

WP21: Data Portal

PI/Lead: Ronan Lyons/ Emma Squires

Overview:

Aim:

Background:

Objective 1: To continue to provide personal and consortia workspaces

Deliverable 1 Develop infrastructure to support the growth in the use of the Data Portal

Deliverable 2 To recruit specialist staff in genetics and imaging to support the development and services for cohorts and researchers.

Deliverable 3 Assist with the development and implementation of the DPUK ontology to cover multi-modal data pipelines in omics and imaging

Deliverable 4 Enhance analysis tools currently available (including for machine learning/AI)

Deliverable 5 Develop solutions to enable rapid replication analyses

Deliverable 6 Enhance computational options for Data Portal users to support analysis of complex, large scale, multi-modal data to include additional storage, omics and imaging analysis toolkits, and enhanced compute power via GPU and HPC support

Deliverable 7 Create knowledge graphs with supporting documentation on Data Portal use and facilities to inform our development

Deliverable 8 Provision of DoIs

Objective 2: Enable centralised data linkage for cohort studies

Deliverable 9 To establish centralised data linkage for cohort studies and further focus existing cohort collaborators' data onto dementia research

Objective 3: Support the development of the Trial Delivery Framework to continue support for Pre-clinical recruitment and extend to Clinical recruitment

Deliverable 10 Continue to develop DPUK data collection platforms for digital phenotyping – Great Minds and the Clinical Studies Register

Deliverable 11 Enable the development of a clinical network database for recruitment and trial delivery

Deliverable 12 Continue to develop DPUK's role on the Alzheimer's Disease Data Initiative after public launch in late 2020 and scale-up phases in 2021

Objective 5: Development of current static and interactive DPUK data discovery and visualisation tools

Deliverable 13: Further updating and development of the DPUK metadata tools to include multi-modal data

Deliverable 14 Re-specification and re-development of the DPUK visualisation tools to include multi-modal data

Objective 6: Development of a Genomics platform

Deliverable 15 Implementation of an initial use case

Deliverable 16 Development of a library of genomics pipelines

Deliverable 17 Systemise the processing of genomics pipelines

Deliverable 18 Development of retention and access capability for genomic data

Objective 7: Support, maintenance and development of the Imaging platform

Deliverable 19 Support and maintenance

Deliverable 20 Systemise the adapted UK Biobank imaging pipelines

Deliverable 21 Systemise the processing of additional imaging pipelines

Deliverable 22 Develop systems to enable harmonisation techniques to produce generalisable IDPs

Objective 8: Development of a Tissue Discovery Platform

Deliverable 23 Implementation of a web-based searchable cell line and clinical information capability

Deliverable 24 Develop application system

Deliverable 25 Further development of the Tissue Discovery Platform

Objective 9 DPUK/MRC project reporting

Deliverable 26 Quarterly reports

Deliverable 27 Annual reports

Deliverable 28 Yearly financial reports

Deliverable 29 Final report

Objective 1: To continue to provide personal and consortia workspaces

Description	How and who	Outcome	Dates
Deliverable 1 Develop infrastructure to support the growth in the use of the Data Portal			
M1.1.1 Monitor usage and increase desktops and storage to ensure availability for users	Compile report to send out each month	Monthly reports shared	Jul '21
M1.1.2 Monitor usage and increase desktops and storage to ensure availability for users	Compile report to send out each month	Monthly reports shared	Jul '22
M1.1.3 Monitor usage and increase desktops and storage to ensure availability for users	Compile report to send out each month	Monthly reports shared	Jul '23
M1.1.4 Monitor usage and increase desktops and storage to ensure availability for users	Compile report to send out each month	Monthly reports shared	Jul '24
M1.1.5 Monitor usage and increase desktops and storage to ensure availability for users	Compile report to send out each month	Monthly reports shared	Jul '25
Deliverable 2 To recruit specialist staff in genetics and imaging to support the development and services for cohorts and researchers.			
M1.2.1 Specialist Research Data Manager member to be added to the Swansea team	Recruit SRDM	SRDM recruited	Jul '21

Deliverable 3 Assist with the development and implementation of the DPUK ontology to cover multi-modal data pipelines in omics and imaging			
M1.3.1 All Imaging and Genomics cohort study data held on DPUK to be re-categorised as per the mature ontology	Re-categorised all cohort data held	All data re-categorised	Dec '22
Deliverable 4 Enhance analysis tools currently available (including for machine learning/AI)			
M1.4.1 Provide suite of ML/AI tooling for Portal users as a Portal library	Source ML/AI tools for the portal library	Suite of tools available for DP users in Portal Library	Dec '22
M1.4.2 Provide further array of software within the virtual environment to cater for emerging analytical methodologies (e.g. Stata/MP, further Python IDEs etc.)	Source further LM/AI tools for the portal library	A further array of tools available for DP users	Dec '22
Deliverable 5 Develop solutions to enable rapid replication analyses			
M1.5.1 Fully integrated Git repository available within DPUK Data Portal	Integrate Git repository	Git repository available for DP users	Jun '21
M1.5.2 Incorporation of a Concept library	Incorporate the concept library	Concept library incorporated within the DP	Dec '22
M1.5.3 Provide access to community tools such as Docker for increased sharing and replication of research	Source community tools, make available	Community tools available on the DP (number?)	Dec '22
Deliverable 6 Enhance computational options for Data Portal users to support analysis of complex, large scale, multi-modal data to include additional storage, omics and imaging analysis toolkits, and enhanced compute power via GPU and HPC support			
M1.6.1 Provide capability to supply services based upon HPC cluster for multi-modal research	Arrange and provide capability	HPC cluster multi-modal research capability available	Jun '21
M1.6.2 Provide capability to supply services based upon GPU for multi-modal research	Arrange and provide capability	GPU multi-modal research capability available	Sep '21
M1.6.3 Fully operational HPC/GPU support for complex analyses, based on a cost recovery model for scale of use	Support process created	HPC/GPU cost recovery model available	Dec '22
Deliverable 7 Create knowledge graphs with supporting documentation on Data Portal use and facilities to inform our development			
M1.7.1 Production of graphs and documentation to highlight how many users are using DPUK desktops, in order to foresee where expansion of desktop pools is most necessary	Create graphs and documentation	Documentation available	Sep '21

M1.7.2 Production of graphs and documentation to monitor and improve data portal operations	Create graphs and documentation	Documentation available	Dec '22
M1.7.3 Incorporate the use of Google analytics to further understand data being accessed and techniques being used	Analyse Google analytics data	Results available	Sep '21
Deliverable 8 Provision of DoIs			
M1.8.1 DOIs available for use by DPUK and researchers to reference and share results and developments	Compile document	Complete list of DoIs available	Jun '21
Objective 2: Enable centralised data linkage for cohort studies			
Deliverable 9 To establish centralised data linkage for cohort studies and further focus existing cohort collaborators' data onto dementia research			
M2.9.1 Identify and negotiate access to linkage information	Developmental work	Process to access linkage information is complete	Sep '21
M2.9.2 Develop a mechanism to link cohort data to specified ONS and NHS digital information	Developmental work	Mechanism available to link cohort data to specific ONS and NHS digital information	Mar '22
M2.9.3 To have 5 cohorts with centralised data linkage	Work to 5 cohorts	5 cohorts data linkage information available	Dec '22
M2.9.4 To have 10 cohorts with centralised data linkage	Work to 10 cohorts	10 cohort's data linkage information available	Dec '23
Objective 3: Support the development of the Trial Delivery Framework to continue support for Pre-clinical recruitment and extend to Clinical recruitment			
Deliverable 10 Continue to develop DPUK data collection platforms for digital phenotyping – Great Minds and the Clinical Studies Register			
M3.10.1 Support for increased cohorts and participants	Provide support	Support provided	Sep '21
M3.10.2 Support for increase selection variables	Provide support	Support provided	Jun '22
M3.10.3 Re-development of visibility and selection tools	Re-develop visibility and selection tools	Improved visibility and selection tools	Dec '22
Deliverable 11 Enable the development of a clinical network database for recruitment and trial delivery			
M3.11.1 Develop a new visualisation tool and selection tool for Clinical recruitment to trials	Develop new tools	New tools created	Jun '22

M3.11.2 Develop a new database for Clinical recruitment to trials	Develop new database	Database created	Sep '22
Objective 4: Development of interoperability with other platforms in the dementia and neuroscience community			
Deliverable 12 Continue to develop DPUK's role on the Alzheimer's Disease Data Initiative after public launch in late 2020 and scale-up phases in 2021			
M4.12.1 Inclusion of further DPUK cohorts on the Alzheimer's Data Workbench (currently 1 DPUK cohort is available to be requested via the Workbench)	Add more cohorts to Alzheimer's Data Workbench	More than one cohort available on the Alzheimer's Data Workbench	Dec '21
M4.12.2 Further development of a secure tokenisation process for approved Alzheimer's Workbench users to be able to analyse DPUK cohort data via API	Develop secure tokenisation process	Approved researchers can analyse DPUK cohort data via API securely	Dec '21
Objective 5: Development of current static and interactive DPUK data discovery and visualisation tools			
Deliverable 13: Further updating and development of the DPUK metadata tools to include multi-modal data			
M5.13.1 User feedback session	Plan, prepare and hold a user feedback session on Directory and Matrix	User feedback session held	Jun '21
M5.13.2 Consistent updating and development of the DPUK Cohort Matrix	Use feedback from users to develop and update Cohort Matrix	Cohort Matrix developed and updated	Dec '21
M5.13.3 Consistent updating and development of the DPUK Cohort Directory	Use feedback from users to develop and update Cohort Directory	Cohort Directory developed and updated	Dec '21
Deliverable 14 Re-specification and re-development of the DPUK visualisation tools to include multi-modal data			
M5.14.1 User feedback session	Plan, prepare and hold a user feedback session on Directory and Matrix	User feedback session held	Dec '21
M5.14.2 Develop tender	Prepare a brief based on user needs for tender	Re-specification and re-development tender offered to suppliers	Dec '21
M5.14.3 Select supplier	Collate bids and select one	Supplier contracted to complete work	Dec '21
M5.14.4 Re-development of the DPUK Cohort Explorer	Supplier re-development of DPUK Cohort Explorer	DPUK Cohort Explorer has been updated to meet users' needs	Jun '22
M5.14.5 Re-development of the Great Minds and CSR Feasibility tools	Supplier re-development of Great Minds and CSR Feasibility tools	Great Minds and CSR Feasibility tools have been updated to meet users' needs	Dec '22
Objective 6: Development of a Genomics platform			
Deliverable 15 Implementation of an initial use case			

M6.15.1 Provide capability to enable the analysis of cross cohort data to develop a polygenic risk score	Create initial use case so a polygenic risk score can be found from analysing cross cohort data	Initial use case implemented	Dec '21
Deliverable 16 Development of a library of genomics pipelines			
M6.16.1 Work with genomics specialists to specify a library of processing pipelines for the interrogation of genomic data	Create a library of processing pipelines	A library of genomics pipelines	Jun '22
Deliverable 17 Systemise the processing of genomics pipelines			
M6.17.1 Develop the capability to incorporate the specified processing pipelines on genomic data	Develop the systems for processing of genomics pipelines	Genomics pipeline process is complete	Dec '22
Deliverable 18 Development of retention and access capability for genomic data			
M6.18.1 Specification and development of a retention capability to enable access to genomic data	Complete the specification and development of retention capability	Genomic data can be accessed	Jun '23
Objective 7: Support, maintenance and development of the Imaging platform			
Deliverable 19 Support and maintenance			
M7. 19.1 Transfer the responsibility of the support and maintenance of the DPUK Imaging platform to UKSeRP	Complete the handing over process of the support and maintenance	DPUK Imaging Platform is supported and maintained by UKSeRP.	Dec '21
Deliverable 20 Systemise the adapted UK Biobank imaging pipelines			
M7.20.1 Adapt UK Biobank processing pipeline to different number of modalities and sequence parameters	Use the UK Biobank processing pipeline to work with other modalities and sequence parameters	A number (2 or 50???) of different modalities and sequence parameters have been created	Mar '22
Deliverable 21 Systemise the processing of additional imaging pipelines			
M7.21.1 Support the application of the pipeline to other datasets	Provide support for the pipeline to be applied to other data sets	Additional imaging pipelines process created (number??)	Sep '22
Deliverable 22 Develop systems to enable harmonisation techniques to produce generalisable IDPs			
M7.22.1 Enable the incorporation of harmonisation techniques in the processing pipelines	Create the system so harmonisation techniques can be used	Harmonisation techniques enabled in the processing pipelines	Dec '23
Objective 8: Development of a Tissue Discovery Platform			
Deliverable 23 Implementation of a web-based searchable cell line and clinical information capability			
M8.23.1 Development of initial capability	Developmental work to create capability	Initial capability created	Dec '22

Deliverable 24 Develop application system			
M8.24.1 Identify application requirements and develop separate application process for the Tissue Development Platform	Work out the requirements and create an application process	Application system created	Mar '23
Deliverable 25 Further development of the Tissue Discovery Platform			
M8.25.1 Specification and development of enhancements to the platform	Using feedback from users develop the Tissue Discovery Platform	Updated Tissue Discovery Platform	Jun '24
Objective 9 DPUK/MRC project reporting			
Deliverable 26 Quarterly reports			
M9.26.1 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Mar '21
M9.26.2 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Jun '21
M9.26.3 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Sep '21
M9.26.4 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Mar '22
M9.26.5 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Jun '22
M9.26.6 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Sep '22
M9.26.7 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Mar '23
M9.26.8 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Jun '23
M9.26.9 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Sep '23
M9.26.10 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Mar '24
M9.26.11 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Jun '24
M9.26.12 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Sep '24

M9.26.13 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Mar '25
M9.26.14 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Jun '25
M9.26.15 Provide Quarterly reports detailing project deliverables and outcomes.	Online quarterly form to be completed for DPUK for MRC meetings	Quarterly report submitted	Sep '25
Deliverable 27 Annual reports			
M9.28.1 Annual reports submitted on time for OB meetings	Annual form to be completed for DPUK for MRC meetings. Annual reports submitted to confirm milestone completion for next payment scheduled	Annual report submitted	Dec '21
M9. 28.2 Annual reports submitted on time for OB meetings	Annual form to be completed for DPUK for MRC meetings. Annual reports submitted to confirm milestone completion for next payment scheduled	Annual report submitted	Dec '22
M.28.3 Annual reports submitted on time for OB meetings	Annual form to be completed for DPUK for MRC meetings. Annual reports submitted to confirm milestone completion for next payment scheduled	Annual report submitted	Dec '23
M9.28.4 Annual reports submitted on time for OB meetings	Annual form to be completed for DPUK for MRC meetings. Annual reports submitted to confirm milestone completion for next payment scheduled	Annual report submitted	Dec '24
Deliverable 28 Yearly financial reports			
M9.28.1 Financial reports submitted at the end of each year	Yearly financial statement on spending	Financial report submitted	Dec '21
M9.28.2 Financial reports submitted at the end of each year	Yearly financial statement on spending	Financial report submitted	Dec '22
M9.28.3 Financial reports submitted at the end of each year	Yearly financial statement on spending	Financial report submitted	Dec '23
M9.29.4 Financial reports submitted at the end of each year	Yearly financial statement on spending	Financial report submitted	Dec '24
M9.28.5 Financial reports submitted at the end of each year	Yearly financial statement on spending	Financial report submitted	Dec '25
Deliverable 29 Final report			
M9.29.1 Final report submitted at end of project	PI to complete final report and send to DPUK /MRC for final payment to be made.	Final report submitted	Dec '25

	2021				2022				2023				2024				2025																														
	Q1		Q2		Q3		Q4		Q5		Q6		Q7		Q8		Q9		Q10		Q11		Q12		Q13		Q14		Q15		Q16		Q17		Q18		Q19		Q20								
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov
WP 21 Data Portal																																															
Objective 1	01, D.1																																														
	01, D.2																																														
	01, D.3																																														
	01, D.4																																														
	01, D.5																																														
	01, D.6																																														
	01, D.7																																														
	01, D.8																																														
Objective 2	02, D.9																																														
Objective 3	03, D.10																																														
		03, D.11																																													
Objective 4	04, D.12																																														
Objective 5	05, D.13																																														
		05, D.14																																													
Objective 6	06, D.15																																														
		06, D.16																																													
			06, D.17																																												
			06, D.18																																												
Objective 7	07, D.19																																														
	07, D.20																																														
		07, D.21																																													
Objective 8		07, D.22																																													
		08, D.23																																													
		08, D.24																																													
		08, D.25																																													
Objective 9	09, D.26																																														
		09, D.27																																													
		09, D.28																																													
		09, D.29																																													