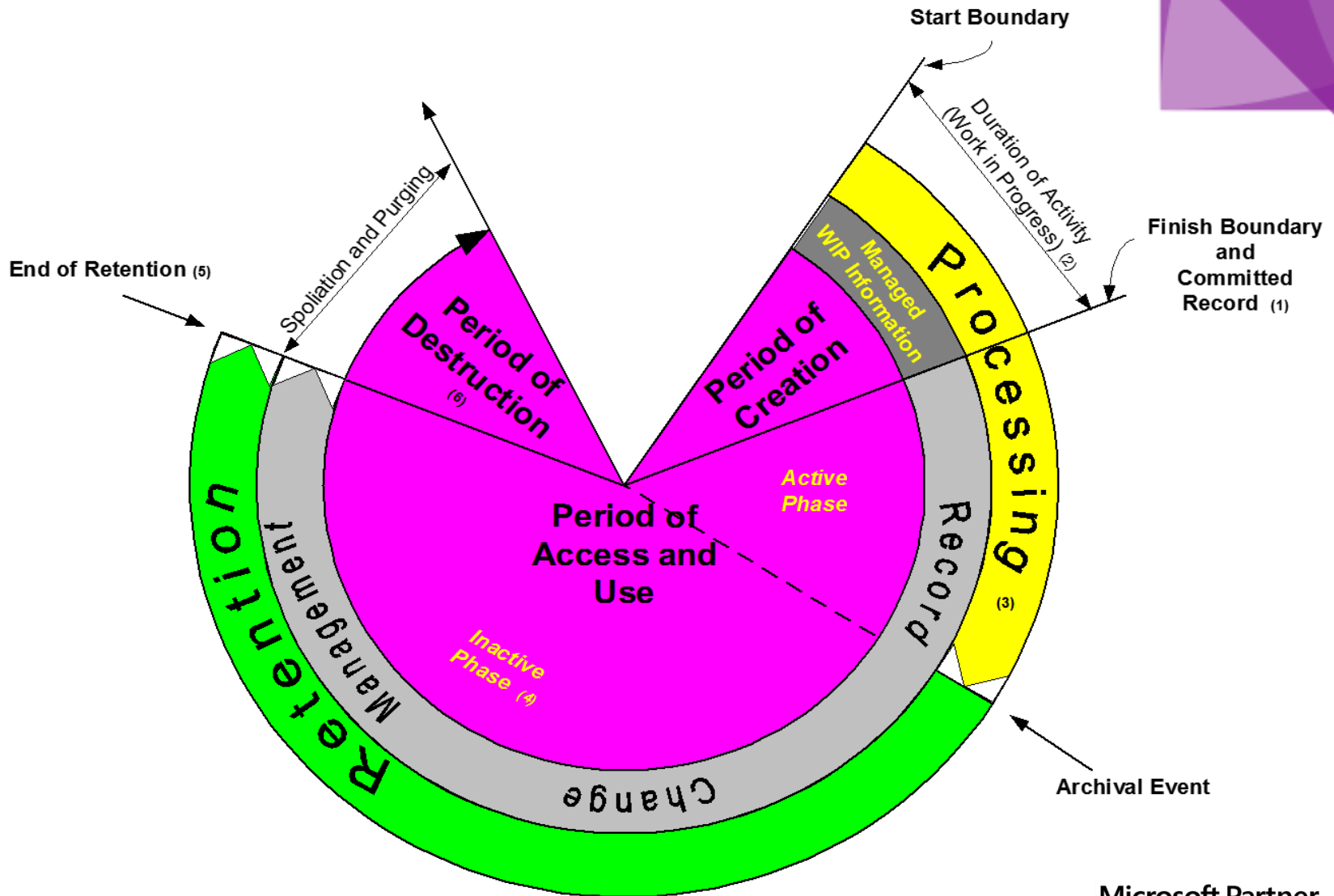
Digital Preservation

1. The risk of obsolescence

2. Factors that affect the likely feasibility and cost of preservation include:

   • The extent to which use of a format is widespread (e.g. JPEG)

   • Whether a format is independent from other technologies (e.g. XML)

   • Whether file format specifications are in the public domain (e.g. TIF)

   • How readily a file can be identified and its contents checked (e.g. CSV)

   • Whether metadata is provided within the format (e.g. PDF/A)

   • Whether the format interoperates with a variety of services (e.g. XML)

Further detail can be found at http://www.digitalpreservation.gov/formats

Microsoft Partner
Gold Collaboration and Content
Gold Hosting
Gold Digital Advertising

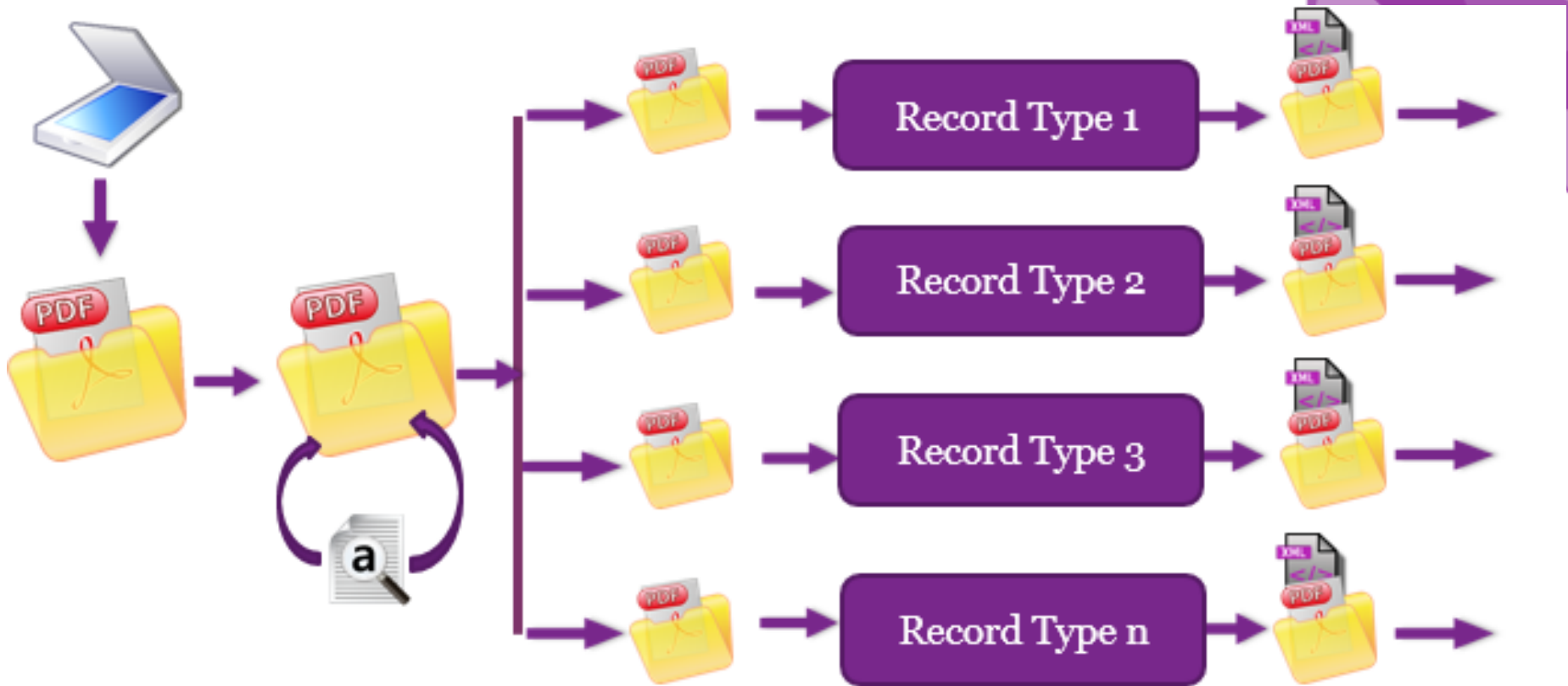# Example Content Types and Preservation Formats

Content Type	Description	Example Application	Native Format	Preservation Format
Formatted document	Sometimes called 'office documents'; made up of a combination of alphanumeric and special characters, sometimes with embedded graphics, video, audio or other types.	Any 'office' suite, email programs, Acrobat	DOC, DOCX, GDOCS, ODF, OOXML, PDF, PPT, PPTX, W51, XLS, XLSX	PDF/A
Alphanumeric (unstructured) data	Files made up of alphanumeric (and other printable) characters that do not depend upon formatting.	XML authoring tools, other specialised text editors	DTD, TXT, XML, XLST	TXT XML
Instrument data	Results from laboratory instruments, typically in proprietary format.	Laboratory instruments	CSV, MI, MX, SMR, SMS, SPA, SPG, WSV	CSV XML
Database	Data structured so as to allow analysis, reporting, etc.	Access, other database packages, statistical analysis software	MDB, DB, DBF, GRDB	CSV XML
Image	Graphics formed from bitmaps (rasters), or vectors, in monochrome, greyscale or color.	Scanners, vector graphics programs, image editors	AI, BMP, CDR , DNG, JPEG, PNG, RAW, TIFF	Several alternatives
Audio	Speech, music or other sounds.	Sound recording and editing software	AAC, DTS, FLAC, MP3, WAV, WMA	Several alternatives
Video	Moving images, with or without sound.	Video recording and editing software	AVC, AVI, FLV, MOV, MP4, WMV	Several alternatives
Website	Internet and intranet sites, including blogs.	Web authoring packages	ASP, CSS, HTM, HTML, MHT	WARC

Some Solutions

For on-boarding paper records

and preserving electronic records
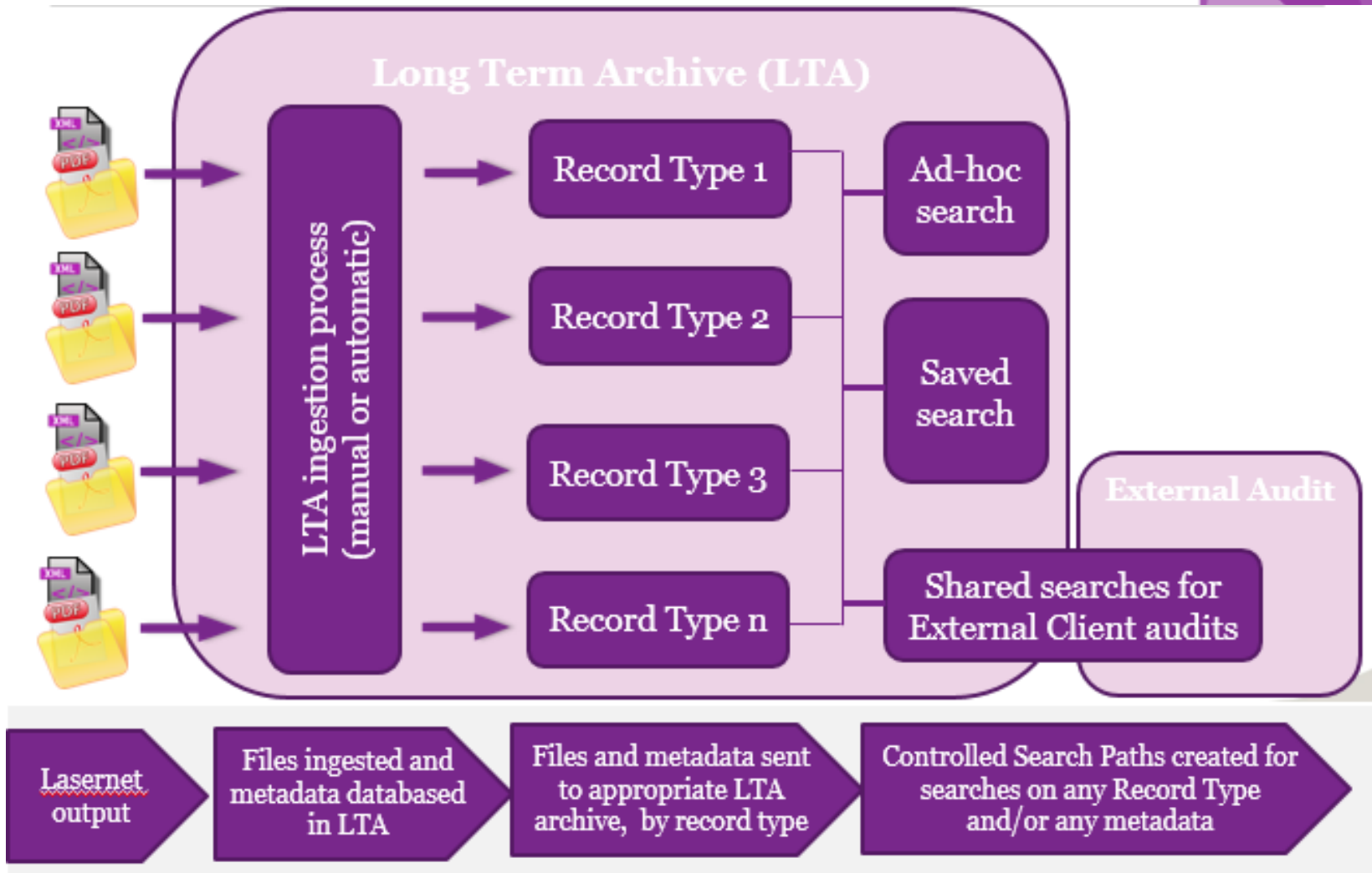
Paper Document Input into the Electronic World

Record Type 1

Record Type 2

Record Type 3

Record Type n

| Scan to source library | Create an OCR library | Documents moved to Lasernet queues, by Record Type | Documents and OCR metadata verified and e-signed | Signed PDF and metadata exported to LTA |

www.formpipe.com/lifescience

Lasernet

- **Capture** information from a wide range of input streams
  - Scanning and OCR
  - Cloud storage
  - Systems (ERP, EDMS…)

- **Manipulate** information
  - Metadata enrichment
  - Forms generation

- Intelligent **Routing** Engine output to
  - Systems / databases
  - Folders
  - Print
  - Email
  - Content management systems

www.formpipe.com/lifescience

Electronic Record Archival & Retrieval

Microsoft Partner
Gold Collaboration and Content
Gold Hosting
Gold Digital Advertising

Long Term Archive (LTA)

- Web based archive solution providing future proofed storage

- Rapid search and retrieval based on preservation object metadata

- Ability to search across archives

- Ability to reference a record as validated / approved once uploaded

- Powerful permissions engine for control of confidentiality backed up by audit trail

- Automated retention and deletion rules

- **Guaranteed document storage for 50 years**

www.formpipe.com/lifescience

Microsoft Partner
Gold Collaboration and Content
Gold Hosting
Gold Digital Advertising
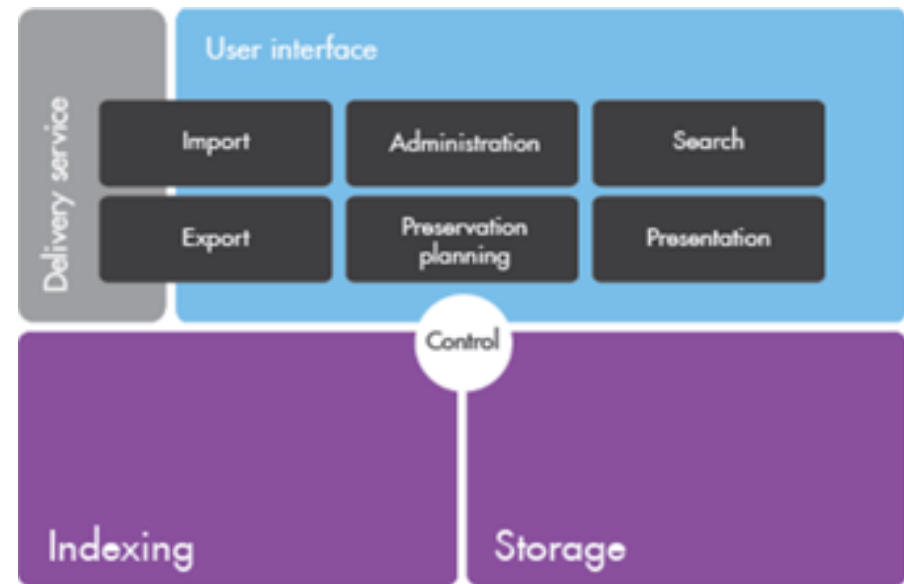
Thanks!!

Keith Williams
keith.williams@formpipe.com
+1 571 220 7645
50 Beaver Avenue
Annandale, NJ 08802

www.formpipe.com/lifescience

Formpipe.

Formpipe Life Science Product map

www.formpipe.com/lifescience

Some of our customers

Microsoft Partner
Gold Collaboration and Content
Gold Hosting
Gold Digital Advertising