

# Digital Humanities & Research Software Engineering Summer School

24–28 July 2023

**13:00 – 14:10** *Research Software Development Life Cycles: Planning and managing sustainable outputs*

**Neil Jakeman**

**14:10 – 14:20** *Break*

**14:20 – 14:50** *Research Themes: Maintaining intellectual agency in research services*

**Neil Jakeman**

**14:50 – 15:40** *Career Pathways in Arts & Humanities RSE in the UK and Beyond*

**James Smithies**

**15:40 – 15:45** *Closing word*

**Mary Chester-Kadwell**

# Neil Jakeman

Senior RSE Analyst

 @kingsdigitallab

neil.jakeman@kcl.ac.uk

kdl-info@kcl.ac.uk

# **Research Software Development Life Cycles: Planning and managing sustainable outputs**



@kingsdh

30 years of activity, against a background of rapid innovation and change in Humanities Computing and Digital Humanities

Centre for Computing in the Humanities (1991) - Prof. Harold Short

Department of Digital Humanities (2011 - present day):

- ~450 students across 5 Masters and 1 Undergraduate degrees
- PhD programme
- ~30 staff including researchers and teaching fellows

Operational since 2015

13 permanent staff: Research Software Lab and Project Managers, Research Software Analysts, Research Software Engineers, Research Software UI/UX Designers, Systems Managers  
+ PostDocs/Research Affiliates/Visiting fellows

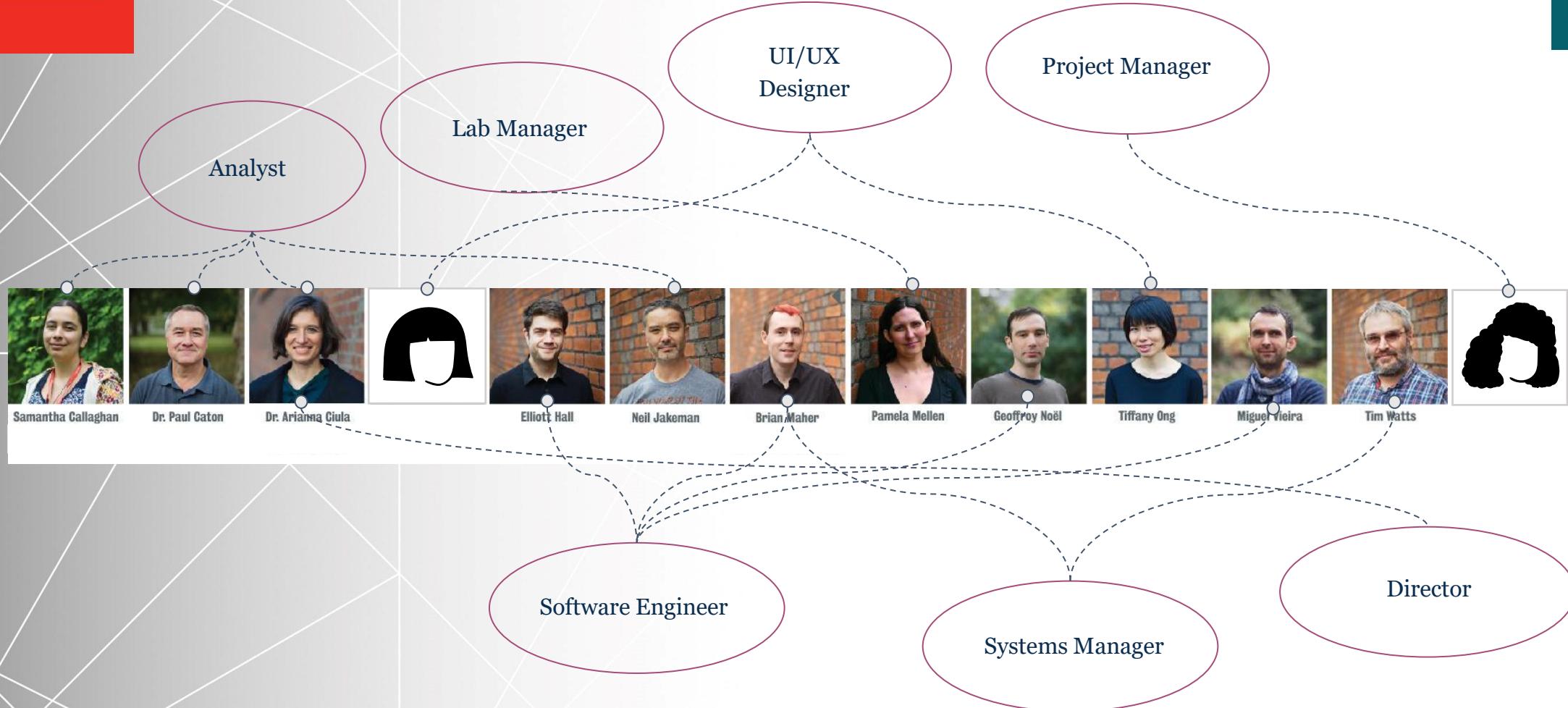
~120 digital projects

5 VMWare Hosts, ~110 VMs, 512GB RAM per host >> migration to central Open Stack infrastructure in progress at KCL e-Research

Supported by external funding, under-written internally

Last year: technical work for ~30 funded projects + ~60 post-projects, total value of ~£13.5m

Faculty of Arts and Humanities				
Interdisciplinary Research Centres and Themes	Departmental Centres	Departments	Themes and Institute	Cross-faculty Research Centres and Institutes
Centre for Early Modern Studies		Classics	Digital Futures Institute	Camões Centre for Portuguese Language and Culture (Based in A&H)
Centre for Enlightenment Studies		Culture, Media & Creative Industries		Centre for Humanities and Health (Based in A&H)
Centre for Hellenic Studies			Environmental Humanities	Gender Studies at King's (Based in SSPP)
Centre for Late Antique and Medieval Studies	Centre for Digital Culture	Digital Humanities		KCL Cybersecurity Centre (Based in NMS)
Centre for Life-Writing Research		English Language and Literature	Global Cultures Institute	King's AI Institute (Based in NMS)
Centre for Modern Literature & Culture		Film Studies		King's Contemporary British History
Jewish Studies at King's	Empires and Decolonizations Research Hub	History	Language Acts and Worldmaking	King's Water Centre (Based in SSPP)
London Shakespeare Centre		Languages, Literatures and Cultures		Quebec and French Canada Research Network (QaFCaRN) (Based in A&H)
Queer@Kings		Liberal Arts		Sir Michael Howard Centre for the History of War (Based in SS)
	Centre for Philosophy and Art	Philosophy		
	Centre for Arts and the Sacred at King's (ASK)	Theology & Religious Studies		
	Religious and Ethnic Diversity in China and Asia Research Unit (REDCARU)			
King's Digital Lab				

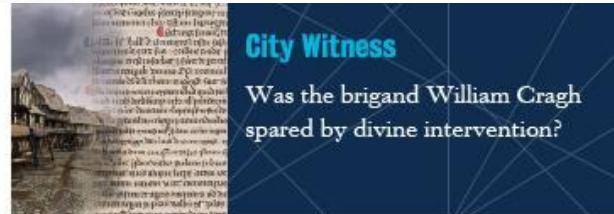


A few years ago a selection of our work looked something like this - historical databases and digital archives, digital scholarly editions, map-based tools, data and text analysis, visualization



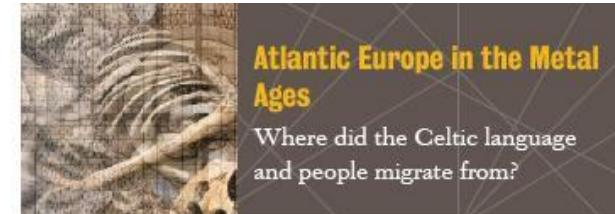
### Archetype

An integrated suite of web-based tools for the study of medieval handwriting, art and iconography.



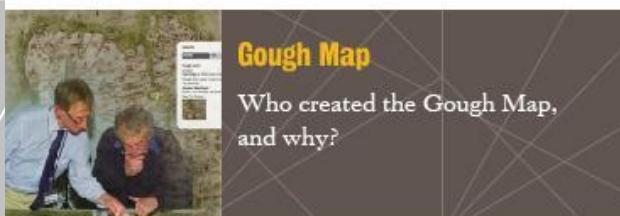
### City Witness

Was the brigand William Cragh spared by divine intervention?



### Atlantic Europe in the Metal Ages

Where did the Celtic language and people migrate from?



### Gough Map

Who created the Gough Map, and why?



### German Screen Studies Network

Designing a platform for online debate on German-language screen culture



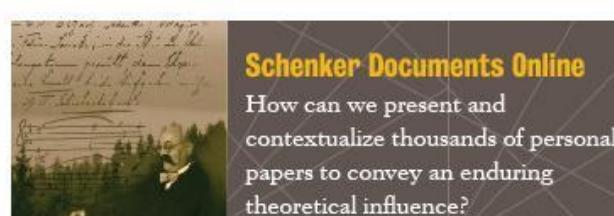
### The Values of French

Interrogating the role of the French language in the emergence of a European identity in the Middle Ages



### Paradox of Medieval Scotland

Examining the roles and relationships of medieval Scottish society



### Schenker Documents Online

How can we present and contextualize thousands of personal papers to convey an enduring theoretical influence?



### Digital Prosopography of the Roman Republic

The remarkable transformation of the Roman Republic from city-state to imperial power

# THE DIGITAL GHOST HUNT





# Once Upon a Time ...

A generative journey through fairyland



There was a **man** made of clockwork who longed to become human so he could watch over his master.



When the master was away he would set himself in a clockwork suit and **go out hunting**, and when he returned the suit would be in motion and he would be caught.



But the huntsman had a faithful servant who was always ready to do his bidding, and when the **suit** was set in motion he would run to the door and give the alarm. Then the man would set himself free, but the huntsman would pursue him, and shoot him down as he ran.



But the **faithful servant**, who was called Dumpling, always **stood by and watched**, and when his master was caught he said: 'It is all over with the man! He is dead!'



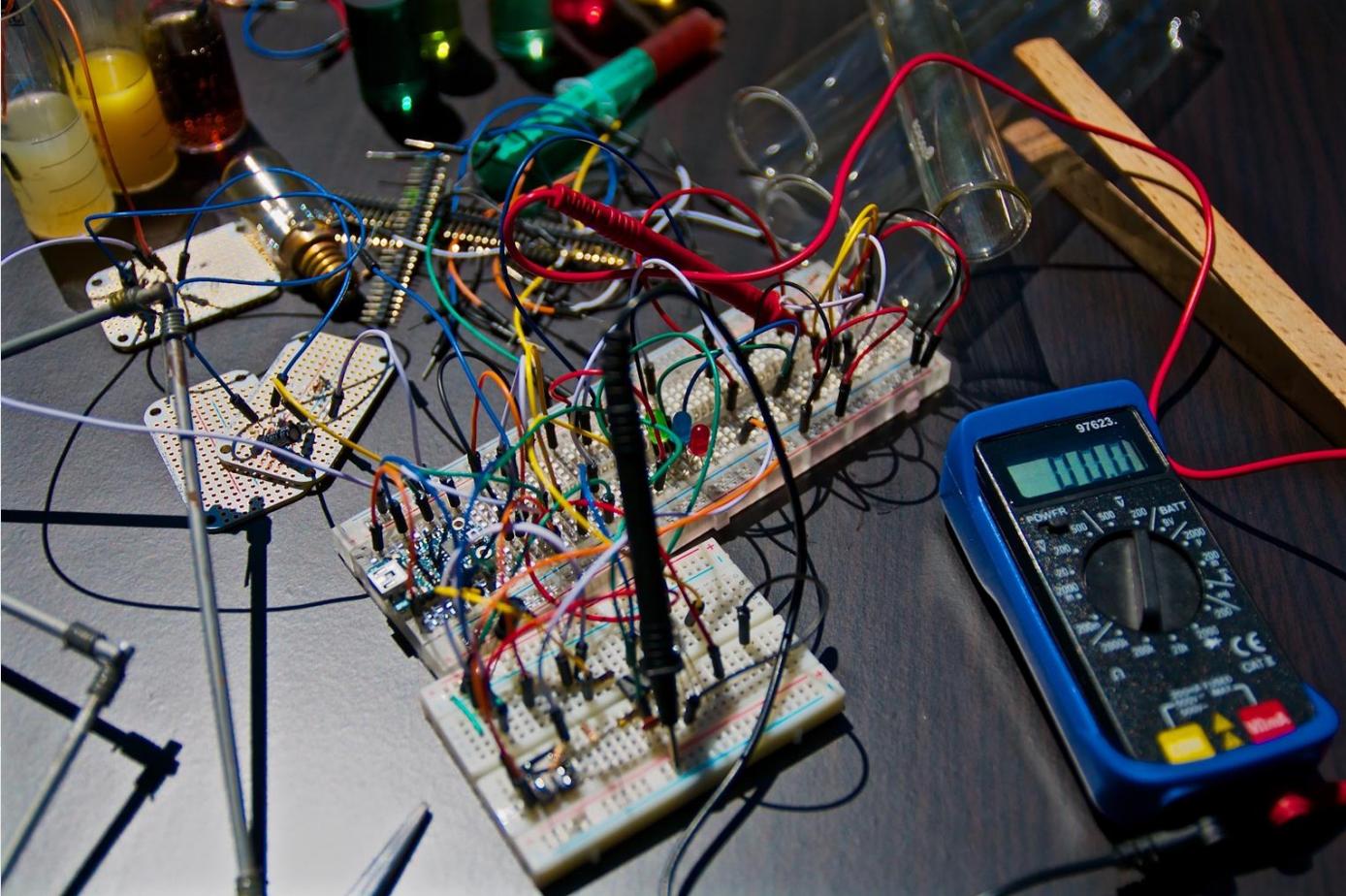
And the **huntsman**, who was in a great hurry, said: 'I only ran into him to get away from you, Dumpling.'



And then the man was so frightened that he stood still and did not move an inch. And the **servant** looked at **each other** for a moment, and thought: 'He is safe,' and went away to their quarters. But the man, who was still in the **clockwork suit**, went outside and called: 'huntsman, give me liberty or I shall set myself free.'

The starting image shows a man, and arguably some clockwork shapes. By §2 a master has been introduced, but there is still a reference to clockwork. By §3 a huntsman and servant appear, and there is one reference to a suit being set in motion (like clockwork?). In §4 the servant is named as Dumpling, and the first truly nonsensical dialogue appears. §5 has more confusing dialogue and the associated image is hard to reconcile with the narrative. At §6 we hear again about the clockwork suit, and two illustrated characters seem to matching the action described.

Just trying stuff out is fun right?



TWO  
MINUTES TO MIDNIGHT



# Minimum Viable (Research) Product

*“A digital research output which meets the essential needs of the project and so, does not compromise the rationale of commencing the project”*

**Neil Jakeman**

## Legacy Issues – Emotional Baggage



# It needs to warp the fabric of space-time, or we've wasted £500k!



<https://onepagelove.com/prevent-feature-creep-and-bad-product-requests>

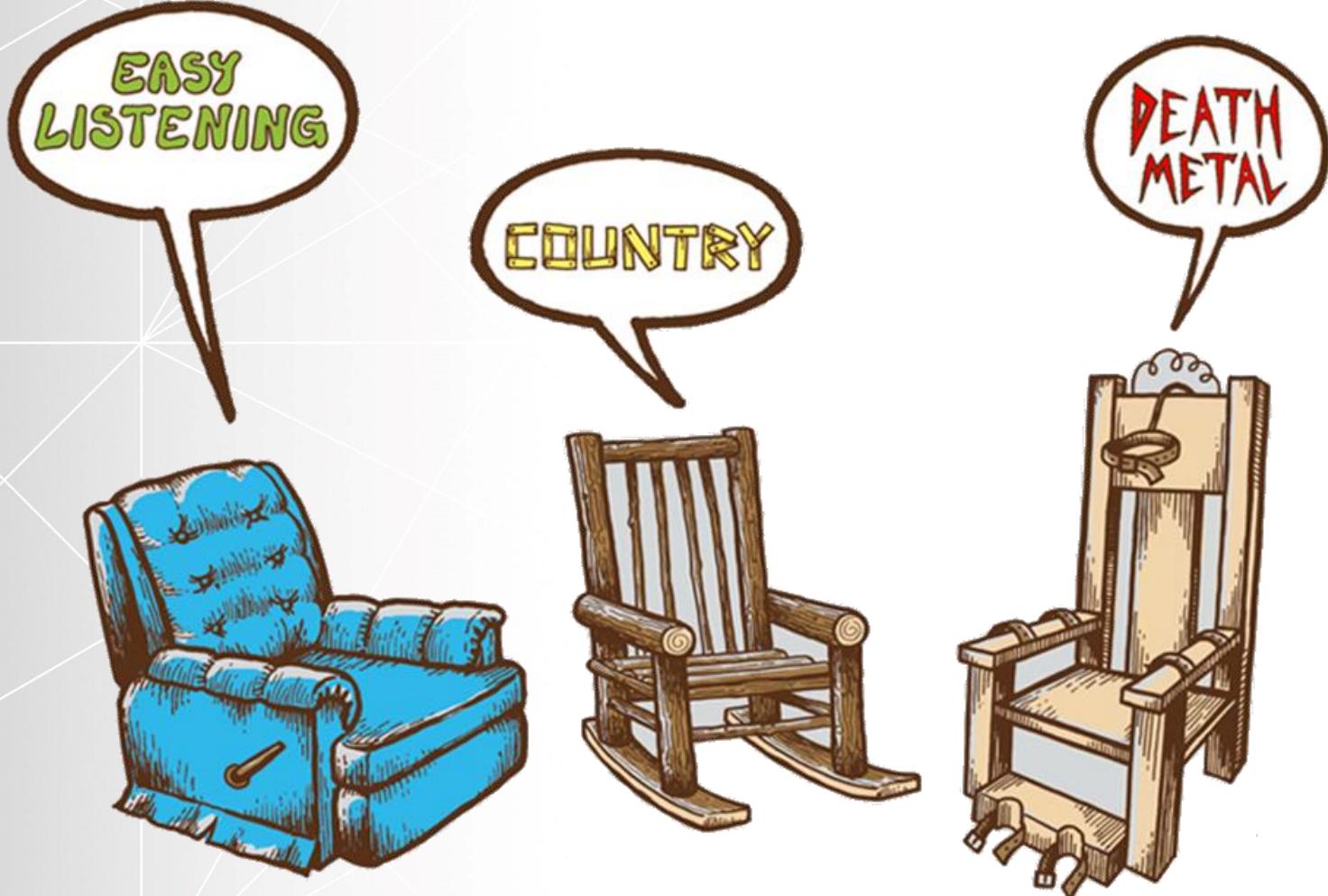
# The myth of the legendary “Basic” website



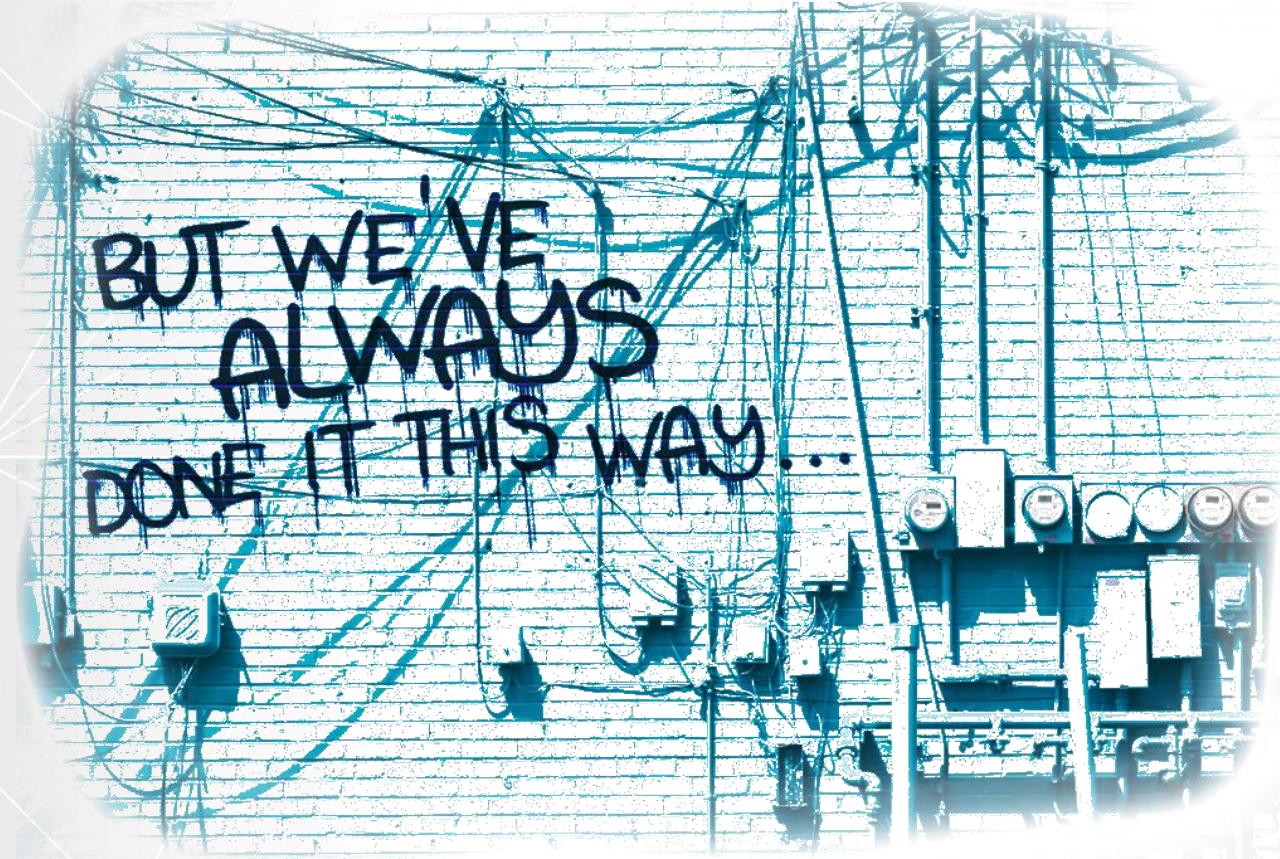
# Niche Design



# Playing Musical Chairs



# Born Digital?



# Die Analogue?

## The SDLC evolved to:

- Ensure a rigorous project **Feasibility** evaluation is undertaken
- Provide a clear product quote in which the **Prioritised Requirements** of the project are listed
- Manage the development period to ensure that **Changing Priorities** can be accommodated
- Validate the requirements defined both at the outset of the project, and to **Iteratively Review** requirements over the development period
- Make sure that a functioning **Minimum Viable Product** is delivered
- Provide a **Service Level Agreement** that defines how the output will be sustained for a defined period
- Provides options for graceful retirement and effective **Archiving** of a digital resource
- Define RSE skills and roles in a way that can be used to define a **Career path and promote Staff continuity**

**Terms of Reference****Initial Contact**

Partner gets in touch with their project idea.

**Feasibility****Internal Assessment**

We review whether the project would be a good fit for KDL.

**Requirements Assessment**

We discuss requirements with the partner and produce a product quote.

**Product Quote****Project****Project****Project Review Record****Evolutionary Development**

Focused on communication, collaboration and flexibility as we develop the project in increments.

**Kick Off**

If funding is approved, we confirm how we'll work with our partner.

**Service Level Agreement****Deployment**

In stages, after each increment, allowing for regular testing and refinement in each development/deployment cycle.

**Release**

The partner signs off a Service Level Agreement and the project goes live.

**Funding Application**

If needed, we assist with incorporating KDL's involvement in the application.

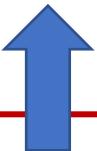
**Post Project**

Ongoing hosting and maintenance for a set period under the terms of the Service Level Agreement.

**bit.ly/30ovYpq**

**Governance Documentation**

(in the form of adaptable templates to be re-used and refined through practical application)



**Document Template  
Guidance**

(useful boilerplate text and question matrices)



**Supporting Processes and Protocols**

(Planning tools, Iteration, Review, Project meetings, Escalation etc.)



**Research Software Roles**

(Clear responsibilities, Career Pathways etc.)

## Home

James Smithies edited this page on Apr 21 · 16 revisions

# A Software Development Toolkit for Software Engineering

## A toolkit for documentation and a research setting.

This repository was established by King's Digital Lab to support the development of a piece of work, enriched with community contributions and research.

The documentation found here represents KDL's initial approach to the toolkit, which has been pragmatically adapted to suit the particular needs of the Software Engineering (RSE) research group.

### How to use this resource

Document guidance and templates have been written in Microsoft Word format, so they can be used most desktop word processing software or converted to PDF for use offline.

#### 1. Getting started

The core component of this toolkit is the collection of document templates, which follow the SDLC. In addition to the document templates, there is also guidance on how to use them and what factors to consider when developing software. This guidance is based on KDL's experiences in many project development cycles.

#### A quick note on naming conventions

The document names are prefixed alpha-numeric codes, such as A1, B1, F1, I1, J1, N1, Q1, L1, E1, and D1. The letter is followed by a number indicating the version of the document. NB. versioning will be required for all projects.

#### Abbreviations used

Single letters are frequently used as shorthand for common terms. For example, A = Analysis, D = Design, P = Planning, etc.

### • Document templates

- [A1: Terms of Reference](#)
- [B1: Project Approach Questionnaire](#)
- [F1: Feasibility](#)
- [I1: Product Quote](#)
- [J1: Statement of Work](#)
- [N1: Web Hosting and Infrastructure Service Level Agreement](#)

### • Document guidance

- [A2: Terms of Reference guidance](#)
- [B2: Project Approach Questionnaire guidance](#)
- [F2: Feasibility guidance](#)
- [I2: Product Quote](#)

ES 42	ne
Document templates	
<a href="#">A1: Terms of Reference</a>	
<a href="#">B1: Project Approach Questionnaire</a>	
<a href="#">F1: Feasibility</a>	
<a href="#">I1: Product Quote</a>	
<a href="#">J1: Statement of Work</a>	
<a href="#">N1: Web Hosting and Infrastructure Service Level Agreement (SLA)</a>	
<a href="#">Q1: Decommissioning Authorisation</a>	
<a href="#">L1: Project Review Record</a>	
<a href="#">E1: Costings table</a>	
Document guidance	
<a href="#">A2: Terms of Reference guidance</a>	
<a href="#">B2: Project Approach Questionnaire guidance</a>	
<a href="#">F2: Feasibility guidance</a>	
<a href="#">I2: Product Quote guidance</a>	
<a href="#">J2: Statement of Work guidance</a>	
<a href="#">Data Management Plan guidance and AHRC template</a>	
<a href="#">L2: Project Review Record guidance</a>	
<a href="#">N2: Web Hosting and Infrastructure Service Level Agreement (SLA)</a>	
<a href="#">Q2: Decommissioning Authorisation guidance</a>	
<a href="#">E2: Costings table</a>	

## EARLY CONVERSATION STARTERS for The Terms of Reference

- Have the RSE team had previous experience of working with this project team or members of the team?
- Will the project team fully engage with the SDLC processes?
- What previous work (if any) is being built upon?
- Has previous work in this area failed to produce quality research? If so, why?
- Is there an existing resource upon which this new work depends?
- What distinguishes this research as particularly ambitious or innovative?
- Does the project align with current internal goals and strategies?
- Does the project align with internal research interests?
- Does the RSE team have access to the necessary expertise to fulfil the project requirements?
- How amenable to digitisation or to computational processing are the data sources being proposed?
  - Are there sample data to access and assess?
  - Where are the datasets hosted if they already exist or where is the project team expecting to host them, and for how long?
  - How amenable to archiving is the project being proposed?
  - Are the ambitions of the team realistic?
  - Does the project present ethical concerns?

## The Terms of Reference Questions and prompts

- What is the timeline?
- Who is the funder?
- Do we have capacity?
- What distinguishes this research as particularly ambitious or innovative?
- Does the project align with current internal goals and strategies?
- Does the project align with internal research interests?
- Does the project offer an opportunity to engage with new technology and approaches?
- Does the project offer an opportunity to develop a mutually beneficial strategic relationship with an - interesting partner?
- Does the project align with your institutional values and ethical considerations?

## SOLUTION ARCHITECTURE

### Questions and prompts

- Does the proposed technical solution take into account data standards?
- Does it, or will it ultimately, adhere to [FAIR principles](#)?
- Is it clear that the RSE team has sufficient expertise to implement the solution?
- Is the solution scalable? Does it need to be?
- Which RSE roles will be performing which tasks?

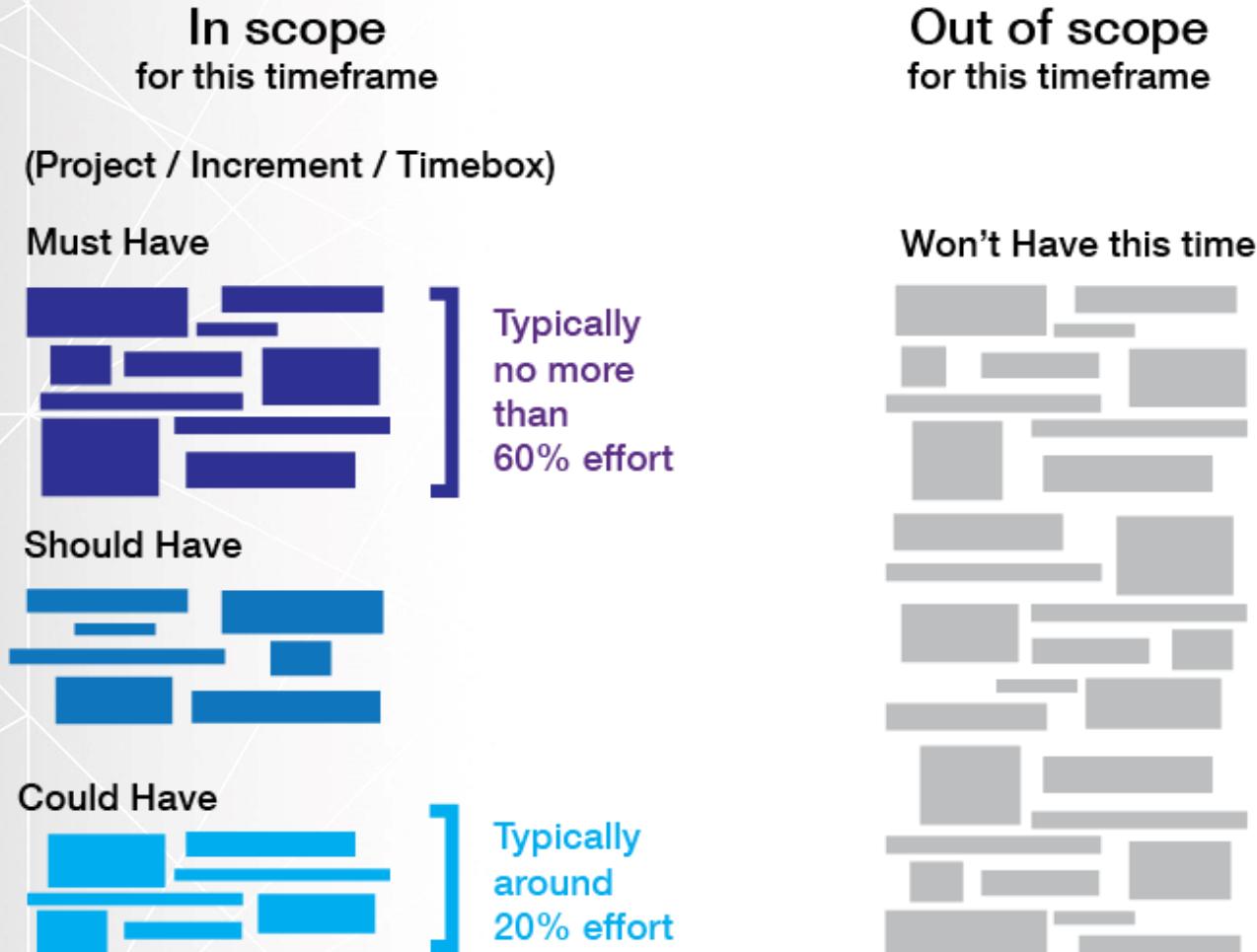
## DEVELOPMENT APPROACH

### Questions and prompts

- Where will the evolving solution be hosted?
- What backup strategies are in place?
- Which channels should the research team use to seek or provide feedback?
- How often will review meetings take place?
- When will *change freeze* be implemented?

M = must have; S = should have; C = could have; W = won't have this time

Priority	Requirement
M	Taxonomic data
M	Metadata schema
M	Site that can store approx 250 objects to be approx 250 objects
M	Site that can store up to 250 objects
S	Admin interface
S	Map functionality
S	Timelines placing items in historical context
C	Integration of 3D visualizations of key rooms at Windsor Castle (creation of 3D images would be by 3rd party)
W	Public interaction with/contributions to site



# Project Management & Communication

Quarterly Timebox meetings

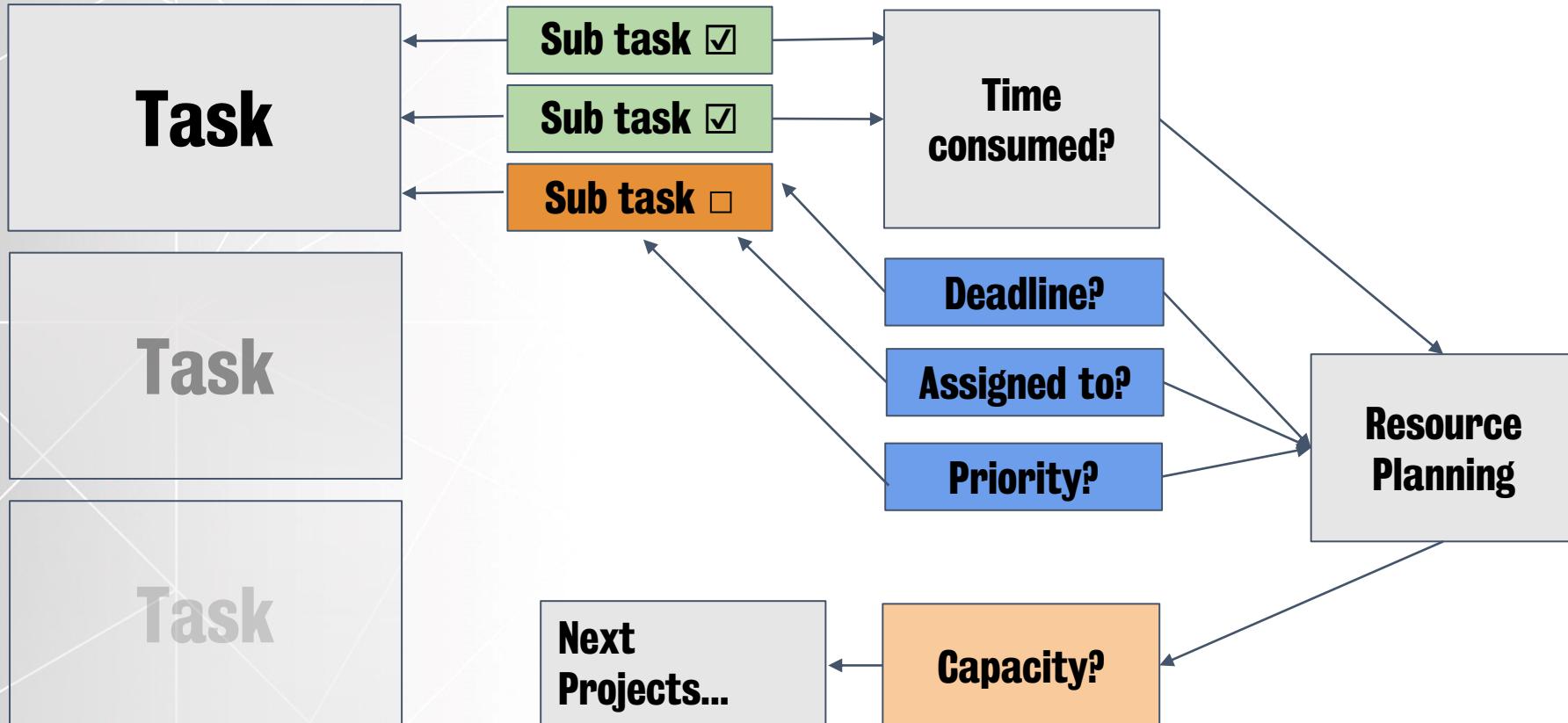
Monthly Team meetings

Fortnightly Timebox meetings

Weekly Project Planning meetings

Daily Standup (Slack channel)

# Task Management



# Archiving & Sustainability Approach

## Maintenance under costed SLA

Usually 5-year

## Migration

College ITS microsite service or to external host (another HEI or commercial provider)

## Static conversion

Maintained public access but reduced functionalities

## Dataset deposit

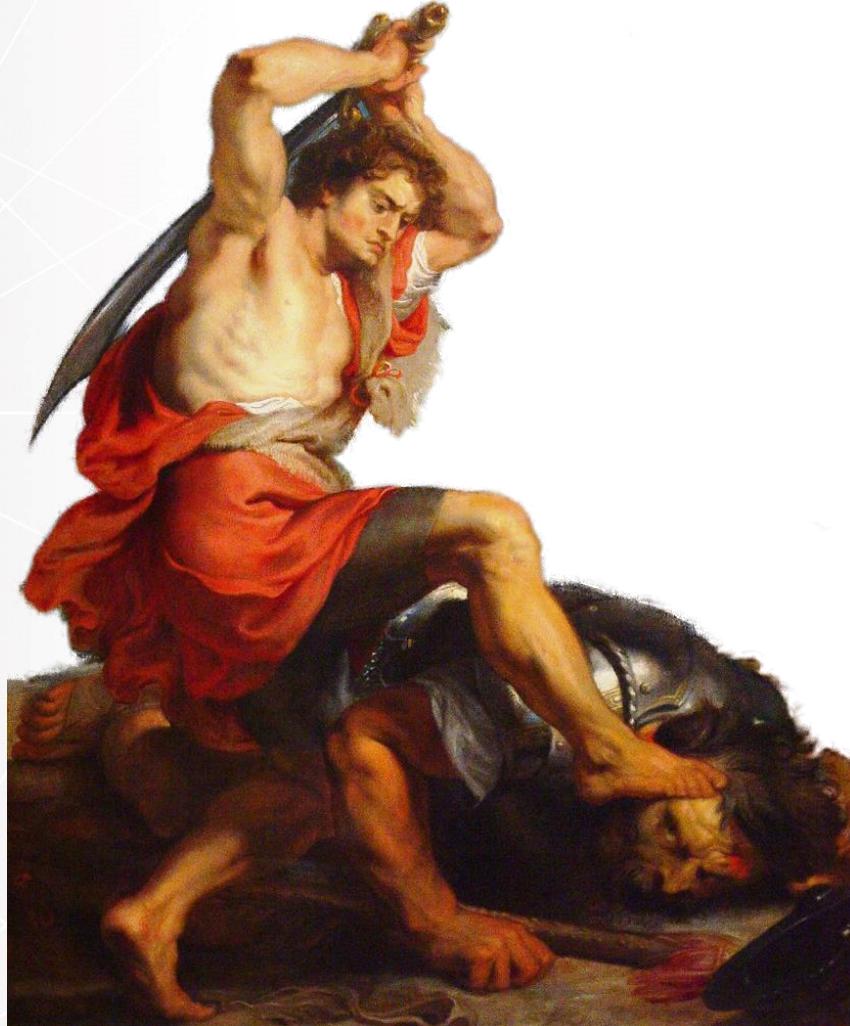
KDL CKAN, institutional technical systems, external repositories

## Minimal archiving & storage

Minimal storage (for two years minimum) for project website (VM) and data on KDL infrastructure as well as web archives. A placeholder page is shown at a project URL with description, metadata, and links to snapshots.



## The Small Project Paradox



# The Small Project Paradox

- **Triple check communications are understood!**
  - The SDLC proved its worth again and again, as a paper trail of the discussions we'd had and what we'd agreed.
  - When things turned fractious occasionally, we could point to the documents we'd all signed to calm things down again
- **Check for competence**
  - Not to be interpreted judgmentally, - you need to have a handle on expectations around training and understanding
- **Quit early?**
  - Trickier - I had plenty of opportunities and warning signs that might've steered us away from starting
  - Sometimes, all you can do is learn from experience!

# The Freeloader



## The Freeloader?

- Playing it forward
  - It was very short-sighted of me not to foresee how this could cause a conflict of interests
- Would I have done it differently?
  - Well, yes. But I might've made clear the risks and still allowed the data entry to happen
  - It didn't seem right to me that someone whose IP was invested in the product shouldn't be free to use it
- Up shot? - I'm still conflicted about it!

**Where is the gallery? When will the sound be added? The map doesn't work? The map does work, but not how I expected? Can you change the blog pages? Can you change the team pages? I don't like the banner. University X is missing from the footer. I can't access the thing. My cousin's can't access the thing. It doesn't work on my brother's iPad...**



- Every project is unique and any SDLC needs to consider the human components, both within the RSE team and the Research team
- Communication is something that needs to be scheduled - don't wait for sporadic contact
- Be firm about your processes - stick to your principles - don't compromise
- You aren't doing anyone a favour by not being consistent
- Ask awkward questions if you need to
- Perform post-project reviews
- Sometimes, you can only do what you can do!

- Ciula A. (2020) *Exposing legacy project datasets in Digital Humanities: KDL experience* <https://www.kdl.kcl.ac.uk/blog/legacy-project-datasets/>
- Ciula A. (2019) *What Makes Good Honey? KDL Checklist for Digital Outputs Assessment in the REF* <https://www.kdl.kcl.ac.uk/blog/checklist-digitaloutputs-ref/>
- Jakeman N. (2020) **Software Development Lifecycle** for Research Software Engineering: Safeguarding an inheritance and ensuring a legacy <https://www.kdl.kcl.ac.uk/blog/sdlc-for-rse/>
- Smithies J. (2019) *The Continuum Approach to Career Development: Research Software Careers in King's Digital Lab* <https://www.kdl.kcl.ac.uk/blog/rse-career-development/>

## Resources

- Agile DSDM Handbook: [https://www.agilebusiness.org/page/ProjectFramework\\_oo\\_welcome](https://www.agilebusiness.org/page/ProjectFramework_oo_welcome) (Agile DSDM Consortium 2014)
- DH2020 paper slides: <http://dx.doi.org/10.17613/789z-9p59> (Ciula et al. 2020)
- FAQ for project partners: <https://www.kdl.kcl.ac.uk/how-we-work/faq-partners/>
- HR Profiles: <https://zenodo.org/record/2564790> (Smithies 2019)
- Managing 100 Digital Humanities Projects: Digital Scholarship & Archiving in King's Digital Lab. *Digital Humanities Quarterly* 13.1 <http://www.digitalhumanities.org/dhq/vol/13/1/000411/000411.html> (Smithies et al. 2019)
- REF Checklist: <http://doi.org/10.5281/zenodo.3361580> (Ciula 2019)
- SDLC Wiki: [github.com/kingsdigitallab/sdlc-for-rse/wiki](https://github.com/kingsdigitallab/sdlc-for-rse/wiki) (KDL 2018-)
- SORSE slides and recording: <https://sorse.github.io/programme/talks/event-027/> (Smithies & Ciula 2020)

