

2023 CIRES 暨 depositar 交流座談

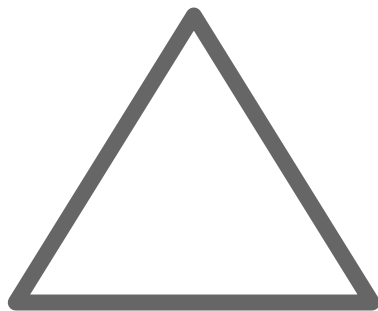
長期社會生態系統研究 的資料管理原則

Data Management in
Long-Term Socio-Ecological Research (LTSER)

2023-01-11
國立臺灣大學圖書資訊研究所
碩士生 黃寶霽

LTSER 成功三要素

持續的資助經費
(Continuous Funding)



彈性且堅實的網絡
(A Loose, yet Stable Network)

可用與能近用的資料與資訊
(Availability of and Access to Data and Information)

長期社會生態觀測研究 的資料管理原則

相關文章可參考： Research Data Management Hub

長期社會生態系統研究的資料管理原則

By 研究資料寄存所編輯群 on 週五, 10/07/2022 - 11:24



研究問題

長期社會生態系統研究之資料：

1. 與一般研究資料有何不同？
2. 在資料管理實務之特別地方？
3. 在資料寄存與發布上需考慮哪些？

● 資料特性 多元異質、稀少、分散

● RDM 原則

資料格式、儀器在長期變動下如何應變？
保留彈性註記空間，妥善完整記錄變動

跨不同研究計畫站之間，該如何將資料整合？
建構支持網絡與確立中央匯集的資料儲存庫
以一致標準儲存資料與其後設資料

LTSER 資料特性

生態學研究 資料特性	Heterogeneous (異質性)
	Small and Diverse (少且多元)
	Non-standard (無標準)

#資料的持續性

「

多元 (diverse)

稀缺 (scarce)

分散 (scattered)

#盡量在海納百川同時，
也確立標準以利後續再用

● 長期現地研究 (Long-term field research)

● 跨領域 (Interdisciplinarity)

● 聚焦問題 (Problem-focused)

● 整合利益關係者 (Stakeholder-integrating)

」

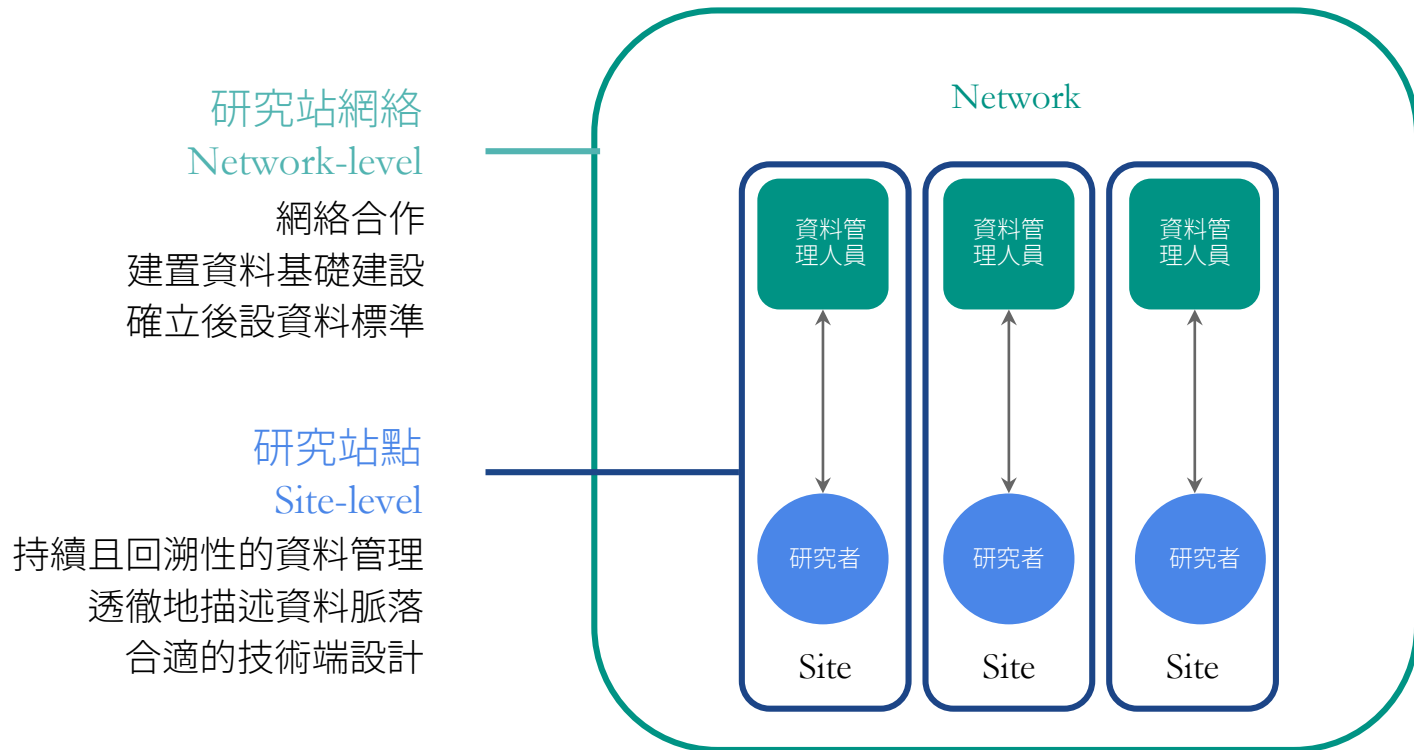
● 資料分享 (Data sharing)

LTSER 資料管理原則

LTSER 仍初步發展中，相關討論與資訊量少
以下統整是奠基於相對發展成熟的
「長期生態系統研究(LTER)於資料管理面的原則與實踐」

LTSER 資料管理可以如何進行？

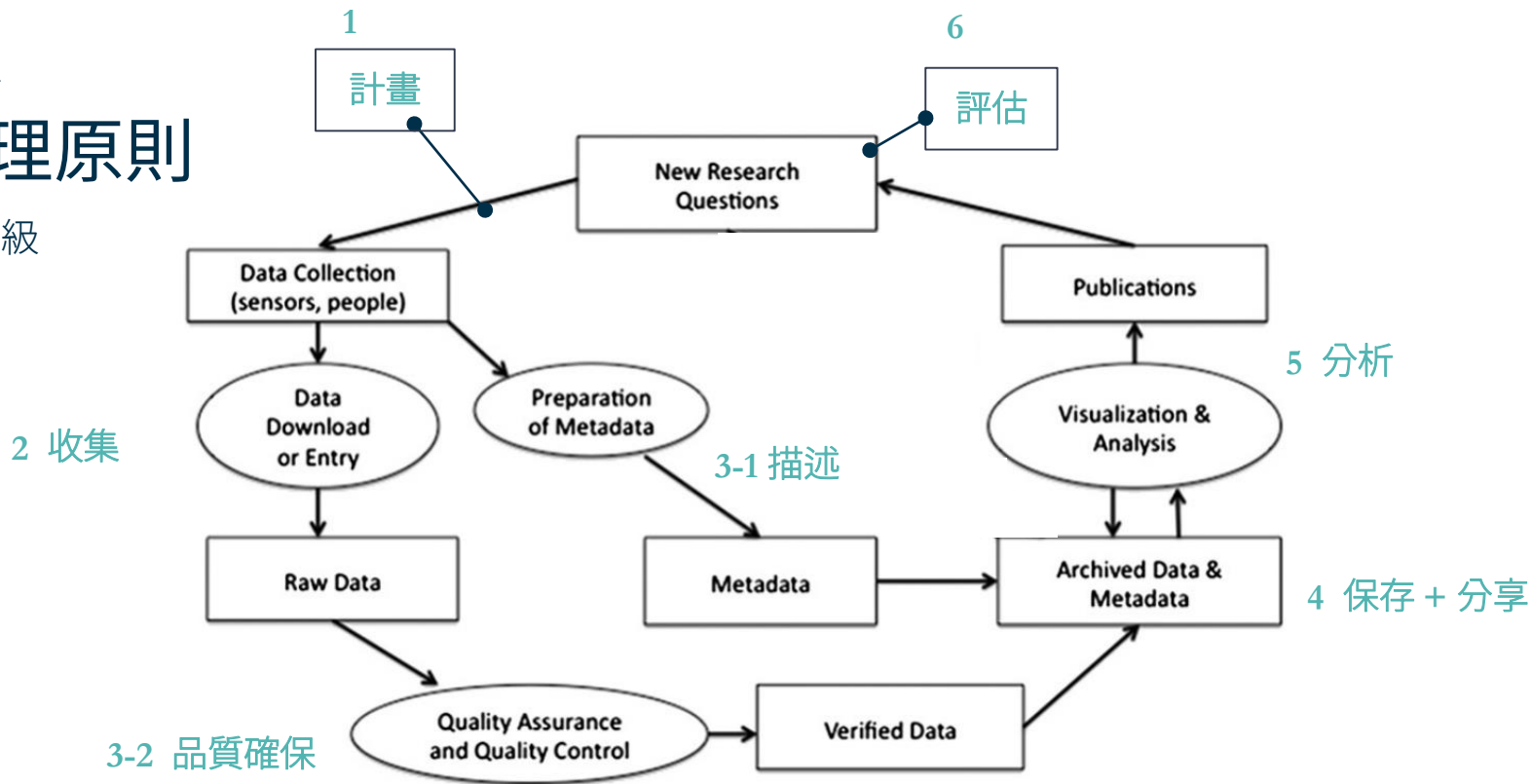
以研究站點與研究站網絡二個階層進行管理



LTSER

資料管理原則

#研究站點層級



LSTER - RDM: 1. 計畫

1. 思考資料再用者的潛在需求
2. 研究設計中的資料面向
3. 資料管理的基礎建設
4. 人員角色與責任、訓練
5. 工作空間與檔案管理

Level	Planned use			
III	Publishable & auditable	Inadequate	Minimal	Good Practice
II	Searchable & third party reuse	Minimal	Good Practice	Excessive
I	Exchange with expert colleague	Good Practice	Excessive	Excessive
		Low Free format, ASCII, narrative, or hard copy	Medium Mixed format, partially parameterized	High Fixed format, highly parameterized, self-documenting and automatically parsable.
		Amount of Structure (Level of effort)		

Figure 2 The degree of metadata format and structure necessary for different levels of projected secondary data utilization. (Figure taken from Michener *et al.*, 1997)

LSTER - RDM: 1. 計畫

建議整體規劃可用表格方式，清楚標示出目標/策略與細項/經費/執掌人員等資訊

Goal	Objective	Metrics	Strategies	Actions Required	By Whom	Location	Status	Time Frame	Financial Resources
1. Provide sources of high-quality, well-documented, and error-checked data at each site that support local science, stimulate synthesis and the creation of new knowledge, and facilitate transformative network-wide research at broad scales.	1A. Fully document site data in accordance with Network standards	Increase in the proportion of datasets collected at each site that are available and NIS compliant	<ul style="list-style-type: none"> - Establish standard and consistent metadata practices for the LTER Network with the expectation that each site will set a reasonable timetable for achieving these standards - Obtain the resources necessary for sites to review and revise each of their LTERfunded data sets and selected legacy data to improve standardization and enhance usability. 	a. Document and create rich EML for data collected and available at each site	Site IM, PI, co-PIs	All sites	Underway	existing data sets now through 2012, ongoing thereafter existing data sets now – end of 2012, ongoing thereafter	Funding for an information management team at each site that will require approximately 3 FTE annually, including a lead information manager trained in ecology and ecoinformatics
				b. Prioritize data sets that are most likely to contribute to Goals, and establish standards reflecting their importance.	Synthesis working groups, site IMs	TBD	Needs planning		

LSTER - RDM: 2. 收集

#完整且一致地收集資料

1. 選擇與定期評估收集方法
2. 考量不同資料類型所延伸出的需求
3. 定期校正儀器
4. 強化完整與小心收集資料的文化慣習

如果要更改紀錄？

不應直接刪除原始值或讓原始值難以辨認
接受加註的可能性

LSTER - RDM: 3-1. 描述

#描述資料的身世脈落

1. 一致的後設資料標準
2. 搭配控制詞彙
3. 決定描述的詳盡度

(1) 辨識

(2) 發現

(3) 評估

(4) 近用

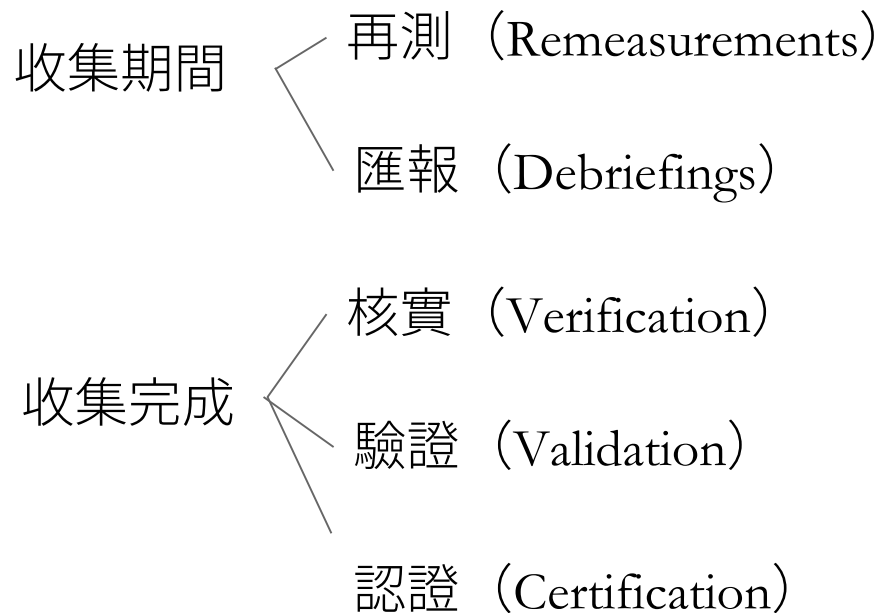
(5) 整合

Completeness Level	Description and Major Elements Added
1: Identification	Minimum content for adequate data set discovery in a general cataloging system or repository (functionally equivalent to LTER DTOC): <ul style="list-style-type: none">• title• creator• contact• publisher• pubDate• keywords• abstract (recommended)• dataset/distribution (i.e. url for general dataset information)
2: Discovery	Level 1 content, plus coverage information to support targeted searches, adding elements: <ul style="list-style-type: none">• geographicCoverage• taxonomicCoverage• temporalCoverage
3: Evaluation	Level 2 content, plus data set details to enable end-user evaluation of the methodology and data entities, adding elements: <ul style="list-style-type: none">• Intellectual Rights• project• methods• dataTable/entityGroup• dataTable/attributes (see issues outlined in the text)
4: Access	Level 3 content plus data access details to support automated data retrieval, adding elements: <ul style="list-style-type: none">• access• physical
5: Integration	Level 4 content plus complete attribute and quality control details to support computer-assisted data integration and re-sampling, adding elements: <ul style="list-style-type: none">• attributeList (full descriptions)• constraint• qualityControl

(2004 EML Best Practices for LTER Sites)

LSTER - RDM: 3-2. 品質確保

#確保資料品質以利無痛使用



資料的完整性、正確性、一致性

#所有的資料都在這裡了嗎？是正確呈現的嗎？

以科學觀點檢視資料的內容與合理性

#資料是合理的嗎？錯誤情況是什麼？

資料符合能被利用與分享的基準

#資料能被利用與分享了嗎？整體品質如何？

LSTER - RDM: 4. 保存+分享

#確保儲存資料能被長期利用

1. 確認須儲存物件、格式
2. 檔案命名管理
3. 決定儲存空間與管理方式
4. 訂定定期儲存資料的時間
5. 若要分享，先確認資料的所有權、機敏性、近用權限

LSTER - RDM: 5. 分析

#確保資料軌跡能被妥善記錄

1. 發展並記錄對應的分析流程
2. 版本控制（資料軌跡）

LSTER - RDM: 6. 評估

#思考目前資料管理原則的狀態

1. 最近一期資料累積的狀態與管理方式
2. 資料收集計畫的可用性與重現性
3. 關乎人員訓練的反饋和建議
4. 資料品質程序的有效性
5. 後設資料紀錄的完整度
6. 設備、工作流程、資料處理流程、基礎建設的效率提昇
7. 資料/研究專案資訊對利益相關者與潛在使用者的影響與價值

LTSER

資料管理原則

#研究站點層級

2 收集

收集的方式
評估、校正

3-2 品質確保

資料收集間：再測、匯報

資料收集後：核實、驗證、認證

1

計畫

研究設計、基礎建設
人員角色

6

評估

人員反饋、
價值檢視

持續且回溯性的資料管理

3-1 描述

透徹地描述資料脈落
設計符合自身需求的技術工具

標準：正確、完整、全面、可被理解

5 分析

工具、流程、
版本控制

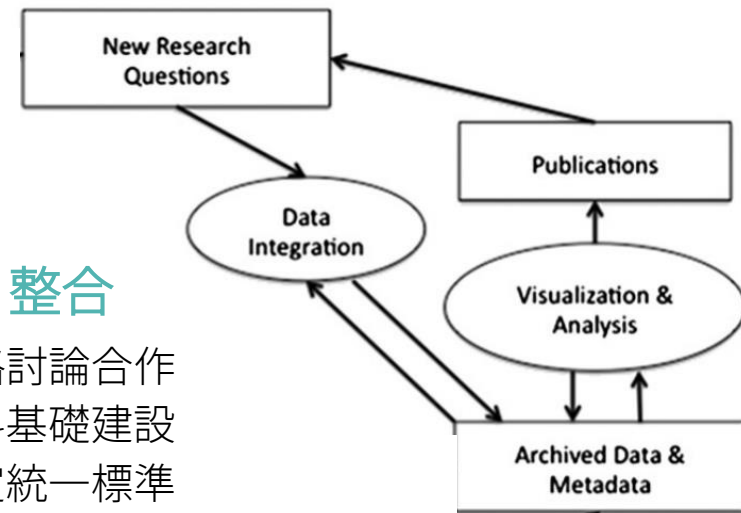
4 保存 + 分享

選用的資料庫類型
資料權限、檔案格式

LTSER

資料管理原則

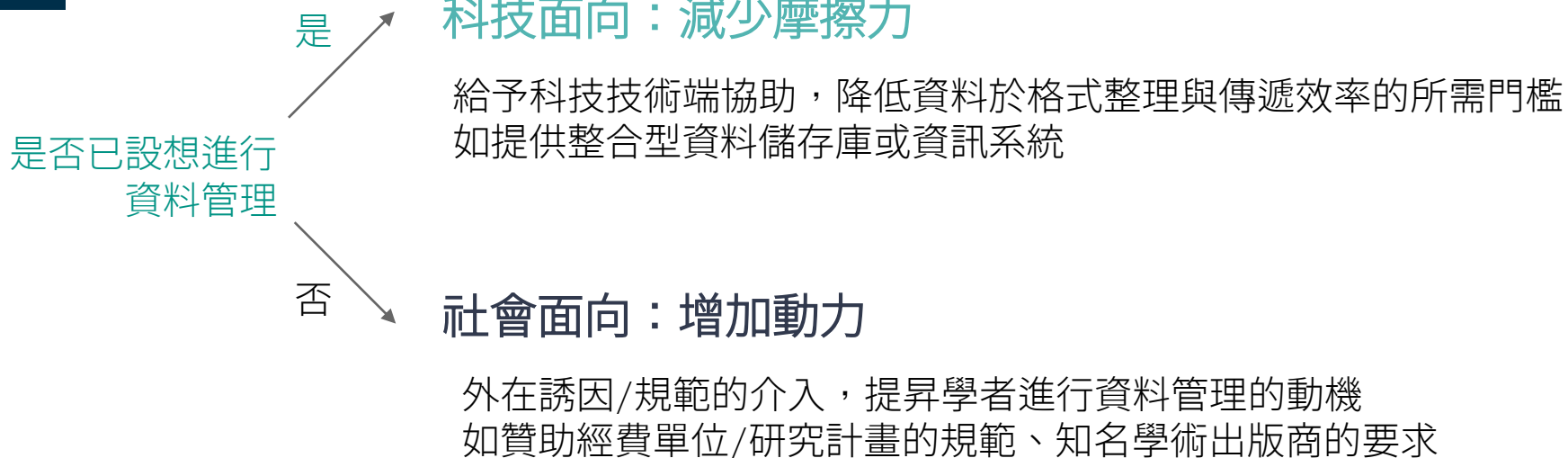
#研究站網絡層級



整合

網絡討論合作
建置共同資料基礎建設
發展與訂定統一標準
(e.g., 後設資料標準、控制詞彙)

兩股力量能推進更完備的資料管理方針



總結

1. 如一般研究，長期社會生態系統研究須具備堅實且彈性的資料管理計畫
在科技、設備、分析方法皆可能變動下，特別須確認新資料能否「嵌入」既有資料集持續累積，因此，更重視於資料與其後設資料的持續性與一致性
2. 由貼近資料的各自研究站點自行建立資料管理規劃，
連結各站的中央資料儲存庫作為站點內/間於資料整合時的重要機制，確立標準與保障資料品質
3. 為提升研究者於資料管理上的動力並減輕其負擔，
資助或主管機構應訂定相關政策，
與此之上，對應支援的專業人力與支持網絡有其存在必要性

THANK YOU

國立臺灣大學圖書資訊學研究所
黃寶霈 r08126018@ntu.edu.tw

Reference:

- Angelstam, P., Manton, M., Elbakidze, M., Sijtsma, F., Adamescu, M. C., Avni, N., Beja, P., Bezak, P., Zyablikova, I., Cruz, F., Bretagnolle, V., Díaz-Delgado, R., Eins, B., Fedoriak, M., Flaim, G., Gingrich, S., Lavi-Neeman, M., Medinets, S., Melecis, V., ... Yamelnets, T. (2019). LTER platforms as a place-based transdisciplinary research infrastructure: Learning landscape approach through evaluation. *Landscape Ecology*, 34(7), 1461–1484. <https://doi.org/10.1007/s10980-018-0757-6>
- Borer, E. T., Seabloom, E. W., Jones, M. B., & Schildhauer, M. (2009). Some Simple Guidelines for Effective Data Management. *Bulletin of the Ecological Society of America*, 90(2), 205–214. <https://doi.org/10.1890/0012-9623-90.2.205>
- Christensen, S. W., Brandt, C. C., & McCracken, M. K. (2011). Importance of Data Management in a Long-Term Biological Monitoring Program. *Environmental Management*, 47(6), 1112–1124. <https://doi.org/10.1007/s10267-010-9576-1>
- Dick, J., Orenstein, D. E., Holzer, J. M., Wohner, C., Achard, A.-L., Andrews, C., Avriël-Avni, N., Beja, P., Blond, N., Cabello, J., Chen, C., Díaz-Delgado, R., Giannakis, G. V., Gingrich, S., Izakovicova, Z., Krauze, K., Lamouroux, N., Leca, S., Melecis, V., ... Van Ryckegem, G. (2018). What is socio-ecological research delivering? A literature survey across 25 international LTER platforms. *Science of The Total Environment*, 622–623, 1225–1240. <https://doi.org/10.1016/j.scitotenv.2017.11.324>
- Gingrich, S., Schmid, M., Dirnböck, T., Dullinger, I., Garstenauser, R., Gaube, V., Haberl, H., Kainz, M., Kreiner, D., Mayer, R., Mirtl, M., Sass, O., Schuppenlehner, T., Stocker-Kiss, A., & Wildenberg, M. (2016). Long-Term Socio-Ecological Research in Practice: Lessons from Inter- and Transdisciplinary Research in the Austrian Eisenwurzen. *Sustainability*, 8(8), 745. <https://doi.org/10.3390/su8080743>
- Kaplan, N. E., Baker, K. S., & Karasti, H. (2021). Long live the data! Embedded data management at a long-term ecological research site. *Ecosphere*, 12(5). <https://doi.org/10.1002/ecs2.3493>
- Karasti, H., Baker, K. S., & Halkola, E. (2006). Enriching the Notion of Data Curation in E-Science: Data Managing and Information Infrastructuring in the Long Term Ecological Research (LTER) Network. *Computer Supported Cooperative Work (CSCW)*, 15(4), 321–358. <https://doi.org/10.1007/s10606-006-9023-2>
- Laney, C. M., Baker, K. S., Peters, D. P. C., & Ramsey, K. W. (2013). Recommendations for Data Accessibility. In *Long-Term Trends in Ecological Systems* (pp. 216–225). United States Department of Agriculture. <https://jomada.nmsu.edu/bibliography/13-029-16.pdf>
- Long-Term Ecological Research Network [LTERnet]. (2022). *2022 LTER Renewal Solicitation*. National Science Foundation. <https://lternet.edu/wp-content/uploads/2022/02/2022-LTER-renewal-solicitation-nsf22543.pdf>
- Long-Term Ecological Research Network [LTERnet]. (2010). *Baltimore Ecosystem Study—Long-Term Ecological Research Renewal Proposal*. <https://lternet.edu/wp-content/uploads/2018/12/BES-2010-LTER03-proposal-clean.pdf>
- Long-Term Ecological Research Network [LTERnet]. (2004). *EMI Best Practices for LTER Sites*. https://lternet.edu/wp-content/uploads/2012/08/emlbestpractices_oct2004_final.pdf
- Long-Term Ecological Research Network [LTERnet]. (2011). *Strategic and Implementation Plan: LTER Research and Education*. https://lternet.edu/wp-content/uploads/2010/12/LTER_SIP_Dec_05_2010.pdf
- Long-Term Ecological Research Network [LTERnet]. (2018). *Guidelines for LTER Information Management Systems*. https://lternet.edu/wp-content/uploads/2018/01/LTER_IMS_Guidelines_V2_1_20181013.pdf
- Long-Term Ecological Research Network [LTERnet]. (2019). *LTER Network Decade Review Self Study*. https://lternet.edu/wp-content/uploads/2019/10/LTER_Self_Study_2019-10-04.pdf
- Long-Term Ecological Research Network [LTERnet]. (2021). *LTER Site Proposals: Florida Coastal Everglades*. <https://lternet.edu/wp-content/uploads/2021/03/FCF-LTER4B-2021.pdf>
- Mauz, I., Peltola, T., Granjou, C., van Bommel, S., & Buijs, A. (2012). How scientific visions matter: Insights from three long-term socio-ecological research (LTER) platforms under construction in Europe. *Environmental Science & Policy*, 19–20, 90–99. <https://doi.org/10.1016/j.envsci.2012.02.005>
- Meléndez-Colom, E. C., & Baker, K. S. (2002). *Common Information Management Framework: In Practice*. Proceedings of the 6th World Multi-Conference on Systemics, Cybernetics and Informatics, US-A, 7, 385–399.
- Miechener, W. K., Porter, J., Servilla, M., & Vanderbilt, K. (2011). Long term ecological research and information management. *Ecological Informatics*, 6(1), 13–24. <https://doi.org/10.1016/j.ecoinf.2010.11.005>
- Mirtl, M., Orenstein, D. E., Wildenberg, M., Peterscil, J., & Frenzel, M. (2013). Development of LTER Platforms in LTER-Europe: Challenges and Experiences in Implementing Place-Based Long-Term Socio-ecological Research in Selected Regions. In S. J. Singh, H. Haberl, M. Chertow, M. Mirtl, & M. Schmid (Eds.), *Long Term Socio-Ecological Research* (pp. 409–442). https://doi.org/10.1007/978-94-007-1177-8_17
- O'Brien, M., Costa, D., & Servilla, M. (2016). Ensuring the quality of data packages in the LTER network data management system. *Ecological Informatics*, 36, 257–246. <https://doi.org/10.1016/j.ecoinf.2016.08.001>
- Peters, D. P. C. (2010). Accessible ecology: Synthesis of the long, deep, and broad. *Trends in Ecology & Evolution*, 25(10), 592–601. <https://doi.org/10.1016/j.tree.2010.07.005>
- Porter, J. (2020). *LTER Information Management*. <https://lter.tari.gov.tw/lterweb/sites/default/files/2020-09/LTER%20Information%20Management.pdf>
- Robbins, R. J. (2011). *Position paper, Data Management for LTER: 1980-2010*. 59.
- Sutter, R. D., Wainscott, S. B., Boetsch, J. R., Palmer, C. J., & Rugg, D. J. (2015). Practical guidance for integrating data management into long-term ecological monitoring projects. *Wildlife Society Bulletin*, 39(3), 451–463. <https://doi.org/10.1002/wsb.548>
- Wohner, C., Peterscil, J., & Klug, H. (2022). Designing and implementing a data model for describing environmental monitoring and research sites. *Ecological Informatics*, 70, 101708. <https://doi.org/10.1016/j.ecoinf.2022.101708>