

Integrated Data Stewardship

Governance, Evidence and Decision Frameworks for Research
Data

Ondřej Ryška

About This Book

This monograph addresses a structural problem in the governance of research data. The normative and procedural layer that governs research data has grown by accretion, and its instruments — FAIR, lifecycle models, data management plans, metadata standards, provenance specifications, preservation reference models, data protection law, repository certification and research assessment reform — are each internally coherent and none was designed to compose with the others. What is missing is the connecting layer that converts the conditions of a specific project or institution into a defensible decision, together with the evidence that the decision was carried out.

The book proposes, operationalizes and assesses an integrated framework for institutional data stewardship. The framework identifies the dimensions across which stewardship decisions are taken, the variables that populate each dimension, the relations between them and the decision rules by which competing requirements are prioritized. It is derived from a mapped corpus of frameworks, standards, institutional policies and literature spanning twenty-three domains of practice, and it is assessed against institutional and project settings in which the competing requirements actually arise.

Intended readership

The book is addressed to researchers in research data management and open science, to data stewards and research support professionals, to those responsible for research policy in universities and research institutes, and to doctoral candidates working on the governance of digital research objects. It presupposes familiarity with the vocabulary of research data management but not with any particular framework.

What the book does not do

It is not a manual, and it does not prescribe procedures for individual institutions. Where practical implications follow from the results, they are derived in Part VI and marked as derivations rather than presented as instructions. The framework bounds its own transferability explicitly:

disciplinary conventions, legal regime, institutional scale and available infrastructure change what any framework of this kind can determine, and Chapter 14 sets out the conditions under which its determinations weaken.

A note on the evidential status of the text

Claims in this book belong to three kinds, and each is signalled where it occurs. Received facts carry a citation to their primary source. Syntheses performed by this study are marked as such in the sentence that states them. Original results of this study are stated as results and carry their evidence.

Where the argument requires an empirical result that the underlying research programme has not yet produced, the text carries an explicit marker in the form [EVIDENCE REQUIRED: ...], naming what must be measured and what claim it would support. Citation markers present in earlier drafts have been resolved against verified sources; the bibliography records the verification. The evidence markers are deliberate. They record the boundary between what the framework asserts and what it has so far demonstrated, and they are to be resolved before the manuscript is submitted for review.

Manuscript status

This document is a working draft. The table below records the state of each chapter.

Table 0-1. Draft status by chapter.

Part	Chapter	Status
I	1. Fragmentation in Research Data Governance	drafted
I	2. Data Stewardship as an Institutional Capability	drafted
I	3. FAIR, CARE and TRUST: A Critical Reading	drafted
II	4. Research Design, Corpus and Analytical Method	drafted
III	5. Governance, Roles and Accountability	drafted
III	6. The Research Data Lifecycle and Provenance	drafted
III	7. Planning as Commitment: Data Management Plans	drafted
III	8. Description, Quality and the Conditions of Reuse	drafted
III	9. Protection, Legal Regime and Controlled Access	drafted
III	10. Reproducibility and Computational Integrity	drafted
III	11. Preservation, Publication and Recognition	drafted
III	12. Operationalizing the Framework	drafted
IV	13. Application to Institutional Settings	drafted
IV	14. Robustness, Sensitivity and Boundary Conditions	drafted
V	15. Critical Discussion	drafted
VI	16. Implications for Institutions and Infrastructures	drafted
VI	17. Conclusion	drafted

All seventeen chapters are drafted, and the end matter now includes the consolidated and verified bibliography, the glossary, the eight figures, the external review protocol (Appendix B) and the rights register (Appendix C). Appendix A indexes every outstanding evidence and citation marker in the manuscript and constitutes the research programme that must be completed before submission. All citation markers have been resolved against verified

sources and the bibliography carries 59 entries. Still outstanding: the empirical programme of Appendix A, the external review itself and the index, which is assembled at production.

PART I

The Problem and Its Theoretical Setting

Chapter 1. Fragmentation in Research Data Governance

1.1 A regulated domain without a governing framework

Research data are governed by a dense and still expanding layer of principles, standards, legal instruments and funder conditions. The FAIR principles describe target properties of digital research objects (Wilkinson et al. 2016); lifecycle models describe the stages through which data pass; data management plans record intentions; metadata standards prescribe description; provenance models record derivation. Preservation is addressed by the OAIS reference model and by repository certification (Consultative Committee for Space Data Systems 2012; CoreTrustSeal 2022), legal and ethical conditions by data protection law and by ethics review (European Parliament and Council 2016), and data publishing by repository selection criteria and persistent identifier infrastructures (Science Europe 2021). Each instrument is competent within the area it was designed for.

The difficulty addressed by this book is not the deficiency of any single instrument. It is that the instruments are not connected into a system that converts the conditions of a given project or institution into a defensible decision. An institution adopting all of them simultaneously acquires a set of partially overlapping obligations addressed to different objects, held by different owners and supported by evidence of unequal strength. Table 1-1 sets out this partial coverage and identifies, for each decision area, what the dominant instrument leaves undetermined.

Table 1-1. Partial coverage of the normative and procedural layer governing research data, with the decisions each dominant instrument leaves undetermined.

Decision area	Dominant instrument	Object addressed	Left undetermined
Principles of description and re-use	FAIR	dataset and its metadata	which properties apply to which data class, and at what cost
Planning	data management plan	project intentions	whether the plan is realized, and who verifies it
Lifecycle	stage and flow models	data objects over time	authoritative copy, versioning threshold, end of life
Description	metadata standards and profiles	discovery and semantics	the boundary between core and domain profile
Provenance	PROV and packaging specifications	derivation of objects	granularity proportionate to risk
Quality	quality assurance frameworks	fitness for use	acceptable residual error for a stated purpose
Protection	data protection law, classification schemes	persons and confidentiality	access regime for data that are neither open nor closed
Preservation	OAIS, certification requirements	archived information	retention period and transfer of custody
Publishing	repository selection criteria	deposit and identifiers	trade-off between disciplinary fit and certified trust
Recognition	research assessment reform	outputs and roles	how stewardship effort enters institutional assessment

The layer described in Table 1-1 grew by accretion, and its shape reflects that history. Each instrument was formulated in response to a distinct failure: data that could not be found, plans that were never written, archives that lost the ability to interpret their own holdings, repositories that closed without succession. Because each was designed against a separate failure mode, each is internally coherent and none was designed to compose with the others. The absence of a connecting layer is thus a predictable consequence of how the field developed rather than a lapse of any of its communities.

The pattern visible across the table is systematic rather than accidental. The instruments are strongest where an object can be described in isolation and weakest where a decision has to be taken under competing constraints. Decisions of the second kind are the ordinary content of institutional and project practice. The gap is therefore located precisely where practice is densest.

1.2 Why fragmentation constitutes a research problem

A first response treats the situation as an implementation gap, remediable by better training, more tools and stricter compliance checks. That response is inadequate for three reasons that this book treats as constitutive of the problem. The instruments are non-commensurable, they carry evidence of unequal strength, and they conflict in ways that cannot be resolved from inside any one of them. Each reason is developed in turn.

Non-commensurability arises because the instruments address different objects. FAIR is predicated of a dataset, TRUST of a repository, a data management plan of a project, data protection law of a processing operation and assessment reform of a career. A judgement that a dataset is FAIR therefore says nothing about whether the institution holding it can preserve it, and a certified repository does not make a deposited dataset fit for reuse. Composition across these objects is not defined by any of the frameworks that constitute them.

Unequal evidential strength compounds the problem. A legal requirement, a ratified standard, a peer-reviewed comparative study and an emerging community practice are routinely cited side by side in institutional policy as though they were of one kind. The revision underlying the corpus analysed here separated sources into three levels of evidential strength and applied that separation consistently across all 23 domains. The comparison reported in this book indicates that most decision-relevant guidance in the field rests on case evidence and consensus documents rather than on comparative measurement, which constrains how strongly any framework built on it can be stated.

Conflict is the third and decisive reason. An open licence maximizes reuse and may simultaneously violate a community protocol; a retention rule and an erasure obligation may both apply to the same records; disciplinary appropriateness and certified trustworthiness may point to different

repositories. Such cases are not failures of implementation, because each competing requirement is correctly applied. They can be settled only by an external rule that assigns priority, names the authority entitled to decide and records the residual risk that the decision leaves in place.

A fourth consideration concerns measurement rather than decision. The indicators in general use count documents and activities: plans written, consultations delivered, training sessions held. Such counts are insensitive to whether the measures described in a plan were carried out, and they cannot distinguish an institution with functioning escalation from one with an organogram. The absence of outcome-level measurement is not a separate deficiency but the same fragmentation seen from the assessment side.

The inadequacy of the implementation reading can also be shown from the supply side. A structured review confined to the life sciences alone analysed more than 300 distinct data stewardship tools (Aksenova et al. 2024), and the FAIRsharing registry records thousands of standards, databases and policies across disciplines (Sansone et al. 2019), which indicates saturation of supply rather than scarcity. Adoption of a further tool digitizes an unclear procedure without settling who decides, under what conditions and with what evidence of completion. What is scarce is not instrumentation but a decision structure into which instrumentation can be placed.

One additional property of the problem determines the form any solution can take. Decisions about research data are taken at two levels that are usually studied apart: the institution sets policy, infrastructure and minimum requirements, while the project takes the concrete decisions about its own data. A framework operating at one level only either produces policy that no project can apply or project practice that the institution cannot assess. The framework developed in Part III is therefore constrained to express both levels through the same variables, so that a project decision remains legible as an instance of institutional policy.

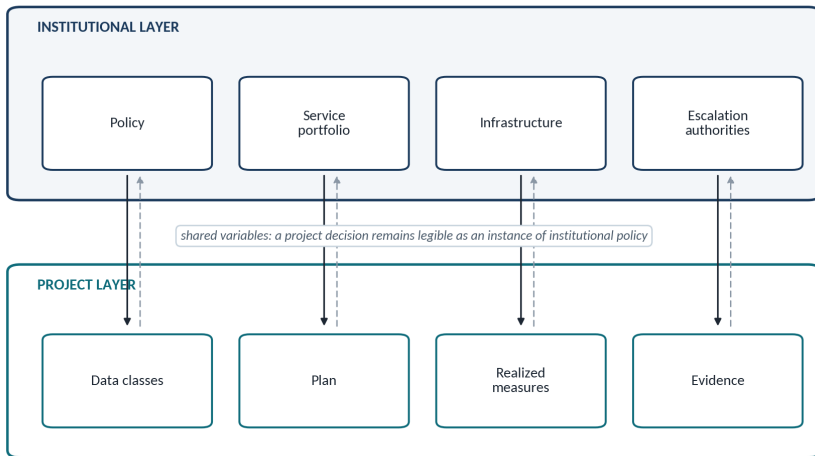


Fig. 1-1. Two-level decision structure, showing the institutional layer (policy, service portfolio, infrastructure, escalation authorities) and the project layer (data classes, plan, realized measures, evidence), with the shared variables that connect them. Author's own elaboration.

[EVIDENCE REQUIRED: an audit of institutional research data policies across a defined population of research organizations, recording for each of the ten decision areas in Table 1-1 which instrument is cited, whether a decision owner is named and whether an escalation path is stated — supporting the claim that coverage is partial and unconnected rather than merely uneven]

[EVIDENCE REQUIRED: paired comparison of data management plan content at submission against realized practice at project close, for a sample of projects, to establish the divergence between documentary compliance and realized practice]

1.3 The research question

The considerations above locate the problem at the level of framework design rather than instruction. What is missing is a mapping from the conditions of a case to a decision that can be defended, together with criteria by which the decision and its realization can be assessed. The book poses one central question and derives its parts from it.

Under what conditions can the partial normative, procedural and technical requirements governing research data be integrated into one framework of variables, decision rules and assessment criteria that yields defensible decisions at institutional and project level?

Six subordinate questions divide this problem and correspond to the parts of the book. First, how is the capability that produces data decisions organized within an institution, and which properties of that organization are decision-relevant rather than cosmetic? Second, what do the principle frameworks of the field determine, and at what point do they cease to determine anything? Third, which variables describe a data situation with sufficient discrimination to drive a decision rule? Fourth, what form of decision rule is warranted by evidence of the strength actually available in this field? Fifth, on what criteria can implementation, rather than documentation, be assessed? Sixth, how far does a framework derived from one national and institutional setting transfer, and what marks the boundary beyond which it does not?

1.4 Material, method and the contribution claimed

The material of the book is a corpus of 23 domains of research data practice, each documented with its stable consensus, its changes over the period 2024–2026, its open questions and its national implementation layer. The corpus was assembled as a comparative reading of legal instruments, ratified standards, infrastructure methodologies, peer-reviewed reviews and institutional case evidence, with the status of volatile sources recorded at a fixed verification date. The analysis proceeds by extracting, for every domain, the decisions actually taken, the variables that condition them and the residual risk that each option leaves. Part II sets out this design in full, including the criteria by which a claim is admitted as a received fact, a synthesis of this study or an original result.

The contribution claimed is threefold and is stated here so that it can be tested against the chapters that follow. The first component is the design of an integrated framework in which decisions across all 23 domains are expressed through a shared set of variables. The second is the operationalisation of that design as decision rules with explicit conditions, priority ordering between conflicting requirements, named decision authority and recorded residual risk, together with assessment criteria that address realization rather than documentation. The third is validation against domain profiles and an integrated case, accompanied by an explicit statement of the boundaries within which the framework transfers.

Three kinds of claim are separated consistently throughout the book, and the separation is part of the contribution rather than an editorial convention. Received facts are attributed to their primary source. Syntheses of this study are marked as such and are traceable to the comparison across the 23 domains. Original results are marked, and where the evidence that would support them has not yet been collected, the requirement is stated explicitly instead of being met by illustration. The density of such statements in the chapters that follow is a deliberate feature of an argument built on a field whose comparative evidence base remains thin.

1.5 Delimitation

The book is not a manual and offers no instructions. It neither assesses nor ranks software tools, whose landscape changes faster than any published analysis can track. It does not constitute legal advice: legal questions are treated as classes of decision with identifiable authority, not as answers valid in a particular jurisdiction at a particular time. It is not a systematic review conducted under PRISMA, and it does not propose national policy, although the Czech implementation layer is used throughout as a test of transferability, notably the introduction of the digital data collection as a recognized result type and the associated national metadata directory (Úřad vlády České republiky 2025; EOSC CZ 2025).

1.6 Structure of the argument

Part I establishes the problem and its setting. It defends the claim that fragmentation is a research problem, defines data stewardship as an institutional capability rather than a post or a service desk, and reads the normative layer of FAIR, CARE and TRUST critically in order to establish precisely what it leaves undetermined.

Part II sets out the research design. It describes the corpus of 23 domains, the grading of evidential strength, the procedure by which decision variables were extracted from documented practice, and the rules by which conflicting sources were weighted. It also states the limits of the design, in particular the predominance of case evidence over comparative measurement in the field.

Part III presents the integrated framework. It defines the variables, the decision rules that connect variable states to options, the priority ordering

applied when requirements conflict, the assignment of decision authority and escalation, and the assessment criteria that distinguish realized measures from documented intentions.

Part IV applies and validates the framework. Domain profiles for technical, natural, social and humanities data test whether the shared variables retain discriminating power under disciplinary variation, and an integrated case traces one project through the full set of decisions. The national implementation layer serves as a second validation setting.

Part V discusses the results. It examines alternative explanations for the observed convergence of institutional practice, states the boundaries of transferability, identifies the empirical work that the framework requires but does not yet possess, and specifies what evidence would falsify its central claims.

The next chapter takes up the first subordinate question. Before a framework can assign decisions to authorities, the capability that holds those authorities has to be characterized as an object of analysis rather than assumed as an organizational fact.

Chapter 2. Data Stewardship as an Institutional Capability

2.1 The unit of analysis

Data stewardship is frequently identified with a post, a department or a service desk, and the presence of any of the three is then read as evidence that an institution manages its research data. This chapter defends a different unit of analysis. Data stewardship is treated here as an institutional capability: a distribution of decision rights, competences and service commitments across several units, of which the named steward is one element among many. The claim defended is that the properties of this distribution, and not the organizational form in which it is packaged, determine whether decisions about research data are taken, recorded and carried out.

The relation to the central research question is direct. A framework that assigns decisions to authorities presupposes that the authorities exist, are identifiable and survive the departure of the persons currently holding them. If organizational form varied without consequence, the framework could treat it as noise; if one form were demonstrably optimal, the framework could fix it as a constant. The analysis in this chapter indicates that neither holds, and that organizational form therefore enters the framework as a variable whose states carry different residual risks.

2.2 From ad-hoc advice to a managed service portfolio

The first documented shift in the material analysed here runs from occasional advice to a described portfolio of services. In the advisory mode, support is delivered on request, its content depends on who is asked and no record connects a query to its resolution. In the portfolio mode, each service is described by its inputs, its outputs, its owner, its expected resolution time and the point at which it hands over to another authority. The difference is not administrative packaging: the second mode makes the support function auditable, and an auditable function can be assessed by outcome rather than by volume.

The shift is coupled to a second one, from stand-alone support to integration with the systems in which research is actually administered. Data management plans, current research information systems, grant administration, storage provisioning, repositories and ethics approval are increasingly connected, so that a decision recorded in one system becomes visible in the others. Integration of this kind changes the nature of the evidence available about an institution's data practice, because realized measures leave traces independent of self-report. The comparison across the 23 domains reported here indicates that this coupling, rather than the existence of a service catalogue as a document, distinguishes institutions that can demonstrate implementation from those that can only describe intentions.

2.3 Organizational models and their trade-offs

Four organizational models recur across the institutional case evidence. The centralized model concentrates methodology, review and service delivery in one unit; the distributed model places support within faculties or institutes; the federated model combines a central methodological and infrastructural core with domain-anchored stewards; the project-embedded model attaches a steward to a specific project team. Table 2-1 states the benefit, the limitation and the fitting context of each, together with the strength of the evidence supporting the pairing.

Table 2-1. Organizational models for data stewardship, with their principal trade-off, the context in which the case evidence locates them and the strength of that evidence.

Model	Principal benefit	Principal limitation	Context in the case evidence	Evidence strength
Centralized	uniform quality and control	distance from disciplinary practice; capacity bottleneck	smaller or methodologically homogeneous institutions; early service	institutional case studies
Distributed	disciplinary relevance	inconsistent rules; duplication	large faculties with mature local support	institutional case studies
Federated	standardisation combined with domain expertise	coordination cost; contested funding lines	research-intensive multidisciplinary universities	European frameworks and practice
Project-embedded	proximity to data and team	loss of knowledge at project end	large or high-risk projects	case evidence

The federated model is the most frequently reported outcome in the European case literature, and the reason given is consistent: a purely central arrangement drifts away from disciplinary practice, whereas a purely local one produces uneven quality and duplicated effort. That convergence, however, is weaker evidence than it appears. It rests on case studies and institutional reports, and comparative studies measuring the long-run effect of different organizational arrangements on research data outcomes remain scarce. The convergence is therefore read here as a plausible institutional response to a coordination problem, not as a demonstrated optimum.

[EVIDENCE REQUIRED: a multi-institution comparison holding institutional size and disciplinary composition constant, measuring outcome indicators (time to resolution, metadata completeness at deposit, proportion of projects with an executed preservation plan) across centralized, distributed and federated arrangements, to test whether organizational model predicts outcome once resources are controlled for]

A further consideration constrains any general recommendation and is treated in this book as a finding rather than a caveat. The models differ not in quality but in the conditions under which their limitations become tolerable. A small institution cannot sustain a fully specialized team and

nevertheless requires identified internal and external competences; a research-intensive university can sustain specialisation and pays for it in coordination. Consequently the framework in Part III does not select a model but requires that whichever model is in place be accompanied by compensating measures for its known limitation and by a named owner of that compensation.

2.4 Accountability across roles

The persistent weakness in the material analysed here is not the absence of roles but the implicitness of accountability. Responsibility for security, metadata, quality, contracts and preservation is frequently assumed to be shared, and shared responsibility that is not decomposed into assignments is functionally unowned. The instrument used to decompose it, a responsibility assignment across responsible, accountable, consulted and informed positions, is treated in this book as an evidential device rather than a management formality, because it converts an organizational chart into a set of testable propositions about who decided what.

The decomposition has a boundary that recurs in every domain of the corpus. The data steward structures the problem, compares options, coordinates the inputs of specialists, documents the decision and verifies that follow-up actions exist. Legal conclusions, ethics approval, security exceptions, acceptance of high residual risk, contract signature and allocation of significant resources belong to the principal investigator, the legal office, the data protection officer, the ethics committee, information technology security or institutional management. Conflation of the advisory and the deciding role is the single most consequential failure mode identified in the analysed material, because it produces decisions that are neither authorized nor traceable.

Two further properties of accountability are decision-relevant. The first is continuity: accountability attached to a person rather than to a role does not survive staff turnover or project closure, and the question of who assumes custody of data, access decisions, keys and documentation after a project ends is answered in the corpus far less often than it is posed. The second is shared function: as institutions rely on federated infrastructures, outsourced storage and external repositories, the demonstrability of

accountability for functions performed elsewhere becomes a condition of trustworthiness rather than a contractual detail (CoreTrustSeal 2022).

[EVIDENCE REQUIRED: retrospective tracing of a sample of completed projects to determine whether a named role held custody of data and access decisions 12 months after project end, supporting the claim that continuity of accountability, not its initial assignment, is the binding constraint]

2.5 Competence, recognition and the supply of capability

Competence frameworks have moved from generic descriptions towards differentiated profiles. The European recommendations on data stewardship skills state that the role is not uniform and that competences vary with institutional, disciplinary, technical and project context (EOSC Association 2024). A baseline curriculum for entering data stewards has since been published (Skills4EOSC 2025), and national curricula organized around competence areas and career expectations have followed, of which the Dutch roadmap towards national implementation of FAIR data stewardship is the most developed instance (Jetten et al. 2021). Advanced practice, by contrast, is characterized in the corpus by risk management, contractual workflow, audit, controlled access environments and domain profiles, none of which is covered by entry-level curricula.

Competence supply is also a function of recognition. Research assessment reform explicitly extends recognition to data curation, research software engineering and technical contributions alongside publications (CoARA 2022; DORA 2013). The consequence for institutional capability is structural rather than symbolic: without career paths, competence profiles and budget lines, the capability depends on the discretionary effort of persons whose contribution is invisible in assessment. Evidence on the effect of the reform on data and software roles specifically remains limited, and the claim advanced here is therefore that recognition is a necessary condition for stable capability, not that it has been shown sufficient.

The competence literature converges on a further point that qualifies the technical reading of the role. Reviews of data librarianship competences identify communication, relationship building and the ability to work alongside researchers as decisive alongside technical knowledge (Federer 2018). Institutional change in data practice proceeds through consultation, facilitation and advocacy, which are precisely the competences least well

captured by curriculum inventories. This asymmetry between what is taught and what is reported as effective is one of the open questions carried forward into Part V.

2.6 Assessing a capability

The measurement problem identified in Chapter 1 takes a specific form here. Counting plans produced, consultations delivered or training sessions attended measures the activity of the support function and is insensitive to its effect. The operational indicators recurring in the analysed material are of a different kind: time of first contact relative to project start, incident frequency, metadata quality at deposit, recoverability of data under test, time to resolution of an access request and the proportion of projects with an executed preservation plan. Such indicators are harder to collect and harder to attribute, which is why the corpus records them as an attribute of mature services rather than of new ones.

The distinction matters for the framework because it fixes what an assessment criterion has to be. A criterion phrased as the existence of a document is satisfied by producing the document; a criterion phrased as the demonstrable execution of a measure is satisfied only by execution and leaves a trace that a third party can inspect. Part III accordingly formulates assessment criteria at the level of realized measures, with the documentary artefact treated as evidence rather than as the object assessed.

2.7 Interpretation and boundary conditions

Two alternative explanations of the observed variation deserve statement. The first attributes the diversity of organizational models to institutional history and local politics rather than to any rational adaptation, in which case model choice would carry no information about performance. The second attributes it to resource level alone, in which case model appears as a proxy for funding. Neither can be excluded on the available evidence, and both are compatible with the position taken here, since the framework treats the model as a variable and requires compensation for its known limitation rather than inferring performance from form.

The boundary condition of this chapter is the evidential asymmetry between description and effect. Organizational arrangements are richly

described in the case literature and thinly measured in their consequences, which limits any claim about optimality to a claim about conditions of adequacy. What follows from the analysis is that institutional capability enters the framework through four decision-relevant properties: assigned decision rights, escalation paths, continuity arrangements and outcome-level measurement.

The next chapter turns to the second subordinate question and examines whether the normative layer that institutions invoke when exercising these decision rights determines the decisions at all.

Chapter 3. FAIR, CARE and TRUST: A Critical Reading of the Normative Layer

3.1 What the normative layer is asked to do

Three principle sets dominate the normative vocabulary of research data: FAIR for the properties of digital objects (Wilkinson et al. 2016), CARE for the collective rights of the communities from whom data originate (Carroll et al. 2020) and TRUST for the qualities of the repositories that hold them (Lin et al. 2020). Institutional policies routinely invoke all three, and the invocation is usually read as settling the question of how data are to be handled. This chapter argues that the reading is mistaken. The three sets address different objects, answer different questions and, taken together, under-determine the decisions that institutions and projects actually face.

The relation to the central research question is precise. If the normative layer determined institutional decisions, an integrated framework would be redundant and the task would reduce to compliance. The analysis presented here indicates that the layer determines the target properties of outcomes while leaving unspecified the conditions under which each property is achievable, the priority between properties in conflict and the authority entitled to settle the conflict. That residue is the object of Part III.

3.2 FAIR: machine-actionability misread as openness

The FAIR principles specify that digital objects be findable, accessible, interoperable and reusable, and their stated purpose concerns the capacity of machines to locate and act on data with minimal human mediation (Wilkinson et al. 2016). Accessibility in this sense is a property of the protocol and of the metadata, not of the licence: the principles require that the conditions of access be retrievable by a standardized procedure and that metadata remain available even where the data themselves are not. Authenticated and controlled access is fully compatible with FAIR. The frequent equation of FAIR with open data is therefore a category error, and one with consequences, because it converts a specification of machine-actionability into a policy of disclosure.

The consequences are visible in three places. Data that carry legitimate restrictions are treated as failures against a principle they do not offend, which discourages description of exactly the material whose discoverability matters most. Conversely, deposit under an open licence is treated as fulfilment of FAIR even where documentation is insufficient for any reuse, which is a failure against the principles under the appearance of compliance. Both errors follow from substituting a licence state for a set of properties.

Assessment practice reproduces the same substitution in a second form. Automated instruments verify what is machine-testable — presence of a persistent identifier, retrievability of structured metadata, declaration of a licence, registration in a searchable index — and are effective within that scope. They cannot assess the factual accuracy of documentation, the appropriateness of an ontology, the sufficiency of provenance for a stated reuse or the legitimacy of disclosure. A score derived from the testable subset then functions as a verdict on the whole, and the property is progressively redefined as whatever the instrument measures (RDA FAIR Data Maturity Model Working Group 2020).

[EVIDENCE REQUIRED: agreement analysis between automated FAIR assessment scores and independent expert audit of fitness for use across a stratified sample of datasets, reporting the proportion of high-scoring datasets judged unusable for their declared purpose, to support the claim that automated scores do not measure reusability]

3.3 CARE: the questions FAIR does not ask

The CARE principles were formulated within Indigenous data governance and address collective benefit, authority to control, responsibility and ethics (Carroll et al. 2020). Their object is not the digital record but the relation between the data, the people they describe and the purposes for which they are used. Where FAIR asks whether an object can be found and reused, CARE asks whether reuse is legitimate, who holds the authority to decide and how the benefit of reuse is distributed. The two are not competing accounts of the same property; they are answers to questions that do not intersect.

The complementarity is genuine but asymmetric in a way that is often overlooked. Compliance with FAIR can be assessed on the object alone, whereas compliance with CARE cannot be assessed without the participation of the community holding the authority. Consequently a CARE assessment that consists of labelling a dataset, adding an ethics statement or citing

the principles in a policy is not an implementation of them. Community-based notices such as Traditional Knowledge and Biocultural Labels express community authority and expectations rather than legal permissions, and applying them without community decision-making inverts their purpose (Local Contexts n.d.; Liggins, Hudson and Anderson 2021).

A second qualification restricts how far CARE can be generalized. Its provisions derive from a specific history of extraction and dispossession, and the transfer of its vocabulary to any dataset concerning any group detaches it from the political context in which authority to control has a determinate holder. The position taken here is that the analytical questions CARE raises — legitimacy of purpose, identifiable authority, distribution of benefit — apply widely, while its normative force applies where a community with recognized authority exists. That distinction is treated in Part III as a variable state rather than as a matter of interpretation.

3.4 TRUST: shifting the object from data to custodian

The TRUST principles address transparency, responsibility, user focus, sustainability and technology, and their object is the repository (Lin et al. 2020). The shift of object matters because the qualities they name cannot be predicated of a dataset at all: a dataset has no funding model, no succession plan and no staff. Certification schemes operationalize these qualities as demonstrable organizational capacity, including governance, finance, personnel, procedures, security and a plan for the orderly termination of the service (CoreTrustSeal 2022). Data that satisfy FAIR in a service without such capacity remain exposed, and the exposure is invisible at the level of the object.

Two limits of this layer are documented in the analysed material. Certification attests that a repository meets a set of requirements; it does not attest that the repository is appropriate for a particular dataset, whose disciplinary fit, format support, access-control capability and community reach are separate questions. And as services become federated and functions are shared across providers, demonstrating accountability for a function performed elsewhere becomes harder precisely where the trust claim rests on it; the certification requirements themselves oblige applicants to declare outsourced and partner arrangements (CoreTrustSeal Standards and Certification Board 2025).

Table 3-1 states the object, the addressee and the question answered by each of the three principle sets, together with the decision each leaves unsettled.

Table 3-1. Object, question and residue of the three principle sets constituting the normative layer.

Prin- ciple set	Object	Question answered	Decision left unsettled
FAIR	dataset and its metadata	can the object be found, accessed, combined and reused by machines and people	whether reuse is legitimate; who bears the cost of achieving the properties
CARE	relation between data, people and purpose	is the use legitimate and equitably distributed	how legitimacy is expressed in a technical access regime
TRUST	repository as an organization	can the custodian sustain its commitments	whether this custodian fits this dataset

3.5 Where the three conflict

The layer is usually presented as a stack of complements, and at the level of stated values it is one. In operation the sets generate conflicts that none of them contains the means to resolve. Four recur across the domains analysed in this book, and their structure is instructive: in each case, two correctly applied requirements point to incompatible actions.

The first is the conflict between an open licence and a community protocol. Maximising reuse and respecting authority to control cannot both be satisfied by the same licence, and the resolution recorded in the corpus is a change of access regime rather than an advisory notice attached to open data. The second is the conflict between reuse maximisation and data minimisation: reusability favours retention of rich records, while data protection law constrains retention to the stated purpose and confers erasure rights (European Parliament and Council 2016). The third arises where disciplinary appropriateness and certified trustworthiness select different repositories, one holding the community and the semantics, the other the demonstrable capacity to persist. The fourth arises between interoperability through general schemas and semantic precision through domain vocabularies, where each gain is bought with a loss in the other.

None of these conflicts is a misapplication of the principles. Each arises from correct application of two of them to the same case, which is the defining signature of an under-determined normative layer. Resolution requires an external rule that ranks the competing requirements for a stated class of case, names the authority entitled to decide, and records what risk remains once the decision is taken. Such rules exist at present as local practice and are seldom stated in a form that can be inspected or compared.

[EVIDENCE REQUIRED: a documented set of conflict cases from institutional practice, coded by conflict type, resolution adopted, deciding authority and residual risk recorded, to establish the distribution of conflict types and to test whether resolutions are consistent within an institution]

3.6 Under-determination as the gap

The result of this reading can be stated compactly. The normative layer fixes the properties that outcomes are expected to have; it does not fix the mapping from the conditions of a case to the action that produces those properties. It provides no ordering between its own components when they conflict, no allocation of decision authority, no account of the cost at which each property is obtained and no criterion distinguishing a realized measure from a declared one. Three separate assessments are consequently required — of the object under FAIR, of the legitimacy and distribution of benefit under CARE, and of the custodian under TRUST — and their composition into a single decision is left to the institution.

This under-determination is not a defect of the principle sets, which were not designed as decision procedures, and the argument of this book is not that they be replaced. It is that the layer above them, in which conditions are mapped to decisions, has no equivalent codification, and that its absence explains why institutions with identical stated commitments diverge in practice. The comparison across the 23 domains reported here indicates that the divergence concentrates in exactly the four conflict types set out above, which suggests that a decision layer of modest size would cover most of the residue.

The three chapters of Part I have thus established the problem, the capability that would carry a solution and the point at which the existing normative apparatus stops. Part II sets out the corpus, the grading of evidence and the procedure by which decision variables were extracted, so

that the framework presented in Part III can be assessed against the material from which it was derived.

PART II

Research Design and Evidence Base

Chapter 4. Research Design, Corpus and Analytical Method

The preceding part identifies a problem that no single normative source resolves. The frameworks that govern research data are authoritative within their own domains and largely silent about one another. That description covers the FAIR, CARE and TRUST principles, data protection law, preservation standards and national research assessment methodology alike (Wilkinson et al. 2016; Carroll et al. 2020; Lin et al. 2020; European Parliament and Council 2016). This chapter settles how such a problem is investigated rather than merely described. The claim it defends is that an integrated framework can be derived from documentary sources by a coding procedure explicit enough to be contested, and that the derivation is distinct from any demonstration that the result holds.

That claim answers the main research question of the book only in part, and deliberately so. The main question asks by what dimensions and variables heterogeneous normative requirements can be integrated into one framework that supports both institutional assessment and project-level decisions. A framework proposed without a stated derivation is an opinion, however plausible; a framework validated without a stated derivation cannot be revised when validation fails, because the origin of each variable is unknown. The design set out here therefore fixes the derivation first, and defers every validating claim to Part IV.

4.1 Research questions and their operationalisation

The main question decomposes into four subordinate questions, each of which is answerable from an identifiable class of source material. The first asks which requirements recur across normative domains under different vocabularies, and is answered by cross-domain term mapping. The second asks where the domains conflict rather than merely differ, and is answered by locating requirement pairs that cannot be satisfied simultaneously without a decision. The third asks which properties of a data activity determine whether a requirement can be verified at all, and the fourth asks what institutional conditions make the framework operable.

Operationalisation converts each question into an observable feature of the corpus. Recurrence is operationalised as the appearance of a functionally equivalent requirement in at least two source strata, one normative and one applied. Conflict is operationalised as an explicit trade-off statement in a source, or as a documented decision point at which one requirement is relaxed in favour of another. Verifiability is operationalised as the presence or absence of a named record — a log, a version, an approval, a metadata field — by which compliance could be demonstrated after the fact.

The fourth question, concerning institutional conditions, resists purely documentary operationalisation. Documents state what an institution intends; they do not establish what it does. This asymmetry is treated throughout as a boundary of the present design rather than as a defect to be argued away, and the empirical layer required to close it is marked where it arises.

[EVIDENCE REQUIRED: institution-level data on the actual execution of stated data policies — service response times, escalation frequency, proportion of projects with an implemented preservation arrangement — to establish whether documentary commitment predicts operational practice.]

4.2 The assignable data activity as the unit of analysis

The unit of analysis is neither the dataset nor the institution but the **data activity**: a bounded operation on research data to which responsibility, an evidentiary record and a residual risk can be attached. Collection under a consent instrument, transformation by a named script, deposit into a repository, approval of controlled access and migration of a preserved object are all data activities in this sense. The choice is consequential, because it makes the framework's dimensions properties of operations rather than attributes of objects. A dataset can be described as FAIR; only an activity can be shown to have been carried out by someone, at a time, under a rule.

This unit also explains why the analysis converges on records rather than on intentions. An activity that leaves no trace cannot be assessed after completion, whatever its quality at the time. The coding therefore treats the evidentiary trace as a coded attribute of every activity, not as a separate topic belonging to audit. The consequence for the framework is developed in Chapter 6, where provenance is shown to be the condition under which the remaining dimensions become assessable at all.

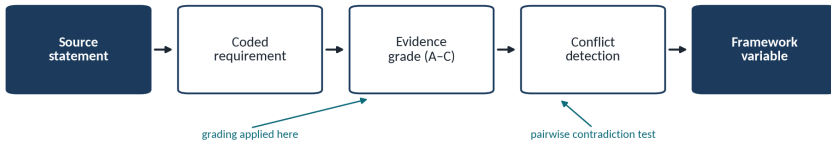


Fig. 4-1. The analytical chain from source statement to coded requirement to framework variable, showing at which step evidence grading and conflict detection are applied. Author's own elaboration.

4.3 The corpus, its strata and its boundaries

The corpus assembled for this study is stratified by the function a source performs rather than by its subject. Four strata are distinguished: binding or normative sources, official methodologies and infrastructure specifications, peer-reviewed synthesis, and applied institutional material. The stratification matters because sources of different function warrant different inferences — a regulation establishes an obligation, a maturity model establishes a comparison, a case study establishes feasibility. Table 4-1 sets out the strata, the inclusion rule applied to each and the inference each licenses.

Table 4-1. Strata of the analysed corpus, with the inclusion rule and the type of inference each stratum licenses.

Stratum	Inclusion rule	Licensed inference
Law and standards in force	Applicable to research data in the European and Czech setting; in force at the verification date	Obligation; non-negotiable constraint
Official methodologies, specifications and infrastructure documentation	Issued by the body responsible for the service or specification; version identified	Reference practice; interoperability requirement
Peer-reviewed reviews, syntheses and empirical studies	Systematic or scoping review, controlled or quasi-experimental design, or established methodology	Generalizable regularity, with stated design limits
Applied institutional and national material	Documented service, policy or guide of a named research organization or national infrastructure	Feasibility and implementation variant only

Exclusion is applied on three grounds. Sources whose status is draft or open for public consultation are retained but never treated as settled, and are labelled by status wherever cited. Sources reporting service counts, user numbers or holdings without a verification date are excluded from any quantitative statement. Promotional or vendor material lacking a technical specification is excluded entirely. Superseded versions are retained only where they explain a stable underlying principle, in which case the version and its limitation are stated. The verification date for time-variable sources across the whole corpus is 27 July 2026.

The applied stratum has two components that require separate treatment. The Czech national component comprises the guidance, metadata recommendations, repository documentation, copyright decision material and core metadata model issued by the national technical library and associated infrastructure. It is analysed as the implementation layer for the Czech setting rather than as an international norm. The international component comprises the documented service portfolios and curricula of research organizations and infrastructures. Both are read as evidence of what an institution has published about its practice, which is weaker evidence than evidence of the practice itself.

[EVIDENCE REQUIRED: a documented search protocol for the applied stratum — databases and registries queried, query strings, date of execution, number of records]

screened and excluded at each step — sufficient for an independent researcher to reconstruct the corpus. The present corpus was assembled by purposive selection against the criteria in Table 4-1, and its completeness is therefore not established.]

4.4 The coding scheme and the derivation of the dimensions

Coding proceeds in three passes over the same material. The first pass extracts requirement statements, defined as any statement that specifies what has to hold for a data activity to be acceptable; each extracted statement retains its source, its stratum and its status label. The second pass assigns each requirement to the activity it constrains and to a mode. The modes distinguish who decides, what is recorded, how the object is described, under what conditions it is accessible and how long it survives. The third pass groups requirements whose mode and target coincide across strata, and it is these groupings, not any prior conceptual scheme, that constitute the dimensions of the framework.

Six dimensions emerged from the third pass and organize Part III. Table 4-2 names them, states the coding mode from which each was derived and identifies the chapter in which each is developed. The ordering is analytical rather than sequential: governance is treated first because it determines who may resolve the conflicts that the later dimensions expose, and lifecycle and provenance second because they determine whether any later claim can be verified.

Table 4-2. The six dimensions of the integrated framework, with the coding mode from which each was derived.

Dimension	Coding mode	Developed in
Governance, roles and accountability	Who decides, who is consulted, who is informed	Chapter 5
Lifecycle and provenance	What is recorded, and about which agents, activities and entities	Chapter 6
Description, metadata and interoperability	How the object is described and by what vocabulary	Part III
Legal, ethical and access conditions	Under what conditions the object is accessible	Part III
Quality, reproducibility and computational reliability	By what evidence a result can be re-obtained	Part III
Preservation, repository trust and sustainability	How long the object survives, and under whose warranty	Part III

Two features of this procedure limit what it can support. Coding was performed by a single analyst, so no measure of coder agreement exists, and the boundary between the second and third dimension in particular rests on an interpretive judgement that another analyst might draw elsewhere. The dimensions are therefore presented as an argued reading of the corpus rather than as a categorization with established reliability.

[EVIDENCE REQUIRED: independent recoding of a random sample of at least 100 extracted requirement statements by a second analyst, with a reported agreement coefficient, to establish whether the dimension boundaries in Table 4-2 are reproducible rather than idiosyncratic.]

4.5 Synthesis, evidence grading and the treatment of conflict

Synthesis within each dimension follows a fixed four-slot structure applied uniformly across all domains analysed: stable consensus, change over the period 2024–2026, open questions and limits of evidence, and conditions of implementation. The uniform structure is what makes the domains comparable, since it forces each to declare not only what is settled but where its

evidence runs out. The comparison across the twenty-three domains reported here indicates that authoritative material is distributed very unevenly. Preservation, data protection and metadata description are supported by standards in force, whereas organizational arrangement and the effectiveness of institutional support rest largely on case reports.

Every claim carried into the framework retains an evidence grade. Grade A denotes a legal provision, a standard in force, an official final recommendation or a broadly accepted consensus framework. Grade B denotes a peer-reviewed study, a systematic review, multiple-case evidence or an established professional methodology. Grade C denotes single-case experience, an emerging community practice or a recommendation for which comparative evidence is absent. The grade attaches to the evidential warrant, not to the quality of any particular implementation, and it is reported wherever a framework variable is stated.

Conflicts between domains are the analytically productive cases and are not resolved by ranking. Open availability stands against disclosure risk, exhaustive provenance against the sensitivity of logs, complete preservation against cost. Where two requirements cannot be satisfied at once, the synthesis records the conditions under which each option is proportionate and the residual risk each leaves unaddressed. This is the reason the framework produces decision structures rather than recommendations: a recommendation conceals the conflict, whereas a conditional structure exposes it and names the authority competent to settle it.

4.6 Internal consistency and the separation of construction from validation

Internal consistency is treated as a testable property of the framework and not as an aesthetic virtue. Three tests are applied. The assignability test requires that every variable be attachable to a named role for a specific data activity. The evidence test requires that every variable name the record by which its state could be demonstrated after the fact. The closure test requires that every conflict identified in Section 4.5 be routed to a dimension containing an authority competent to decide it. A variable failing any test is either reformulated or reported as an open element of the framework.

Validation, designed in Part IV, is a separate operation on a separate body of material, and no part of Part III is presented as validated by the sources

from which it was built. The design has two components. Expert construct review separates technical correctness from operational usability, and distributes the review across an RDM specialist, a repository or technical specialist and a legal or ethics reviewer, with acceptance conditional on the absence of uncorrected substantive error. Application testing then applies the framework to authentic project material in at least two disciplines and assesses whether the resulting decisions are documented, defensible and traceable to framework variables.

Two properties of that design deserve emphasis in advance. Participant satisfaction is not treated as evidence of validity, since a framework can be experienced as helpful while producing decisions that fail an evidential test. And the assessment of applied outputs is designed to be conducted against a stated rubric by an assessor blind to the origin of the output, in order to separate the framework's contribution from the skill of the person applying it.

[EVIDENCE REQUIRED: execution of the expert construct review — number and disciplinary composition of reviewers, the claims reviewed, and the disposition of each critical comment — to support any claim that the framework is technically correct.]

[EVIDENCE REQUIRED: execution of the application test with 12–20 participants from at least two disciplines, including baseline assessment, blind rubric scoring of the produced artefact and a follow-up check of transfer into the participant's own project after six to eight weeks, to support any claim that the framework changes practice rather than knowledge.]

4.7 Ethical and legal conditions, and the limits of the design

The construction phase analyses published documents and infrastructure specifications, and involves no personal data and no human participants; no ethical approval is therefore engaged for Part III. The validation phase does involve participants and their project material, and rests on three conditions. Participation is voluntary and separable from any assessment of employment or study. Project material is analysed in a form from which neither project nor participant can be identified in reporting, and any personal data processed remains subject to the applicable data protection regime (European Parliament and Council 2016). Where participants bring material

governed by a third-party agreement, the design excludes that material rather than seeking a derogation.

Three limits qualify everything that follows. The corpus is weighted towards the European and Czech regulatory setting, so the legal and infrastructural dimensions transfer less readily than the governance and provenance dimensions. Time-variable sources were verified at a single date, and specifications identified here as current will be superseded; the framework variables are formulated to survive such changes, but the illustrations are not. And documentary evidence establishes what is required and what is claimed, never what is done, which is precisely the gap that the markers in this chapter identify and that Part IV is designed to begin closing.

With the design fixed, Part III turns to the framework itself, beginning with the dimension that determines who is competent to decide when the remaining dimensions come into conflict.

PART III

The Integrated Framework

Chapter 5. Governance, Roles and Accountability

Governance is the first dimension of the integrated framework because it settles a question the other dimensions presuppose. Every conflict that the later dimensions expose — between openness and disclosure risk, between the cost of preservation and its evidential value, between contractual restriction and reuse — terminates in a decision taken by someone with the standing to take it. The claim this chapter defends is that data stewardship is a governed function rather than a support service, and that the properties which make it governable are assignable to each data activity individually rather than to the institution as a whole. Where those properties are not assigned, the appointment of specialist staff redistributes work without redistributing authority.

This bears on the main research question in a specific way. If the framework is to support assessment as well as decision-making, it needs variables whose state can be established without asking the institution what it intends. Organizational charts, policies and job titles are weak evidence, since they record intention rather than practice. The governance variables derived here are therefore formulated as properties of activities and records: who decided, on what basis, within what limit and with what trace.

5.1 The decision layer and the advisory layer

The most consequential distinction in institutional data stewardship separates the layer that advises from the layer that decides. Advisory work identifies the problem, structures the options, gathers the documentation, records the decision and checks that dependent obligations follow. Decision authority binds the institution: it approves legal conclusions, grants security exceptions, accepts high residual risk, signs contracts, issues ethical approval and allocates resources. The analysis of the corpus indicates that these two layers are routinely conflated in practice, and that the conflation is the mechanism by which a data steward comes to carry risk that the role has no mandate to carry.

The conflation is not a matter of misconduct but of structural silence. A researcher facing a question about personal data receives an answer from

whoever is available, and an orienting assessment then functions as an approval because nothing in the process marks it as provisional. The framework accordingly treats the advisory-decision boundary as an explicit variable of each data activity, expressed as the threshold beyond which the matter passes to the data protection officer, legal office, security function, ethics committee or curator. A technical possibility is not thereby converted into a legal, ethical or scientific entitlement, and the boundary is what keeps the two apart.

Three consequences follow for the design of roles. The value of the data steward lies in problem identification, option structuring, coordination, documentation and verification of follow-through, none of which requires decision authority. Decision authority that is exercised without documentation is indistinguishable, after the fact, from no decision at all. And a role description that states only competences, without stating limits, leaves the boundary to be negotiated case by case under time pressure.

5.2 Accountability as a property of the data activity

The framework treats responsibility, accountability, consultation and information as four assignable properties of each data activity rather than as a description of an organizational unit. Applied to the master copy of a long-term sensor measurement, this yields separate assignments for calibration, backup, metadata, publication and eventual transfer, each with a named responsible party and a single accountable party. The distinction between responsible and accountable is what survives delegation: work moves, answerability does not. Consultation and information assignments make explicit which specialists are engaged before the fact and which functions learn of the outcome afterwards.

This treatment resolves a recurring ambiguity in the literature on data management plans. Where the plan is treated as the governance instrument, activities not covered by the plan fall outside governance entirely, and the plan's approval becomes a proxy for institutional control that it cannot sustain. Where the plan is treated as one evidentiary record among several, it documents intentions while logs, approvals, access reviews and deposit records document execution. The framework adopts the second reading, and the DMP is carried forward as an evidence variable rather than as a control variable.

Assignment also has to be legible at the moment of the activity, not reconstructible afterwards from correspondence. An assignment recorded only in a project proposal is unavailable to the person who performs the activity three years later. For this reason the framework's assignability test, stated in Section 4.6, requires that each variable be attachable to a named role for a specific activity in a record that the performer of the activity can reach.

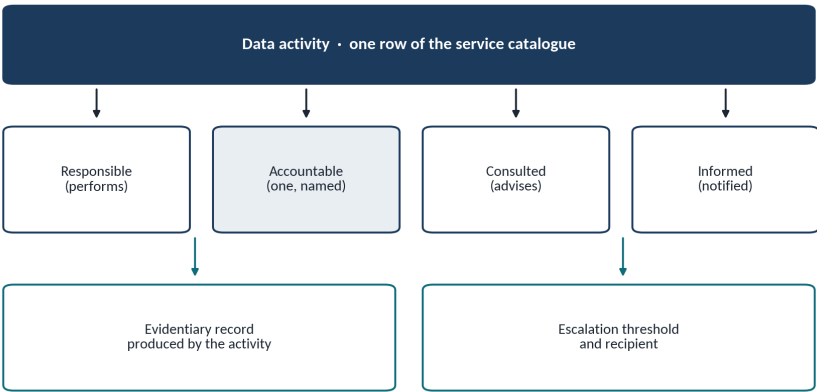


Fig. 5-1. The assignment matrix for a single data activity, showing responsible, accountable, consulted and informed roles alongside the evidentiary record produced and the escalation threshold. Author's own elaboration.

5.3 Escalation as the structural test of the model

Institutional arrangements for data stewardship are conventionally compared as organizational models, and the corpus supports the familiar typology of centralized, distributed and federated arrangements, alongside project-embedded stewardship. The comparison across the corpus indicates a convergence on federated arrangements in research-intensive universities, where central methodology and infrastructure are combined with discipline-embedded stewards (Science Europe 2021; EOSC Association 2024). That convergence rests on institutional case reports rather than on comparative outcome studies, and the framework accordingly grades it at level B. Table 5-1 states the conditions under which each arrangement is proportionate and the residual risk each leaves.

Table 5-1. Organizational arrangements for institutional data stewardship, with proportionate conditions and residual risk.

Arrange- ment	Proportionate condition	Residual risk	Evidence grade
Centralized	Smaller or methodologically homogeneous institution; early service maturity	Distance from disciplinary practice; capacity bottleneck at the centre	B
Distributed	Strong faculties with established local support	Inconsistent rules and duplicated effort	B
Federated	Multidisciplinary research-intensive institution	Coordination cost; unclear funding of the shared layer	B
Project- embedded	High-risk or data-intensive single project	Knowledge loss at project end	C

The typology is nevertheless the weaker part of the analysis, because the models differ less in what they determine than the literature suggests. What distinguishes a functioning arrangement from a nominal one is the presence of a traversable escalation path: a defined threshold, a named recipient, an expected response interval and a record of the outcome. An institution with a federated chart and no escalation path performs worse than a centralized institution in which every unresolved question reaches a competent authority within a known interval. The framework therefore carries escalation as a variable in its own right and treats organizational form as a context condition.

[EVIDENCE REQUIRED: comparative outcome data across institutions operating different organizational arrangements — time from request to resolution, proportion of requests escalated, proportion of escalations closed with a recorded decision — to test whether arrangement type predicts governance performance once escalation quality is controlled.]

5.4 The service catalogue as the operational form of governance

Governance becomes observable at the point where it takes the form of described services. A service catalogue specifies, for each service, its inputs, its outputs, its owner, the expected resolution interval and the escalation point at which it hands over. The analysis reported here indicates that the

catalogue is the operational expression of the governance model rather than an administrative accessory: consultation on a plan, security classification, metadata design, anonymization support and curatorial checking of a deposit are distinct services with distinct owners, and describing them separately is what makes ownership assignable. Where a catalogue is absent, the institution offers a contact point instead of a service, and the boundary of what is provided is negotiated in each case.

The catalogue also disciplines the relation between tools and governance. Systematic mapping of the tool landscape shows substantial fragmentation — a review confined to the life sciences analysed more than 300 tools (Aksenova et al. 2024) — and the introduction of a platform does not by itself allocate authority. A tool deployed over an undefined process reproduces the ambiguity of that process in a machine-readable form. The order of operations that the evidence supports is therefore the reverse of the common one: service definition, process ownership, responsibility boundary and escalation route precede tool selection.

A catalogue carries one characteristic failure mode worth stating. Services described without an expected resolution interval permit indefinite delay while remaining formally compliant, and delay is the most common way in which data governance fails without any rule being broken.

5.5 Continuity after the project and after the person

Accountability that is attached to a project or to an individual expires with it. The corpus treats project end, staff departure and service discontinuation as three instances of a single problem: the transfer of accountability for data, access decisions, software, keys and documentation to a named successor. Continuity is a governance variable rather than an archival one, because the question it answers is who is answerable next rather than where the bits reside. An institution that cannot name the successor for a completed project's controlled-access decisions has an unassigned obligation, whatever the state of its repository.

Two structural conditions determine whether continuity holds. The first is budgetary: curation and preservation that are not funded within the project budget become unfunded obligations of the unit that inherits them, and unfunded obligations are discharged selectively. The second is documentary: a successor who cannot reconstruct the basis of earlier access

decisions either refuses all subsequent requests or grants them without a basis, and both outcomes are failures of governance rather than of judgement. The framework carries both conditions as variables, since each is assessable from records that an institution already holds.

[EVIDENCE REQUIRED: longitudinal data on datasets whose responsible person has left the institution — proportion with a documented successor, proportion for which subsequent access requests were answered, and time to first unanswerable request — to establish the practical decay rate of unassigned accountability.]

5.6 Measuring institutional support without counting plans

The measurement of institutional support is the point at which governance frameworks most often become self-confirming. Counting data management plans, training sessions or consultations measures the activity of the support function and not the condition of the data, and a plan that is written and never implemented raises the count while lowering the reliability of everything downstream. The analysis of the corpus indicates that outcome-oriented indicators are both available and rarely used: the timeliness with which support is engaged relative to the data activity, the recovery of data under an actual restoration test, the completeness of metadata against a stated profile, the resolution interval for access requests and the proportion of projects with an implemented preservation arrangement. These indicators share a property that counting does not: each names a record that could contradict the institution's own account of itself.

Outcome indicators nevertheless carry a known distortion. Indicators that determine resources reshape the behaviour they measure, and a target expressed as a proportion of projects with an approved plan produces approved plans rather than managed data. The framework mitigates this by pairing every indicator with the evidentiary record from which it is derived, so that the indicator cannot be satisfied without the record existing. This pairing is also what connects governance to audit: an audit that verifies the existence of documents establishes only that documents exist, whereas an audit that samples files, permissions, restorations and metadata establishes what was done (National Institute of Standards and Technology 2024).

[EVIDENCE REQUIRED: paired institutional data on plan-based indicators and outcome indicators for the same set of projects, to test the study's claim that the two diverge and that plan counts overstate the state of data.]

5.7 The governance variables carried forward

Table 5-2 states the variables that this dimension contributes to the integrated framework, each with the record from which its state is read. The formulation is deliberately austere: every variable is a property of an activity, and every variable names a trace. Variables that could not be tied to a trace were reported in Section 5.3 as context conditions rather than admitted to the framework.

Table 5-2. Governance variables of the integrated framework and their evidentiary records.

Variable	Definition	Evidentiary record
Decision boundary	The threshold at which an activity passes from advice to binding decision	Role description with stated limits; escalation entry
Accountability assignment	The single accountable party for each data activity	Assignment matrix held with the project record
Escalation traversability	Existence of a named recipient and expected interval for each threshold	Escalation log with outcome
Service definition	Inputs, outputs, owner and resolution interval of each service	Service catalogue entry
Continuity assignment	Named successor for data, access decisions, software and keys	Handover record; funded preservation arrangement
Outcome measurement	Indicators derived from records rather than from counts of documents	Restoration test, access request log, metadata validation report

Two boundary conditions limit the transfer of this dimension. Small institutions cannot sustain differentiated roles, and for them the variables are satisfied by named external competence rather than by internal units; the variables are indifferent to whether the competent party is employed locally. And in federated infrastructures the accountable party may sit outside the institution altogether, which makes contractual specification of shared responsibility the only available evidentiary record (CoreTrustSeal 2022). Neither condition weakens the variables, but both change where their evidence is found.

The dimension developed here determines who is answerable for each data activity. Whether that answerability can be discharged depends on whether the activity left a record at all, which is the subject of the following chapter.

Chapter 6. The Research Data Lifecycle and Provenance

The governance dimension settles who is answerable for a data activity. It does not settle whether the activity left anything with which that answerability can be discharged. This chapter takes up the second dimension of the integrated framework and defends two connected claims: that the research data lifecycle is a structure of layers, transitions and change regimes rather than a sequence of stages, and that provenance is the property on which the assessability of every other dimension depends. Where derivation is unrecorded, statements about quality, reproducibility and rights remain assertions that no assessment can confirm or refute.

The bearing on the main research question is direct. A framework that supports assessment as well as decision-making requires variables whose state can be read from records that the activity itself produces, rather than from what a project reports about its own practice. Lifecycle and provenance variables have exactly this character, because each of them names a trace: a designated authoritative copy, a versioning threshold, a derivation record, a retention decision, a transfer of custody. The dimension developed here therefore supplies the evidentiary substrate that the governance variables of Chapter 5 presuppose but cannot themselves generate.

6.1 The failure of the stage model

The model that divides research data work into planning, collection, processing, analysis, publication and archiving is didactically economical and descriptively false for most projects. Research is iterative: a preliminary result redirects collection, a revised cleaning rule invalidates a derived table, a reviewer's objection reopens a settled model. A representation adequate to this pattern captures repetition, branching and return, which curation lifecycle models already accommodate in principle (Higgins 2008). The defect of the linear model is not that it simplifies but that it simplifies in the wrong direction.

The direction of the error can be stated precisely. A stage model assigns each object a position in time, whereas stewardship decisions depend on its status: what may change it, what it is evidence for, how long it is required

and which copy of it binds. Two files created in the same week may occupy one stage while differing in every respect that matters to a decision. The framework accordingly replaces the stage with the pair of layer and transition, and treats time as a property recorded within the derivation rather than as the organizing axis.

6.2 Layers and the authoritative copy

Research data separate into layers that differ in evidential function and in the changes each admits. Original acquisition, working copies, harmonized layers, analytical datasets and released packages are governed by distinct regimes: the first is not rewritten under any circumstances, the middle layers are edited under a change log and the last is frozen at the point of citation. Retaining every intermediate output is expensive and produces an aggregate in which no copy is authoritative; deleting them all removes the possibility of audit. Table 6-1 sets out the layers, the function each performs and the change regime that follows from that function.

Table 6-1. Data layers, their evidential function and the change regime each admits.

Layer	Evidential function	Permitted change	Principal retention driver
Original acquisition	Establishes what was measured or observed	None; corrections recorded in a derived layer	Irreplaceability
Working copy	Supports exploration and repair	Free, under a change log	Regenerability from input and code
Harmonized layer	Establishes comparability across sources	Rule-based, with mapping recorded	Cost of re-harmonization
Analytical dataset	Supports the published claim	Frozen at the point of analysis	Reproducibility of the result
Released package	Carries citation and re-use	New release only	Persistence commitment of the repository

The decision the table forces is the designation of an authoritative copy for each layer. Where no copy is designated, divergence occurs silently: two harmonized versions circulate, both are plausible and neither is wrong in any way that inspection reveals. The analysis of the corpus indicates that

this designation is rarely made explicitly, and that its absence is more consequential than the choice of storage or format, because it removes the reference point against which any later discrepancy could be adjudicated. The framework therefore carries the authoritative copy as a variable of each layer rather than as an attribute of a project.

6.3 Version, release and working revision

Three notions are routinely conflated and are separable. A working revision is a change internal to a layer, a version is a citable state of an object and a release is a version to which a persistence commitment attaches. The distinction becomes operational at one point: the threshold at which a change generates a new persistent identifier rather than an update under the existing one. That threshold is a decision about what the identifier is taken to denote, and it cannot be delegated to a repository setting without deciding it implicitly.

Dynamic datasets make the threshold unavoidable. A continuously updated database supports no stable citation unless the cited object is specified as a time slice, a query or a frozen release, and the recommendations for citing evolving data resolve this by timestamping and query identification rather than by freezing the source (Rauber et al. 2015). What reproducibility requires in such cases is not the whole database but the retrievability of the state the analysis actually used (National Academies of Sciences, Engineering, and Medicine 2019). The framework consequently records the cut date, the filter and the query as provenance elements of the analytical dataset, and treats a citation lacking them as incomplete irrespective of whether it resolves.

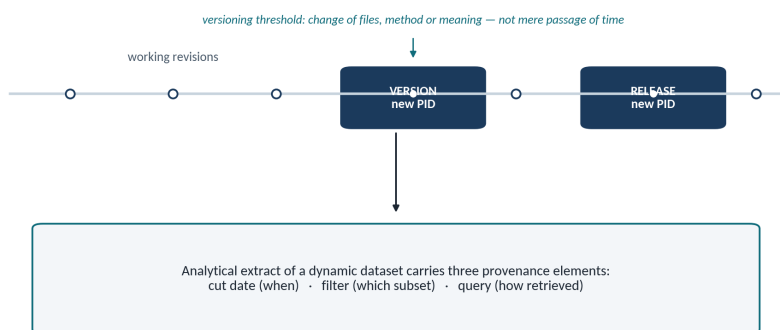


Fig. 6-1. The relation between working revision, version and release for a continuously updated dataset, showing the versioning threshold at which a new persistent identifier is minted and the three provenance elements — cut date, filter, query — attached to each analytical extract. Author's own elaboration.

6.4 Why provenance conditions the other dimensions

Provenance answers where an object came from, by what activity, from which inputs, through which agent and at what time. It is broader than a file history and narrower than a methodological description of the study, and the entity–activity–agent model represents it neutrally across manual import, scripted transformation, instrument calibration and formal approval (World Wide Web Consortium 2013). Packaging specifications extend the same structure to transportable objects carrying data, code, workflow and context together (Soiland-Reyes et al. 2022). The apparatus is thus stable and available; what the literature leaves open is how much of it a given project has reason to instantiate.

The argument this chapter advances concerns the standing of provenance among the dimensions rather than its representation. Claims about data quality, about reproducibility and about rights share a common form: each asserts something about how an object came to be as it is. A quality claim asserts that stated procedures were applied to stated inputs, a reproducibility claim asserts that a result follows from identified data under identified code and conditions, and a rights claim asserts that the entitlements attaching to the inputs permit what is now proposed. None of the three is verifiable against the object itself, since an object carries no visible mark of its own derivation.

The asymmetry that follows is the substance of the claim. Provenance can be recorded for an object whose quality is poor, whose result does not reproduce and whose licence forbids redistribution; the record remains informative and is in fact what makes those deficiencies demonstrable. The converse does not hold, because no amount of quality assurance, reproducibility engineering or licence documentation establishes anything verifiable about an object whose derivation is unknown. This study therefore orders provenance ahead of the dimensions it conditions, and treats the presence of a derivation record as a precondition for admitting any assessment claim about the other dimensions rather than as one criterion among them.

6.5 Granularity proportionate to risk

Recording every event in a computational pipeline is feasible and, for most research, disproportionate. The practical question is the minimum evidential chain, which the corpus consistently resolves into six elements: the identifier and version of each input, the code and configuration applied, the time, the responsible role, the identifier of the output and a checksum. Table 6-2 states the granularity options, the condition under which each is proportionate and the residual risk it leaves. The grading reflects that the options are documented in practice reports rather than compared by outcome measurement.

Table 6-2. Provenance granularity options, proportionate conditions and residual risk.

Option	Proportionate condition	Residual risk	Evidence grade
Full event provenance	Regulated or high-consequence processing	Volume and complexity obscure the decisive steps	B
Checkpoint provenance	Ordinary research project with defined transformation points	Small changes between checkpoints remain invisible	B
Automated workflow log	Computational pipeline on a stable platform	Dependence on the platform; loss of record on migration	B
Narrative history file	Small project with few transformations	Unstructured and vulnerable to omission	C

Manual steps are the systematic weakness of every option in the table. Copying cells between spreadsheets, interactive filtering and image

adjustment leave no trace unless one is created deliberately, and an automated log records faithfully the portion of the work that passed through the pipeline while remaining silent about the portion that did not. Where a step resists automation, its provenance is carried by a structured protocol that states parameters rather than only the name of the software. The comparison across the corpus indicates that the gap between recorded and actual derivation is concentrated almost entirely in these manual transitions.

[EVIDENCE REQUIRED: for a sample of published analytical datasets, a reconstruction attempt from the deposited provenance record alone, measuring the proportion of transformation steps recoverable and the proportion of unrecoverable steps that were performed manually — supporting the claim that manual transitions account for most of the recorded-to-actual gap.]

Provenance also carries a legal component, frequently separated from the technical one to the detriment of both. A dataset combining sources of different origin can be redistributed only if the licence, consent scope and access conditions attaching to each input remain traceable through the transformations applied. Where that trace is absent, the derived dataset is not thereby unlawful to hold, but its redistribution becomes a decision taken without the information the decision requires. The framework accordingly treats the licence manifest as a provenance element rather than as separate legal documentation.

6.6 Appraisal and the limits of regeneration

Selective preservation rests on the proposition that an intermediate output which can be regenerated need not be retained. The proposition is sound and its condition is demanding: regeneration presupposes that the computation is deterministic, that the code and configuration are preserved and that the execution environment remains available or reconstructible. Retaining inputs and code in place of outputs therefore substitutes a dependency on environment executability for a dependency on storage volume, and the substitution is favourable only where the first dependency is actually managed. Appraisal in the framework is consequently expressed through four criteria: evidential value, replacement cost, demonstrated regenerability and any binding retention obligation.

The word demonstrated carries the weight. A claim that an output is regenerable is a hypothesis about a future execution, and periodic re-execution is the only evidence converting it into a fact about the present environment. Preservation planning in the reference model treats this as a monitoring obligation of the archive rather than as a property of the deposit (Consultative Committee for Space Data Systems 2012). Appraisal resting on untested regenerability claims is therefore recorded as a decision taken under an unquantified assumption.

[EVIDENCE REQUIRED: repeated re-execution tests on a defined set of deposited computational datasets at fixed intervals after deposit, recording the proportion of successful regenerations and the causes of failure — establishing how quickly the regenerability assumption underlying selective preservation decays.]

6.7 End of life as a stage of the lifecycle

Lifecycle models that terminate at deposit into an archive cannot express the events that most reliably destroy data holdings. Projects end, infrastructures are decommissioned, staff leave and institutions reorganize, and each of these transfers custody of data, of access decisions, of software, of encryption keys and of the documentation without which none of the preceding is interpretable. For long-running observatories and infrastructure holdings, the arrangement for succession is a condition of the sustainability claim rather than an administrative appendix to it, and repository certification treats succession planning as a criterion of trustworthiness for precisely this reason (CoreTrustSeal 2022). The framework therefore represents end of life as a stage with its own decisions, its own accountable party and its own evidentiary record.

Treating termination as a stage has one consequence worth stating separately. It converts deletion from an event that happens when nobody objects into a decision with a stated basis, an authority and a record, which is the same status the framework assigns to preservation. Deletion and retention thereby become symmetrical outcomes of one appraisal decision rather than an action and a default. The asymmetry in current practice — where retention is decided and deletion occurs by attrition — is the mechanism through which holdings are lost without any decision to lose them having been taken.

6.8 The lifecycle and provenance variables carried forward

Table 6-3 states the variables this dimension contributes, each with the record from which its state is read. As in the governance dimension, every variable names a trace, and properties that could not be tied to one were retained as context conditions.

Table 6-3. Lifecycle and provenance variables of the integrated framework and their evidentiary records.

Variable	Definition	Evidentiary record
Layer designation	The layer to which an object belongs and the change regime that follows	Data registry entry with layer and change rule
Authoritative copy	The copy that binds when copies diverge	Registry designation with location and custodian
Versioning threshold	The condition under which a change generates a new persistent identifier	Versioning policy; identifier minting record
Derivation record	Inputs, activity, agent, time, output and checksum for each transformation	Provenance graph, workflow log or structured protocol
Manual step protocol	Structured description of transformations performed outside the pipeline	Laboratory or transformation protocol with parameters
Rights trace	Licence, consent scope and access conditions carried through derivation	Licence manifest linked to the provenance record
Regenerability status	Whether a non-retained output can currently be reproduced	Dated re-execution test result
Succession arrangement	Named recipient of data, access decisions, software, keys and documentation	Transfer agreement; funded preservation arrangement

Two boundary conditions limit the transfer of this dimension. Where research is conducted in trusted research environments, the derivation record sits inside a boundary that external assessment cannot cross, and the variable is then satisfied by an attested audit performed within the environment rather than by an inspectable graph. And in disciplines whose material is not computational, the derivation of an analytical claim is carried by scholarly apparatus rather than by transformation logs, which changes the form of the record without removing the requirement that one exist.

The dimension established here determines what a data activity leaves behind. Whether the activity is shaped in advance by decisions that later records can be checked against is the question the following chapter addresses.

Chapter 7. Planning as Commitment: Data Management Plans and Machine- Actionable Planning

The preceding dimension establishes what a data activity leaves behind. This chapter takes up the point at which the activity is shaped in advance, and defends a claim about what planning contributes: a data management plan has evidential value only where it records decisions later checkable against the traces the activity produces. So understood, the plan is neither a compliance artefact nor a control instrument but a dated statement of commitments, some of which cannot be revised without losing what they were meant to secure. The chapter identifies which commitments have that irreversible character, and specifies when structuring a plan for machine processing adds assurance rather than formalizing an unclear procedure.

The relation to the main research question turns on assessability. Chapter 5 established that documentary compliance and realized practice diverge, and that document counts cannot detect the divergence. A plan nevertheless remains the only instrument that fixes intentions before the activity, which makes it the point at which the framework's variables acquire their initial state. This dimension therefore treats the plan as the origin of variable states and the records of Chapter 6 as the evidence against which those states are later confirmed or falsified.

7.1 Two readings of the plan

The plan as administrative deliverable is written once, at the moment of application, in a register calibrated to reviewers rather than to the project. Its statements are general because generality is safe: data will be stored securely, metadata will follow community standards, data will be deposited in a suitable repository. Such statements are unfalsifiable, and their unfalsifiability is the reason they are written, since a specific commitment can be shown to have failed while a general one cannot. Assessment of plans against structured criteria of specificity and feasibility responds to exactly this problem (Science Europe 2021).

The plan as a record of decisions has a different grammar. Each item names the system, the frequency, the responsible role and the record by which fulfilment can be shown, converting a promise into a proposition with a truth value. The analysis of the corpus indicates that this conversion, rather than any increase in length or coverage, separates plans that support later verification from those that do not. The framework adopts the second reading and treats the first as a documented failure mode rather than an inferior variant of one instrument.

The distinction also fixes the plan's relation to neighbouring documents. A protocol governs scientific procedure, a preregistration fixes hypotheses and analysis, an impact assessment addresses risk to persons and a security plan addresses threats to systems (European Parliament and Council 2016). A plan that absorbs all of them becomes unmaintainable, and one that ignores them repeats their content inconsistently. The framework accordingly treats the plan as the linking document identifying which decision is recorded where, rather than as the container of every decision.

7.2 Commitments that cannot be made retrospectively

Most planning decisions are revisable at moderate cost: storage can be migrated, metadata enriched, a repository changed before deposit. A small number are not, and their irreversibility is what gives planning its evidential force. Table 7-1 identifies these commitments, states why each resists retrospective repair and names the record in which the commitment is made.

Table 7-1. Planning commitments that cannot be established retrospectively, with the reason and the record of commitment.

Commitment	Why it resists retrospective repair	Record of commitment
Scope of consent	Consent is given for stated purposes before collection; later reuse beyond that scope cannot be authorized by the researcher	Information sheet and consent form, dated
Rights position of inputs	Terms accepted at acquisition bind downstream use; renegotiation after integration is rarely available	Data use agreement; licence manifest
Provenance capture	A transformation that left no trace cannot be reconstructed once the environment changes	Workflow configuration; protocol template
Storage and security class	Class determines where data may reside from the first copy onward; reclassification does not undo prior exposure	Classification decision with approval
Identifier and versioning policy	Citations already issued cannot be re-pointed without breaking the record	Versioning policy fixed before first release

The commitments in Table 7-1 share one structural feature: each governs the moment of acquisition or of first release, and each constrains a population of later decisions unknown at the time of commitment. Consent scope is the clearest case, since a narrow formulation forecloses reuse that no one anticipated, while an expansive one may fail to be specific enough to be valid. The framework does not resolve that tension, which belongs to the legal and ethical dimension, but it records where the tension has to be resolved and which authority is competent to resolve it. What planning contributes here is not foresight but the placement of a decision at the only moment when it can still be taken.

7.3 The anatomy of a verifiable plan item

A plan item that supports later assessment has four components beyond its substantive content. It names the system or service in which the activity occurs, the frequency or trigger at which it recurs, the role accountable for it and the record that would demonstrate its performance. An item lacking the fourth component may still be true, but its truth cannot be established

without asking the project, which returns assessment to self-report. The framework therefore treats the evidence component as the defining feature of a commitment rather than as a refinement of it.

This structure also exposes commitments the project cannot meet. An item specifying quarterly integrity checks on a storage service with no reporting interface fails at the point of naming its record, and it fails during planning rather than at the final review. Linking plan items to the budget and to the service catalogue described in Section 5.4 performs the same function for resources, since a commitment to a service the institution does not offer is detectable when it is written. Planning in this reading is a feasibility test carried out on paper, which is the least expensive place to conduct it.

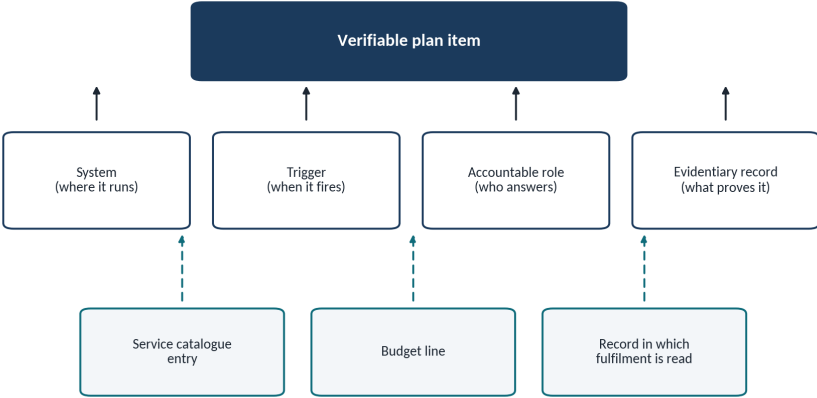


Fig. 7-1. The structure of a verifiable plan item, showing the four components — system, trigger, accountable role, evidentiary record — and their linkage to the service catalogue entry, the budget line and the record in which fulfilment is later read. Author's own elaboration.

7.4 Machine-readable and machine-actionable

The two properties are routinely equated and are distinct in a way that matters for what can be assured. A machine-readable plan is one whose content a system can parse: a structured document, a form export, a schema-valid file. A machine-actionable plan is one whose content a system can act on — validate against a rule, transmit to another service, use to trigger a

request or check against an observed state (Miksa et al. 2019). Readability is a property of the format; actionability is a property of the format together with the semantics that make its values interpretable by a system other than the one that produced them.

The distinction has an operational consequence. Structuring a plan into fields yields readability immediately and actionability only where the values are drawn from authoritative sources that other systems recognize: a person identified by a researcher identifier, a funder and grant by a registry entry, a repository by its registry record, a published dataset by its persistent identifier. Common data models for plan exchange exist for this purpose and define an interoperable minimum rather than a complete institutional record; the RDA DMP Common Standard, version 1.1, is the reference specification (Miksa, Walk and Neish 2020). The framework consequently separates the interoperable minimum from local extension, and treats only the former as carrying cross-system commitments.

A second consequence concerns what structure does not deliver. A syntactically valid but outdated value propagates unchanged to every system consuming it, and its structure lends it an appearance of verification it has not received. Validation of syntax is therefore not validation of content, and the framework records for each interoperable value its source, the time of last update and its verification state. This study treats the conflation of structure with truth as the principal governance risk introduced by automation rather than as a defect of particular tools.

7.5 The conditions under which automation adds assurance

Tooling for planning ranges from text templates through structured forms to knowledge-based systems that adapt questions to previous answers and to fully integrated arrangements that exchange data with grant systems, current research information systems, storage services and repositories. Table 7-2 states the conditions under which each is proportionate and the residual risk each leaves. The evidence grading reflects that the comparison rests on implementation reports and specification documents rather than on outcome studies of plan quality or realization.

Table 7-2. Implementation options for data management planning, proportionate conditions and residual risk.

Option	Proportionate condition	Residual risk	Evidence grade
Text template	Small standalone project; no institutional exchange requirement	Weak validation; items unlinked to records	C
Structured form	Institutional standardization across many projects	Mechanical completion; fields satisfied without decisions	B
Knowledge-based system	Complex service portfolio; heterogeneous data classes	Maintenance burden of the knowledge model	B
Fully integrated machine-actionable plan	Mature infrastructure with owned integrations	Propagation of erroneous values; governance of integrations	B

The condition separating assurance from mere digitization can be stated independently of the option chosen. Automation adds assurance where the process already has a defined owner, defined inputs, a defined output, a defined method of checking and stated conditions for update; where any of these is missing, the automated version reproduces the ambiguity at higher speed and with greater apparent authority. The analysis across the corpus indicates that this pattern recurs in every domain in which automation was examined, which suggests a property of automation in weakly specified processes rather than a feature of data planning.

One further boundary applies specifically to planning. A system can identify a risk, request the information a decision requires and route the case to the competent authority, more consistently than a human process does. It cannot determine a legal basis, judge whether anonymization is sufficient or assess scientific adequacy, because each is a judgement whose authority the governance dimension has already assigned. Automation therefore transfers information and never answerability, and an arrangement in which a system's output functions as an approval has relocated a decision without relocating accountability for it.

[EVIDENCE REQUIRED: comparison of plan items produced under text, structured-form and knowledge-based tooling for comparable projects, scored against a specificity and

feasibility rubric, to test whether tooling affects the specificity of commitments or only the completeness of fields.]

7.6 Revision as evidence rather than as failure

A plan identical at project close to the version submitted with the application evidences one of two things: a project in which nothing unforeseen occurred, or a plan that was never consulted. The second is considerably more common, and the corpus reports the same pattern across domains. Revision triggered by events — a completed pilot, a change of method, a new agreement, a change of storage or of personnel — records the project's actual trajectory, and a version history with reasons is a stronger evidentiary object than any single version. The framework accordingly treats the revision record as a variable and the absence of revision as a signal requiring explanation.

This reading resolves a tension in how funders and institutions treat deviation. Where the plan is read as a promise, deviation is non-compliance and the rational response is to write plans vague enough that no deviation is detectable. Where the plan is read as a dated record of decisions, deviation that is justified, approved and versioned is evidence of management, and the failure mode is undocumented deviation rather than deviation as such. The incentive structures created by the two readings differ sharply, and the first predicts precisely the generality observed in submitted plans.

[EVIDENCE REQUIRED: longitudinal record of plan versions across a cohort of projects, coded for revision trigger, approval and justification, to test the study's claim that event-triggered revision correlates with realized measures more strongly than plan completeness at submission does.]

7.7 Why counting plans does not measure stewardship

Indicators reporting the number of plans written, or the proportion of projects holding one, measure the reach of an obligation rather than the state of data. Their insensitivity has three separable sources. A plan can be complete without any item naming a record, so completeness is satisfied by unfalsifiable statements. A plan can be accurate at submission and superseded within months, so its content describes a state that has ceased to obtain. And a plan written by a support unit on the project's behalf

evidences the capacity of that unit rather than any decision taken by the project.

What can be measured instead follows from the variables. A plan item naming a record makes a paired observation possible: the commitment on one side, the record on the other, and a determination of whether the record exists and shows what the item claimed. That comparison is expensive relative to counting, and its cost is the reason it is rarely performed rather than a reason it is unavailable. This study treats the paired comparison as the minimal outcome-level measurement for the planning dimension, and plan counts as a coverage statistic with no outcome interpretation.

7.8 The planning variables carried forward

Table 7-3 states the variables this dimension contributes. Each is a property of the plan as an object, and each names the record against which its state is read.

Table 7-3. Planning variables of the integrated framework and their evidentiary records.

Variable	Definition	Evidentiary record
Item specificity	Whether each item names system, trigger, accountable role and record	Plan item as written
Irreversible commitment coverage	Whether the commitments of Table 7-1 are fixed before the activity they govern	Dated consent, agreement, classification and versioning decisions
Plan linkage	Identification of which decision is recorded in which neighbouring document	Cross-references to protocol, impact assessment and security plan
Interoperable minimum	Values drawn from authoritative registries, with source and update time	Structured plan export; identifier resolution record
Value verification state	Whether an exchanged value has been checked for currency, not only for syntax	Validation log with timestamp
Revision trace	Event-triggered versions with reason and approval	Plan version history
Realization evidence	Paired comparison of each commitment with its named record	Final plan version with deviations stated

Two boundary conditions qualify the dimension. Consortium and doctoral settings raise the question of whether one plan, a plan with sub-project annexes or separate linked plans is appropriate, and the corpus resolves it by how far data, roles, agreements and infrastructure differ between the parts rather than by the size of the consortium. And publication of plans engages security and contractual risk that a redacted public version addresses while the full internal version remains the object of assessment; the variables above are read from the internal version, and assessment performed on public versions alone measures something narrower than it appears to.

Planning fixes commitments about description, quality and access before the data exist. Whether those commitments yield data another researcher can interpret and reuse is the subject of the following chapter.

Chapter 8. Description, Quality and the Conditions of Reuse

The two preceding dimensions establish who is answerable for a data activity and what record it leaves. Neither settles whether a dataset can be understood by anyone who did not produce it. This chapter takes up description and quality as the dimension that determines interpretability, and defends two claims: that fitness for use is a relation between a dataset and a declared purpose rather than a property the dataset possesses, and that description generated during the activity is systematically stronger than description reconstructed afterwards. The second claim is stated with its boundary conditions, since it is an empirical generalization rather than a necessary truth.

The bearing on the main research question is that description and quality are the dimension in which the cost of stewardship is most visible and most contested. Requirements to describe are easily multiplied and are experienced by researchers as unbounded, which makes the boundary between what description covers and what it leaves out a decision the framework has to support rather than avoid. The variables derived here therefore express description as a bounded obligation defined by an intended use and an intended user, and quality as a declared residual error against a stated purpose.

8.1 Metadata as the carrier of interpretability

Metadata are not one description but several, addressed to different readers and answering different questions. Citation and discovery metadata identify an object and connect it to persons, funders, projects and related outputs; disciplinary metadata describe variables, measurements, samples, instruments and methods; technical metadata support preservation; provenance metadata explain derivation. A general schema serves the first layer well and is not intended to carry the second, and a disciplinary standard carries the second while integrating weakly into cross-domain discovery (DataCite Metadata Working Group 2024). Treating either as sufficient alone is the recurrent error the framework is designed to prevent.

The layers differ in more than content, because they differ in who bears the cost of omission. Omitted citation metadata are noticed immediately, since the object cannot be found or credited, and the omission is therefore self-correcting. Omitted disciplinary metadata are noticed only by a reuser, often years later, and by then the information required to supply them may be irrecoverable. Description obligations that are enforced uniformly across the layers therefore concentrate effort where failure is least consequential, which is one mechanism through which compliance-driven description produces formally complete and substantively empty records.

8.2 The boundary between core and disciplinary profile

An application profile resolves the layering question by selecting elements from one or more standards and fixing, for each, whether it is mandatory, in what format it is expressed, from which vocabulary its values are drawn and by what rule it is validated. Without a profile, identically named fields acquire divergent meanings across projects, and records become formally complete while remaining mutually incomparable. The profile is thus the artefact in which the boundary between the common core and the disciplinary layer is actually drawn, and drawing it is a decision with an owner rather than a technical configuration. Table 8-1 states the options for combining general and disciplinary description with the condition under which each is proportionate.

Table 8-1. Options for combining general and disciplinary metadata, with proportionate conditions and residual risk.

Option	Proportionate condition	Residual risk	Evidence grade
General schema only	Deposit in a generalist repository with documentation attached	Scientific detail insufficient for reuse	B
Disciplinary standard only	Deposit in a specialized repository serving one community	Weak connection to cross-domain discovery	B
Dual layer, rich record plus general export	Reuse expected both inside and outside the discipline	Crosswalk maintenance and loss at export	B
Free-text description	Supplementary explanation only	Not interoperable; not machine-processable	C

The dual layer carries a cost that the literature identifies but rarely quantifies. An analysis of fourteen crosswalks from disciplinary research data schemas to the general-purpose Schema.org vocabulary shows that descriptive elements map well while discipline-specific, rights and structural elements are left unmapped or mapped inconsistently (Wu et al. 2023), and the elements lost are typically those carrying the scientific meaning of the variables. Retaining the original record alongside the export is the arrangement the corpus consistently reports as preferable, on the reasoning that a lossy export is recoverable while a lossy original is not. The framework accordingly treats the authoritative description as the disciplinary record and the general layer as a derived view, which parallels the treatment of authoritative copies in Section 6.2.

8.3 Controlled vocabularies as the condition of interoperability

Description becomes comparable across the boundary of a project only where its values are drawn from terms that exist independently of the project. A controlled vocabulary unifies values, and an ontology additionally represents relations between concepts in a form that supports inference; the second is not automatically preferable to the first, since its maintenance

requirements are considerably greater. What makes either usable is a set of conditions independent of expressive power: persistent identifiers for each concept, documentation, an explicit licence, a community that sustains the resource and a stated procedure for managing change (World Wide Web Consortium 2009). A vocabulary failing these conditions produces terminological agreement that dissolves the moment its host service does.

Two consequences follow for the framework. Local terms are unavoidable in specialized work and are not defects, but a local term without a persistent identifier and a definition is indistinguishable to a reuser from a typographical error. And multilingual settings require parallel titles, descriptions and language tags rather than translation of values, since translated values silently break the correspondence that controlled vocabularies exist to guarantee. The framework therefore records vocabulary use as a pair of variables: which vocabulary governs each field, and whether its terms resolve to stable identifiers.

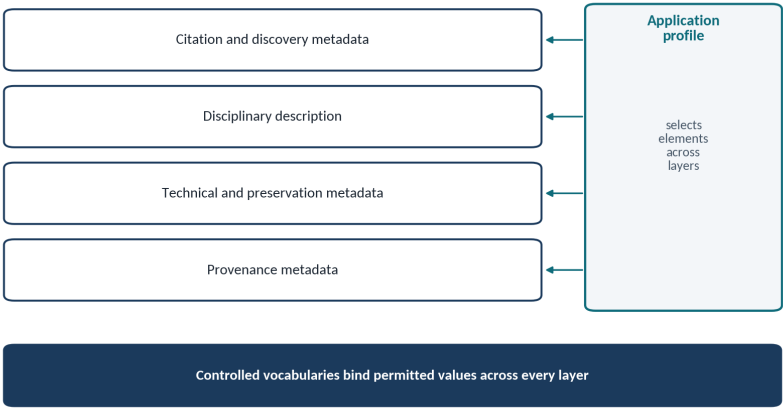


Fig. 8-1. The layered description architecture, showing citation and discovery metadata, disciplinary description, technical and preservation metadata and provenance metadata, with the application profile selecting elements across layers and the controlled vocabularies binding permitted values. Author's own elaboration.

8.4 Assurance as design and control as detection

Quality assurance and quality control are separated in the field's guidance and merged in most practice. Assurance operates before data exist:

instrument design, form design, calibration, training, standard operating procedures and validation rules embedded in collection. Control operates afterwards: range checks, duplicate detection, distributional inspection, double entry and independent review. The distinction matters because the two have different cost structures and different limits, and because control cannot recover information that assurance would have preserved.

The asymmetry is worth stating precisely. Control detects that a value is implausible; it does not establish what the value was, and where the original observation is unrecoverable the detected error becomes missing data rather than corrected data. Assurance prevents the loss but requires decisions before collection, at the moment when the eventual analysis is least well understood. This study treats the balance between them as a variable conditioned by the replaceability of the observation: where re-observation is possible, control is a reasonable substitute for assurance, and where the observation is unique, it is not.

Correction procedure is the point at which quality and provenance meet. Original values are not overwritten; a correction is made in a derived layer and carries its reason, its date and its author, which preserves the ability to distinguish a recording error from a valid extreme value. Missing values are equally heterogeneous — not measured, instrument failure, refusal, inapplicable, below detection limit — and a single undifferentiated code discards a distinction on which analysis frequently depends. The framework consequently treats the reason code for missing values and the correction log as description elements rather than as internal working practice.

8.5 Fitness for use as a purpose-relative property

Quality dimensions such as accuracy, completeness, consistency, timeliness, representativeness, validity and traceability are meaningful only when predicated of a use. A measurement series adequate for detecting a seasonal trend may be inadequate for calibrating a model that is sensitive to short-term variance, and no property of the series itself distinguishes the two cases. Fitness for use is therefore a relation, and a quality statement that omits the purpose against which it was assessed conveys less than it appears to. The framework requires the purpose to be declared alongside the quality statement, which is what converts the statement from a claim about the data into a claim that a reuser can test against a different purpose.

The corollary concerns residual error. Every dataset carries error that its producers know about and have decided not to remove, either because removal is impossible or because it is disproportionate to the declared purpose. Stating that residual error explicitly — its magnitude where quantifiable, its direction where known, its mechanism where identified — is more informative than a claim of accuracy, and it is also falsifiable, which a claim of accuracy generally is not. Table 8-2 illustrates the relation between purpose, the dimensions that become decisive and the form the residual statement takes.

Table 8-2. The purpose-relative form of quality statements, with the dimensions each purpose renders decisive.

Declared purpose	Decisive dimensions	Form of the residual statement
Verification of a published result	Traceability, consistency	Deviations between the analytical dataset and the deposited version
Reuse in secondary analysis	Representativeness, completeness	Selection mechanism and non-response structure
Integration into a combined dataset	Validity, comparability of units and categories	Mapping rules and unresolved category conflicts
Long-term trend analysis	Timeliness, consistency over time	Method and instrument changes and their effect on the series

Aggregation deserves separate mention because it is routinely misread. Combining sources does not raise quality, since the combined dataset inherits the weakest guarantees of its inputs unless harmonization records original units, category mappings and uncertainty. Volume does not compensate for systematic bias, which persists undiminished as the sample grows and becomes harder to detect as the standard errors shrink. The framework therefore records harmonization decisions as description elements and treats an unrecorded harmonization as a quality defect irrespective of the quality of the sources.

8.6 Why description in process outperforms description afterwards

The central argument of this dimension concerns the moment at which description is generated. Description produced during the activity is

generated by an agent who holds the information, in a context where the values are directly observable and where errors are correctable against the state of the work. Description produced afterwards is a reconstruction, and reconstruction is subject to three failures that do not affect contemporaneous description: information held only in working memory decays, salience determines what is recalled rather than what a reuser needs, and no trace remains against which the reconstruction can be checked. This study advances the claim that retrospective description is systematically weaker for these reasons, not because it is written under time pressure at the end of a project, though it usually also is.

The claim is empirical and its evidence in the field is indirect. What is documented is the outcome — deposits with sparse disciplinary description, and clarification requests from reusers on points the deposited record leaves unanswered — rather than the mechanism connecting the moment of generation to the outcome. The framework therefore carries the claim at a moderate strength and treats the timing of description as a decision variable rather than as a settled finding.

[EVIDENCE REQUIRED: paired comparison of description completeness and of reuser clarification requests between datasets described during collection and datasets described at deposit, controlling for discipline, data complexity and repository, to test whether the timing of description generation predicts reuse cost.]

Three boundary conditions limit the claim and are part of it. Where an instrumented pipeline emits parameters, versions and configurations automatically, the description is generated in process even if the document that presents it is assembled at deposit, and the claim does not apply because the relevant record was contemporaneous. Where a dataset is small, stable and produced by one operator over a short interval, decay is negligible and the difference between the two moments approaches zero. And where the applicable description standard did not exist during collection, retrospective description against the current standard is not weaker than the alternative, since the alternative is no description at all. The claim thus concerns the moment at which the information is captured rather than the moment at which the document is written, and conflating the two is the most likely source of apparent counter-examples.

The relation between description quality and the cost of reuse is what gives the claim its practical weight. Every element a record omits transfers

work to the reuser, who must either infer the missing information, request it from a producer who may be unreachable or abandon the reuse. That transfer is invisible in the accounts of the producing project and appears in no indicator of stewardship, which means that the cost of thin description is borne entirely outside the unit that decides how much description to produce. The framework treats this externality as the structural explanation for persistent under-description, and consequently locates the remedy in institutional requirements at deposit rather than in exhortation addressed to producers.

[EVIDENCE REQUIRED: measurement of reuser effort — time to first analysis, number of clarification requests, proportion of abandoned reuse attempts — as a function of application profile completeness, establishing the magnitude of the cost transferred by incomplete description.]

8.7 The description and quality variables carried forward

Table 8-3 states the variables this dimension contributes. Each names a record, and each is readable without asking the producing project what it intended.

Table 8-3. Description and quality variables of the integrated framework and their evidentiary records.

Variable	Definition	Evidentiary record
Application profile	The elements required, their format, vocabulary and validation rule	Documented profile with version
Layer separation	Whether an authoritative disciplinary record exists alongside the general export	Repository record and its derived export
Vocabulary binding	Which vocabulary governs each field and whether its terms resolve	Field definition with vocabulary URI
Description timing	Whether description elements are captured in process or reconstructed	Timestamps of description records against activity records
Assurance measures	Preventive measures fixed before collection	Validation rules, calibration records, procedure documents
Control measures	Detection measures applied after collection	Validation reports; correction log
Declared purpose	The use against which fitness was assessed	Quality statement accompanying the deposit
Residual error statement	Known error retained, with magnitude, direction or mechanism	Quality statement; harmonization documentation

Two boundary conditions qualify the dimension. In disciplines whose primary material is textual or interpretive, the disciplinary layer is carried by scholarly apparatus and by source description rather than by variable-level documentation, and the variables above are satisfied in a different form rather than waived. And in fields where description standards remain unsettled, the profile variable records a local profile with a stated maintenance arrangement, which is weaker than community adoption but stronger than an undocumented convention.

Description and quality determine whether a dataset can be interpreted by someone who did not produce it. Whether it may be released, to whom and under what conditions is determined by the protection dimension, which the following chapter develops.

Chapter 9. Protection, Legal Regime and Controlled Access

Description determines whether a dataset can be interpreted; it does not determine whether the dataset may leave the environment in which it was produced. This chapter takes up protection as the dimension that governs storage, access, transfer and release, and defends two connected claims. The first is that data classification is the conditioning variable of the dimension, in the sense that the storage environment, the access model, the egress control and the publication regime are all functions of it rather than independent choices. The second is that the maxim "as open as possible, as closed as necessary" operates as a decision rule only when the necessity condition is decomposed into assessable variables, since in its undecomposed form it licenses any outcome.

The bearing on the main research question is that protection is the dimension in which conflicting requirements are densest and in which the cost of an unowned decision is highest. A retention obligation and an erasure obligation may apply to the same records, an open licence may be legally available and ethically incompatible with a community protocol, and a security classification may prohibit what a funder mandate requires. Such conflicts cannot be settled inside any of the instruments that generate them, which is why the framework treats protection as a set of variables with named decision authority and recorded residual risk rather than as a compliance checklist.

9.1 Classification as the conditioning variable

Classification is frequently reduced to a binary distinction between personal and non-personal data, and the reduction is the principal source of misclassification. Confidentiality arises equally from contractual restriction, trade secrecy, critical infrastructure protection and cultural sensitivity, none of which is visible in a personal-data test. Classification restricted to confidentiality is incomplete in a second respect, since integrity and availability requirements differ across objects of identical sensitivity (National Institute of Standards and Technology 2024). The framework

accordingly treats classification as a vector over impact on persons, on the institution, on public safety and on the evidential value of the research.

Two properties of classification determine how it enters a decision rule. Classification attaches to data objects and in some cases to variables, not to projects, because a single project routinely combines public sensor records, restricted infrastructure drawings and identifiable contact data. And classification is time-dependent: material confidential before a patent application may be publishable after it, aggregation may lower identification risk and linkage may raise it. A classification recorded once at project start therefore describes a state that no longer obtains, which is the mechanism by which over-protection of ordinary data and under-protection of sensitive fragments arise together. Table 9-1 states what a class determines once assigned.

Table 9-1. Consequences determined by an assigned data class, with the residual risk each combination leaves in place.

Class	Approved environ- ment	Access model	Egress control	Release regime
Public	Ordinary institution- al storage	Open within the team	None	Open deposit
Internal	Institutional net- work storage	Role-based, time-limited	Transfer logged	Deposit with attribu- tion conditions
Confid- ential	Institutional storage with audit, or ap- proved cloud	Role-based, approval re- quired	Export approved by a named role	Controlled access
Highly sensitive	Trusted research environment	Named, as- sessed, pro- ject-bound	Output checking be- fore release	Metadata only, or con- trolled access under agreement

The residual risks differ in kind and not only in degree. Ordinary storage leaves weak audit and export control, an approved cloud platform leaves contractual and jurisdictional exposure, and a trusted research environment leaves cost and workflow restriction borne by the research rather than by the institution. Since every option in Table 9-1 leaves a residual risk, the defensibility of the choice depends on stating which risk was accepted and by whom.

9.2 The legal regime and the limits of consent

Personal data are not defined by the presence of a direct identifier, and a combination of location, age, occupation and event may identify a person in the absence of any name (European Parliament and Council 2016). Pseudonymized data remain personal data, because the additional information permitting attribution continues to exist. The research provisions of the Regulation are conditional rather than permissive: they presuppose a legal basis, purpose limitation, minimization, transparency, security and the safeguards required under Article 89. Treating the research context as an exemption in itself is the most consequential legal error the corpus records.

Consent occupies a particular position that the framework separates into three layers. Ethical consent to participate in a study, the legal basis for processing under Article 6 and the condition permitting processing of special categories under Article 9 are distinct instruments answering distinct questions, and a signed participation form settles only the first. Consent as a legal basis is withdrawable, which places long-term preservation and secondary reuse in a structurally unstable position, and public-interest or legitimate-interest bases combined with Article 89 safeguards are frequently more robust for research conducted by public institutions. The framework therefore records the legal basis as a variable with a named owner, ordinarily the data protection officer, rather than as a field completed by the research team.

The evidential status of current guidance in this area requires explicit statement. European guidance on processing for scientific research purposes and on anonymization was in consultation during 2026 and had not reached final form at the verification date of the corpus (European Data Protection Board 2026). Guidance in draft indicates an interpretive direction without constituting a rule, and the framework carries that distinction as a property of the source rather than as a caveat in the text.

9.3 Anonymization as a contextual risk assessment

Anonymization is a state in which a person is not identifiable by means reasonably likely to be used, and it is therefore relative to an environment rather than intrinsic to a file. Removal of direct identifiers does not establish that state, and the assessment proceeds through three tests: whether a record can be singled out, whether it can be linked to another record or

dataset, and whether an attribute can be inferred with sufficient confidence. Rich qualitative, genetic, spatial and longitudinal data commonly fail at least one test after direct identifiers are removed, and for such data absolute zero risk cannot ordinarily be demonstrated. The assessment consequently depends on what external sources exist, which changes over time and is not a property the producing project controls.

An asymmetry follows that the framework treats as a decision rather than as a technical detail. Anonymization strong enough to satisfy all three tests may destroy the analytical value the data were collected to carry, and the defensible alternative is pseudonymization combined with a stricter access environment rather than undocumented weakening of the protection. Synthetic derivatives support development, teaching and code testing, and are not equivalent to anonymization of the underlying records. The recorded variable is the anonymization decision itself, with the tests applied, the external sources considered and the information loss accepted.

[EVIDENCE REQUIRED: re-identification testing applied to a sample of deposited research datasets declared anonymous, reporting the proportion failing the singling-out, linkability and inference tests, to establish the frequency with which direct-identifier removal is treated as sufficient]

9.4 Controlled access and the Five Safes

The open-or-closed dichotomy leaves unplaced the large class of data that carry high analytical value and non-trivial identification risk. Controlled access is the regime for that class, and it is a service rather than a setting: public metadata, stated conditions, assessment of the applicant and the proposed purpose, a data use agreement, a secure setting, audit and checking of outputs. The Five Safes framework provides the assessment structure by separating safe people, projects, data, settings and outputs, and its analytical contribution is the claim that protection arises from the combination rather than from any one layer (UK Data Service 2025). A dataset too risky to release under one configuration of the five may be releasable under another, which converts an apparently binary judgement into a design problem.

Table 9-2 sets out the access regimes the framework distinguishes, with the condition under which each is proportionate. They are not ordered by strictness alone, since a metadata-only record preserves discoverability

where no access is possible and a trusted research environment preserves analysis where transfer is not.

Table 9-2. Access regimes for data carrying identification or confidentiality risk, with proportionate conditions and residual risk.

Regime	Proportionate condition	Residual risk	Evidence grade
Open anonymous release	Realistic identification risk low and tested	Future external sources may enable linkage	A/B
Download after approval	Moderate risk; identifiable applicant and agreement	Limited control of subsequent use	B
Trusted research environment	High risk or sensitive linkage	Cost; restriction of the analytical workflow	A/B
Metadata only	Access not lawfully or ethically possible	Discoverable but not directly usable	B

A data availability statement is the point at which the chosen regime becomes visible, and its informational content distinguishes a regime from a refusal. A statement naming the repository, the conditions of access, the deciding body and the reason for restriction describes a regime that can be assessed; an undifferentiated formula offering data on reasonable request describes none.

9.5 Continuity as a protection property

Confidentiality dominates discussion of protection, while the failures that terminate research projects are more often failures of availability and integrity. A backup is a separated copy intended for restoration, synchronization propagates deletion and encryption as fast as it propagates content, and an archive pursues long-term interpretability rather than operational recovery. Continuity extends beyond files to knowledge and authorization, since the loss of an encryption key, the departure of the only administrator or the termination of a cloud contract produces the same outcome as physical destruction. The framework therefore treats continuity as part of protection rather than as an adjacent infrastructure service.

The measured quantity is the decisive point. The existence of a backup job establishes nothing about whether the copy is complete, readable and

usable with the software and keys required to interpret it, and the corresponding variables are the recovery point objective, the recovery time objective and the date and outcome of the last restore test. Both objectives are properties of a data class rather than of an institution, since irreplaceable observational input and regenerable intermediate output warrant different values, and a uniform daily schedule expresses no assessment at all. At least one administratively separated or immutable copy addresses the failure mode that a compromised account or an administrator error creates, which copies inside one administrative domain do not.

[EVIDENCE REQUIRED: audit of research projects recording whether recovery objectives are stated per data class and whether a full restore test was performed and documented within a defined period, to establish the divergence between backup provision and demonstrated recoverability]

9.6 Rights as layers rather than as ownership

The expression data ownership describes no single legal relation and obscures the layering that actually governs release. One research package may combine unprotected facts, authored text, third-party images, a database structure attracting the *sui generis* right where substantial investment can be shown, personal data and software, each with a different rights holder (European Parliament and Council 1996). Employment relations determine who may grant permissions for material produced in the course of duties, and contractual research adds provisions on background, results, access rights, publication and the state of affairs after termination. For stewardship purposes the operative questions are narrower than ownership: who may authorize archiving, who may authorize publication and under what conditions further use is permitted.

Licence selection follows from that inventory rather than from a default. A licence conveys only the permissions the grantor actually holds and cannot create rights in unprotected facts, remove obligations under data protection law, or override third-party rights and community authority (Carroll et al. 2020). Objects of different kinds require licences of different families, so that one undifferentiated licence file covering data, documentation and code creates ambiguity precisely where machine processing requires certainty; a licence manifest recording, per logical part, the rights holder, the source of the permission and the excluded material resolves it.

The framework records the licence as a decision with a stated consequence, since the permissive dedication that maximizes legal reuse is unavailable where the depositor lacks full rights, and an attribution requirement accumulates across derived compilations.

9.7 Making the necessity condition assessable

The maxim "as open as possible, as closed as necessary" is widely adopted and, taken alone, is not a decision rule, because it contains no test for the necessity it invokes. Any level of restriction can be justified by asserting necessity, and any level of openness by denying it, so that the maxim ratifies whichever practice an institution already follows. The contribution this study makes at this point is to decompose necessity into six variables, each independently assessable and each with an owner: legal basis, residual identification risk, third-party rights, contractual restriction, community protocol and security classification. Openness is then the outcome that follows when every variable stands in its null state, and each departure from openness is attributable to a named variable rather than to a general appeal to sensitivity.

The decomposition produces two consequences that the undecomposed maxim cannot. A restriction attributed to a variable is reviewable, so that publication of a patent or completion of an anonymization assessment triggers reclassification instead of leaving a restriction in place by inertia. And a restriction attributable to none of the six variables is exposed as an unstated risk preference rather than a requirement. Table 9-3 states the protection variables carried forward into the integrated framework.

Table 9-3. Protection variables of the integrated framework and their evidentiary records.

Variable	Definition	Evidentiary record
Data class	Impact-based class assigned per data object	Classification register with assignment date
Legal basis	The basis for processing, distinguished from ethical consent	Record of processing; data protection officer's assessment
Identification risk	Result of the singling-out, linkability and inference tests	Anonymization assessment with information loss stated
Access regime	Open, approved download, controlled environment or metadata only	Repository access setting; data availability statement
Egress control	Whether outputs are checked before leaving the environment	Output checking log
Recovery objectives	Acceptable data loss and restoration time per class	Recovery plan; dated restore test report
Rights inventory	Rights holder and permission source per object layer	Licence and attribution manifest
Restriction attribution	The necessity variable producing each departure from openness	Decision record naming variable, owner and review date

Three boundary conditions qualify the dimension. Legal conclusions are jurisdiction-dependent, and the variables identify the class of decision and its authority rather than the answer; the framework therefore locates the escalation point and does not substitute for legal assessment. In fields working with community-held knowledge, legal availability and ethical permissibility diverge, and the community protocol variable is not reducible to the rights inventory (Carroll et al. 2020). And in small research settings, a trusted research environment may be unavailable at any acceptable cost, in which case the framework records a metadata-only regime with a stated reason rather than an unsupported claim of anonymity.

Protection determines what may be released and to whom. Whether what is released supports the claims made from it is determined by the reproducibility dimension, which the following chapter develops.

Chapter 10. Reproducibility and Computational Integrity

The protection dimension determines what may be released; it does not determine whether the released material supports the claims drawn from it. This chapter takes up reproducibility as the dimension that connects a stated result to the data, code and conditions that produced it, and defends one principal claim: reproducibility is a property of the evidence rather than an ethical supplement to it. A result that cannot be regenerated from the material accompanying it is not a result whose correctness has been left unverified, but a result whose content is partly undetermined, since the undocumented choices made in producing it are constitutive of the number reported. The chapter therefore states what the framework measures instead of what practitioners are urged to do.

The bearing on the main research question is that reproducibility is the dimension in which the object of stewardship extends furthest beyond the dataset. Data alone do not reconstitute an analysis, and code alone does not run; the analytical object comprises input versions, transformation code, configuration, dependency versions, random seeds and an execution environment, distributed across systems governed by different services and different owners. The framework consequently expresses reproducibility through variables that attach to the computational object as a whole, which is the only level at which the property can be either satisfied or assessed.

10.1 Two properties frequently conflated

Reproducibility and replicability are separate properties with separate evidential force, and their conflation weakens both. Reproducibility obtains when the same data, the same code and the same conditions yield the same result, and it tests the computational chain between observation and claim. Replicability obtains when a new study collecting new data under comparable conditions reaches a consistent conclusion, and it tests the empirical generalization (National Academies of Sciences, Engineering, and Medicine 2019). Failure of the first locates a defect in the pipeline or its documentation; failure of the second may indicate a genuine limit of the phenomenon, and treating either as a verdict on the other misassigns the diagnosis.

The distinction has a consequence for what stewardship can deliver. Reproducibility is within the control of the producing team, since it depends on artefacts the team holds, and it is therefore assessable at deposit; replicability depends on independent effort and later evidence, and no deposit can establish it. The framework accordingly claims jurisdiction over the first and treats the second as an outcome the first enables, which also explains why reproducibility is not a proxy for correctness. Code that faithfully reproduces an erroneous transformation reproduces the error with complete fidelity, so that the property demonstrates traceability and not validity.

10.2 The computational environment as part of the research object

An analysis executes against a stack of library versions, an operating system, a hardware architecture and, increasingly, external services, and each layer can alter the output. Documentation of this stack ranges from a package list through a lock file to a container image or a virtual machine, and the options differ in what they capture, not only in how much effort they cost. A container raises portability substantially and does not guarantee long-term executability, since it depends on a registry, a base image, an architecture and any external service the analysis calls. Table 10-1 sets out the options with the condition under which each is proportionate.

Table 10-1. Options for recording the computational environment, with proportionate conditions and residual risk.

Option	Proportionate condition	Residual risk	Evidence grade
Documented package list	Small analysis in a common language	Versions drift; results not regenerable later	C
Lock file with resolved versions	Standard package ecosystem	Operating system and external services uncaptured	B
Container image with build definition	Complex dependencies or system libraries	Dependence on runtime, registry and architecture	B
Virtual machine or emulation	Historically critical or unmaintainable workflow	High cost; preservation burden transferred forward	B/C

Two consequences follow for the framework. Bit-identical output is not attainable across hardware, parallel execution orders and library versions for many numerical workloads, which makes an acceptable numerical tolerance a variable of the analysis rather than an admission of failure. And an image retained without its build definition, its source and its metadata preserves execution while destroying the ability to reconstruct or amend the environment, so that the framework records the environment specification and the image as distinct artefacts. Retention of both is the arrangement the corpus consistently reports as preferable, on the same reasoning that governs authoritative copies in Section 6.2.

10.3 Degrees of freedom in the path from data to result

Between raw observation and reported estimate lies a sequence of defensible choices: which cases to exclude, which variables to derive, which transformation to apply, which covariates to include, which model to fit and when to stop collecting. Each choice is separately reasonable, and the set of them defines many paths from the same data to different numbers. Where a path is selected after the results of several are known and reported as though it had been fixed in advance, the evidential strength of the reported number is overstated. The effect arises without any intention to deceive (Gelman and Loken 2014; Simmons, Nelson and Simonsohn 2011). Repeated testing during accumulating collection, terminating at the point of significance, alters the error rate of the test in the same way and is admissible only within a sequential design that accounts for it.

The framework treats analytical flexibility as a recordable quantity rather than as a matter of researcher virtue. The recordable element is the pairing of each decision point with the time at which it was made and with the information available at that time, which distinguishes a choice fixed before the data were seen from one conditioned on them. Multiverse and specification-curve analyses extend the record by reporting the distribution of results across the defensible specifications, and their contribution is diagnostic: they show whether a conclusion depends on one analytical choice (Steegen et al. 2016). They do not substitute for a substantively justified primary model, since a curve of results without a reasoned primary specification reports variation without interpreting it.

10.4 Pre-specification as a separator of claim types

Pre-specification separates the analytical plan from knowledge of the results, and its function is not the suppression of exploration but the labelling of it. A plan recorded before collection, with a time stamp and an unalterable record, permits a later reader to distinguish a confirmatory claim from an exploratory one; it does not prohibit the second, and exploration remains the mechanism by which most hypotheses originate (Nosek et al. 2018). Deviation from a plan is frequently justified by conditions encountered during collection, and a transparently reported deviation is stronger evidence than adherence to a plan that has become inapplicable. What the plan makes possible is the assignment of each reported claim to one of two categories with different evidential weight.

Three boundary conditions limit the reach of the device. Pre-specification does not confer quality, and a poorly reasoned pre-registered model may be less appropriate than a transparently corrected analysis. Not every field tests hypotheses, and qualitative, exploratory, modelling and simulation research requires a different form of analytical plan describing the intended procedure, the criteria of interpretation and the conditions for revision. And pre-specification interacts with the planning dimension of Chapter 7, since a statistical plan that is not connected to the data management plan produces two documents describing one project with no defined relation between them.

[EVIDENCE REQUIRED: comparison of pre-registered analytical plans against the corresponding published analyses for a sample of studies, recording the proportion of deviations that are reported as such, to establish whether pre-specification functions as a separator of claim types in practice]

10.5 Uncertainty as a reported quantity

Uncertainty in a research result originates in measurement, sampling, model specification, imputation, calibration and future variability, and reporting a standard error addresses one of these sources. A single point estimate accompanied by a threshold decision suppresses the remaining sources rather than resolving them, and the framework treats the sources acknowledged and the sources set aside as elements of the result. Sensitivity analysis against defensible alternative assumptions is the operational form of that

acknowledgement, and its output belongs with the primary estimate rather than in supplementary material. Precision-oriented planning, in which sample size is derived from the width of interval the conclusion requires, extends the same logic backwards into the design.

Graphical communication is where suppression of uncertainty most often occurs, and where it is least visible to the reader. Summary statistics can conceal multimodality, clustering and outlying structure that a plot of the underlying values reveals, which is why visualization functions as part of the analysis rather than as its decoration (Anscombe 1973). Error bars carry no fixed meaning, since they may denote a standard deviation, a standard error, a confidence interval, a prediction interval or a credible interval, and an undefined bar is uninterpretable rather than approximately interpretable. The evidence indicates that no single graphical form is optimal for communicating uncertainty across tasks, so the framework records the interval definition, the sample size and the smoothing or binning parameter instead of prescribing a chart type.

Two further properties connect visualization to the dimensions already established. A figure generated by script from a versioned dataset carries provenance, whereas a figure adjusted by hand in a graphical editor breaks the chain between the displayed value and the data of record. The framework accordingly treats manual adjustment as an event to be documented rather than as a neutral act. And accessibility is a property of the evidence and not a courtesy, since meaning encoded solely in colour is unavailable to part of the readership, and direct labelling, contrast, shape and alternative text carry it where colour cannot (World Wide Web Consortium 2024).

Provenance is broken wherever a step cannot be re-derived from record alone

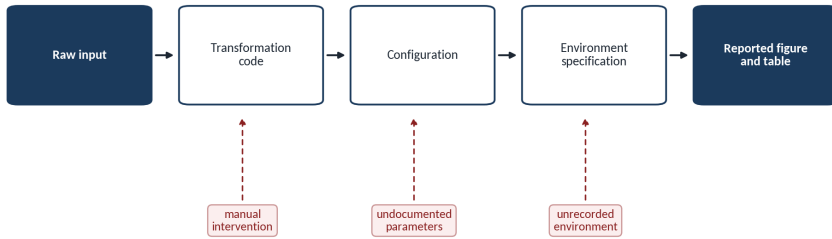


Fig. 10-1. The chain from raw input through transformation code, configuration and environment specification to reported figure and table, showing the points at which provenance is broken by manual intervention, undocumented parameters or an unrecorded environment. Author's own elaboration.

10.6 Research software as an output with its own identity

Research software is simultaneously part of the method and an output in its own right, and treating it only as an appendix to a dataset misdescribes both. It differs from data in being executable, composite and continuously changing, so that the rules governing static objects transfer imperfectly. Findability requires identification of specific releases rather than of a project, and reusability requires a licence, documentation, dependency information and a statement of limitations (Barker et al. 2022). A commit in a version control system is not a usable citation target for a reader; whereas a tagged release, archived and bound to a persistent identifier, is. Software citation practice supplies the mechanism, and a machine-readable citation file makes the intended citation explicit rather than inferred (Smith, Katz and Niemeyer 2016).

Identification of code has acquired dedicated infrastructure, and its limits are as important as its capabilities. A content-addressed identifier for archived source code identifies exactly the state of the code that was run. Metadata schemas now admit such identifiers as a related identifier type, which permits a dataset, an article and a code state to be linked as one connected object (DataCite Metadata Working Group 2026; Soiland-Reyes et al. 2022). Archived source code is nevertheless not automatically buildable or executable, since dependencies, configuration, test data and documentation

are required to reconstruct behaviour. Maintenance after a project ends requires a named owner and funding, and an open licence by itself produces neither a community nor corrections, which is why the framework records a maintenance horizon and a procedure for marking software unmaintained.

The identifier scheme itself has meanwhile been standardized: the SoftWare Hash IDentifier, developed for the Software Heritage archive, was published as an international standard in 2025 (ISO/IEC 2025), which converts a community convention into a citable normative reference.

10.7 What the framework measures

The variables of this dimension are chosen because each names a record that can be inspected without asking the producing team what it intended. Table 10-2 states them. Two carry most of the discriminating power. The first is the clean-run test, in which the workflow is executed from an empty environment from beginning to end and its result recorded. The second is the separation of immutable input, transformation code, configuration and generated output, which makes the direction of derivation visible in the structure of the deposit itself.

Table 10-2. Reproducibility variables of the integrated framework and their evidentiary records.

Variable	Definition	Evidentiary record
Object separation	Whether input, code, configuration and output occupy distinct layers	Deposit structure; manifest
Environment specification	The form in which the execution stack is recorded	Lock file, build definition or image with metadata
Version binding	Which data version, code release and dependency set produced a result	Release identifiers cited in the output
Numerical tolerance	The deviation accepted as a successful reproduction	Statement accompanying the reproduction package
Clean-run result	Outcome of end-to-end execution from an empty environment	Dated execution log
Decision timing	When each analytical choice was fixed relative to the data	Analytical plan with time stamp; deviation record
Uncertainty statement	Sources acknowledged and sources set aside	Sensitivity analysis; interval definitions
Figure provenance	Whether outputs are generated by script from the data of record	Generating script; documented manual intervention
Software identity	Whether releases are identified, licensed, documented and citable	Release identifier; citation file; maintenance horizon

Protected data constrain the dimension without suspending it. Where input data cannot be released, complete code, a documented workflow and a minimal or synthetic example dataset preserve inspectability of the procedure, and verification against the real inputs can be performed inside the controlled environment described in Section 9.4. Transparency of this kind does not require breach of a legal restriction, and its absence is therefore not explained by sensitivity alone. What the framework records in such cases is which component was withheld, under which access regime it can be examined and by whom that examination was performed.

Two boundary conditions qualify the whole dimension. In fields whose analysis is interpretive rather than computational, the chain from material to claim runs through coding schemes, analytical memos and source apparatus, and the variables are satisfied in that form rather than waived. And full environment preservation carries financial and environmental cost that

scales with the number of preserved objects, so that the framework treats the depth of environment capture as proportionate to the irreplaceability of the result rather than as a uniform requirement.

Reproducibility establishes that a result can be regenerated while the artefacts remain intact and interpretable. Whether they remain so over the periods for which research evidence is expected to hold, and whether the effort of maintaining them is recognized, is determined by the preservation and recognition dimension, which the following chapter develops.

Chapter 11. Preservation, Publication and Recognition

The dimensions established so far determine whether a result can be interpreted, released and regenerated at the moment of deposit. None of them determines whether the material remains interpretable once the project that produced it has ended, or whether the effort of keeping it so is visible to anyone assessing the institution. This chapter takes up preservation, publication and recognition as one dimension, and defends two claims. The first is that preservation is an active commitment distinct from both backup and publication. The second is that recognition is a structural variable of the framework rather than a matter of professional morale, because effort that is not assessable is not sustained.

The bearing on the main research question is that this dimension carries the framework's longest time horizon and its weakest ownership. Decisions about retention, format and repository take effect years after the project team has dispersed, and the costs fall on an institution that did not participate in the decision. Where the effort involved is not registered by any assessment system, the equilibrium outcome is deposit at the lowest cost that satisfies a funder check, which is a rational response to the incentives rather than a failure of diligence. The framework therefore treats recognition as a variable that conditions the other variables in the dimension.

11.1 Three functions frequently merged

Backup, preservation and publication answer different questions and fail in different ways, and their conflation is the most common structural error in institutional practice. Backup restores an operational state after loss and is measured by recovery objectives, as Section 9.5 established. Preservation maintains the ability of a future community to find, verify and interpret an object, and it is therefore an organizational commitment held by people, systems and procedures rather than a storage arrangement (Consultative Committee for Space Data Systems 2012; International Organization for Standardization 2025). Publication makes an object citable and discoverable, and a published object with a persistent identifier may be neither backed up nor preserved in any active sense.

The reference model contributes two constructs that the framework adopts and one term that it guards. Information packages separate what is submitted, what is stored and what is disseminated, and preservation planning is a continuing function rather than an event at ingest. The designated community fixes the knowledge base assumed of future users, and it converts the otherwise unbounded question of how much documentation is enough into an assessable one, since sufficiency is relative to a stated community. The word open in the model's name denotes open development of the standard and not unrestricted access, and the framework keeps preservation and access as separate variables for precisely this reason.

11.2 Appraisal and retention as decisions with owners

Retention of everything is neither financially nor environmentally sustainable at institutional scale, and it degrades preservation by dispersing effort away from the objects that warrant it (National Digital Stewardship Alliance 2026). Appraisal is therefore unavoidable, and the question is whether it occurs as a documented decision or as attrition. The criteria the corpus consistently reports are scientific and evidential value, irreplaceability, the legal position, the cost of retention and the possibility of regenerating the object, and these can be weighed only against a stated retention period. Disposal decided by these criteria and recorded is a stewardship act; disposal by expiry of a storage contract is a loss with no decision behind it.

Two structural features determine whether appraisal survives contact with practice. Retention periods arise from several sources at once — funder conditions, legal obligations, disciplinary norms and the institution's own policy — and where the sources conflict, an ordering rule and a named authority are required rather than a longer default. And an erasure obligation arising under data protection law may apply to material that a retention obligation also covers, a conflict that cannot be resolved inside either instrument and that the framework routes to the escalation authority identified in Chapter 5. The recorded variables are consequently the retention period, its source, the appraisal decision and the role that took it.

[EVIDENCE REQUIRED: institutional audit recording, for a defined population of completed projects, whether a retention period was stated, whether an appraisal

decision was documented and what proportion of project data remained on unmanaged storage after project close]

11.3 Obsolescence, migration and integrity

Formats fail in two ways that require different responses: the software able to render a format ceases to exist, and the specification of the format is lost or was never public. Migration to an open or widely supported format addresses the first and may alter precision, functionality or meaning, particularly for proprietary instrument output; emulation preserves original behaviour at a cost that recurs indefinitely. The choice between them is determined by the significant properties that the object has to retain, which is a scientific judgement about what the object is evidence for rather than a technical preference. For proprietary scientific formats the arrangement most consistently reported as defensible retains the original, a documented export and the conversion documentation together.

Integrity verification is frequently taken to settle more than it does. A checksum recomputed and compared demonstrates that the bit sequence has not changed, and it demonstrates nothing about whether the object remains interpretable or whether a migration preserved its meaning. The two properties therefore require separate evidence: fixity checks with dated results, and validation that the significant properties survive a format transformation. A preservation metadata vocabulary supplies the structure in which such events, the agents that performed them and the rights governing them are recorded, so that the preservation history of an object is itself an inspectable record (Library of Congress 2015).

Long-term management extends beyond the object to the service holding it. Version records, tombstone records for withdrawn objects, changes of licence and a succession plan describe what happens when an object or a repository changes state. A repository unable to state what becomes of its holdings if it closes has not addressed the failure mode that most threatens its depositors. Certification requirements have moved in this direction, strengthening the evidence expected for sustainability, shared responsibility and cessation planning (CoreTrustSeal 2022). This study treats an exit arrangement as a property of the deposit decision rather than as an operational matter internal to the repository.

11.4 Repository selection as an assessment

Repository selection is commonly settled by habit, by the choice a colleague made or by whichever service issues an identifier with least friction. Treated as an assessment, it resolves into criteria that can be stated in advance and applied to candidates. These comprise disciplinary fit, the mandatory requirements of funder, institution and journal, supported access regimes, persistent identification, curation and metadata quality, size and format limits, versioning, cost, service commitments and the arrangements that apply when the service ends. Registration and certification are distinct signals and are frequently conflated: a registry entry indicates that a service exists and describes itself, whereas certification indicates that documented requirements were assessed by an external body (re3data.org 2026; CoreTrustSeal 2022). Table 11-1 states the criteria and the evidence that satisfies each.

Table 11-1. Criteria for repository selection, with the evidence that satisfies each and the residual risk if it is absent.

Criterion	Satisfying evidence	Residual risk if absent
Disciplinary fit	Community-adopted metadata standard; specialist curation	Record found but scientifically uninterpretable
Mandatory requirements	Funder, institutional and journal conditions checked	Deposit does not discharge the obligation
Certification	Current external certification, verified at the issuing body	Trustworthiness asserted rather than assessed
Persistent identification	Identifiers minted and resolved, with versioning	Citation breaks; no stable reference to a version
Access regimes	Open, embargoed, controlled and metadata-only supported	Sensitive data forced into an unsuitable regime
Exit arrangement	Documented succession or transfer plan; metadata export	Holdings dispersed without custody transfer

The criteria conflict, and the framework's contribution here is to name the conflict rather than to conceal it in a ranking. The disciplinary repository frequently offers the richest metadata and the community most likely to reuse the data, and it may lack certification, size capacity or support for controlled access; the certified generalist repository offers assessed

trustworthiness and weaker scientific description. Where the conflict cannot be resolved, split deposit connected by persistent identifiers is available, at the cost of maintaining the connection over time. The evidence supporting these options is graded B, since it rests on case reports and service documentation rather than on comparative measurement of reuse outcomes.

[EVIDENCE REQUIRED: comparative measurement of reuse, citation and long-term availability for datasets deposited in disciplinary against certified generalist repositories, controlling for field, to establish whether the trade-off between disciplinary fit and certified trustworthiness has measurable consequences]

11.5 Publication of data as a distinct scholarly act

Deposit and publication are separated in the framework because they carry different obligations. A dataset published as an output acquires an identifier, a citation, a licence, a described version and an author list, and it becomes an object that other work can reference precisely (Wilkinson et al. 2016). The act is scholarly rather than administrative. Selection of what constitutes the publishable object, description sufficient for an unknown reuser and the decision on which version is authoritative are analytical judgements, and they are the judgements Chapters 6 and 8 assign to lifecycle and description. A dataset attached as supplementary material to an article has none of these properties and cannot be cited, versioned or assessed independently.

Two features of data publication carry implications for the assessment layer. Citation of data is possible only where the identifier resolves and the metadata record supports it, so that credit for a dataset depends on infrastructure the author does not control (Cousijn et al. 2019). And the contribution to a dataset is typically distributed across collection, curation, description and software, which the single authorship convention of the article represents poorly; contributor role taxonomies exist to make those contributions separately visible. The framework records the published object, its identifier, its version and the contributor roles as variables, on the reasoning that credit which cannot be attributed cannot be assessed.

11.6 Recognition as a structural variable

Research assessment reform proceeds from the position that a narrow set of publication metrics represents research contribution inadequately and that

data, software, methods, curation and team roles warrant recognition in their own right (CoARA 2022; DORA 2013). The reform's own hazard is visible in the corpus: replacing an article count with a dataset count reproduces the mechanism it was intended to correct, since a number of deposits is as insensitive to quality, documentation and reuse as a number of papers. Responsible assessment therefore requires qualitative evidence of contribution interpreted in disciplinary context, supported by indicators rather than replaced by them. Usage and citation statistics for data are useful and are not comparable across fields, and their manipulability is well enough understood to prevent their use as a sole basis.

The Czech implementation layer provides a concrete instance of the general movement and a test of its consequences. From 2026 a digital data collection is recognized as a distinct result type in the national assessment system, with registration in the national metadata directory as the associated condition (Úřad vlády České republiky 2025; EOSC CZ 2025). Two effects follow that are analytically separable from the national setting. The first is that a data output becomes an assessable object, so that curation and description effort enters the record through which institutions are funded. The second is that a metadata registration requirement converts description quality from a professional norm into a condition of recognition, which changes who within an institution has an interest in it.

This study treats the Czech arrangement as an instance rather than as guidance, and its analytical interest lies in the mechanism it exposes. Recognition operates as a variable that conditions the others in this dimension, because appraisal, format decisions, repository assessment and description all consume effort that is invisible to an assessment system built on publication counts alone. Where recognition is present, the framework's assessment criteria describe activity the institution has reason to sustain; where it is absent, the same criteria describe an unfunded expectation, and the boundary condition on the whole framework is correspondingly sharp. The prediction this study derives from the mechanism is testable, and the evidence required to test it does not yet exist.

[EVIDENCE REQUIRED: measurement of deposit volume, metadata completeness and controlled-access provision at Czech research organizations before and after the introduction of the data collection result type, against a comparison group without such

recognition, to test whether recognition changes stewardship behaviour rather than reporting behaviour]

11.7 The variables carried forward

Table 11-2 states the variables this dimension contributes to the integrated framework. Each names a record, and each is inspectable after the producing project has ended, which is the condition the dimension exists to satisfy.

Table 11-2. Preservation, publication and recognition variables of the integrated framework and their evidentiary records.

Variable	Definition	Evidentiary record
Designated community	The knowledge base assumed of future users	Preservation policy statement
Retention period	The period, its source and the resolving authority	Retention register; escalation record
Appraisal decision	What is retained, what is disposed of and on what criteria	Dated appraisal record with the deciding role
Significant properties	The properties an object has to retain through transformation	Format assessment; migration documentation
Integrity evidence	Dated fixity results and validation after migration	Preservation event records
Repository assessment	Candidates assessed against stated criteria, with the choice reasoned	Selection record referenced in the data management plan
Exit arrangement	What becomes of holdings if the service ends	Documented succession plan; metadata export provision
Published object	Identifier, version, licence and contributor roles	Repository record; citation metadata
Recognition path	How stewardship effort enters institutional assessment	Assessment rules; result registration records

Three boundary conditions qualify the dimension. Preservation obligations for very large data volumes cannot be met by retaining everything, and the framework's response is a documented appraisal rather than an unmet universal commitment. Certification status is time-limited and revisable, so that a selection record citing certification without a verification date states less than it appears to. And recognition arrangements are jurisdiction-

specific and recent, so that the mechanism described in Section 11.6 is stated as a conditional prediction rather than as an established regularity, pending the measurement the marker above specifies.

The five dimensions developed in Part III complete the variable set of the framework. Whether these shared variables retain discriminating power when applied to data of markedly different kinds, and where their application breaks down, is the question the following part addresses through disciplinary profiles and an integrated case.

Chapter 12. Operationalizing the Framework: Decision Rules and Maturity Levels

The seven preceding chapters establish a set of variables. They do not yet establish a framework, since a set of variables becomes one only when its structure is stated, its relations of dependence are declared, its behaviour under conflict is specified and its states are observable by someone other than the party described. This chapter performs that assembly and carries the principal original contribution of the book. The claim it defends is that the 55 variables of Chapters 5–11 compose into one structure of four levels, ordered by a defensible relation of conditioning, closed under five decision rules and readable as a profile of observable states rather than as a score.

Integration was posed in Chapter 1 as the question whether partial requirements can be converted into defensible decisions. A structure leaving the priority between competing requirements unstated does not answer it, since the conflicts identified in Section 4.5 are exactly where the partial instruments fail. What follows is therefore falsifiable at the point of conflict rather than at the point of description, and it remains a construction whose validation is the business of Part IV.

12.1 Four levels: dimension, component, variable, indicator

The framework distinguishes four levels, and the distinction is not presentational. A dimension is a grouping of requirements sharing a coding mode, and conditioning holds at that level. A component is a set of variables attaching to one decision object and one accountable party, and an institution allocates work at that level. A variable is a property of a data activity with a defined state space and a named evidentiary record, and a decision is taken at that level.

The fourth level is new to this chapter and supplies what Chapters 5–11 left open. Those chapters name, for every variable, the record from which its state is read, and a record does not by itself fix a reading. An indicator is the record together with the operation performed on it, the sampling unit and the date of observation. One arrangement supports readings of unequal

strength — a restore plan, a dated restore report and its outcome — and the framework's assessment properties depend on which is specified.

The framework as assembled carries seven dimensions rather than the six groupings of Table 4-2. Planning was separated because its variables are commitments fixed before an activity rather than descriptions of one, and because their accountable party differs from that of the governance variables with which coding had grouped them. Quality was moved to the description grouping, since fitness for use and interpretability are read from the same records. Both bear on the reliability of the derivation procedure and strengthen the case for the recoding in Section 4.4. Table 12-1 sets out the result.

Table 12-1. The seven dimensions, their components, the variables each contributes and their position in the conditioning order.

Dimension	Components	Vari- ables	Position
D1 Governance and ac- countability	Authority; assignment and continu- ity; service and measurement	6	conditioning
D2 Lifecycle and proven- ance	Object identity; derivation; persist- ence of capability	8	conditioning
D3 Planning as commit- ment	Commitment specificity; exchange; revision	7	intermediate
D4 Description and qual- ity	Description structure; quality de- termination	8	conditioned
D5 Protection and access	Classification and legal ground; access and containment; rights	8	conditioning
D6 Reproducibility	Object and environment; execution evidence; analytical discipline	9	conditioned
D7 Preservation, publica- tion and recognition	Appraisal; interpretability; custody and release; recognition	9	conditioned, with a back-edge

12.2 The conditioning order and its defence

Three dimensions condition the others, each in a different currency. Governance transmits authority, since a framework without a named au-
thority produces analyses rather than determinations. Lifecycle and
provenance transmit assessability, since every variable names a record and

records exist only where object identity, derivation and custody are managed. Protection transmits the option set, because the data class determines which storage environments, access regimes and publication routes remain available.

Planning occupies an intermediate position that Table 12-2 represents rather than smooths away. Its variables are conditioned by those three dimensions, since a commitment cannot be made without authority, verified without a record or exceed what the data class permits, and its irreversible commitments then condition the dimensions that follow.

Table 12-2. Conditioning relations between the dimensions, with what each relation transmits.

Dimension	Conditioned by	Conditions	Transmitted
D1	—	D2-D7	authority to decide and to accept residual risk
D2	D1	D3-D7	existence and linkage of the records other variables name
D5	D1, D2	D3, D4, D6, D7	the admissible option set, through the data class
D3	D1, D2, D5	D4, D6, D7	commitments not revisable without loss
D4	D1-D3, D5	D6, D7	interpretability, without which regeneration is undefined
D6	D1-D5	D7	the evidential content of what is preserved
D7	D1-D6	all, through recognition	resources and attention

The ordering is defended by three tests rather than asserted. Removing a conditioning dimension leaves the conditioned dimension without a determinate state, and the converse does not hold. No conditioned variable changes the state of a conditioning variable, so the relation is asymmetric wherever it is claimed. And the closure test of Section 4.6 is satisfied, because D1 has no predecessor.

One relation runs against the order deliberately. The recognition component of D7 conditions every dimension, because effort invisible to institutional assessment is not sustained, and appraisal, description, curation and environment capture consume effort of that kind (CoARA

2022). The back-edge creates no cycle in decision dependency, since recognition determines whether an activity is resourced rather than what it determines. Its presence nevertheless means that the assessment criteria describe an unfunded expectation wherever the recognition path is absent.

[EVIDENCE REQUIRED: cross-sectional measurement of variable states within a set of projects, testing whether conditioned variables covary with their conditioning variables as Table 12-2 asserts — the observation that would falsify the ordering.]

12.3 Decision rules for competing requirements

The rules form an ordered procedure applied to conflicts between correctly applied requirements. Rule 1 separates admissibility from priority: a non-derogable requirement — a legal obligation, a binding retention period, the scope of a valid consent — removes options rather than competing within the set. Rule 2 orders the remainder by irreversibility, placing disclosure, deletion and loss of an executable environment above thin metadata and delayed registration. Rule 3 applies where both options remain reversible and prefers the requirement carrying the higher evidence grade of Section 4.5.

Two further rules dispose of the residue. Rule 4 applies where grades are equal and prefers the requirement attaching to the higher data class, with cost assessed against the impact of that class rather than against a budget line. Rule 5 governs what remains: where no preceding rule discriminates, the framework adopts the option preserving the larger future option set — retention over disposal, controlled access over open release — and records the decision as taken under an unresolved conflict with a stated review condition. Rule 5 is a default rather than a determination, and treating it as one is the most likely misuse of the framework. Table 12-3 states the rules and their authorities.

Table 12-3. The decision rules, the authority entitled to apply each and the record produced.

Rule	Applies when	Determination	Deciding authority	Record
1 Admissibility	A requirement is non-derogable	Removes options from the set	Legal office, data protection officer or ethics committee	Assessment naming the provision
2 Irreversibility	Options differ in reparability	Prefers the reparable breach	Accountable party for the activity	Decision record stating the ground
3 Evidence grade	Both options reparable	Prefers the higher-graded requirement	Accountable party, on advice	Decision record citing grade and source
4 Class proportionality	Grades equal	Prefers the higher data class	Owner of the classification register	Classification reference in the record
5 Option preservation	No rule discriminates	Adopts the reversible option provisionally	Escalation recipient named in D1	Unresolved-conflict record with review condition

The lexical ordering carries a substantive commitment, in that reversibility outranks evidential strength. A weakly evidenced requirement whose breach is permanent thus outranks a strongly evidenced requirement whose breach can be repaired, which inverts the ranking a purely evidential ordering produces. The inversion is defended by the conditions of the decision: future use is uncertain, and the option set is the asset being managed.

Every application of Rules 1–5 produces a residual risk, which the framework treats as an output with its own record. That record carries six fields: the variables in conflict, the option adopted, the option foregone, the risk it leaves, the role accepting that risk and the condition under which the decision is revisited. A decision recorded without the fifth field is indistinguishable afterwards from no decision, and the sixth prevents a restriction adopted under Rule 5 from persisting after its ground has lapsed.

[EVIDENCE REQUIRED: application of Rules 1–5 to a fixed set of documented conflict cases by several assessors working independently, reporting the proportion receiving the same determination and the same residual risk attribution.]

12.4 Maturity as observable state

Maturity levels in this field are commonly stated as aspirational labels whose criteria describe intent, and such scales are satisfied by an institution that describes itself well (Paulk et al. 1993; Crowston and Qin 2011). The framework instead defines five levels as observable states of a single variable, each distinguished from the level below by one observable an assessor can seek and fail to find. Levels attach to variables rather than to institutions, and institutional maturity is the distribution of the 55 variables across them. A profile with three variables at level 0 warrants a different response from a uniformly mediocre profile of the same mean, and one number discards that difference. Table 12-4 states the five levels.

Table 12-4. Maturity levels as observable states, with the observable distinguishing each level from the one below.

Level	State	Distinguishing observable	Typical record
0 Unassigned	The variable has no readable state	—	none
1 Asserted	A document states the requirement for the activity class	A document naming the requirement and applying to the activity	Policy; template; plan item
2 Assigned	A named role and a threshold exist for the activity	A record naming a role and a date rather than a unit or an intention	Assignment matrix; catalogue entry
3 Evidenced	The named record exists for sampled activities and matches the assertion	A paired comparison of commitment against trace, succeeding on a sample	Log; validation or deposit record
4 Tested	The determination has survived an event the institution did not schedule	A dated outcome of restoration, re-execution, an access request or an incident	Test report; incident record with resolution

Most institutional self-descriptions terminate between levels 1 and 2, and the framework’s assessment claim is concentrated between levels 2 and 3. Level 3 is reachable only through a record produced as a by-product of the activity rather than as an artefact of assessment, which is what makes it costly to simulate. Level 4 adds the condition that the timing of the observation was not controlled by the party observed, and it is the only level testing a determination against reality rather than against documentation. The

discriminating power of these thresholds is a claim about measurement and has not been measured.

[EVIDENCE REQUIRED: assessment of the 55 variables at institutions selected to differ on independent indicators of research data performance, testing whether the 2-to-3 and 3-to-4 thresholds separate institutions that independent evidence distinguishes.]

12.5 Automation and machine learning within the framework

Automation applied to research data, including machine learning as method and as object of stewardship, does not constitute a further dimension, and the reason is structural. A dimension is warranted where a class of requirement addresses a decision object no existing dimension addresses. The requirements arising here each attach to an object the seven dimensions already carry, as Table 12-5 sets out (Geburu et al. 2021; Mitchell et al. 2019; National Institute of Standards and Technology 2023). What changes is the state of existing variables and the location of their evidence.

One property does change structurally, and the framework's response to it is the strongest argument against a separate dimension. Automation moves the agent of a data activity from a person to a system, and the derivation record of D2 already carries an agent field. The advisory-decision boundary of D1 then transfers without amendment: an automated classification, a suggested metadata value or a recommended access determination is advice, and the decision is its acceptance by a party with authority. Treating a machine output as a decision repeats the category error by which an orienting assessment comes to function as an approval.

Table 12-5. Where automation and machine learning requirements attach.

Requirement	Dimension and variable
Provenance and permission for training material	D2 rights trace; D5 legal basis, rights inventory
Dataset composition, exclusions and limitations	D4 application profile, declared purpose, residual error statement
Separation of training, validation and test partitions	D6 object separation, decision timing
Versioning of model, parameters, code and environment	D6 environment specification, version binding
External service used on non-public inputs	D5 data class, egress control; D1 decision boundary
Model assessment, drift and adversarial testing	D4 control measures; D6 uncertainty statement

Two consequences follow for assessment. Instruments addressed to artificial intelligence raise the evidence grade of certain requirements and shift the escalation point for certain use classes, and neither effect requires a new variable (European Parliament and Council 2024). Automated controls monitoring changes the cost of observation rather than the level observed, making level 3 cheap while producing level 4 only where an unscheduled event has been survived.

[EVIDENCE REQUIRED: coding of requirement statements in the artificial intelligence governance and documentation corpus against the 55 variables, reporting the proportion assignable to an existing variable and characterizing any residue — the test of whether Table 12-5 is exhaustive rather than illustrative.]

12.6 The status of the construction

What Part III presents is a construction with properties of two kinds. Internal properties are established by inspection, and the three tests of Section 4.6 return the following: every variable is attachable to a named role, every variable names a record, and every conflict routes to D1. External properties cannot be established by inspection at all, since the discriminating power of an indicator, the reliability of a rule and the separating power of a maturity threshold are measurements rather than derivations. No claim of

validity is made here, and the markers above state what would have to be measured for each such claim to become available.

Part IV begins that work. The following chapter applies the framework to institutional settings and distinguishes throughout between reconstruction against documented sources and empirical application against an institution's own records.

PART IV

Validation

Chapter 13. Application to Institutional Settings

Part III closes with a construction whose properties have been established by inspection alone. This chapter opens the part in which the construction meets material it did not generate, and it defends a claim about what such a meeting can establish. Applying a framework to a setting shows whether its variables are expressible in that setting, whether its procedure terminates and where its rules leave a conflict open; it does not show that the framework is valid, and no application reported here is presented as showing that. The chapter therefore separates two operations that the literature on framework development routinely merges, and it states which of them has been carried out.

The bearing on the main research question is one of sequence. Chapter 1 asks under what conditions partial requirements can be integrated into a framework yielding defensible decisions, and a construction answers only the first half of that question. The second half is answered by contact with settings in which the requirements actually compete, and the value of such contact depends entirely on whether the reader can tell reconstruction from measurement. That distinction is the organizing principle of what follows.

13.1 The logic of case selection

Settings were selected to maximize variation on the four context conditions that Chapter 14 identifies as the sources of sensitivity: disciplinary convention, legal regime, institutional scale and available infrastructure. Selection of this kind is theoretical rather than statistical, and it supports claims about the range across which a framework remains expressible rather than claims about how institutions are distributed (Seawright and Gerring 2008). A design selecting only favourable settings cannot produce a negative result, and the set therefore includes one setting in which several of the framework's variables are expected to be unreadable. Table 13-1 states the four settings, what each varies and what each is capable of testing.

Table 13-1. Settings selected for application, with the condition each varies and the determination each tests.

Setting	Varies	Tests	State
S1 Research-intensive multidisciplinary university, federated arrangement, national recognition scheme (Úřad vlády České republiky 2025)	scale and infrastructure at their most favourable	whether all seven dimensions are readable where every record type exists	reconstruction
S2 Small institution without local repository, data protection function or preservation service	scale and infrastructure at their least favourable	whether variables survive when satisfied by external competence	reconstruction
S3 International consortium project combining sensor data, restricted structural drawings, interview material, third-party licensed sources and an in-house classification model	legal regime and object heterogeneity	whether the decision rules terminate under simultaneous conflict	reconstruction
S4 Paired disciplinary settings, one interpretive and one instrumented	disciplinary convention	whether variables retain discriminating power where record forms differ	reconstruction

The choice of S3 requires a word of justification, since a project rather than an institution occupies the unit of analysis. The framework was constrained in Chapter 1 to express the institutional and project levels through the same variables, and a setting in which many conflicts arise at once is the only setting in which that constraint can fail visibly. A project combining public sensor measurement, restricted infrastructure drawings, personal interview material, externally licensed inputs and a model used in a safety-relevant determination engages all five decision rules. Settings in which one conflict arises at a time do not discriminate between a framework with an ordering and a framework without one.

13.2 The application procedure

The procedure has seven steps and follows the conditioning order of Table 12-2 rather than the chapter order of Part III. Data objects are first enumerated in a register with stable identifiers, since every variable attaches either to an object or to an activity performed on one. The conditioning

variables are then assigned in order: data class, authoritative copy and layer, and the accountable party for each activity. Variables of the conditioned dimensions are read only afterwards, because a description or preservation determination taken before the class is fixed is not a determination but a preference.

The remaining four steps produce the output. Each variable's state is read through its indicator, which specifies the operation, the sampling unit and the date rather than merely naming the record. Variables whose state cannot be read are recorded with the reason, distinguishing absence of a record from absence of access to one, since the two carry different consequences and only the first is a finding about the setting. Conflicts are then resolved through Rules 1–5 with the deciding role and residual risk recorded, a maturity level is assigned per variable, and the closure test of Section 4.6 is applied to confirm that every conflict terminated with a named authority.

Two reading modes are available at step three, and the difference between them is the subject of the next section. In reconstruction, states are read from documented material describing settings of the given type. In empirical application, states are read from an institution's own records under negotiated access, with a stated sampling frame and independent coding. The procedure is identical; the evidential status of its output is not.

13.3 Reconstruction and empirical application

No application of this framework has been carried out against the records of a named institution. The four settings in Table 13-1 are reconstructions built from documented material, and they are reported here as reconstructions. Narrating them as though they were institutions whose records had been examined would produce exactly the illustrative case evidence that Chapter 4 identifies as the weakest stratum of this field, with the additional defect of being invented. The book therefore states the design that would convert each reconstruction into a result rather than supplying its appearance.

What a reconstruction can establish is limited but not empty. It can show that every variable of the framework is expressible in the setting, that the procedure terminates, that conflicts route to an authority and that certain conflicts are not settled by the rules. These are properties of the construction observed under a demanding description of a setting, and they are reported

below as such. What a reconstruction cannot establish is the distribution of variable states, the discriminating power of any indicator, the level of agreement between assessors or the frequency with which the rules leave a conflict open.

[EVIDENCE REQUIRED: empirical application at three to five research organizations selected on the criteria of Table 13-1, comprising negotiated access to project records, a sampling frame of data activities defined by the assessor rather than by the institution, independent coding of all 55 variables by two assessors, a reported agreement coefficient, the distribution of maturity levels and a register of conflicts with the rule applied and the residual risk recorded — the study that would convert the reconstructions below into findings about institutions.]

[EVIDENCE REQUIRED: application of the framework to an active multi-partner project across its lifetime, with variable states recorded at three points and with the paired comparison of plan commitments against realized records performed at project close, to establish whether the framework's determinations remain stable as a project's conditions change.]

13.4 What the reconstructions indicate about completeness

Three completeness findings follow from the reconstructions, and each is a property of the variable set rather than of any setting. The first concerns an absent object. No variable among the 55 names an incident or a corrective action, although loss of confidentiality, integrity or availability, erroneous publication, unjustified deletion and loss of provenance are documented events with identifiable owners and records (National Institute of Standards and Technology 2024). The maturity scale of Section 12.4 relies on such events at level 4 while the variable set carries no place to record them, which is an internal inconsistency rather than an omission of detail.

The gap is repairable within the framework's own criterion for admitting a dimension. Incident handling addresses no decision object that governance does not already address, since it concerns who decides, within what interval and with what record, and it therefore belongs in D1 as a further variable rather than as an eighth dimension. Its evidentiary record is an incident register carrying classification by impact type, the assessment of any notification obligation, the corrective and preventive measures with owners

and dates, and the outcome. The framework as stated in Chapter 12 is incomplete in exactly this respect, and the reconstruction is what exposed it.

The second finding concerns apparent duplication. Three variables address the transfer of accountability under different names: continuity assignment in D1, succession arrangement in D2 and exit arrangement in D7. The reconstructions indicate that they are distinguishable by the party whose disappearance triggers them — a person, a project and a service respectively — and that the records differ correspondingly. Retaining them separately is therefore defensible, and the framework states the distinguishing trigger rather than merging the three into one weaker variable.

The third finding concerns cost. Rules 4 and 5 both invoke proportionality, and proportionality presupposes a judgement about what an institution can sustain, yet no variable in the framework carries capacity, funding or unit cost. The corpus treats cost as a constraint on decisions rather than as a property with a record, and the framework inherits that treatment. The consequence is precise and is carried into Chapter 14: the framework determines what is proportionate to a risk and does not determine what is affordable, so that a determination it produces may be correct and unimplementable at once.

13.5 The interpretability of the variables

The reconstructions separate the variable set into two groups whose difference matters more than any aggregate statement about usability. One group is read without interpretation, in that two assessors examining the same record would return the same state: the designated authoritative copy, the access regime set on a deposit, the outcome of a clean run, the identifier and licence of a published object, the date of the last fixity check. Variables of this group carry the framework's assessment weight, because their states are contestable only by producing a different record.

The second group requires a judgement that the framework structures without eliminating. Provenance granularity proportionate to risk, the significant properties of an object, the declared purpose against which fitness is assessed, the specificity of a plan item, the attribution of a restriction to a necessity variable and the knowledge base assumed of a designated community all involve an assessment before a state can be assigned. This is not a defect to be engineered away, since each of these

variables encodes a decision that only a competent party can take. It does mean that maturity levels for the second group are secure only to the extent that assessors agree, which is the condition Chapter 14 treats as prior to the usability of the scale.

[EVIDENCE REQUIRED: independent assignment of states and maturity levels for all 55 variables by at least three assessors working from the same project material, with agreement reported separately for the two groups identified above, to establish which variables are usable in assessment and which require adjudication.]

13.6 Conflicts the decision rules did not settle

Three classes of conflict were constructed in which Rules 1–5 return no determination, and their existence is a finding about the rules rather than about the settings. In the first, two non-derogable requirements meet: a binding retention obligation and an erasure request addressed to the same records. Rule 1 removes options from both sides and empties the set, so the framework produces an escalation with a documented conflict instead of a decision. This is the behaviour the closure test requires, and it is worth stating plainly that a framework which produced a determination here would be overreaching.

In the second class both breaches are irreversible. Publication of an infrastructure dataset cannot be undone, and a withholding period exceeding the useful life of the measurement destroys the value the measurement was collected for. Rule 2 does not discriminate because irreversibility is present on both sides, Rule 3 discriminates only where the evidence grades differ, and Rule 4 refers to a data class defined on one side of the conflict alone. The framework's output is then Rule 5 with a review condition, which preserves the option set at the cost of leaving the decision provisional.

The third class is the most consequential and the least tractable. Where a legal permission and a community protocol conflict, Rule 1 does not engage, because the protocol is not non-derogable in the legal sense, and Rule 3 places the legal instrument above it by evidence grade (Carroll et al. 2020). The determination that follows is one that a substantial part of the field would reject, and the framework's response is to record the residual risk and the accepting role rather than to conceal a value judgement inside a grade. Table 13-2 states the three classes and what the framework produces in each.

Table 13-2. Conflict classes in which the decision rules return no determination.

Class	Why the rules do not dis-criminate	Framework output
Two non-derogable re-quirements	Rule 1 empties the option set	Escalation with documented con-flict; no determination
Both breaches irrevers-ible	Rule 2 inapplicable; Rule 4 defined on one side only	Provisional option under Rule 5, with review condition
Legal permission against community protocol	Rule 3 grades the protocol lower than the instrument	Determination with recorded resid-ual risk and named accepting role

[EVIDENCE REQUIRED: a register of authentic conflict cases collected from research projects, coded for class and for the rule that settled them, reporting the proportion falling into the three classes of Table 13-2 — the measurement on which any claim about the coverage of the rules depends.]

The reconstructions thus return a mixed result that is stated without softening. The framework is expressible across settings of markedly different scale and discipline, its procedure terminates and its conflicts route to an authority; it is incomplete in respect of incidents, silent on cost and indeterminate for a characterizable class of conflicts. Whether its determinations hold when the setting itself changes, and which of its variables move when they do, is the question the following chapter takes up.

Chapter 14. Robustness, Sensitivity and Boundary Conditions

A framework that produces determinations is answerable for the stability of those determinations under conditions it did not choose. This chapter assesses that stability and defends a claim with two parts: the framework's sensitivity to setting is patterned rather than diffuse, and the pattern follows the conditioning order established in Chapter 12. Variables in the conditioning dimensions hold their definition and their reading across settings, while variables in the conditioned dimensions change either the form of their record or their state as the setting changes. Transferability is therefore asymmetric in a way that can be stated precisely, and the statement is a result of this study rather than a caution attached to it.

The bearing on the main research question is the sixth subordinate question of Chapter 1, which asks how far a framework derived from one national and institutional setting transfers and what marks the boundary beyond which it does not. That question is answered here at the level of variables rather than at the level of the framework as a whole, since a framework does not transfer or fail to transfer in one piece. What follows identifies which components move, which hold and under what conditions the determinations weaken to the point of being uninformative.

14.1 Four sources of sensitivity

Disciplinary convention changes the form of the record without changing the requirement it satisfies. In instrumented fields derivation is carried by workflow logs and configuration files, and in interpretive fields by scholarly apparatus, coding schemes and analytical memos; both satisfy the derivation variable, and neither substitutes for the other. What shifts under this condition is the description profile, the granularity at which provenance is proportionate and the form of reproducibility evidence. What holds is the assignment of accountability, the escalation threshold, the classification of the material and the retention decision.

Legal regime operates differently, because it acts on admissibility and therefore on the option set rather than on the form of evidence. The determinations of the protection dimension are jurisdiction-relative by

construction, and the framework locates the class of decision and the authority competent to take it rather than supplying the answer (European Parliament and Council 2016). Institutional scale acts on the location of competence: small institutions satisfy the governance and protection variables through named external competence, and their states are then read from contracts rather than from organizational charts. Available infrastructure acts on attainable levels, since an institution without a repository, a controlled environment or a preservation service cannot raise several variables above level 1 whatever its intentions. Table 14-1 states what shifts and what holds under each condition.

Table 14-1. Sources of sensitivity, with what shifts and what holds under each.

Condition	What shifts	What holds	Effect on determinations
Disciplinary convention	Form of the record; profile; provenance granularity	Assignment, escalation, classification, retention	form of evidence changes; requirement unchanged
Legal regime	Admissibility; legal basis; identification risk; retention period	Structure of the rights inventory; restriction attribution	option set changes; decision class unchanged
Institutional scale	Location of competence; internal or contracted	Existence of a named accountable party	evidence moves to contracts
Available infrastructure	Attainable maturity level	Definition of every variable	ceiling on levels, not on definitions

The last row carries a distinction that maturity instruments in this field routinely confound. An institution whose infrastructure makes level 3 unattainable for a preservation variable is in a different position from an institution that has the infrastructure and does not use it, and a scale reporting one number for both describes them identically. The framework separates the two by recording the reason a state could not be read, as specified in the application procedure of Section 13.2. This separation is the reason an infrastructural ceiling is reported here as a finding about the setting rather than as a deficiency of the institution.

14.2 Which variables are stable and which are setting-dependent

The reconstructions of Chapter 13 support a classification of the variable set into three classes, and the classification is an analytic judgement of this study rather than a measured property. Invariant variables hold their definition, their record type and their reading across all four conditions, and they include accountability assignment, escalation traversability, the authoritative copy, the existence of a derivation record, the revision trace and the attribution of a restriction to a necessity variable. Form-variant variables retain the requirement while the record changes with the setting, and they include the application profile, vocabulary binding, the manual step protocol, environment specification, significant properties and the uncertainty statement. Setting-determined variables have states fixed largely by the setting rather than by the case, and they include the access regime, recovery objectives, repository assessment, retention period, the interoperable minimum and the recognition path (Úřad vlády České republiky 2025).

The distribution of these classes across the dimensions is the substantive finding. Invariant variables concentrate in governance and in lifecycle and provenance, which are the two dimensions that condition all others; setting-determined variables concentrate in protection and in preservation, publication and recognition, which are the dimensions closest to legal regime and infrastructure. The framework is therefore most stable exactly where it carries authority and assessability, and least stable where it carries the answers. This is what makes a decision structure exportable while its outputs remain local.

An alternative explanation of that pattern deserves statement, since it would weaken the result considerably. The conditioning variables may be invariant because they are abstract rather than because they are robust, in that a requirement to name an accountable party is satisfiable in any setting precisely by saying little about what the party does. On this reading the framework's stable core would be stable in the way a truism is stable, and its apparent transferability would be purchased by vacuity. The two readings are distinguishable empirically, because an abstract variable takes the same state everywhere while a robust one discriminates between settings that differ.

[EVIDENCE REQUIRED: independent classification of all 55 variables into the three classes above by two analysts working from the definitions alone, with an agreement coefficient reported, together with the distribution of states for the invariant variables across settings — the observation that separates robustness from abstraction.]

14.3 Conditions under which the determinations weaken or fail

Four conditions weaken the framework's determinations, and the first is absolute. Where a data activity leaves no record whatever, every variable attaching to it stands at level 0 and the framework returns a statement of unreadability rather than an assessment. It cannot diagnose what it cannot read, and an unmanaged environment is therefore invisible to it in the same way it is invisible to an audit. The framework assesses managed environments, and its silence about the remainder is a boundary of the instrument rather than a finding about the practice it fails to see.

The second condition is custody outside the assessed institution. In federated infrastructures, commercial services and trusted research environments the evidentiary record sits behind a boundary that external assessment cannot cross, and the variable is then satisfied by attestation rather than by inspection (CoreTrustSeal 2022). The evidence grade falls accordingly, from direct observation to third-party assurance, and the framework records the substitution rather than treating the two as equivalent. Where neither inspection nor attestation is available, the honest output is an unread variable and not a favourable default.

The third condition is change in the normative and infrastructural layer. The variables were formulated to survive such change, but indicators reference specific artefacts — certification schemes, metadata registries, identifier services — whose currency decays, so that a selection record citing a certification without a verification date states less than it appears to. The fourth condition is legal indeterminacy: where the applicable law is contested or newly in force, the admissibility determination of Rule 1 cannot be made and the ordering stalls at its first step. The framework's output is then an escalation, which is correct behaviour and is not a determination.

14.4 Inter-rater agreement as the condition of usability

The maturity scale of Section 12.4 is usable only on a condition that is prior to any of its content. If two assessors examining the same material assign different levels to the same variable, the resulting profile is not a measurement, and comparisons between institutions or across time have no interpretation. Agreement is thus not a desirable property of the scale but the condition under which the scale exists as an instrument. This study makes no claim that the condition is satisfied, because the measurement establishing it has not been performed.

The reconstructions nevertheless indicate where disagreement is concentrated, and the pattern is informative. Agreement is expected to be high at the boundaries between levels 0, 1 and 2, since the observables at those thresholds are the existence of a document and the presence of a named role and date. It is expected to be lower at the boundaries between levels 2, 3 and 4, where a judgement about the sufficiency of a record enters, and lowest for the interpretive variables identified in Section 13.5. A scale reporting one agreement figure across all variables would conceal exactly this structure.

[EVIDENCE REQUIRED: an agreement study with at least three assessors independently assigning maturity levels to all 55 variables for the same project material, reporting agreement per variable and per level boundary against a threshold stated in advance, with variables falling below it reported as adjudicated rather than assessed.]

14.5 The risk that the instrument is gamed

Any instrument that determines resources reshapes the behaviour it measures, and a maturity instrument is gamed by producing the record rather than the practice. The level definitions of Chapter 12 were constructed against this mechanism: levels 3 and 4 are reachable only through records generated as by-products of the activity and through events whose timing the assessed party did not choose. The design raises the cost of simulation to approximate the cost of the work, which is a design property rather than a proof, and three residual vulnerabilities remain.

The first is the sampling frame. Where the institution selects the activities examined, by-product records can be curated in advance, and the response available to the framework is an assessor-defined frame drawn from an object register the institution maintains for other purposes. The

second is the staged event: a restore test performed on a small, recently written dataset satisfies the letter of level 4 while testing nothing, which is why the indicator specifies the sampling unit and not only the operation. The third is the register kept for assessment: an incident register with no entries is indistinguishable from a setting without incidents unless it is paired with independent traces such as access logs and support requests.

No maturity instrument is immune to these strategies, and a claim of immunity would be the clearest evidence that an instrument had never been used in a setting where something depended on the result. The framework's claim is narrower and testable. It holds that the gap between a self-assessed profile and an assessor-verified profile is a measurable quantity, that it is concentrated in the interpretive variables and that it grows with the consequences attached to the result.

[EVIDENCE REQUIRED: paired self-assessment and assessor verification of the same variables at the same organizations, reporting the direction and magnitude of divergence by variable class, to establish how far self-reported maturity can be relied upon and where independent verification is indispensable.]

14.6 What the framework does not determine

The framework's limits follow from the property that makes it assessable, and stating them is part of the result rather than a qualification of it. A framework whose variables are read from records can determine only what records can carry. Legal conclusions are outside it, since it locates the class of decision and the competent authority without supplying the answer. The scientific significance of a body of data is likewise outside it, because appraisal requires a disciplinary judgement that no record supplies and that the framework routes to a named role instead of resolving.

Four further exclusions are worth stating explicitly. The framework does not determine what is affordable, as Section 13.4 establishes, so a determination it produces may be correct and unimplementable at once. It does not determine whether a decision was right, only whether it was taken by a competent party on a stated basis with the residual risk recorded, which is a weaker and more defensible property. It determines neither tool selection nor organizational form, both of which Chapter 5 treats as context conditions rather than as variables. And it does not determine the ethical

acceptability of a research purpose, which is settled by bodies whose authority the framework presupposes.

14.7 Transferability stated as a finding

The transferability of this framework is asymmetric, and the asymmetry is its most consequential property. The decision structure — the variables of the conditioning dimensions, the ordering, the rules and the requirement that residual risk be recorded — transfers across disciplinary convention, legal regime, scale and infrastructure, because the variables carrying it are invariant under all four. The determinations do not transfer, because they are produced by conditioned variables whose states the setting largely fixes. A framework of this kind can consequently be exported as a structure and cannot be exported as a set of answers.

That conclusion has a direct consequence for how such work is published. Publishing the determinations as recommendations would reproduce the failure diagnosed in Chapter 1, in which instruments designed for one setting are applied in another as though their conclusions travelled with them. The appropriate output is the structure together with the conditions under which each determination holds, which is what Part III states and what this chapter bounds. The remaining question is whether the convergence of institutional practice that the corpus reports admits explanations other than the one this book has given, and Part V takes it up.

PART V

Critical Discussion

Chapter 15. Critical Discussion

Part IV closes with a framework that is expressible across settings of markedly different scale and discipline, whose procedure terminates and whose transferability is asymmetric. None of those properties settles whether the framework is worth having. The claim this chapter defends has two parts: the framework occupies a position that no neighbouring family of instruments occupies, and the alternative explanations capable of dissolving that position survive examination in weakened but non-empty form. What remains at the end is narrower than what Chapter 1 announced, and the narrowing is reported rather than absorbed into a summary.

The bearing on the main research question concerns standing rather than content. Chapter 1 asks under what conditions partial requirements can be integrated into a framework yielding defensible decisions, and Parts III and IV supply a construction with its boundaries. Whether that construction answers the question depends on whether integration has been achieved rather than simulated. The matter is settled below with qualifications stated exactly, since a qualification expressed as a general caution is indistinguishable from none.

15.1 Position among instruments of the same family

The framework shares with the capability maturity tradition a staged structure and an institutional profile as its output (Paulk et al. 1993; Crowston and Qin 2011; Cox et al. 2019). It departs where that tradition locates its levels. Levels here attach to variables rather than to institutions, and institutional maturity is a distribution across 55 variables rather than a stage (Section 12.4). Each level is separated from the one below by an observable an assessor can seek and fail to find, which removes the aspirational language by which stage descriptions are commonly satisfied.

With FAIR assessment instruments the framework shares the ambition of a reading that does not depend on the producer's account of itself (Wilkinson et al. 2016; RDA FAIR Data Maturity Model Working Group 2020). The departure lies in the object assessed. Such instruments predicate their scores of a digital object and are silent about who decided the form that object takes, whereas every variable here is a property of a data activity and

of the record it left, so that a dataset can score highly while the decision producing it was taken by nobody identifiable.

Repository certification is the closest methodological relative, since it assesses a custodian against documented requirements through records rather than assertions (CoreTrustSeal 2022; Lin et al. 2020). The same evidentiary logic is applied here to a different unit: certification addresses the service holding data at the end of a lifecycle, while the framework addresses the institution and project in which the decisions determining what reaches that service are taken.

Research assessment reform shares the diagnosis that counting outputs is insensitive to what the counted objects contain (CoARA 2022; DORA 2013). Reform treats recognition of data work as an end, whereas the framework treats recognition as a variable conditioning every dimension through the back-edge of Section 12.2, on the ground that effort invisible to institutional assessment is not sustained. Assessment reform is on that reading a precondition of stewardship policy rather than a parallel agenda, which is a more falsifiable position than the reform literature ordinarily takes. Table 15-1 states the four relations, and the costs in its fourth column are not incidental: an instrument yielding no single number cannot serve the comparative purposes for which institutions most often commission assessment.

Table 15-1. *The framework's relation to four neighbouring families of instrument, with the cost of each departure.*

Family	Shared	Departure	Cost of the departure
Capability and maturity models	Staged development; institutional profile	Levels attach to variables, and each threshold names an observable	No single score, and therefore no ranking
FAIR assessment instruments	Producer-independent reading	Unit is the data activity and its record	Not automatable; costlier than a parsed score
Repository certification	Record-based assessment of a custodian	Unit is the deciding institution	No accrediting body, and no external warrant
Research assessment reform	Counting is insensitive to content	Recognition is a conditioning variable	Criteria become conditional on policy the institution may not control

15.2 Fragmentation as a transitional condition

The strongest alternative to the diagnosis of Chapter 1 holds that the fragmentation described there is a phase rather than a structure. On this reading the instruments are unconnected because the field is young, and infrastructural convergence — federated cataloguing, common metadata models, national directories — will supply the connecting layer the book supplies analytically (EOSC Association 2024; EOSC CZ 2025). Institutional practice would then converge without any framework, and a decision structure formulated now would document a transitional state. The reading deserves to be taken seriously, because the period 2024–2026 covered by the corpus is one in which such convergence is visibly under way.

The reading is nevertheless incomplete in a way that can be located precisely. Convergence resolves conflicts of format, vocabulary and exchange, since these are coordination problems whose solution benefits every party. It does not resolve the conflicts of Section 4.5 instantiated in Section 13.6, because those arise between requirements that are each correctly applied and whose satisfaction is mutually exclusive. A common metadata model determines neither whether a retention obligation or an erasure request prevails, nor whether a legal permission or a community protocol governs release.

Convergence may in fact sharpen the second class of conflict, since where deposit is easy and cataloguing federated, decisions about disclosure, appraisal and controlled access are taken more often and faster. The mechanism is the one identified for automation in Section 12.5: moving an activity into infrastructure changes the cost of the record and not the location of the authority. The transitional reading is therefore accepted for the interoperability layer and rejected for the decision layer.

[EVIDENCE REQUIRED: repeated measurement of the ten decision areas of Table 1-1 at a fixed population of institutions across at least five years, recording whether the proportion with a named decision owner rises with infrastructural convergence — the observation supporting the transitional reading against the structural one.]

15.3 Integration by abstraction

Section 14.2 raises a counter-hypothesis that would deprive the book of its principal claim, and it is confronted here rather than left where it was

stated. The stable core may be stable because it is abstract, in that a requirement to name an accountable party is satisfiable anywhere precisely by saying nothing about what that party does. Integration across seven dimensions would then be an artefact of describing each at a level general enough for differences to disappear. The hypothesis is not answerable by inspection, since an abstract variable and a robust one look identical in their definitions.

Two considerations bear against it without refuting it. The first is that several conditioning variables are defined by an observable that a substantial proportion of settings would fail: escalation traversability requires a log with a recorded outcome and an interval, and the authoritative copy a registry designation that binds when copies diverge. A variable is not vacuous if the record it names is frequently absent, and the framework's own account predicts that these records are absent more often than present. The second is that the rules produce determinations some parties reject, as Section 13.6 documents for the conflict between legal permission and community protocol, whereas a vacuous structure produces no determination anyone objects to.

Both considerations are weaker than they look, since each rests on a property of the construction rather than on observation. The decisive test is the one Section 14.2 names: an abstract variable takes the same state in every setting, while a robust one separates settings that differ on independent evidence. Until the distribution of states for the invariant variables has been measured, the counter-hypothesis remains live and the integration claim is held open to that extent.

[EVIDENCE REQUIRED: measurement of the states of the invariant variables of Section 14.2 across at least 20 institutions differing on independent indicators of research data performance, reporting whether those variables separate the institutions or take the same state throughout — the test distinguishing robustness from vacuity.]

15.4 Whether the framework formalizes existing authority

A decision framework that assigns authority can improve decisions or merely record who already decides, and the second possibility is not remote. Each data activity is assigned to a single accountable party and each conflict routes to a named recipient, and in most institutions the parties available for such assignment already hold institutional power. Formalization then

converts an informal distribution into a documented one, adding legitimacy without redistributing anything. The objection is sound, and the framework contains no mechanism of redistribution.

What it supplies instead is asymmetric visibility. A decision recorded with its basis, its residual risk and the role accepting that risk is contestable in a way an unrecorded decision is not, because contesting it requires a different record rather than access to the occasion on which the decision was taken. The separation of the advisory from the deciding function operates in the same direction, preventing an orienting assessment from functioning as an approval nobody has taken responsibility for issuing. Neither property redistributes authority, and both make its exercise attributable.

The sharpest form of the objection concerns the value judgements embedded in the rule ordering itself. Rule 3 places a legal instrument above a community protocol by evidence grade, and that ordering encodes a position on a contested question rather than a neutral procedure (Carroll et al. 2020). The response available is to record the residual risk and the accepting role rather than present the ordering as technical, which makes the judgement visible and locates a party answerable for it. Visibility is a weaker remedy than neutrality, and the only one available to a structure that terminates in a determination.

15.5 Whether the instrument serves audit better than practice

Instruments that read records are attractive to audit functions, and the assessment properties developed here make the framework usable by an internal auditor with no stewardship competence at all. The risk is genuine rather than hypothetical, because the incentive structure of Section 14.5 applies to this framework as much as to the instruments it criticizes. An institution adopting it under audit pressure would rationally produce the registers the variables name while leaving the practices those registers trace unchanged. The result would reproduce, at a higher level of sophistication, the plan-counting failure diagnosed in Section 7.7.

The design contains one property that resists this outcome and one that invites it. Resistance comes from the level definitions, since levels 3 and 4 are reachable only through by-product records and through events whose timing the assessed party did not choose. Invitation comes from the breadth of the

variable set, because 55 variables across seven dimensions constitute a burden an institution discharges most cheaply by treating assessment as a documentation exercise. Which tendency prevails depends on whether the sampling frame is defined by the assessor, which is a condition of use rather than a property of the framework, and no claim that the framework improves decisions is available in any case.

[EVIDENCE REQUIRED: comparison of decision records and documentation volume at institutions adopting the framework for decision support against those adopting it for internal audit, to establish whether the mode of adoption determines whether practice or documentation changes.]

15.6 Established, supported and conjectural claims

The separation of received facts, syntheses and original results maintained throughout does not by itself tell a reader which results are secure. A second separation is therefore stated here, cutting across the first. Established claims follow from inspection of the construction or from the corpus as mapped; supported claims rest on reconstruction or argument rather than measurement; conjectural claims are predictions the framework generates and that nothing in the book tests. Table 15-2 assigns the principal claims to the three classes.

Table 15-2. The principal claims of the book, with the evidential class of each and what would move it to a stronger class.

Claim	Class	What would strengthen it
The ten decision areas of Table 1-1 are partially covered and unconnected	Established for the mapped corpus	An audit of institutional policies
The 55 variables satisfy the assignability, evidence and closure tests	Established by inspection	Nothing; the property is derivational
The conditioning order of Table 12-2 is asymmetric	Established by inspection	Covariation measurement across projects
Rules 1-5 leave the three conflict classes of Table 13-2 open	Established by construction	A register of authentic conflicts
The variables are expressible across four varying conditions	Supported by reconstruction	Empirical application at three to five organizations
Transferability is asymmetric between conditioning and conditioned dimensions	Supported by analytic judgement	Independent classification with an agreement coefficient
The thresholds between levels 2, 3 and 4 separate institutions that differ	Supported by design argument	Assessment against independent performance indicators
Recognition conditions whether stewardship effort is sustained	Conjectural	Before-and-after measurement against a comparison group
Framework use changes decisions rather than documentation	Conjectural	A longitudinal adoption study

The difference between the second and third classes is the one most easily lost in reading. A supported claim has survived contact with material the framework did not generate and could have failed that contact, whereas a conjectural claim is derived from the framework's own structure and has met nothing. Statements of the second kind warrant the hedge *indicates* and statements of the third *suggests*, and the hedging of Parts III and IV was assigned on that basis.

15.7 The standing of the evidence base

A framework derived from a mapped corpus and assessed against reconstructed settings supports weaker claims than one validated against institutional records, and the difference lies in what could have gone wrong. The

corpus was assembled by purposive selection rather than by a documented search protocol, and coded by a single analyst, so neither its completeness nor the reproducibility of its dimension boundaries is established (Sections 4.3 and 4.4). A reconstruction cannot contradict the framework as an institution can, since the material from which it is built is the material the framework was derived from. What the book reports is consequently a design study with an internal-consistency result and a specified validation programme rather than a validated instrument.

That standing determines the form of every claim the book is entitled to make. It licenses claims about what the framework determines, what it leaves undetermined and which variables are readable without interpretation, since these are properties of a construction. It does not license claims about the distribution of institutional practice, the reliability of any assessment or the effect of adoption, all of which are measurements. The markers throughout the text register that boundary, and their number reflects the field's evidence base as much as this study's design.

One consequence follows for how the result is used, and it is a consequence rather than a caution. A framework in this evidential position can be applied, contested and tested, and it cannot be cited as a warrant. An institution invoking it to certify its own arrangements would treat a construction as an authority, which is the error Section 3.6 identifies in the use of principle frameworks as compliance instruments.

What follows from these results for institutions, for the roles holding the decisions, for infrastructures and for the systems that recognize this work is the subject of the following chapter, in which every implication is derived from a stated result and marked where the result is itself not yet demonstrated.

PART VI

Implications

Chapter 16. Implications for Institutions, Data Stewards and Infrastructures

Part V establishes what the framework is, what it determines and what its claims currently rest on. This chapter derives what follows from those results for the parties whose decisions the framework describes, and the derivation is subject to one discipline throughout. Every implication below is stated together with the result it follows from, and where that result belongs to the supported or conjectural classes of Table 15-2 the implication inherits that status and is marked accordingly. An implication whose supporting result has not been demonstrated is a prediction about what would follow if the result held, and presenting it otherwise would convert a design study into guidance.

The bearing on the main research question is the final step of the argument rather than an appendix to it. Chapter 1 poses the question of integration in terms of decisions that can be defended at institutional and project level, and a framework whose consequences for those levels remain unstated leaves the question half answered. What follows therefore identifies the points at which the framework's structure, rather than its determinations, changes what an institution, a role or a service is doing. The determinations do not transfer, as Section 14.7 establishes, and no determination is offered here.

16.1 What the conditioning order implies for research institutions

The conditioning order of Table 12-2 places governance without a predecessor, and the closure test of Section 4.6 is satisfied for that reason alone. From this a consequence follows for the sequence in which institutional effort has effect rather than for the sequence in which it is usually expended. Effort placed on conditioned dimensions before the conditioning dimensions have determinate states produces variables whose states cannot be read, since a description profile, a preservation arrangement or a reproducibility requirement adopted before the data class, the authoritative copy and the

accountable party are fixed is a preference rather than a determination. The order is a property of the framework's dependency structure, and it holds independently of whether the covariation predicted in Section 12.2 is confirmed.

The most consequential instance concerns tools. Chapter 1 reports saturation rather than scarcity in the supply of stewardship tools and services, and Section 5.2 formulates the governance variables as properties of activities and records rather than of systems. A tool acquired before decision authority is assigned digitizes an undetermined procedure, with the result that the records it generates satisfy no variable, since a variable is read from a record that names a role, a date and an outcome. Acquisition of instrumentation therefore has assessable effects only after the assignment on which its records depend, which is a statement about the order of dependencies and not about procurement.

Two further consequences follow from findings of Chapter 13, and both concern what an institution holds after a determination has been reached. The incident gap of Section 13.4 places incident handling within governance rather than in a separate function, since it concerns who decides, within what interval and with what record, and the corresponding institutional consequence is that an incident register belongs with the escalation structure rather than with a security service that holds no authority over research data. The cost silence of the same section has a sharper consequence: the framework determines what is proportionate to a risk and not what is affordable, so that an institution can hold a correct determination it cannot implement. The party competent to accept an unimplemented determination as a standing residual risk is a governance question the framework raises and does not answer.

[EVIDENCE REQUIRED: comparison of institutions that assigned decision authority before acquiring stewardship tooling against institutions that reversed the order, measuring the proportion of variables readable at level 2 or above in each group, to establish whether the dependency order has the practical effect derived here.]

16.2 The data steward's decision rights derived from the framework

A role derived from a framework differs from a role derived from a job description in what determines its boundaries. Section 5.1 separates the advisory layer from the decision layer and treats the decision boundary as a variable with its own record, from which it follows that the data steward's rights are fixed by the thresholds at which advice becomes determination rather than by an enumeration of tasks. The steward's decision rights on this derivation cover the readings the framework requires — the state of a variable, the record from which it is read, whether a conflict has arisen and which rule applies — and stop at the point where a determination allocates risk. Determinations of that second kind belong to the accountable party for the activity, to the classification register owner or to the escalation recipient named in D1, as Table 12-3 assigns them.

The finding that the advisory and deciding functions cannot be merged without loss has a consequence easily read as a limitation and in fact the opposite. Section 12.5 states the same boundary for automated outputs: a suggested classification or a recommended access determination is advice, and the decision is its acceptance by a party with authority. A steward who both advises on a disclosure question and approves the release holds a position in which no external party can afterwards distinguish the assessment from the approval, which removes the contestability that Section 15.4 identifies as the framework's only remedy against formalized power. The separation therefore protects the steward's assessment as much as it constrains the steward's authority.

What the framework does place on the role is an escalation duty with a determinate discharge condition. Escalation traversability is defined in Table 5-2 by a named recipient, an expected interval and a logged outcome, so the duty is discharged by the routing and the record rather than by the result, and an escalation that receives no answer is a finding about the institution rather than a failure of the steward. Chapter 13 adds a second duty of the same form: where a variable's state cannot be read, the reason is recorded, distinguishing absence of a record from absence of access to one. Both duties are satisfiable by a role holding no authority at all, which is what makes them derivable from the framework rather than from an organizational position.

Table 16-1 sets out the principal implications of this chapter, the result from which each is derived and the evidential class the implication inherits.

Table 16-1. Implications derived in this chapter, with the supporting result and the class it carries under Table 15-2.

Implication	Derived from	Class inherited
Governance effort precedes effort in conditioned dimensions	Conditioning order, Table 12-2	Established by inspection
Tool acquisition has assessable effects only after authority is assigned	Sections 1.2 and 5.2	Supported
Incident handling belongs with the escalation structure	Incident gap, Section 13.4	Established
An institution can hold an unimplementable determination	Cost silence, Sections 13.4 and 14.6	Established
Steward rights end where a determination allocates risk	Decision boundary, Section 5.1	Established
The escalation duty is discharged by routing and record, not by outcome	Escalation traversability, Table 5-2	Established
Institutional assessment requires machine-actionable state from repositories	Level definitions, Section 12.4; custody, Section 14.3	Supported
Recognition arrangements condition whether criteria describe funded activity	Back-edge, Section 12.2; Section 11.6	Conjectural
Stewardship curricula addressing decision structure outperform tool curricula	Sections 1.2 and 5.1	Conjectural

16.3 What institutional assessment requires infrastructures to expose

The framework's assessment claim is concentrated between maturity levels 2 and 3, and level 3 is reachable only through a record produced as a by-product of the activity (Section 12.4). Where the activity takes place inside a repository, a metadata registry or a national service, that by-product record sits with the service rather than with the institution, and Section 14.3 states the consequence: the variable is then satisfied by attestation rather than by inspection, and its evidence grade falls. An institution cannot raise such a variable above level 1 by any effort of its own, since the ceiling is a property

of what the service exposes. The interfaces at issue are therefore a condition of institutional assessment rather than a convenience.

Five classes of exposure follow from variables already stated, and each is derived rather than proposed. The published object variable of Table 11-2 requires identifier, version, licence and contributor roles to be retrievable per deposit and not only displayed on a landing page. The access regime and egress control variables of Table 9-3 require the regime set on a deposit to be readable by the depositing institution, since an institution that cannot read the access state of its own holdings cannot assess its own restriction attributions. The integrity evidence variable requires dated fixity and migration events to be retrievable rather than asserted in a policy document, and the exit arrangement variable requires the succession provision and metadata export route to be stated in a form that a selection record can cite with a verification date.

Exposure class	Framework variable served	Maturity made attainable	Fallback where absent
Machine-readable metadata export	description; recognition	Level 3 — measured	manual registry extract
Access-decision log	protection; governance	Level 4 — verified	periodic attestation
Fixity and audit report	preservation	Level 4 — verified	sampled manual check
PID resolution and versioning	lifecycle; publication	Level 3 — measured	local identifier map
Plan-to-record linkage (maDMP)	planning	Level 5 — improving	clerical reconciliation

Interfaces repositories and registries must expose for institutional assessment to be possible

Fig. 16-1. The interface layer between institutional assessment and external services, showing for each of the five exposure classes the framework variable it serves, the maturity level it makes attainable and the fallback where exposure is absent. Author's own elaboration.

The fifth class concerns aggregation and is the one current services satisfy least often. Assessment at institutional level requires the states above to be retrievable for the institution's holdings as a set rather than one deposit at a time, since a sampling frame defined by the assessor presupposes an enumerable population (Section 14.5). Where no such query exists, the sampling frame is necessarily supplied by the institution being assessed,

and the first gaming vulnerability of Section 14.5 is structurally open. National metadata directories occupy a distinctive position in this respect, since registration requirements give them an enumerable population by construction (EOSC CZ 2025).

[EVIDENCE REQUIRED: a survey of the machine-actionable interfaces exposed by widely used generalist and disciplinary repositories against the five exposure classes derived here, reporting for each variable of Tables 9-3 and 11-2 the maximum maturity level an external institution can attain.]

16.4 The consequence of treating recognition as a conditioning variable

Section 12.2 admits one relation running against the conditioning order, in which recognition conditions every dimension because effort invisible to institutional assessment is not sustained. Treating recognition as a variable rather than as a policy aspiration produces a consequence for the standing of the framework's own criteria. Where a recognition path exists, the assessment criteria describe activity the institution has reason to sustain; where it does not, the same criteria describe an unfunded expectation, and a low profile is then a finding about the assessment system rather than about the institution's competence. The distinction is available only because recognition is carried as a variable with a record, namely the assessment rules and result registration records of Table 11-2.

A second consequence concerns the design of recognition arrangements themselves. Section 11.6 identifies the hazard that replacing an article count with a dataset count reproduces the mechanism the reform was intended to correct, and the framework's variables locate precisely where the substitution occurs: a deposit count reads the published object variable alone and touches none of the description, appraisal, access or preservation variables that determine what the deposit is worth. A recognition arrangement that reads several variables rather than one is on this derivation less susceptible to the substitution, which is a structural claim and not a measured one. The Czech arrangement from 2026 provides an instance in which registration in a national directory couples recognition to description quality, and its effects on behaviour are the subject of a marker in Chapter 11 that remains open (Úřad vlády České republiky 2025).

[EVIDENCE REQUIRED: comparison of recognition arrangements that read a single deposit count against arrangements coupling recognition to metadata or curation conditions, measuring description completeness and controlled-access provision in each, to test whether multi-variable recognition resists the substitution mechanism.]

16.5 What follows for the teaching of stewardship competence

The framework locates the failures it addresses in decision structure rather than in skill, and the location is derived rather than asserted. Chapter 1 reports saturation of tool supply alongside persistent failure, Section 5.1 attributes the recurrent category error to the merging of advisory and deciding functions, and Section 13.5 shows that the variables carrying the framework's assessment weight are read from records without interpretation. On that reading, a curriculum organized around tools and formats addresses the layer at which failure is least concentrated. The inference is strong at its first step and weaker at its second, since the concentration of failure in decision structure is a synthesis of the corpus rather than a measurement, and the implications below are marked as conjectural in Table 16-1 for that reason.

Three consequences follow for the content of stewardship education if the location holds. Competence in reading a record for what it evidences — distinguishing a plan from a realized measure, an assertion from a by-product record, an attestation from an inspection — is the operational core, since every variable state is such a reading. Competence in routing rather than resolving follows from the escalation duty of Section 16.2, since the ability to identify which authority a conflict belongs to is separable from the authority to settle it. And case material carrying conflicts that the rules do not settle, of the three classes in Table 13-2, teaches what a curriculum of worked examples cannot: that a defensible outcome is sometimes an escalation with a documented conflict rather than a determination.

A fourth consequence concerns assessment of learning and mirrors the framework's own distinction. A curriculum assessed by what participants can describe measures documentation of competence in the same way that plan counts measure documentation of stewardship, and the analogous outcome measure is whether a determination taken afterwards in the participant's own project is traceable to framework variables. The design for such

a measurement is stated in Section 4.6 and has not been carried out. Its absence is why the educational implications here are the weakest in the chapter.

[EVIDENCE REQUIRED: an application test with 12–20 participants from at least two disciplines, comparing a decision-structure curriculum against a tool-centred curriculum by blind rubric scoring of an artefact produced in the participant's own project after six to eight weeks, to establish whether the location of failure derived here holds.]

The implications derived above share one property that determines how they are to be read. Each follows from a stated result and carries that result's evidential class, so that the strongest of them are as secure as an inspection of the framework's structure and the weakest are predictions awaiting the measurements named in their markers. What the book has established, what it has left open and what the research programme now requires is the subject of the final chapter.

Chapter 17. Conclusion

17.1 The problem and the answer given

The problem this book addresses is not a deficiency of any instrument governing research data but the absence of a layer connecting them. FAIR is predicated of a dataset, TRUST of a repository, a data management plan of a project, data protection law of a processing operation and assessment reform of a career, and composition across these objects is defined by none of them (Wilkinson et al. 2016; Lin et al. 2020; European Parliament and Council 2016). Chapter 1 established that this partial coverage is systematic rather than accidental: the instruments are strongest where an object can be described in isolation and weakest where a decision has to be taken under competing constraints, which is where practice is densest. The question that follows is whether the partial requirements can be integrated into one framework of variables, decision rules and assessment criteria yielding decisions that can be defended at institutional and project level.

The answer the book gives is conditional and its conditions are stated rather than implied. Integration is achievable at the level of decision structure and is not achievable at the level of determinations, because the variables carrying authority and assessability hold across settings while the variables carrying the answers have their states largely fixed by legal regime, discipline, scale and infrastructure. A framework of this kind is therefore exportable as a structure and not as a set of answers, which is the finding of Section 14.7 and the reason the book contains no recommendations. What integration delivers, on this reading, is not agreement about what to do but a procedure that terminates in a determination attributable to a named party, with the residual risk recorded and the condition of revisiting stated.

17.2 The contribution in the order in which it was built

The argument was built in a sequence in which each step supplied what the next presupposed. Part I established the problem, characterized data stewardship as an institutional capability rather than a post or a service desk, and read the normative layer critically in order to determine precisely where FAIR, CARE and TRUST cease to determine anything (Carroll et al. 2020). Part II fixed the research design: a corpus of 23 domains stratified by

the function each source performs, a three-level grading of evidential strength, a three-pass coding procedure from which the dimensions were derived rather than imposed, and three internal consistency tests requiring that every variable attach to a role, name a record and route its conflicts to an authority. The separation of construction from validation was fixed at the same point, and it governs everything that follows.

Part III produced the framework itself in seven chapters. Six dimensions were developed as sets of variables, each defined as a property of a data activity and each naming the record from which its state is read: governance and accountability, lifecycle and provenance, planning as commitment, description and quality, protection and access, reproducibility, and preservation with publication and recognition. Chapter 12 then performed the assembly that converts a set of variables into a framework, adding the indicator level, establishing the conditioning order among the seven dimensions, stating five decision rules for competing requirements and redefining maturity as an observable state of a variable rather than an aspirational stage of an institution. The residual risk record with its six fields, including the role accepting the risk and the condition under which the decision is revisited, is the output that distinguishes a determination from a preference.

Part IV brought the construction into contact with material it did not generate, and separated two operations that framework development routinely merges. Four settings were reconstructed to vary scale, discipline, legal regime and infrastructure, and the reconstructions returned a mixed result reported without softening: the framework is expressible and its procedure terminates, and it is incomplete in respect of incidents, silent on cost and indeterminate for three characterizable classes of conflict. Chapter 14 established that sensitivity to setting is patterned rather than diffuse and follows the conditioning order, which is the asymmetry Section 17.1 restates. Part V positioned the result against maturity models, FAIR assessment instruments, repository certification and assessment reform, took the alternative explanations seriously and assigned every principal claim to one of three evidential classes.

Part VI derived the consequences of these results for institutions, for the roles holding the decisions, for infrastructures, for research assessment and for education, and derived them under one discipline: every implication carries the evidential class of the result it follows from. That discipline

produces an uneven set, in which the consequences of the conditioning order are as secure as an inspection of the framework's dependency structure while the consequences for teaching are predictions awaiting an application test. The unevenness is reported rather than levelled, because a chapter of implications presented at uniform strength would be indistinguishable from a set of recommendations.

The contribution claimed in Chapter 1 can accordingly be assessed against what was delivered. The integrated design exists and its internal properties are established by inspection. The operationalisation exists, with conditions, priority ordering, named authority and residual risk, and its external properties — discriminating power, reliability, separating power — are measurements that have not been made. The validation is a reconstruction accompanied by a stated boundary on transferability, which is less than validation and more than illustration.

17.3 The limits of the result

The limits of the result are of a different kind from the limits of the framework, and the distinction matters at the close. Chapter 14 bounds what the framework determines: not what is affordable, not whether a decision was right, not the scientific significance of a body of data and not the ethical acceptability of a purpose. What is at issue here is narrower and concerns the standing of the book's own claims. A framework derived from a purposively assembled corpus, coded by a single analyst and assessed against reconstructions supports claims about a construction and not claims about institutions, and no sentence in the preceding chapters was written to carry more than that.

Two consequences of this standing are worth restating in their own terms. The first is that the density of evidence markers throughout the text is a property of the field as much as of this study, since comparative measurement in research data stewardship remains rare enough that any framework of comparable scope would meet the same constraint. The second is that the counter-hypothesis of Section 14.2 — that the framework's stable core is stable because abstract rather than because robust — remains live, and the book's integration claim is held open to exactly that extent. A conclusion that closed it by assertion would demonstrate the failure the counter-hypothesis alleges.

One further limit concerns the use to which a result in this position can be put. A construction whose failure conditions are stated can be applied to a setting, contested where its determinations are unwelcome and tested where its predictions are specific, and it cannot serve as a warrant for the arrangements of any institution that adopts it. The distinction is the one Chapter 3 draws between a principle framework read as a design constraint and the same framework read as a compliance instrument, and it applies to this framework with no less force. Its adoption establishes that an institution has a structure and establishes nothing about the decisions taken within it.

17.4 The research programme the book leaves open

The evidence markers distributed through the preceding chapters are not a list of caveats but a specification of the empirical programme this framework requires, and they fall into four groups. The first group is descriptive and would establish the state of the object: an audit of institutional research data policies across a defined population, recording for each decision area which instrument is cited, whether a decision owner is named and whether an escalation path is stated. The second group is measurement of the instrument itself, comprising empirical application at three to five organizations with negotiated access, an assessor-defined sampling frame, independent coding of all 55 variables and a reported distribution of maturity levels. Both groups convert reconstructions into findings about institutions, which is the step Part IV declines to take.

Inter-rater agreement occupies the third group and holds a position prior to the others. A maturity scale on which two competent assessors assign different levels to the same variable is not a measurement, and comparisons between institutions or across time then have no interpretation. Section 14.4 predicts a structure in the disagreement rather than a uniform level of it — high agreement at the boundaries between levels 0, 1 and 2, lower between 2, 3 and 4, and lowest for the interpretive variables of Section 13.5 — and a study reporting a single coefficient across all variables would conceal exactly that structure. The design is stated: at least three assessors, the same project material, agreement reported per variable and per level boundary against a threshold fixed in advance, with variables below it reported as adjudicated rather than assessed.

The fourth group is longitudinal and carries the question on which the framework's practical value ultimately turns. Adoption of any instrument that determines resources changes the behaviour it measures, and the change most cheaply produced is documentary. Whether framework adoption alters the decisions institutions take, or only the records they keep, is answerable by observing variable states, conflict registers and realized measures at intervals across several years, with institutions adopting the framework for decision support distinguished from those adopting it under audit pressure. The prediction the framework generates is specific enough to fail: if adoption changes documentation alone, the gap between self-assessed and assessor-verified profiles will not close, and level 4 observations will remain concentrated in scheduled events.

Two further questions sit outside these groups and would extend rather than test the framework. The incident variable identified in Section 13.4 requires formulation and assessment as part of the governance dimension, and the cost silence of the same section marks a boundary at which a framework of this type meets a question it was not built to answer, namely how a determination that is proportionate to a risk and unaffordable to the institution holding it is to be disposed of. Whether capacity belongs in the framework as a variable with a record, or remains a context condition against which determinations are assessed, is an open design question rather than a defect to be repaired by addition.

[EVIDENCE REQUIRED: a repeated-measures study across at least five years at institutions adopting the framework and at matched institutions that do not, recording variable states, conflict registers, realized measures and the divergence between self-assessed and verified profiles, to establish whether adoption changes decisions or documentation.]

17.5 What the result amounts to

The condition on which everything in this book depends is stated most economically as a property of records. A variable is usable in assessment when its state is read from a record that the activity produced for its own reasons, because such a record can contradict the account an institution gives of itself, and no assertion, policy or plan has that property. This is why the framework counts nothing that an institution can produce by describing itself well, why maturity level 3 is separated from level 2 by a by-product

record and level 4 by an event the institution did not schedule, and why plan counts are treated throughout as coverage statistics with no outcome interpretation. The distinction between documentation and practice is not a theme of the book but its operative criterion.

The second thing the result amounts to is a location for the field's scarcity. Several hundred tools and services address research data stewardship, and their number continues to grow, while the decisions that determine what data are worth after a project ends remain unassigned in a large proportion of settings. Instrumentation is abundant, and what is scarce is a structure into which instrumentation can be placed: an assigned authority, a threshold at which advice becomes determination, an escalation that terminates and a record of the risk somebody accepted. A framework that supplies that structure does not make decisions better, and it makes them attributable, which is the weaker property and the one on which any subsequent improvement depends.

The third and least comfortable point concerns what the framework does when requirements genuinely conflict. It does not dissolve the conflict, and in three identified classes it returns no determination at all, escalating instead with the conflict documented. That behaviour is not a shortfall of the construction but the honest terminus of any structure that respects the fact that a retention obligation and an erasure request, or a legal permission and a community protocol, are both correctly applied. Frameworks that produce a determination in such cases conceal a value judgement inside a procedure, and the contribution of this one is that the judgement, its accepting role and the risk it leaves are on the record where they can be contested.

References

Aksenova, Anna, Anoop Johny, Tim Adams, Phil Gribbon, Marc Jacobs, and Martin Hofmann-Apitius. 2024. "Current State of Data Stewardship Tools in Life Science." *Frontiers in Big Data* 7: 1428568. <https://doi.org/10.3389/fdata.2024.1428568>.

Anscombe, F. J. 1973. "Graphs in Statistical Analysis." *The American Statistician* 27 (1): 17–21. <https://doi.org/10.1080/00031305.1973.10478966>.

Barker, Michelle, Neil P. Chue Hong, Daniel S. Katz, Anna-Lena Lamprecht, Carlos Martinez-Ortiz, Fotis Psomopoulos, Jennifer Harrow, Leyla Jael Castro, Morane Gruenpeter, Paula Andrea Martinez, and Tom Honeyman. 2022. "Introducing the FAIR Principles for Research Software." *Scientific Data* 9: 622. <https://doi.org/10.1038/s41597-022-01710-x>.

Carroll, Stephanie Russo, Ibrahim Garba, Oscar L. Figueroa-Rodriguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons, Kay Raseroka, Desi Rodriguez-Lonebear, Robyn Rowe, Rodrigo Sara, Jennifer D. Walker, Jane Anderson, and Maui Hudson. 2020. "The CARE Principles for Indigenous Data Governance." *Data Science Journal* 19 (1): 43. <https://doi.org/10.5334/dsj-2020-043>.

CoARA (Coalition for Advancing Research Assessment). 2022. *Agreement on Reforming Research Assessment*. Brussels: CoARA. <https://coara.eu/agreement/the-agreement-full-text/>.

Consultative Committee for Space Data Systems. 2012. *Reference Model for an Open Archival Information System (OAIS)*. Recommended Practice CCSDS 650.0-M-2 (Magenta Book). Washington, DC: CCSDS Secretariat. <https://public.ccsds.org/pubs/650x0m2.pdf>. Superseded for standardization purposes by ISO 14721:2025.

CoreTrustSeal Standards and Certification Board. 2022. *CoreTrustSeal Requirements 2023–2025* (Trustworthy Data Repositories Requirements). Version 01.00. Geneva: Zenodo. <https://doi.org/10.5281/zenodo.7051012>.

CoreTrustSeal Standards and Certification Board. 2025. *CoreTrustSeal Requirements 2026–2028*. Version 01.00. Geneva: Zenodo. <https://doi.org/10.5281/zenodo.17660463>.

Cousijn, Helena, Patricia Feeney, Daniella Lowenberg, Eleonora Presani, and Natasha Simons. 2019. "Bringing Citations and Usage Metrics Together

to Make Data Count." *Data Science Journal* 18 (1): 9. <https://doi.org/10.5334/dsj-2019-009>.

Cox, Andrew M., Mary Anne Kennan, Liz Lyon, Stephen Pinfield, and Laura Sbaifi. 2019. "Maturing Research Data Services and the Transformation of Academic Libraries." *Journal of Documentation* 75 (6): 1432–1462. <https://doi.org/10.1108/JD-12-2018-0211>.

Crowston, Kevin, and Jian Qin. 2011. "A Capability Maturity Model for Scientific Data Management: Evidence from the Literature." *Proceedings of the American Society for Information Science and Technology* 48 (1): 1–9. <https://doi.org/10.1002/meet.2011.14504801036>.

DataCite Metadata Working Group. 2024. *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs*. Version 4.6. Hannover: DataCite e.V. <https://doi.org/10.14454/mzv1-5b55>.

DataCite Metadata Working Group. 2026. *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs*. Version 4.7. Hannover: DataCite e.V. <https://doi.org/10.14454/qdd3-ps68>.

DORA. 2013. *San Francisco Declaration on Research Assessment*. <https://sfdora.org/read/>.

EOSC Association. 2024. *Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC)*. Version 1.3. Brussels: EOSC Association. <https://doi.org/10.5281/zenodo.17582648>.

EOSC CZ. 2025. "Datové sady jako nový druh výzkumného výsledku" [Datasets as a new type of research output; result type T "Digitální kolekce dat"]. EOSC CZ, December 16, 2025. <https://www.eosc.cz/novinky-akce/clanky/vysledek-t>.

EOSC CZ. 2026. "Národní metadatový adresář (NMA)" [National Metadata Directory]. EOSC CZ. Accessed August 13, 2026. <https://www.eosc.cz/sluzby/vyhledavani/nma>.

European Data Protection Board. 2026. *Guidelines 02/2026 on Anonymisation*. Version for public consultation. Brussels: EDPB. https://www.edpb.europa.eu/public-consultations/guidelines-022026-on-anonymisation_en.

European Parliament and Council. 1996. "Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal

Protection of Databases." *Official Journal of the European Communities* L 77 (27 March 1996): 20–28. <https://eur-lex.europa.eu/eli/dir/1996/9/oj>.

European Parliament and Council. 2016. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)." *Official Journal of the European Union* L 119 (4 May 2016): 1–88. <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.

European Parliament and Council. 2024. "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)." *Official Journal of the European Union* L 2024/1689 (12 July 2024). <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

Federer, Lisa. 2018. "Defining Data Librarianship: A Survey of Competencies, Skills, and Training." *Journal of the Medical Library Association* 106 (3): 294–303. <https://doi.org/10.5195/jmla.2018.306>.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2021. "Datasheets for Datasets." *Communications of the ACM* 64 (12): 86–92. <https://doi.org/10.1145/3458723>.

Gelman, Andrew, and Eric Loken. 2014. "The Statistical Crisis in Science: Data-Dependent Analysis—a 'Garden of Forking Paths'—Explains Why Many Statistically Significant Comparisons Don't Hold Up." *American Scientist* 102 (6): 460–465. <https://doi.org/10.1511/2014.111.460>.

Higgins, Sarah. 2008. "The DCC Curation Lifecycle Model." *International Journal of Digital Curation* 3 (1): 134–140. <https://doi.org/10.2218/ijdc.v3i1.48>.

International Organization for Standardization. 2025. *ISO 14721:2025. Space Data System Practices — Reference Model for an Open Archival Information System (OAIS)*. Geneva: ISO. <https://www.iso.org/standard/87471.html>.

ISO/IEC (International Organization for Standardization and International Electrotechnical Commission). 2025. *ISO/IEC 18670:2025. Information Technology — SoftWare Hash IDentifier (SWHID) Specification V1.2*. Geneva: ISO. <https://www.iso.org/standard/89985.html>.

Jetten, Mijke, Marjan Grootveld, Annemie Mordant, Mascha Jansen, Margreet Bloemers, Margriet Miedema, and Celia W. G. van Gelder. 2021. *Pro-*

Professionalising Data Stewardship in the Netherlands: Competences, Training and Education. Dutch Roadmap towards National Implementation of FAIR Data Stewardship. Version 1.1. Amsterdam: Dutch National Programme Open Science. <https://doi.org/10.5281/zenodo.4320504>.

Library of Congress. 2015. *PREMIS Data Dictionary for Preservation Metadata*. Version 3.0. Prepared by the PREMIS Editorial Committee. Washington, DC: Library of Congress. <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>.

Liggins, Libby, Māui Hudson, and Jane Anderson. 2021. "Creating Space for Indigenous Perspectives on Access and Benefit-Sharing: Encouraging Researcher Use of the Local Contexts Notices." *Molecular Ecology* 30 (11): 2477–2482. <https://doi.org/10.1111/mec.15918>.

Lin, Dawei, Jonathan Crabtree, Ingrid Dillo, Robert R. Downs, Rorie Edmunds, David Giarretta, Marisa De Giusti, Hervé L'Hours, Wim Hugo, Reyna Jenkyns, Varsha Khodiyar, Maryann E. Martone, Mustapha Mokrane, Vivek Navale, Jonathan Petters, Barbara Sierman, Dina V. Sokolova, Martina Stockhause, and John Westbrook. 2020. "The TRUST Principles for Digital Repositories." *Scientific Data* 7: 144. <https://doi.org/10.1038/s41597-020-0486-7>.

Local Contexts. n.d. "About the Labels." Accessed August 13, 2026. <https://localcontexts.org/labels/about-the-labels/>.

Miksa, Tomasz, Paul Walk, and Peter Neish. 2020. *RDA DMP Common Standard for Machine-actionable Data Management Plans*. Version 1.1. Research Data Alliance, DMP Common Standards Working Group. <https://doi.org/10.15497/RDA00039>.

Miksa, Tomasz, Stephanie Simms, Daniel Mietchen, and Sarah Jones. 2019. "Ten Principles for Machine-Actionable Data Management Plans." *PLOS Computational Biology* 15 (3): e1006750. <https://doi.org/10.1371/journal.pcbi.1006750>.

Mitchell, Margaret, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. "Model Cards for Model Reporting." In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*, 220–229. New York: Association for Computing Machinery. <https://doi.org/10.1145/3287560.3287596>.

National Academies of Sciences, Engineering, and Medicine. 2019. *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25303>.

National Digital Stewardship Alliance. 2026. *Levels of Digital Preservation*. Version 2.1. NDSA. <https://www.ndsa.org/publications/levels-of-digital-preservation/> (materials at <https://osf.io/rcs38>).

National Institute of Standards and Technology. 2010. *Contingency Planning Guide for Federal Information Systems*. NIST Special Publication 800-34, Rev. 1, by Marianne Swanson, Pauline Bowen, Amy Wohl Phillips, Dean Gallup, and David Lynes. Gaithersburg, MD: NIST. <https://doi.org/10.6028/NIST.SP.800-34r1>.

National Institute of Standards and Technology. 2023. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1. Gaithersburg, MD: NIST. <https://doi.org/10.6028/NIST.AI.100-1>.

National Institute of Standards and Technology. 2024. *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*. NIST AI 600-1. Gaithersburg, MD: NIST. <https://doi.org/10.6028/NIST.AI.600-1>.

Nosek, Brian A., Charles R. Ebersole, Alexander C. DeHaven, and David T. Mellor. 2018. "The Preregistration Revolution." *Proceedings of the National Academy of Sciences* 115 (11): 2600–2606. <https://doi.org/10.1073/pnas.1708274114>.

Paulk, Mark C., Bill Curtis, Mary Beth Chrissis, and Charles V. Weber. 1993. "Capability Maturity Model, Version 1.1." *IEEE Software* 10 (4): 18–27. <https://doi.org/10.1109/52.219617>.

Rauber, Andreas, Ari Asmi, Dieter van Uytvanck, and Stefan Pröll. 2015. *Data Citation of Evolving Data: Recommendations of the Working Group on Data Citation (WGDC)*. Research Data Alliance. <https://doi.org/10.15497/RDA00016>.

RDA FAIR Data Maturity Model Working Group. 2020. *FAIR Data Maturity Model: Specification and Guidelines*. Research Data Alliance. <https://doi.org/10.15497/rda00050>.

Sansone, Susanna-Assunta, Peter McQuilton, Philippe Rocca-Serra, Alejandra Gonzalez-Beltran, Massimiliano Izzo, Allyson L. Lister, and Milo Thurston. 2019. "FAIRsharing as a Community Approach to Standards, Repositories and Policies." *Nature Biotechnology* 37 (4): 358–367. <https://doi.org/10.1038/s41587-019-0080-8>.

Science Europe. 2021. *Practical Guide to the International Alignment of Research Data Management — Extended Edition*. Brussels: Science Europe. <https://doi.org/10.5281/zenodo.4915861>.

Seawright, Jason, and John Gerring. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61 (2): 294–308. <https://doi.org/10.1177/1065912907313077>.

Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." *Psychological Science* 22 (11): 1359–1366. <https://doi.org/10.1177/0956797611417632>.

Skills4EOSC. 2025. *Data Steward: Minimum Viable Skills Profile*. Version 3, by Karsten Kryger Hansen, Angus Whyte, Laurence Horton, Bojana Koteska, Päivi Rauste, and Nida van Leersum. Skills4EOSC project. Zenodo. <https://doi.org/10.5281/zenodo.16760455>.

Smith, Arfon M., Daniel S. Katz, Kyle E. Niemeyer, and the FORCE11 Software Citation Working Group. 2016. "Software Citation Principles." *PeerJ Computer Science* 2: e86. <https://doi.org/10.7717/peerj-cs.86>.

Soiland-Reyes, Stian, Peter Sefton, Mercè Crosas, Leyla Jael Castro, Frederik Coppens, José M. Fernández, Daniel Garijo, Björn Grüning, Marco La Rosa, Simone Leo, Eoghan Ó Carragáin, Marc Portier, Ana Trisovic, RO-Crate Community, Paul Groth, and Carole Goble. 2022. "Packaging Research Artefacts with RO-Crate." *Data Science* 5 (2): 97–138. <https://doi.org/10.3233/DS-210053>.

Steegen, Sara, Francis Tuerlinckx, Andrew Gelman, and Wolf Vanpaemel. 2016. "Increasing Transparency through a Multiverse Analysis." *Perspectives on Psychological Science* 11 (5): 702–712. <https://doi.org/10.1177/1745691616658637>.

UK Data Service. 2025. "What Is the Five Safes Framework?" UK Data Service. Accessed August 13, 2026. <https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/>.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3: 160018. <https://doi.org/10.1038/sdata.2016.18>.

World Wide Web Consortium. 2009. *SKOS Simple Knowledge Organization System Reference*. W3C Recommendation, 18 August 2009, edited by Alistair Miles and Sean Bechhofer. <https://www.w3.org/TR/skos-reference/>.

World Wide Web Consortium. 2013. *PROV-O: The PROV Ontology*. W3C Recommendation, 30 April 2013, edited by Timothy Lebo, Satya Sahoo, and Deborah McGuinness. <https://www.w3.org/TR/prov-o/>.

World Wide Web Consortium. 2024. *Data Catalog Vocabulary (DCAT) — Version 3*. W3C Recommendation, 22 August 2024, edited by Riccardo Albertoni, David Browning, Simon Cox, Alejandra Gonzalez-Beltran, Andrea Perego, and Peter Winstanley. <https://www.w3.org/TR/vocab-dcat-3/>.

Wu, Mingfang, Stephen M. Richard, Chantelle Verhey, Leyla Jael Castro, Baptiste Cecconi, and Nick Juty. 2023. "An Analysis of Crosswalks from Research Data Schemas to Schema.org." *Data Intelligence* 5 (1): 100–121. https://doi.org/10.1162/dint_a_00186.

Úřad vlády České republiky, Rada pro výzkum, vývoj a inovace. 2025. *Metodika hodnocení výzkumných organizací (Metodika 2025+)* [Methodology for the Evaluation of Research Organisations 2025+]. Approved by Government Resolution No. 458 of 18 June 2025. Praha: Úřad vlády ČR. <https://vyzkum.gov.cz/FrontClanek.aspx?idsekce=1064052>.

Verification notes

Year corrections (manuscript key → verified):

- CoreTrustSeal (2023) → **2022**. The *CoreTrustSeal Requirements 2023–2025* were published 5 September 2022 (v01.00, DOI 10.5281/zenodo.7051012); "2023–2025" is the validity period, not the publication year.
- EOSC CZ (2026) → **2025** for the article on result type T ("Datové sady jako nový druh výzkumného výsledku," published 16 December 2025). The NMA service page itself is undated and is cited as EOSC CZ 2026 with an access date (the new NMA version launched 19 January 2026), so the key 2026 remains usable only for the NMA entry.
- European Commission (2016) → author corrected to **European Parliament and Council** (GDPR, Regulation (EU) 2016/679); year 2016 is correct.
- W3C (2023) → **2024** for DCAT 3: it became a W3C Recommendation on 22 August 2024 (in 2023 it was only a working draft/candidate).

- NIST backup/contingency (SP 800-34 Rev. 1) → **2010**, not 2023/2024. The NIST (2023) key matches AI 100-1 (AI RMF 1.0, January 2023) and NIST (2024) matches AI 600-1 (Generative AI Profile, July 2024); if the manuscript's backup/contingency citation points to SP 800-34, it must carry the year 2010 (entry included above).
- EDPB (2026): verified as *Guidelines 02/2026 on Anonymisation* — currently the public-consultation version; if a final adopted version appears before the book goes to press, update the entry. (The companion *Guidelines 01/2025 on Pseudonymisation* exist should the manuscript also need a 2025 EDPB key.)
- NDSA (2026): verified — Version 2.1 of the Levels of Digital Preservation was announced 23 March 2026.
- ISO 14721:2025 (OAIS revision): verified as published; the CCSDS (2012) Magenta Book CCSDS 650.0-M-2 remains the correct source for the 2012 key.
- EOSC Association (2024): verified as SRIA v1.3 (dated 31 October 2024, deposited on Zenodo 11 December 2024); the formal corporate author on Zenodo is the "Horizon Europe Co-programmed Partnership for the European Open Science Cloud," commonly cited under EOSC Association.
- DataCite Metadata Working Group (2024) = Schema 4.6 (released 5 December 2024) and (2026) = Schema 4.7 (released 3 March 2026) — both verified with DOIs.
- Skills4EOSC (2025): verified as *Data Steward: Minimum Viable Skills Profile*, v3, 7 August 2025 (extract from the annex to deliverable D2.6).
- UK Data Service (2025): the Five Safes page is an undated webpage; cited with the manuscript's year and an access date.
- Federer (2018): verified as the JMLA data-librarianship competencies survey (not the PLOS ONE data-reuse study, which matches the "data librarian / data services" description less well).

UNVERIFIED entries: none.

All [CITATION REQUIRED] markers in the manuscript were resolved on 13 August 2026; fourteen entries above were added in that pass and verified against Crossref, Zenodo, ISO and publisher records.

Glossary

Definitions are working definitions for the purposes of this book. Where a term is defined by a legal instrument or a ratified standard, the primary source governs.

Accountability. The property of a data activity that a single named role answers for its outcome. Distinguished from responsibility, which designates who performs the activity.

Anonymization. Processing after which a person is no longer identifiable by means reasonably likely to be used. Anonymized data fall outside the scope of data protection law; the singling-out, linkability and inference tests govern whether the threshold is met.

Appraisal. The assessment by which objects are selected for preservation or scheduled for destruction. An appraisal decision has an owner and a recorded ground.

Concept identifier. A persistent identifier that resolves to all versions of an object, as distinct from the identifier of one specific version.

Controlled access. The access regime in which metadata are public while access to data requires an assessed application. The regime for data that are neither open nor closed.

Controlled vocabulary. An enumerated set of permitted values for a metadata element, maintained by an identifiable authority.

Data class. The protection category assigned to a data object, determining its permitted storage, access and sharing regime.

Data management plan (DMP). A versioned record of the decisions taken about the creation, organization, protection, sharing and preservation of research data. In this book a plan is treated as a set of verifiable commitments rather than as an administrative deliverable.

Data steward. The professional role that coordinates responsible data management, advises researchers and escalates risk. The framework requires the advisory function of the role to remain separate from deciding authority.

Data stewardship. The institutional capability by which research data are governed across their lifecycle; the practice this book models. Research data management denotes the operational activity within it.

Dimension. The highest level of the framework hierarchy: one of the seven areas — governance, lifecycle and provenance, planning, description and quality, protection, reproducibility, preservation and recognition — across which stewardship decisions are taken.

Escalation. The recorded transfer of a decision from the role that encountered it to the authority entitled to take it, triggered by a stated threshold.

Evidence grade. The classification of a source statement by evidential strength: legal or normative requirement; ratified standard or certified scheme; case evidence, consensus document or emerging practice.

FAIR. Findable, Accessible, Interoperable, Reusable — target properties of digital research objects and their metadata. FAIR does not entail open.

Fixity. The property that the bit-level content of an object has not changed, verified by checksums.

Indicator. The lowest level of the framework hierarchy: a record, an operation on it, a sampling unit and an observation date, by which the state of a variable is observed.

Machine-actionable. Structured such that another system can read the content and act on it. Stronger than machine-readable, which requires only that the content be parsed.

Maturity level. One of five observable states of a framework variable or dimension, distinguished from the level below by a stated observable rather than by an aspirational label.

Metadata. Structured statements describing the identity, meaning, origin, technical properties, rights or preservation of an object.

Persistent identifier (PID). An identifier whose resolution to its object is maintained independently of changes to location or custody.

Preservation. The managed activity by which selected objects remain usable over time. Distinguished from backup, which serves operational recovery, and from publication, which serves access.

Provenance. The recorded history of the agents, activities and entities that produced an object in its present form. In the framework, the property that makes quality, reproducibility and rights assessable.

Pseudonymization. The replacement of direct identifiers such that attribution to a person requires separately held information. Pseudonymized data remain personal data.

Reproducibility. Obtaining consistent results from the same data, the same code and the same analytical conditions. Distinguished from replicability, which concerns new data or a new study.

Residual risk record. The six-field record produced by every application of the decision rules: the variables in conflict, the option adopted, the option foregone, the risk left in place, the role accepting it and the condition for revisiting the decision.

Trusted research environment. A controlled technical and organizational setting for analysis of sensitive data, with assessed people, projects and outputs.

Variable. The unit of the framework between component and indicator: a named property of institutional or project practice whose state can be observed and compared.

Versioning threshold. The stated change of files, method or meaning at which a working revision becomes a new version with its own persistent identifier.

Appendix A. Index of Outstanding Evidence

This appendix indexes every point at which the manuscript states an empirical requirement that the underlying research programme has not yet met. It is generated from the markers embedded in the chapters and is therefore complete by construction. Each entry names a measurement or collection that must be carried out before the associated claim can be stated as a result. The citation markers present in earlier drafts were resolved against verified sources on 13 August 2026, and the eight figures have been drawn; both are therefore no longer indexed here.

Table A-1. Outstanding evidence requirements by chapter (45 entries).

Ch.	Requirement
1	an audit of institutional research data policies across a defined population of research organizations, recording for each of the ten decision areas in Table 1-1 which instrument is cited, whether a decision owner is named and whether an escalation path is stated — supporting the claim that coverage is partial and unconnected rather than merely uneven
1	paired comparison of data management plan content at submission against realized practice at project close, for a sample of projects, to establish the divergence between documentary compliance and realized practice
2	a multi-institution comparison holding institutional size and disciplinary composition constant, measuring outcome indicators (time to resolution, metadata completeness at deposit, proportion of projects with an executed preservation plan) across centralized, distributed and federated arrangements, to test whether organizational model predicts outcome once resources are controlled for
2	retrospective tracing of a sample of completed projects to determine whether a named role held custody of data and access decisions 12 months after project end, supporting the claim that continuity of accountability, not its initial assignment, is the binding constraint
3	agreement analysis between automated FAIR assessment scores and independent expert audit of fitness for use across a stratified sample of datasets, reporting the proportion of high-scoring datasets judged unusable for their declared purpose, to support the claim that automated scores do not measure reusability
3	a documented set of conflict cases from institutional practice, coded by conflict type, resolution adopted, deciding authority and residual risk recorded, to establish the distribution of conflict types and to test whether resolutions are consistent within an institution
4	institution-level data on the actual execution of stated data policies — service response times, escalation frequency, proportion of projects with an implemented preservation arrangement — to establish whether documentary commitment predicts operational practice.
4	a documented search protocol for the applied stratum — databases and registries queried, query strings, date of execution, number of records screened and excluded at each step — sufficient for an independent researcher to reconstruct the corpus. The present corpus was assembled by purposive selection against the criteria in Table 4-1, and its completeness is therefore not established.
4	independent recoding of a random sample of at least 100 extracted requirement statements by a second analyst, with a reported agreement coefficient, to establish whether the dimension boundaries in Table 4-2 are reproducible rather than idiosyncratic.
4	

Ch. Requirement	
	execution of the expert construct review — number and disciplinary composition of reviewers, the claims reviewed, and the disposition of each critical comment — to support any claim that the framework is technically correct.
4	execution of the application test with 12–20 participants from at least two disciplines, including baseline assessment, blind rubric scoring of the produced artefact and a follow-up check of transfer into the participant's own project after six to eight weeks, to support any claim that the framework changes practice rather than knowledge.
5	comparative outcome data across institutions operating different organizational arrangements — time from request to resolution, proportion of requests escalated, proportion of escalations closed with a recorded decision — to test whether arrangement type predicts governance performance once escalation quality is controlled.
5	longitudinal data on datasets whose responsible person has left the institution — proportion with a documented successor, proportion for which subsequent access requests were answered, and time to first unanswerable request — to establish the practical decay rate of unassigned accountability.
5	paired institutional data on plan-based indicators and outcome indicators for the same set of projects, to test the study's claim that the two diverge and that plan counts overstate the state of data.
6	for a sample of published analytical datasets, a reconstruction attempt from the deposited provenance record alone, measuring the proportion of transformation steps recoverable and the proportion of unrecoverable steps that were performed manually — supporting the claim that manual transitions account for most of the recorded-to-actual gap.
6	repeated re-execution tests on a defined set of deposited computational datasets at fixed intervals after deposit, recording the proportion of successful regenerations and the causes of failure — establishing how quickly the regenerability assumption underlying selective preservation decays.
7	comparison of plan items produced under text, structured-form and knowledge-based tooling for comparable projects, scored against a specificity and feasibility rubric, to test whether tooling affects the specificity of commitments or only the completeness of fields.
7	longitudinal record of plan versions across a cohort of projects, coded for revision trigger, approval and justification, to test the study's claim that event-triggered revision correlates with realized measures more strongly than plan completeness at submission does.
8	paired comparison of description completeness and of reuser clarification requests between datasets described during collection and datasets described at deposit, controlling for discipline, data complexity and repository, to test whether the timing of description generation predicts reuse cost.

Ch.	Requirement
8	measurement of reuser effort — time to first analysis, number of clarification requests, proportion of abandoned reuse attempts — as a function of application profile completeness, establishing the magnitude of the cost transferred by incomplete description.
9	re-identification testing applied to a sample of deposited research datasets declared anonymous, reporting the proportion failing the singling-out, linkability and inference tests, to establish the frequency with which direct-identifier removal is treated as sufficient
9	audit of research projects recording whether recovery objectives are stated per data class and whether a full restore test was performed and documented within a defined period, to establish the divergence between backup provision and demonstrated recoverability
10	comparison of pre-registered analytical plans against the corresponding published analyses for a sample of studies, recording the proportion of deviations that are reported as such, to establish whether pre-specification functions as a separator of claim types in practice
11	institutional audit recording, for a defined population of completed projects, whether a retention period was stated, whether an appraisal decision was documented and what proportion of project data remained on unmanaged storage after project close
11	comparative measurement of reuse, citation and long-term availability for datasets deposited in disciplinary against certified generalist repositories, controlling for field, to establish whether the trade-off between disciplinary fit and certified trustworthiness has measurable consequences
11	measurement of deposit volume, metadata completeness and controlled-access provision at Czech research organizations before and after the introduction of the data collection result type, against a comparison group without such recognition, to test whether recognition changes stewardship behaviour rather than reporting behaviour
12	cross-sectional measurement of variable states within a set of projects, testing whether conditioned variables covary with their conditioning variables as Table 12-2 asserts — the observation that would falsify the ordering.
12	application of Rules 1–5 to a fixed set of documented conflict cases by several assessors working independently, reporting the proportion receiving the same determination and the same residual risk attribution.
12	assessment of the 55 variables at institutions selected to differ on independent indicators of research data performance, testing whether the 2-to-3 and 3-to-4 thresholds separate institutions that independent evidence distinguishes.
12	coding of requirement statements in the artificial intelligence governance and documentation corpus against the 55 variables, reporting the proportion assignable to an existing variable and characterizing any residue — the test of whether Table 12-5 is exhaustive rather than illustrative.

Ch.	Requirement
13	empirical application at three to five research organizations selected on the criteria of Table 13-1, comprising negotiated access to project records, a sampling frame of data activities defined by the assessor rather than by the institution, independent coding of all 55 variables by two assessors, a reported agreement coefficient, the distribution of maturity levels and a register of conflicts with the rule applied and the residual risk recorded — the study that would convert the reconstructions below into findings about institutions.
13	application of the framework to an active multi-partner project across its lifetime, with variable states recorded at three points and with the paired comparison of plan commitments against realized records performed at project close, to establish whether the framework's determinations remain stable as a project's conditions change.
13	independent assignment of states and maturity levels for all 55 variables by at least three assessors working from the same project material, with agreement reported separately for the two groups identified above, to establish which variables are usable in assessment and which require adjudication.
13	a register of authentic conflict cases collected from research projects, coded for class and for the rule that settled them, reporting the proportion falling into the three classes of Table 13-2 — the measurement on which any claim about the coverage of the rules depends.
14	independent classification of all 55 variables into the three classes above by two analysts working from the definitions alone, with an agreement coefficient reported, together with the distribution of states for the invariant variables across settings — the observation that separates robustness from abstraction.
14	an agreement study with at least three assessors independently assigning maturity levels to all 55 variables for the same project material, reporting agreement per variable and per level boundary against a threshold stated in advance, with variables falling below it reported as adjudicated rather than assessed.
14	paired self-assessment and assessor verification of the same variables at the same organizations, reporting the direction and magnitude of divergence by variable class, to establish how far self-reported maturity can be relied upon and where independent verification is indispensable.
15	repeated measurement of the ten decision areas of Table 1-1 at a fixed population of institutions across at least five years, recording whether the proportion with a named decision owner rises with infrastructural convergence — the observation supporting the transitional reading against the structural one.
15	measurement of the states of the invariant variables of Section 14.2 across at least 20 institutions differing on independent indicators of research data performance, reporting whether those variables separate the institutions or take the same state throughout — the test distinguishing robustness from vacuity.

Ch.	Requirement
15	comparison of decision records and documentation volume at institutions adopting the framework for decision support against those adopting it for internal audit, to establish whether the mode of adoption determines whether practice or documentation changes.
16	comparison of institutions that assigned decision authority before acquiring stewardship tooling against institutions that reversed the order, measuring the proportion of variables readable at level 2 or above in each group, to establish whether the dependency order has the practical effect derived here.
16	a survey of the machine-actionable interfaces exposed by widely used generalist and disciplinary repositories against the five exposure classes derived here, reporting for each variable of Tables 9-3 and 11-2 the maximum maturity level an external institution can attain.
16	comparison of recognition arrangements that read a single deposit count against arrangements coupling recognition to metadata or curation conditions, measuring description completeness and controlled-access provision in each, to test whether multi-variable recognition resists the substitution mechanism.
16	an application test with 12-20 participants from at least two disciplines, comparing a decision-structure curriculum against a tool-centred curriculum by blind rubric scoring of an artefact produced in the participant's own project after six to eight weeks, to establish whether the location of failure derived here holds.
17	a repeated-measures study across at least five years at institutions adopting the framework and at matched institutions that do not, recording variable states, conflict registers, realized measures and the divergence between self-assessed and verified profiles, to establish whether adoption changes decisions or documentation.

The distribution of entries across chapters is itself informative. Concentrations in the methodological chapter and in Part IV reflect the fact that the framework is at present a construction supported by a mapped corpus rather than a validated instrument, which is the standing assigned to it in Chapter 15. Resolution of the entries in Table A-1 is the condition under which the claims classed as supported in Table 15-2 could be reclassified as established.

Appendix B. External Review and Validation Protocol

The result type against which this book is to be assessed requires an external expert review of the completed manuscript. Assessment of the book proposal by the publisher does not meet that requirement, because it evaluates a plan rather than the executed argument. The protocol below governs the review of this manuscript and produces the documentation that the assessment requires.

B.1 Review protocol

1. The version of the manuscript submitted for review is frozen and a checksum of the file is recorded, so that the reviewed text is identifiable afterwards.
2. Reviewers are selected from outside the workplace of the author, and their expertise and independence are documented. Two reviews are sought: one in data stewardship and open science, one in methodology or institutional governance.
3. Each reviewer receives the complete manuscript, an assessment form and an explicit request for an expert review of a scholarly monograph.
4. Each review is returned as a dated, signed document containing a recommendation to publish, to publish subject to revisions, or not to publish.
5. A comment table is produced recording every reviewer comment, the author's response and the change made, with a reference to the location of the change.
6. The post-review version is frozen with its own checksum, and the reviews, the responses and a difference protocol against the reviewed version are archived.
7. Only after these steps is the final manuscript released to production.

B.2 Documentation retained

Table B-1. Records produced by the review process.

Record	Content	Retained where
Frozen pre-review manuscript	file and checksum	project archive
Reviewer dossier	selection ground, affiliation, independence statement	project archive
Signed reviews	dated recommendation and full text	project archive
Comment table	comment, response, change, location	appendix to archive; summary citable in the book
Frozen post-review manuscript	file, checksum, difference protocol	project archive

B.3 Relation to the evidence markers

The review required by this appendix assesses the argument of the book. It does not discharge the evidence markers indexed in Appendix A, which name empirical work that no review can substitute. The two tracks are deliberately separate: the markers are resolved by research, the manuscript by review, and the assessment of the result requires both.

Appendix C. Rights Register for Figures and Tables

A register of rights is maintained for every figure and table, because citation of a source does not by itself constitute permission to reproduce third-party material. The present state of the register is simple: every figure and every table in this manuscript is the author's own work, drawn or compiled from the author's synthesis, and no third-party image, screenshot, logo or map is reproduced. The register must be revisited whenever material is added.

Table C-1. Rights register.

ID	Item	Origin	Rights holder	Regime	Evidence
FIG 1-1	Two-level decision structure	own diagram	author	own work	source file in project archive
FIG 4-1	Analytical chain from source to variable	own diagram	author	own work	source file in project archive
FIG 5-1	Assignment matrix for a data activity	own diagram	author	own work	source file in project archive
FIG 6-1	Revision, version and release	own diagram	author	own work	source file in project archive
FIG 7-1	Structure of a verifiable plan item	own diagram	author	own work	source file in project archive
FIG 8-1	Layered description architecture	own diagram	author	own work	source file in project archive
FIG 10-1	Reproducibility chain and its break points	own diagram	author	own work	source file in project archive
FIG 16-1	Interface layer for institutional assessment	own diagram	author	own work	source file in project archive
TAB (all)	Tables 0-1 through C-1	own compilation	author	own work	manuscript sources

Diagrams that redraw a third-party original, adapted material, and any material previously published in a thesis or article of the author require a separate row with the permission obtained and its evidence, before the manuscript is submitted.