King's Digital Lab

KING'S College LONDON

Preserving 30 years of Digital Humanities Work: The Experience of King's Digital Lab

James Smithies
King's College London
@jamessmithies

Anna-Maria Sichani
Huygens-KNAW
@amsichani

Carina Westling
King's College London
@CarinaWestling

DPASSH 2017, 'Preserving Abundance: The Challenge of Saving Everything', 14-15 June 2017, University of Sussex.

Operational Context

Operational Context: DH at Scale

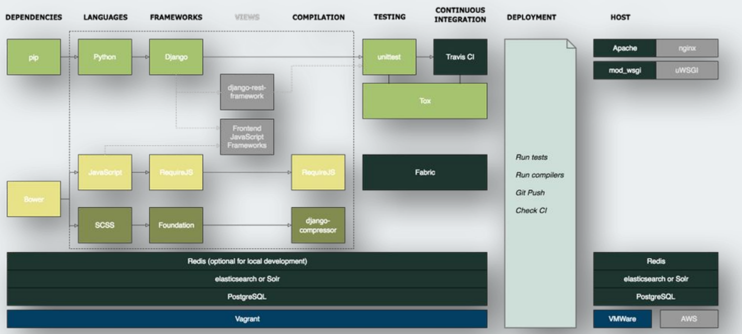
@kingsdh:

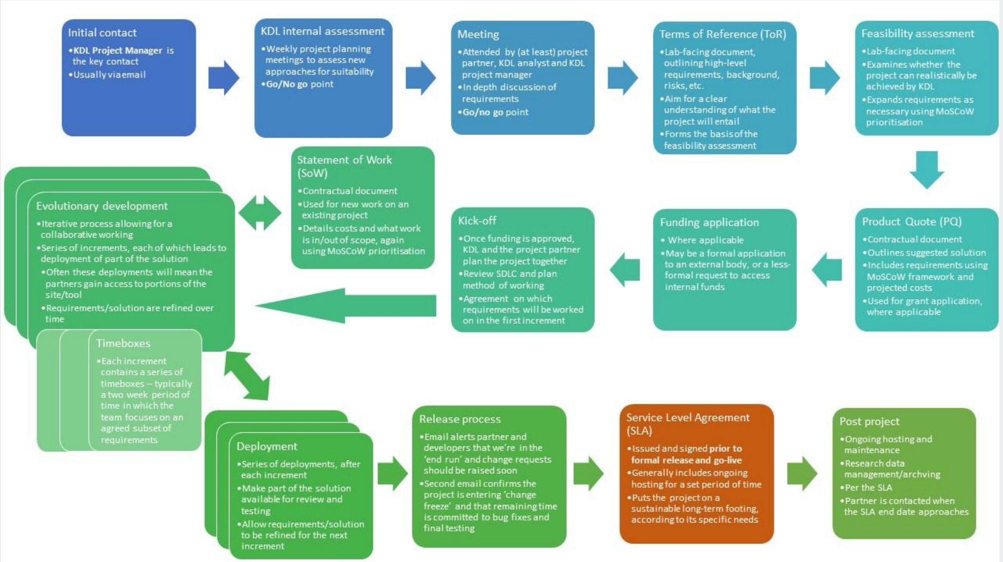
- 30 years of activity, against a background of rapid innovation and change in Humanities Computing and Digital Humanities.
- Centre for Computing and the Humanities (1991); Centre for eResearch in the Humanities (2008).
- Department of Digital Humanities (2011-):
- ~300 students across 5 Masters and 1 Undergraduate degrees
- ~33 staff (fte academic; fte research and fte teaching fellows)

King's Digital Lab (2015):

- 14 staff (includes 1 project-specific).
- 180+ virtual machines, 400GB RAM, 27TB data + AWS / Azure.
- 90 inherited projects, 20 ongoing. ~5 million digital objects.
- Business, HR, Operational Plans for 2016 – 2020 complete.
- Supported by external funding, under-written internally.

Operational Context: Technical Control





Operational Context: ALM & RDM

ALM = Application life-cycle management
RDM = Research Data Management

Priority	Requirement
M	Scoping data model to support the output of data extraction and the input for data visualisation. Note that this is based on the outcome of the first project meeting in Month 1 when more detailed specification of the actual visualisation and data extraction should be defined.
M	Software prototype to extract: • Frequency data for selected lemmas contained in a selected small corpus or achievable sample of already lemmatised processable texts following a data model suited for the analysis and visualisation as above. The prototype will be based on an algorithm developed together with Informatics (the aim is to track down strings i.e. word forms of word content i.e. lemmas of non-contiguous data within an original text; each repetition string should hence have a unique identifier). • Spatio-temporal information such as quantifiable distance between repetitions. Note that given the time constraints and the experimental nature of the project, the online environment hosting the project outputs will be very minimal and ready-made.
M	Max 3 meaningful and coherent types of data visualizations (at least 1 interactive one allowing the user to move from global to local view; others in static form) for comparative purposes across languages (to match translators' choices for repetition strings - omission, variation, retention - across parallel

FORWARD PLANNING

If outputs from the work quoted for here will include an online resource and/or digital data, then KDL will contact the partner(s) in good time before completion of the work to discuss available post-completion options. These options are summarized below.

Web Hosting and Infrastructure Service Level Agreement (SLA)

If the output is an online resource, after delivery and sign-off of the final increment KDL can - if required and funding is available - host the resource on KDL's server infrastructure under a Service Level Agreement (SLA).

Archiving and Sustainability

If the output is an online resource and either KDL hosting is not required or the resource has reached end-of-life, KDL will discuss available archiving options with the project partners during the requirements gathering phase. Such options might include, for example, generating a static site from the resource or packaging the resource in a suitable contained for transferral to another host. KDL will contact the project owner before the SLA expires, to remind them of their options, help them seek additional funding, or scope their chosen archiving option. When an option is chosen, then KDL will draw up a Statement of Work specifying the requirements and costings for the work involved to be agreed and signed-off by the partner, and discuss funding options with the partner. Note that code created as part of the project will also be accessible on the Github-based repository that KDL will set up as part of the development process.

Research Data Management

As a major research institution King's College London (KCL) treats preserving and promulgating research data as a high priority. It has an official Research Data Management Policy and to support the policy the KCL library service has established a comprehensive Research Data Management service (RDMS) including facilities for depositing project-generated data and

Scholarly Context

Scholarly Context: (re)defining "sustainability" in DH

We're at an inflection point:

- Experimentation *as well as* institutional responsibility to integrate with Library & IT, and manage legal and financial risk.
- The DH community needs to be able to illustrate innovation *but also* continuity and value for funding £.
- Scale & complexity of infrastructure requires application of industry-standard methods.

Scholarly Context: Developing tools, methods, standards

We're at an inflection point:

- Develop assessment procedures in order to document community and brand impact as well as potential (technical, financial, data-related) risks of a DH project
- Commit to adopt open technical frameworks and standards in order to enable interoperability of data and long-term maintenance
- Transparent documentation throughout the project's process
- Develop research data management workflows in close collaboration with ULibrary/ Data center

Results

King's Digital Lab

KING'S College LONDON

KDL Archiving & Sustainability Project

Goals:

1. Assess: Scholarly value, Impact, Security, Maintenance cost *[1]
2. Triage: Risk, Funding, PI & KDL availability *[2]
3. Maintain or archive *[3]

*[1] 'Fit' with FA&H objectives and KDL model (operational and technical process dimensions)

*[2] Prioritisation for phases 1 and 2, informing approach and allocation of resources

*[3] Integration within KDL production cycle (post-project) or authorised archival

Project name (in full)	Acronym	Cost code	Status	SLA Y/N	ITS hosting	Start date	End date	Extensions	Grant income	Pre-KDL costs	Cost pre-project	Cost project staff	Cost project non-staff	Cost post-project	Disk space	RAM	CPUs	Annual hosting	SLA End date	Total income	Total staff cost	Tot staff w o/heads	Total cost	Absorbed costs
CEGC	CEGC		Post-project	N/A	Yes				£0.00	£0.00	£0.00	£0.00	£0.00	£181.76						£0.00	£181.76	£236.29	£236.29	£236.29
Early Modern Dress and Textiles			Post-project	N/A	Yes				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
ERC Multipal			Post-project	N/A	N/A				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
History Blog			Post-project	N/A	Yes				£0.00	£0.00	£0.00	£0.00	£0.00	£35.96						£0.00	£35.96	£46.75	£46.75	£46.75
Out Of the Wings	OOTW		Post-project	N/A	Yes	1/6/2006	31/7/2012		£0.00	£0.00	£0.00	£0.00	£0.00	£227.68						£0.00	£227.68	£295.98	£295.98	£295.98
Supporting Productive Queries for Research	SPQR		Post-project	N/A	N/A						£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
UK Research Information Shared Service Project	UKRISS		Post-project	N/A	N/A						£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
3D Visualisation in the Arts Network	3DVisA		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£53.94						£0.00	£53.94	£70.12	£70.12	£70.12
AHDS: Enabling Digital Resources for the Arts and Humanities			Post-project	No	N/A				£0.00	£0.00	£0.00	£0.00	£0.00	£35.96						£0.00	£35.96	£46.75	£46.75	£46.75
AHRC Crowd Sourcing Study: report			Post-project	No	No	1/2/2012	31/10/2012		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Anglo-Saxon Charters	ASChart		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	17	2	1			£0.00	£0.00	£0.00	£0.00	£0.00
Anglo-Saxon Cluster	ASCluster		Post-project	No	No	1/10/2006	30/6/2010		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	14	1	1			£0.00	£0.00	£0.00	£0.00	£0.00
Aphrodisias in Late Antiquity			Post-project	No	No																			
Breaking of Britain	BoB	AAVFHPR	Post-project	No	No	1/9/2010	31/8/2013		£8,875.13		£0.00	£8,875.13	£0.00	£0.00						£8,875.13	£8,875.13	£11,537.67	£11,537.67	£2,662.54
British Printed Images to 1700	BPI1700		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	19	1	1							
Catalogue of English Literary Manuscripts 1450-1700	CELM		Post-project	No	No	2005	2010		£0.00	£0.00	£0.00	£0.00	£0.00	£43.22						£0.00	£43.22	£56.19	£56.19	£56.19
City Witness		AAVFHXR	Post-project	No	No	1/2/2012	31/7/2013		£21,666.00	£21,666.00	£0.00	£0.00	£0.00	£0.00	19	2	1			£21,666.00	£0.00	£0.00	£0.00	£0.00
Commodities and Culture website			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Cony			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Creative King's			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£17.98						£0.00	£17.98	£23.37	£23.37	£23.37
Cultures of the Lusophone Black Atlantic	CLBA		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Digital Exposure of English Place-names	DEEP		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Digital Humanities: Works in Progress			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Digital Image Archive of Medieval Music	DIAMM		Post-project	No	No	1/1/2007	30/8/2011		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	853	2	1			£0.00	£0.00	£0.00	£0.00	£0.00
Discover Medieval Chester		AAVFHWR	Post-project	No	No	1/2/2012	31/7/2013		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	31	2	1			£0.00	£0.00	£0.00	£0.00	£0.00
Durham Liber Vitae Project			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Early Modern Spain	EMS		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Electronic Sawyer	ESAWYER		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	24	1	1			£0.00	£0.00	£0.00	£0.00	£0.00
Forensic Investigation of digital objects	FIDO		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Girls and Digital Culture: Transnational Reflections on Girlhood 2012	GDC		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Golden Pages. Conference listings in musicology and other links for musicologists			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Greek Bible in Byzantine Judaism	GBBJ		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	19	2	1			£0.00	£0.00	£0.00	£0.00	£0.00
Henry III Fine Rolls Project	FRH3		Post-project	No	No	1/4/2008	31/3/2011		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	21	2	1			£0.00	£0.00	£0.00	£0.00	£0.00
Heritage Gazetteer of Cyprus	HGC		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	28	1	1	£510.00	4/4/2021	£510.00	£0.00	£0.00	£0.00	-£510.00
Historic Weather			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Hofmeister XIX			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	23	2	1			£0.00	£0.00	£0.00	£0.00	£0.00
Inscriptions of Aphrodisias	InsAph		Post-project	No	No	1/2/2004	31/1/2007		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	34	2	1			£0.00	£0.00	£0.00	£0.00	£0.00
Inscriptions of Libya			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	33	4	1			£0.00	£0.00	£0.00	£0.00	£0.00
Inscriptions of Roman Tripolitania	IRT		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Inventory of Byzantine Churches on Cyprus	IBCC		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
JAINpedia			Post-project	No	N/A				£0.00	£0.00	£0.00	£0.00	£0.00	£17.98	74	4	1			£0.00	£17.98	£23.37	£23.37	£23.37
Jane Austen's Fiction Manuscripts			Post-project	No	No	1/5/2006	30/4/2009		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	19	2	1			£0.00	£0.00	£0.00	£0.00	£0.00
Jonathan Swift Archive			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
La Entretenida			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	18	1	1			£0.00	£0.00	£0.00	£0.00	£0.00
Language of Landscape	Langscape		Post-project	No	No	1/10/2004	30/9/2007		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00	17	2	1			£0.00	£0.00	£0.00	£0.00	£0.00
Late-Life Creativity and the New Old Age			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Linking Parliamentary Records through Metadata	LIPARM		Post-project	No	N/A				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
London Association of Classical Teachers Online Records	LACTOR		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Map of Shakespeare's London			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Mapping Migration	MMKM		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Masonic Periodicals			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Medieval Francophone Literary Culture Outside France		AAVFHRR	Post-project	No	No	1/4/2011	30/9/2014		£57,864.00	£57,864.00	£0.00	£0.00	£0.00	£0.00						£57,864.00	£0.00	£0.00	£0.00	£0.00
Methods Network	Methnet		Post-project	No	No	1/4/2005	31/3/2008		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Modern Languages Open	MLO		Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Nineteenth-Century Serials Edition	NCSE		Post-project	No	No	1/1/2005	31/12/2007		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
People of Medieval Scotland 1093-1286	POMS		Post-project	No	No	1/9/2007	31/8/2010		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Personal website of the musicologist J. P. E. Harper-Scott			Post-project	No	No				£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Preservation Exemplar at King's	PEKin		Post-project	No	No	1/4/2009	31/10/2010		£0.00	£0.00	£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Schenker Documents Online	SDO	AAVFHQR	Post-project	No	No	1/10/2010	30/9/2013	2016 applic	£0.00	£0.00	£2,482.20	£0.00	£0.00	£2,411.00	50	6	3			£0.00	£7,375.40	£9,588.02	£9,588.02	£9,588.02
Sharing Ancient Wisdoms	SAWS		Post-project	No	No	15/6/2010	14/6/2013				£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Standards for Networking Ancient Prosopographies	SNAP DRGN	AAVFIER	Post-project	No	No	1/1/2014	31/12/2014		£22,500.00	£22,500.00	£0.00	£0.00	£0.00	£0.00						£22,500.00	£0.00	£0.00	£0.00	£0.00
Stet Journal	STET		Post-project	No	No						£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Strandlines Digital Community			Post-project	No	Yes	1/6/2010	30/4/2011				£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
The AHRC Research Centre for the History and Analysis of Recorded Music	CHARM		Feasibility	No	No				£99,999.00		£0.00	£0.00	£0.00	£108.61	145	2	1			£0.00	£108.61	£141.19	£141.19	£141.19
The Art of Making in Antiquity			Post-project	No	No	1/1/2011	31/12/2013				£0.00	£0.00	£0.00	£0.00	225	2	1			£0.00	£0.00	£0.00	£0.00	£0.00
The Commonplacer			Post-project	No	No						£0.00	£0.00	£0.00	£0.00	15	1	1			£0.00	£0.00	£0.00	£0.00	£0.00
The Complete Works of James Shirley	SHIRLEY	AAVFIDR	Evolutionary	No	No	1/4/2013	31/12/2014		£14,133.00	£14,133.00	£0.00	£0.00	£0.00	£1,789.44	21	4	1			£14,133.00	£1,789.44	£2,326.27	£2,326.27	£2,326.27
The Corpus of Romanesque Sculpture in Britain and Ireland	CRSBI	AAVFAER	Deployment	No	N/A	1/4/2011	31/3/2012	2/9/2016	£2,500.00			£345.76			1181	4	1			£2,500.00	£345.76	£449.49	£449.49	-£2,050.51
The Digital Classicist			Post-project	No	No										51	2	1							
The Fortune Theatre Records			Post-project	No	No	1/1/2012	31/12/2012				£0.00	£0.00	£0.00	£0.00	45	4	1			£0.00	£0.00	£0.00	£0.00	£0.00
The London Charter			Post-project	No	No						£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
The Making of Charlemagne's Europe	MKCHEUR	AAVFHTR	Post-project	No	No	30/11/2011	29/6/2014		£71,642.00		£0.00	£0.00	£0.00	£6,321.00						£71,642.00	£6,321.00	£8,217.30	£8,217.30	-£63,424.70
The People of Northern England database 1216-1286	PoNE		Post-project	No	No						£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
The Pompey Project			Post-project	No	No	18/6/2007	17/7/2007				£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
The Sanskrit Knowledge-Systems Project (CeRCH)			Post-project	No	No	1/3/2004	28/2/2007		£0.00	£0.00	£0.00	£0.00	£0.00	£35.96						£0.00	£35.96	£46.75	£46.75	£46.75
The Skenographia Project			Post-project	No	No						£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Theatron 3			Post-project	No	No						£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Theoretical Roman Archaeology Conference	TRAC		Post-project	No	No						£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
UX Pattern Library			Post-project	No	No						£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
Village economy and society: early Roman Tebtunis			Post-project	No	No	1/9/2011	31/8/2014				£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00
William Hogarth's Analysis of Beauty (1753)			Post-project	No	No						£0.00	£0.00	£0.00	£0.00						£0.00	£0.00	£0.00	£0.00	£0.00

KDL Archiving & Sustainability Project

Academic disciplines

Classics	23
History	23
Digital Humanities	20
English	14
Music	5
Cultural Studies	5
Spanish	4
Other	5
Total	**99**

Age of resource

5 > old	22
5 < old	77
Total	**99**

Platform

Java (Tomcat)	26
Python (Django)	52
Wordpress	7
Omeka	9
Typo3	5
Total	**99**

Operating System

Windows (2003)	2
Windows (2008)	9
Linux (Debian 4)	13
Linux (Debian 5)	32
Linux (Debian 6)	33
Linux (Debian 7)	10
Total	**99**

KDL Archiving & Sustainability Project

Triage Results

Pilot phase	29
Phase 1	35
Phase 2	35
Total inherited	**99**

Anticipated Results

Service Level Agreement	49
Migration - ITS	10
Migration - other	20
Archived	20
Total	**99**

Questions?