Introduction to LEARN KPIs

In this document, LEARN provides a list of Key Performance Indicators (KPIs) that should help each individual research institution measure how successful they are being in implementing the recommendations of the LERU Research Data Roadmap and the Toolkit[1] produced by LEARN.

The first set of indicators is aimed at **preparing** the institution for managing research data, whilst the second set should be used to measure the **implementation** of an institutional policy on research data.

Each KPI has at least one direct connection with one of the LEARN Toolkit sections: *Advocacy; Costs; Model RDM Policy; Open Data; Policy and Leadership; Research Data Infrastructure; Roles, Responsibilities and Skills; Subject Approaches; Tool Development.*

These KPIs are based on the LEARN Model RDM Policy and the expected values defined here should be taken as a good indication that a suitable implementation of RDM policy has been reached. For instance, it is important to know the uptake of the facilities provided, and levels of researcher engagement. Success in both of these areas can be measured using the proposed KPIs.

However, before adopting these indicators and establishing their expected values, the institution must consider what the goals are that are defined in its RDM policy and how it would like to achieve them.

Among the KPIs for implementation of the policy, LEARN suggests measuring the use of the designated facilities provided to researchers by the institution (KPI I3). This figure might be expected to grow quickly after the adoption of a RDM policy, and it is advisable to measure take-up on a monthly basis. The increase in the number of datasets stored and published in the designated facilities could reveal a need to improve or enlarge the facilities, and it could also affect the costs of the institution's overall RDM activity.

LEARN proposes measuring the number of datasets stored in the designated infrastructure even if, afterwards, they are just archived and not published (KPI I4). In cases where data cannot be shared openly, the publication of the related metadata is advisable to let people know the existence of a dataset and how to access to it, if possible, by means of closed procedures.

LEARN also recommends the use of indicators to monitor the institutional use of external facilities (KPI I6), such as disciplinary repositories, e.g. the SAO/NASA Astrophysics Data System (ADS), Crystallography Open Database, bepress Legal Repository etc. The institution should

[1] http://learn-rdm.eu/en/research-data-management-toolkit-now-available/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654139.

ascertain the reasons for any unexpected increase in these numbers, which may lead to a redefinition of their RDM policy.

Another important indicator by which to measure implementation is the engagement of researchers – not only their use of the facilities provided, but also their attendance at training sessions (KPI I9). This indicator can be split by discipline, to ascertain the level of adoption of best practice by individual subject areas, enabling the design of specific actions as a result.

The list of KPIs is appended, together with a scorecard that can be used on an annual basis to check the performance of the research institution as a whole in the RDM process. LEARN also suggests scoring by use of a Traffic Lights system of Red, Amber Green. A Green result indicates that the institution has met that KPI; Amber that the institution expects to meet the KPI but that it has not yet done so; Red indicates that the KPI has not or cannot be met. The more Green scores, the better the institution is performing in terms of Research Data Management activity.

The expected values that are proposed in the LEARN KPI lists are suggested indications of achievement. Some of the given expected values are minima or maxima; other indicators, especially the ones related to preparation, have a binary YES/NO format, which LEARN expects to be positive before the institution proceeds to measurement of the implementation indicators.

LEARN also expects that each institution could modify/increase the expected values associated with these KPIs over time, according to policy and feasibility. As stated before, each institution should have a clear definition of the goals of its policy at the outset. Ongoing review of policy and related goals should in turn inform the ongoing definition and re-definition of the KPIs, as services mature.

.

learn

List of KPIs for preparation

Number	KPI	LEARN Toolkit Theme	Measurement	Expected Value	Rationale for measure
P1	Institutional policy	Policy and Leadership	Policy exists	YES	Any research performing institution should have a policy on RDM in place
P1.1	Alignment with the LEARN RDM Model	Model RDM Policy	Review of items in Policy which are included in the LEARN model	90%	To avoid multiplicity of policy models, LEARN suggests a comparison with the LEARN model
P2	Steering committee dedicated to RDM	Advocacy; Policy and Leadership	Steering Committee exists	YES	Any research performing institution should have a steering committee or a designated group to lead the institutional policy on RDM
P3	Services created to work on RDM	Roles, Responsibilities and Skills	Number of new services created	>1	Any research performing institution should have at least a service for data stewardship and to help researchers on RDM tasks and duties
P4	Staff involved in RDM	Advocacy; Roles, Responsibilities and Skills	FTE staff dedicated to RDM	>2	Any research performing institution should assign part of its staff for data stewardship and to help researchers on RDM tasks and duties
P5	Job profiles dedicated to RDM	Roles, Responsibilities and Skills	Number of new job profiles created or updated	>1	Any research performing institution should create new job profiles for working on data stewardship and to help researchers on RDM tasks and duties
P6	Information point on RDM	Advocacy	Information Point exists	YES	Any research performing institution should have at least one information point, physical and virtual, on RDM.
P7	Training sessions on RDM	Advocacy; Roles, Responsibilities and Skills	Number of sessions developed in a year	>5	Any research performing institution should train its staff, researchers and students on RDM best practices

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654139.

List of KPIs for implementation

Number	KPI	LEARN Toolkit Theme	Measurement	Expected Value	Rationale for measure
I1	Monitoring of the institutional policy	Policy and Leadership, Tool Development	Monitoring activity exists	YES	Any research performing institution with a RDM policy in place should have a way to monitor it
I1.1	Review of the policy	Policy and Leadership, Tool Development	Number of reviews in a year	1	A yearly review is advised as best practice to monitor the policy
I1.2	Updates of the policy	Policy and Leadership, Tool Development	Number of updates since its enforcement	< 1/year	An update of the policy means that some of the initial expectations are not fulfilled. It is advisable to make as many updates as needed, but several updates in a year would mean that the policy is not well designed
I2	Disciplines engaged in RDM within the research institution	Advocacy; Subject approaches	Percentage of disciplines engaged in RDM activities	90%	Any research performing institution expects to engage all its researchers in a policy. Therefore, is advisable to monitor the engagement of all disciplines, even if some have difficulties in starting work on RDM
I3	Datasets stored in the policy-designated infrastructure(s)	Research Data Infrastructure	Number of datasets stored	Increasing year after year	If the institution establishes an infrastructure to store datasets, it must monitor the growth of the infrastructure and measure the number of datasets available
I4	Datasets published on the policy-designated platform(s) with a clear statement on terms of use	Open Data; Research Data Infrastructure; Tool development	Number of datasets published with a clear statement (I4.1) / Number of all datasets published (I4.2)	100%	Although open data could be a final goal, many datasets cannot be shared openly due to some restrictions such as confidentiality, security, privacy. Nevertheless, it is required that at least metadata is publicly available and there is an indication about the degree of openness of each dataset

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654139.

Number	KPI	LEARN Toolkit Theme	Measurement	Expected Value	Rationale for measure
I5	Persistent identifiers for published research data	Research Data Infrastructure; Tool development	Number of persistent identifiers	I4.1	In order to follow the FAIR principles, the institution should provide/advocate for a persistent identifier for any dataset published in its facilities
I6	Datasets stored outside the policy-designated infrastructure	Research Data Infrastructure	Number of datasets stored	< I3	If the institution establishes an infrastructure to store datasets, the number of resources stored outside policy-designated infrastructure should decrease year after year. However, it will never reach zero because in some disciplines researchers have consolidated data repositories
I6.1	Number of datasets shared outside the policy-designated infrastructure	Open Data; Research Data Infrastructure	Number of datasets shared outside the policy-designated infrastructure	< I4.2	If the institution establishes an infrastructure to store datasets, the number of resources stored outside the policy-designated infrastructure should decrease year after year. However, it will never reach zero because in some disciplines researchers have consolidated repositories.
I7	Active researchers using policy-designated facilities	Advocacy; Research Infrastructures; Roles, Responsibilities and Skills	Percentage of active researchers using facilities	90%	Any research performing institution expects to engage all its researchers in a policy and that they use the facilities provided. This indicator can be applied to individual disciplines, if required.
I8	Amount of research income dedicated to RDM activities	Costs	Percentage of research income dedicated to RDM. The costs should include infrastructure (I8.1), staff (I8.2), and activities (I8.3)	>5%	Recommendation of High Level Expert Group's Report on the European Open Science Cloud

Number	KPI	LEARN Toolkit Theme	Measurement	Expected Value	Rationale for measure
I9	Training sessions on RDM	Advocacy; Roles, Responsibilities and Skills	Number of sessions developed in a year	>5	Any research performing institution should train its staff, researchers and students in RDM best practice. Same as P8 but measured after the adoption of a policy
I9.1	Active researchers attending training sessions	Advocacy; Roles, Responsibilities and Skills; Subject approaches	Percentage of active researchers attending training sessions in a year	>10%	Every year around, at least 10% of researchers should attend training sessions to know how to manage research data according to the adopted policy
I10	Queries for support received	Advocacy; Tool Development	Number of queries in a year	>50	A value that is expected to be high in the first years after the adoption of a policy and lower after its consolidation
I11	Data Management Plans (DMPs) created	Tool Development	Number of plans created	>20	The policy should include a provision requiring the elaboration of DMPs; accordingly it should be monitored
I11.1	Data Management Plans published	Advocacy; Tool Development	Number of plans published	90% of I11	The publication of DMPs represent best practice as they can serve as an example for beginners in RDM
I12	Incidents	Tool Development	Number of incidents in RDM activities in a year	< 50	A good policy and best practice will allow the institution to reduce incidents in RDM, such as losing data because of negligence
I12.1	Datasets deleted	Tool Development	Number of datasets deleted from the designated storage facility	<1% of I3	It is expected that the deletion process will occur, but the number of datasets deleted will be nearly zero. If the number is high, then the policy must be reviewed
I12.2	Datasets withdrawn	Tool Development	Number of datasets deleted from the designated storing facility	<1% of I3	It is expected that the deletion process will occur, but the number of datasets withdrawn will be nearly zero. If the number is high, then the policy must be reviewed

learn

KPI Scorecard

Number	KPI	Measurement	Expected Value	Score	Red/Amber/ Green
P1	Institutional policy	Policy exists	YES		
P1.1	Alignment with the LEARN RDM Model	Review of items in Policy which are included in the LEARN model	90%		
P2	Steering committee dedicated to RDM	Steering Committee exists	YES		
P3	Services created to work on RDM	Number of new services created	>1		
P4	Staff involved in RDM	FTE staff dedicated to RDM	>2		
P5	Job profiles dedicated to RDM	Number of new job profiles created or updated	>1		
P6	Information point on RDM	Information Point exists	YES		
P8	Training sessions on RDM	Number of sessions developed in a year	>5		
I1	Monitoring of the institutional policy	Monitoring activity exists	YES		
I1.1	Review of the policy	Number of reviews in a year	1		
I1.2	Updates of the policy	Number of updates since its enforcement	< 1/year		
I2	Disciplines engaged in RDM within the research institution	Percentage of disciplines engaged in RDM activities	90%		
I3	Datasets stored in the policy-designated infrastructure(s)	Number of datasets stored	Increasing year after year		
I4	Datasets published in the policy-designated platform(s) with a clear statement on terms of use	Number of datasets published with a clear statement (I4.1) / Number of all datasets published (I4.2)	100%		
I5	Persistent identifiers for published research data	Number of persistent identifiers	I4.1		

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654139.

Number	KPI	Measurement	Expected Value	Score	Red/Amber/ Green
I6	Datasets stored outside the policy-designated infrastructure	Number of datasets stored	< I3		
I6.1	Number of datasets shared outside the policy-designated infrastructure	Number of datasets shared outside the policy-designated infrastructure	< I4.2		
I7	Active researchers using policy-designated facilities	Percentage of active researchers using facilities	90%		
I8	Amount of research income dedicated to RDM activities	Percentage of research income dedicated to RDM. The costs should include infrastructure (I8.1), staff (I8.2), and activities (I8.3)	>5%		
I9	Training sessions on RDM	Number of sessions developed in a year	>5		
I9.1	Active researchers attending training sessions	Percentage of active researchers attending training sessions in a year	>10%		
I10	Queries for support received	Number of queries in a year	>50		
I11	Data Management Plans created	Number of plans created	>20		
I11.1	Data Management Plans published	Number of plans published	90% of I11		
I12	Incidences	Number of incidences in RDM activities in a year	> 50		
I12.1	Datasets deleted	Number of datasets deleted from the designated storage facility	<1% of I3		
I12.2	Datasets withdrawn	Number of datasets deleted from the designated storing facility	<1% of I3		