PublicnEUro：用于受控访问人类神经影像数据的 FAIR 与 TRUST 可信基础设施

Cyril R Pernet[1,2], Stephan Heunis[3], Michael Hanke[3,4], Alyssa Dai[5],Jean Baptiste Poline[5,6,7], Peter S Jensen[1], Martin Nørgaard[8], Melanie Ganz[1,8]& Gitte M Knudsen[1,9]

1 神经生物学研究单元，Rigshospitalet，哥本哈根，丹麦2 丹麦技术大学健康技术系，林比，丹麦 3 神经科学与医学研究所，于利希，德国 4 杜塞尔多夫海因里希·海涅大学，杜塞尔多夫，德国 5 麦康奈尔脑成像中心，The Neuro（蒙特利尔神经病学研究所医院），蒙特利尔，加拿大 6 麦吉尔大学神经病学与神经外科系，蒙特利尔，加拿大 7 麦吉尔大学计算机科学学院，蒙特利尔，加拿大 8 哥本哈根大学计算机科学系，哥本哈根，丹麦9 哥本哈根大学临床医学系，哥本哈根，丹麦

ORCID
Cyril R Pernet https://orcid.org/0000-0003-4010-4632
Stephan Heunis https://orcid.org/0000-0003-3503-9872
Michael Hanke https://orcid.org/0000-0001-6398-6370
Alyssa Dai https://orcid.org/0000-0003-0174-6800
Jean Baptiste Poline https://orcid.org/0000-0002-9794-749X
Peter S Jensen https://orcid.org/0000-0001-7058-6930
Melanie Ganz https://orcid.org/0000-0002-9120-8098
Gitte M Knudsen https://orcid.org/0000-0003-1508-6866

摘要

共享脑成像数据有助于二次分析和可复现性，但隐私风险以及不同的法律和治理要求，可能使不受限制的发布并不适宜。PublicNeuro知识库旨在支持在多样化法律要求下进行发现和再利用。本文介绍了 PublicNeuro 的治理、架构、基于BIDS的数据存放与数据整理、元数据和标识符服务，以及开放和受控检索工作流。随后，我们依据 FAIR子原则评估该存储库已发布的数据集，并依据 TRUST 原则评估该存储库的公开信息。FAIR 评估显示出较强的可发现性和公开元数据可访问性。TRUST 评估记录了治理、托管、连续性保障措施和技术保障措施。总体而言，这表明 PublicNeuro 通过将开放发现与受治理的数据检索相分离，在无需不受限制地披露敏感脑成像数据的情况下，支持符合 FAIR 的可信再利用。

Keyw术语： PublicnEUro、脑成像、存储库、资源、数据共享、FAIR、TRUST

1. 引言

共享数据能够使一次数据采集的科学价值超越其最初的研究问题。二次分析可以整合不同队列、检验新的假设、开发方法，并独立评估已发表研究结果的稳健性（Botvinik-Nezer等人，2019；Poline等人，2012）。这一机会对于脑成像尤其宝贵：数据采集需要专门设备、受试者招募和表型分析，而正电子发射断层成像（PET）等成像模态还需要放射性示踪剂生产以及针对特定采集过程的测量；这些成本可能限制样本量，并使某些队列难以复现（Knudsen等人，2020；Milham等人，2018）。因此，要充分利用现有数据集，仅保留图像文件是不够的。数据存储库必须保存可发现的数据集记录、持久标识符以及足够的标准化元数据，以便研究人员理解、引用数据，并申请再利用这些数据（Gorgolewski等人，2016；Wilkinson等人，2016）。

研究人员要共享其数据，数据存储库需要提供能够满足法律、伦理和知识产权要求的服务。人类神经影像数据是在全球隐私法环境中共享的。神经影像学研究活动较为活跃的司法管辖区——包括欧盟（EU）、美国、加拿大、澳大利亚、中国、日本和韩国——都对可识别健康信息的某些方面进行监管，但其定义和法律要求各不相同（DLA Piper Intelligence，2026）。根据司法管辖区的不同，受到更高程度保护的信息可能包括种族或族裔、宗教、政治信仰，以及健康、遗传或生物特征信息。对于神经影像数据而言，身心健康信息尤其相关：脑扫描图像及相关临床或表型测量可能会揭示健康信息，而移除直接标识符并不一定能消除识别参与者的可能性（Kibsgaard 等，2025；Schwarz 等，2019）。因此，数据存储库需要提供适当的保护措施，包括访问条件、披露记录，以及在需要时对国际传输进行控制。受控访问还可服务于法律合规以外的目的。首先，数据集特定协议和审批程序可以落实针对每个数据集制定的许可和限制，帮助数据作者尊重参与者的权利和期望。其次，受控访问可以为敏感或脆弱人群提供额外保护，包括罕见诊断患者或具有独特属性的人群（Kibsgaard 等，2025；Machanavajjhala 等，2007）。第三，禁运期可以为作者提供时间，在更广泛地发布数据之前发表计划中的分析结果，而访问条件则可以保护正当的知识产权利益。

在欧盟，不同存储库为数据存储提交和访问提供不同条件，以反映研究人员需求的多样性（Lefort-Besnard 等人，2025）。PublicnEUro（Heunis & Pernet，2024）作为专注于脑成像的研究数据存储库而创建（re3data.org，2026；研究资源标识符：SCR_027770），提供在不同访问要求下共享数据的服务，包括因欧盟《通用数据保护条例》（GDPR）而产生的要求。本技术报告介绍了 PublicnEUro 的架构、治理和运营工作流程，评估其与 FAIR（Wilkinson 等人，2016）和 TRUST（Lin 等人，2020）原则的一致性，并使用开源 PolicyModels 工具包（Murphy 等人，2021）总结相关结果。

2. 材料与方法

2.1. 材料

PublicnEUro 由三个相互连接的功能层组成（图1）。（i）*治理与仓储管理层*规定机构责任、法律协议、整理流程、数据集生命周期管理和访问决策，并由运营数据库提供支持。（ii）*安全数据服务层*在托管环境内提供存储、保存、身份验证、授权、审计和受控数据检索。（iii）*发现与互操作层*公开公共目录、持久标识符、数据集级元数据、机器可读使用条件和存储库级表示。治理贯穿两个技术层，而持久数据集标识符和状态记录则维持受保护的数据资产与其公共表示之间的一致性。

治理与管理：PublicnEUro 由哥本哈根大学医院里格医院下属的神经生物学研究单元（NRU）运营，并接受丹麦首都大区治理。该基础设施作为 OpenNeuroPET 项目（Ganz 等人，2026 年）的目标之一而建立，资金由诺和诺德基金会提供。对于已存储提交的数据集，数据来源机构仍是数据控制者，并保留其所有权以及对允许用途的决策权。NRU 作为数据处理者，根据数据控制者记录在案的指示以及适用的服务协议和数据处理协议，接收、整理、存储和分发数据。上述服务按成本收费（用于补偿行政工时和计算/存储成本）。PublicnEUro 还单独作为控制者，负责管理存储库账户和访问记录所需的个人信息。治理职责由监督委员会和管理委员会分担。监督委员会负责制定战略方向、审查政策和科学决策，并监督存储库活动。管理委员会将战略方向落实为技术规范和操作规程，在需要监督时审查访问申请，审计数据集使用情况，并向监督委员会报告存储库活动。数据保护官（DPO）为两个委员会提供建议，参与数据保护影响评估（DPIA）和风险登记册的编制，审查访问记录，并在涉及个人数据的事件发生时协调监管通报或向数据控制者通报。

特定于数据集的共享规则在由 SQL 支持的运营数据库中实现，该数据库连接用户、机构、协议、数据集、访问请求和审计记录。在概念层面，三个二元政策决策——数据访问是开放的还是受控的、地域资格是否受到限制，以及审批是自动进行还是需要监督——产生五种数据共享配置。具体实现提供了更细致的表示：数据集记录区分私有、仅限欧盟、欧盟及充分性国家、全球范围和开放访问，并区分自动审批或基于监督的审批。在发放临时检索链接之前，访问检查会核验用户和机构状态、地域资格、审批要求，以及任何适用的标准合同条款。
PublicnEUro 数据库实现文档记录了这些规则以及相关的浏览器和命令行服务，并根据 MIT 许可证以开放方式发布。保留和保存承诺在每个数据集的服务协议中定义，而不是通过单一的通用保留期限规定。在

协议结束时，控制者和 NRU 可以续订现有托管服务，将文件移至冷归档存储，将整理后的 BIDS 数据集返还给控制者，或删除文件，同时保留公开元数据和数字对象标识符（DOI）着陆页，并更新其生命周期状态。因此，数据集级别的继承通过合同进行管理，其中规定了返还、删除或归档的相关指示。

数据服务：存储提交始于数据作者或负责机构与 NRU 之间签订的合同。作者仍负责依法收集数据、进行适当的假名化处理，并提供特定于数据集的许可证或数据使用协议（DUA）。文件通过安全途径传输，并且必须按照脑成像数据结构（BIDS，Poldrack 等，2024）进行组织，包括磁共振成像（MRI）、PET、脑电图（EEG）、脑磁图（MEG）及相关模态。BIDS Validator 作为主要技术验收测试工具，但也接受类 BIDS 数据集（例如，实现了尚未发布的 BIDS 扩展的数据集）。数据集层面的信息通过 PublicnEUro 元数据记录提供，并在数据整理过程中进行补充。增强工作可包括修正 BIDS 组织结构和元数据、目录检查、参与者摘要、Neurobagel 注释（Urchs 等，2026），以及将权威许可证或数据使用协议转换为机器可读的数字使用条件（DUC；Jeanson 等，2024）。PublicnEUro 分配内部 PublicnEUro（PN）标识符、注册数字对象标识符，并创建公共目录记录。更新后的数据发布可以作为具有独立数字对象标识符的新版本发布，同时与早期版本或已取代的版本保持关联。数据集文件托管于 Computerome 托管环境，这是 PublicnEUro 批准的用于存储和计算的次级处理者。Computerome 位于丹麦，旨在处理敏感研究数据和健康数据。托管和数据库服务组件均在符合 ISO 27001 认证要求的控制措施下运行。

开放数据集可直接通过浏览器访问，但通过命令行检索则需要注册。对于受控数据集，申请人需创建账户、确认电子邮件地址和机构隶属关系，并接受 PublicnEUro 团队的身份验证。访问服务还会根据数据集的地理政策评估该机构所在国家/地区。向欧盟/欧洲经济区以外的地区以及受充分性决定覆盖的国家/地区传输数据时，用户机构必须先签署适用的欧盟标准合同条款，之后才能启用访问权限。每个数据集都需单独申请访问权限，并须接受或签署其数据使用协议。根据数据控制者的政策，符合要求的申请可在在线同意后自动获得审批，也可能需要由数据控制者或授权管理员审核并批准已签署的协议。审批通过后，系统会生成一个有时限的浏览器检索链接；已注册且获得授权的用户还可以通过命令行界面检索完整数据集或选定的文件夹。系统会记录申请、审批及后续访问情况，以便数据控制者审计谁获取了哪些数据集，以及获取时所依据的条件。

PublicnEUro 结合组织和技术控制措施，保护保密性、完整性、可用性和问责性。个人账户、人工注册核查、密码认证和一次性链接支持身份与访问管理。授权权限针对特定于数据集的需求设置，并可予以撤销；工作人员访问权限仅限于业务需要，且会记录个人操作。授权审查和数据保护官的监督将

技术控制措施与 GDPR 职责相衔接。当保密性、完整性或可用性可能受到损害时，事件处理程序要求 NRU 在无不必要延迟的情况下通知相关数据控制者，并提供监管评估、遏制和恢复所需的信息。

发现与互操作性：公共目录的生成方式是，将数据集的每个文件和文件夹表示为 JSON 文件，并使用 DataLad 数据目录将元数据合并到一个数据集级文件中（Halchenko 等人，2021）。随后，该文件会自动呈现为特定于数据集的着陆页，展示全部信息（这些信息也可在 GitHub 上公开获取）。数据集级 JSON 记录呈现描述性、技术性、参与者、资金、作者身份、许可或访问条件、生命周期以及检索元数据，同时不暴露参与者级文件。每个数据集都会获得一个稳定的 PN 标识符，每个已发布版本均以不同的 Crossref DOI 明确表示。版本记录可分别标记为活跃、已归档、已退役、已撤回或已取代。状态与检索模式相互分离，以便持续的数据整理、物理存储以及当前访问路径保持可区分。用户可使用数据集描述、作者、关键词和结构化元数据搜索并浏览目录，同时在需要时对底层文件保持保护；用户也可使用 Neurobagel 查询工具，按影像数据类型、疾病状态、年龄范围等条件，在 PublicnEUro 上跨数据集检索参与者，甚至可跨通过 Neurobagel 关联的数据存储库进行检索。

PublicnEUro 元数据存储库根据目录为每个数据集及其版本生成治理记录。这些记录保留来源路径和内容哈希、生命周期状态、检索信息、数字对象标识符、DUC 配置文件，以及许可到 DUC 映射的来源信息。数据集表示由 JSON Schema 规范，仓储范围内的 re3data 和 OpenAIRE CERIF XML 导出则支持与外部信息系统交换。

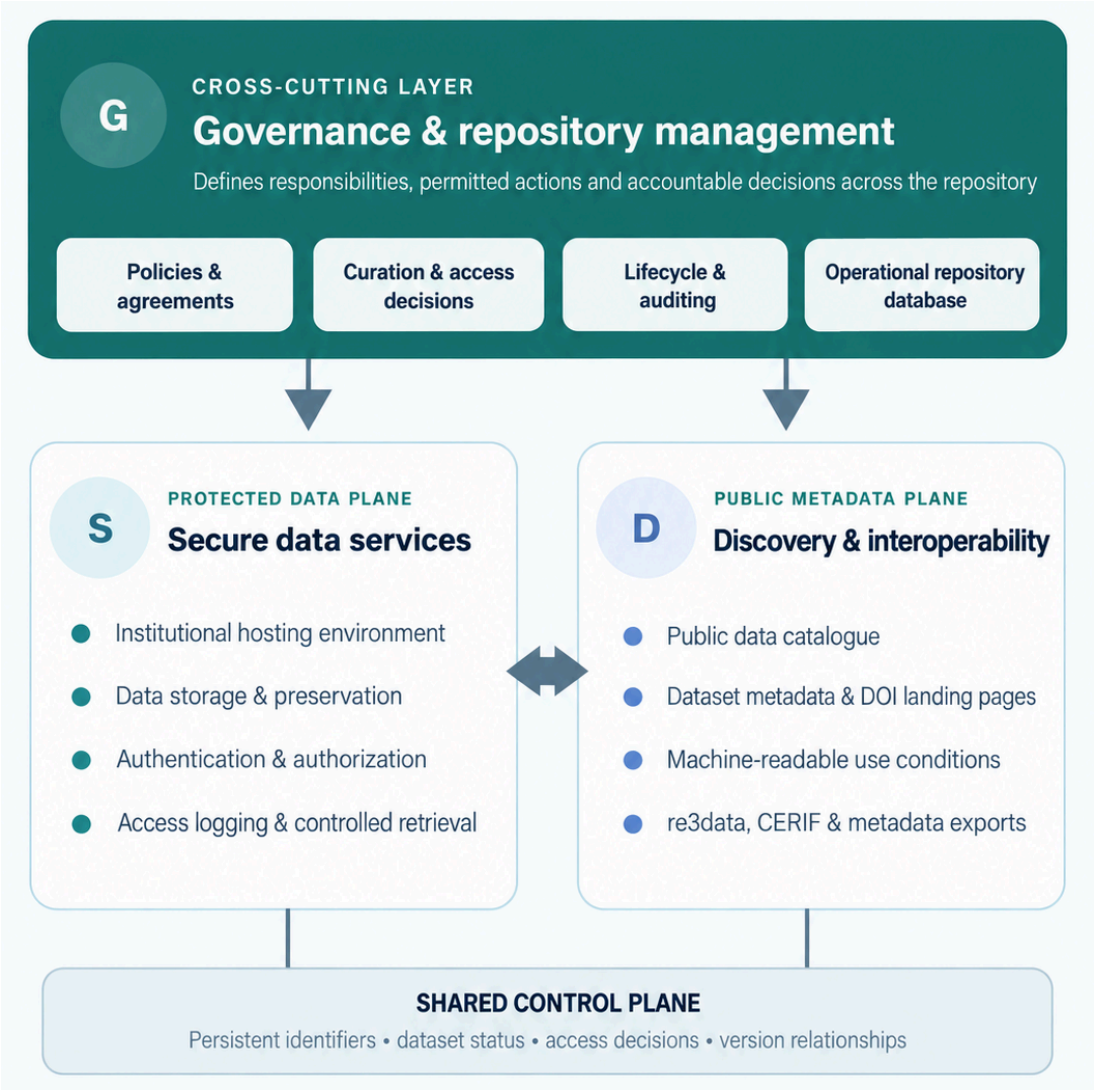


图*1. PublicnEUro* 的功能架构。 治理与仓储管理层制定存储库的政策、职责、整理流程、生命周期管理和访问决策。安全数据服务层提供受保护存储、保存、身份验证、授权、日志记录和受控检索。发现与互操作层提供公共目录、持久标识符、数据集元数据、机器可读使用条件和仓储级元数据导出。持久标识符、数据集状态、访问决策和版本关系确保受保护数据与公共元数据平面之间的一致性。

2.2 方法

FAIR 评估： 依据 FAIR 指南的 15项子原则，使用结构化自我评估对 FAIR 一致性进行了评估。评估单位为已发布的数据集版本及其公开元数据和访问机制。证据来自 PublicnEUro 网站常见问题、政策与标准仓储（也可在网站上查阅）、PublicnEUro DataLad 目录和 DOI 落地页，以及数据集级 JSON Schema、JSON 记录和仓储范围内的 XML 导出。各项子原则按以下方式评级：

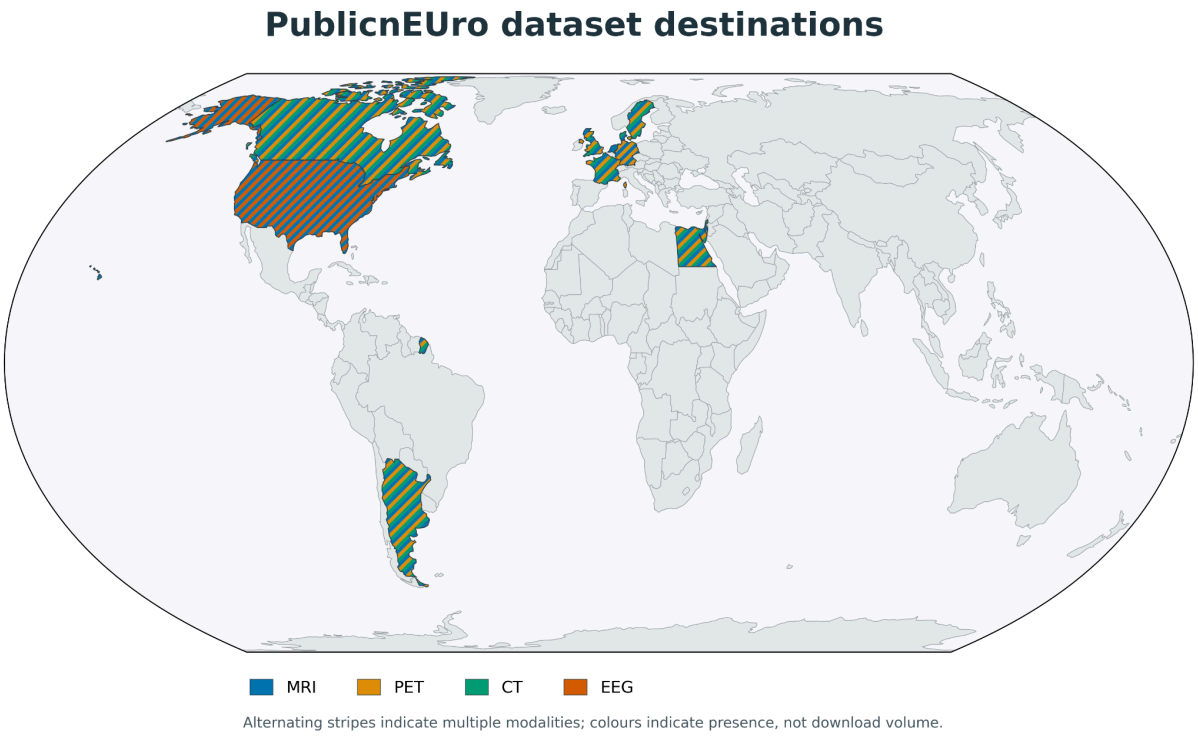
- 当所需行为既有文档记录且可供检查时，认定为原则已体现；- 当实施或证据不完整或内部不一致时，认定为部分体现；- 当未发现任何公开证据时，认定为未体现。- 评级关注数据和元数据的可发现性、可访问性、互操作性和可复用性，而非科学有效性或数据质量。敏感数据的身份验证和授权在 A1.2 项下进行评估，不视为可访问性方面的失误。

TRUST 评估： TRUST 对齐情况依据 TRUST 五项原则——透明度、责任、用户导向、可持续性和技术（Lin 等，2020）——通过结构化自我评估进行评估。评估单位是作为组织与技术存储库服务的 PublicnEUro，而非单个数据集版本。针对每项原则所描述的要求均被具体化为证据问题，涵盖公开披露使命、范围、条款和保存承诺；分配并履行管理责任；与目标社群保持一致并对其需求作出响应；资金、连续性、风险和继任规划；以及采用适当的标准、基础设施和安全控制措施。证据取自 PublicnEUro 网站和常见问题、其版本控制的政策与标准存储库、公开的 DataLad 目录、数据集级治理元数据以及存储库级 XML 导出文件。当相关能力和支持性公开证据足够完整时，将该原则评定为已体现；当证据或实施不完整时，评定为部分体现；当未发现任何公开证据时，评定为未体现。

开放、FAIR、可引用且可信（OFCT）摘要：在基于原则的 FAIR 和 TRUST 评估基础上，我们回答了由 Murphy 等人（2021 年）开发的 31 个问题的 OFCT评估工具。OFCT评估工具将存储库特征转化为可观察的标准，增加了开放性和数据可引用性的明确维度，并生成结构化档案，以识别文档记录和实施缺口。

3. 结果

截至 2026 年 10 月，PublicnEUro 托管了 24 个数据集（25 个版本），涵盖 2130 名参与者，总规模为 5.07 TB。目录记录代表了 MRI、PET、EEG 和 CT 神经影像数据。MRI 包括结构成像和功能成像、灌注成像以及扩散数据。来自世界各地机构的经验证用户共有 31 名，目前每月约新增 5 名用户。需要监督的数据集通常可在 2 周内完成审批。图2展示了已下载的访问受控数据集的地理分布情况，因而也反映了注册用户的地理分布（开放访问数据未被追踪）。



*F*图 *2*：*PublicNeuro* 数据集按成像模态着色的地理分布 *y.*

3.1 评估

3.1.1 FAIR 数据集

PublicnEUro 提供 FAIR共享（https://doi.org/10.25504/FAIRsharing.4c07be），其遵循的原则超越了“开放获取”。

可发现性 依据四项原则进行评估：**(F1)** 为（元）数据分配全球唯一且持久的标识符；**(F2)** 使用丰富的元数据描述数据；**(F3)** 元数据清晰、明确地包含其所描述数据的标识符；以及 **(F4)** 在可搜索的资源中注册或索引（元）数据。每个已发布的数据集均分配了一个 DOI（数字对象标识符），包括同一数据集的不同版本，因此满足 F1。对于 F2、F3 和 F4，公共目录和数据集级 JSON 记录提供丰富的描述性、参与者、技术、访问、生命周期和引文元数据；每条版本记录均明确包含其数据集标识符和 DOI；这些记录可通过可搜索目录发现，而可选的 Neurobagel 注释则进一步支持参与者级数据可用性的联合发现。

可访问性 是指：能够通过标准化通信协议，使用标识符检索 （元）数据，其中（**A1.1)**该协议开放、免费且可由任何人实现，并且（**A1.2)，**在必要时，该协议允许执行身份验证和授权流程。此外，即使数据不再可用，元数据（**A2)**也应保持可访问。在 PublicnEUro 中，数据集元数据和着陆页可通过 HTTPS，借助数字对象标识符解析和公共目录进行检索。开放数据

可直接下载，而受控数据则需完成有文档记录的注册、身份验证、数据集特定数据使用协议，并在需要时获得数据控制者批准及国际传输保障措施。系统同时支持浏览器检索和命令行检索。保存政策和元数据模型通过独立于文件可用性维护元数据和数字对象标识符着陆页，支持数据在活跃、已归档、已退役、已撤回和已取代状态下的保存。该承诺已形成文档并以机器可表示的形式记录，尽管当前发布的集合尚未提供非活跃记录。

互操作性 根据以下标准定义：**(I1)**（元）数据使用正式、可访问、共享且广泛适用的知识表示语言；**(I2)**（元）数据使用遵循 FAIR 原则的词汇表；以及 **(I3)**（元）数据包含指向其他（元）数据的限定引用。在此，I1 和 I3 已得到证明，而 I2 仅得到部分证明。PublicnEUro 使用正式的机器可读表示形式，包括 BIDS、由 JSON Schema 管理的 JSON、兼容 DUC 的 JSON、OpenAIRE CERIF XML 和 re3data XML。数据集记录还在数据集版本、数字对象标识符、目录落地页、确切的源记录、权威许可证或 DUA、DUC 来源信息以及（如适用）替代版本之间建立限定链接。然而，互操作性（I2）仍不完整，因为若干 DUC 条件术语以及存储库状态和检索值均以本地标签而非持久性本体标识符表示；此外，Neurobagel 语义（基于本体的）注释虽受到鼓励，但并非每个数据集都强制要求。

可复用 （元）数据是经过丰富描述的数据，包含多个准确且相关的属性：**(R1.1)**（元）数据发布时附有清晰且可访问的使用许可，**(R1.2)**（元）数据关联有详细的来源信息，且**(R1.3)**（元）数据符合与领域相关的社区标准。PublicnEUro 的记录包含描述性、技术性、参与者、访问、生命周期（每个生成的版本均记录其目录来源和哈希值）及引文属性；开放数据集附有许可，而受控数据集则保留权威的数据使用协议，并附有机器可读的 DUC 摘要，记录版本、来源协议和条件级证据；存储提交的数据则使用社区 BIDS 标准进行整理和验证。

3.2 TRUST 可信存储库

透明度。PublicnEUro 的使命、机构运营方、范围、目标数据类型、存入者和用户服务、开放与受控访问模式以及资金来源均可在线查阅。版本控制政策说明了治理、存储提交、访问、隐私与安全、保存、生命周期状态、可持续性和技术标准。最后，数据集记录公开许可或数据使用协议、经审查的 DUC、数字对象标识符、状态和检索途径。这些内容共同体现了该存储库的透明度。

责任。已发布的治理文件明确规定了 NRU、丹麦首都大区、数据控制者、作者、注册用户、监督与管理委员会以及数据保护官的角色。数据管理包括 BIDS 验证和可选的数据整理、带版本的目录元数据、数据集版本 DOI、受控访问、数据使用协议和传输保障措施的管理、访问日志记录、事件响应，以及在整个生命周期的各个阶段持续保留公开元数据。参与者级别数据的持续托管责任

文件取决于合同和状态，明确不会将其表述为存储库的无限期托管责任。

用户导向。该服务面向人类脑成像数据提交者和再利用者，支持 MRI、PET、EEG、脑磁图及相关 BIDS 数据。用户可以通过公共目录和开放元数据发现并评估可用数据集，在申请前查看访问和再利用条件，并通过浏览器或命令行获取符合条件的数据。存入者可获得数字对象标识符分配、受控访问管理、BIDS 验证以及可选的数据整理支持；Neurobagel 注释支持联合发现。目前尚未建立正式的公共流程，用于收集发布后的数据质量反馈或再利用结果。

可持续性。PublicnEUro 依托 NRU 的机构归属，提供数据集级服务协议，并采用成本回收模式，涵盖安装、管理、维护和存储费用，同时明确规定在线、冷归档、外部托管和不可用状态之间的转换。相关政策涵盖地理冗余备份、风险管理、合同续签、文件和元数据的返还或删除、数字对象标识符持久性以及仓储关闭流程。

技术。PublicnEUro 集成了 BIDS 和 BIDS Validator、DataLad 目录版本控制、Crossref DOI、JSON Schema、经审查的 DUC 配置文件、re3data XML 导出以及可选的 Neurobagel 注释。受控访问采用个人账户、人工身份验证、密码和一次性密码认证、授权及日志记录。相关文档所述的托管安排将文件存储于丹麦的 Computerome，并规定了加密传输、地理冗余备份、访问控制、风险评估和事件响应。ISO 27001 适用于已确定的服务组件。

3.3 OFCT 评估

Q	评估问题	PublicnEUro 的回答和证据	P
1	存储库是否提供对 数据的访问，且限制最少或没有限制?	在法律允许的情况下，是的。大多数数据集都是 然而，为保护敏感的 人类数据。	O
2	用户可以自由下载和再利用 数据，而不受限制或仅受最低限度的限制吗?	每个数据集都有自己的使用条款。有些 数据集使用知识共享署名许可等许可；其他数据集 在其数据使用协议中对使用有更多限制。	O, A
3	存储库是否提供了明确的 数据再利用许可?	数据集层面：每个数据集都有开放 许可或数据集特定数据使用协议。	A, R
4	数据是否受 符合 Commons 要求的许可证约束?	不，每个数据集都有自己的使用条款	O
5	存储库平台是否让 数据易于使用?	是：符合条件的数据可以通过浏览器检索 和命令行工作流。	O
6	该存储库是否要求或提供 支持规范再利用的文档?	是：必须提供 BIDS 元数据和文档； 许可证或数据使用协议以及经审查的 DUC 配置文件说明了再利用条件。	R
7	该存储库是否提供搜索 数据和元数据的功能?	是：公开的 DataLad 目录支持数据集 发现和搜索。Neurobagel 查询还支持 参与者级检索。	F

Q	评估问题	PublicnEUro 答案和证据	P
8	存储库是否分配全球唯一 且持久的标识符?	是：每个已发布的 数据集版本。	F
9	用户能否自行向一个 数据集或贡献分配标识符?	不适用：Q8 的回答为“是”。	F
10	持久标识符或数据集标识符是否 解析到描述该数据的着陆页? 数据?	是：每个数字对象标识符都解析到一个特定版本的公共目录 记录。	F、A、 C
11	元数据是否明确包含其所描述的 数据的标识符?	是：数据集版本记录同时包含 PublicnEUro 标识符和数字对象标识符。	F
12	作者能否将 ORCID iD 与 数据集关联?	支持但非必需：作者元数据 模式接受 ORCID 标识符。	C
13	该存储库是否支持丰富的 用于搜索和再利用的元数据?	是：记录包括描述性元数据、参与者元数据、 技术元数据、访问元数据、生命周期元数据、来源信息和引文 元数据。	F
14	数据和元数据是否附有详细的来源信息 包含详细的来源信息?	是，完整提供：源记录和哈希值、元数据 来源、版本历史和 DUC 映射证据 均有记录。	R
15	该存储库是否提供 数据引用所需的元数据?	是，完整提供：标题、创建者、版本、数字对象标识符、日期和 提供了访问或许可信息。	C
16	元数据是否包含限定 对其他数据或元数据的引用?	是，完整包含：记录链接了版本、数字对象标识符、源 记录、协议、数据使用条件证据以及 替代项。	I
17	对于与论文相关联的数据，是否支持 双向关联?	部分支持：出版物标识符可以在 机器可读元数据中表示，但系统性的反向 论文关联并无保证。	I
18	该存储库是否强制要求或允许 数据格式的社区标准或 元数据?	是：存储提交的神经影像数据经过整理，并 使用 BIDS 进行验证。	R
19	该存储库是否接受 适用于其目标社群的元数据 学科?	是：支持人类神经影像学数据集级和参与者级元数据 ；针对人类神经影像学； 针对 dkNET 的措辞已调整为适用于 PublicnEUro 社群。	R
20	政策能否确保元数据持续保留 当数据不再可用时?	根据既定政策，元数据和 DOI 落地页 在已归档、已退役、已撤回和 已取代状态下仍然保留。	A
21	元数据是否使用了 遵循 FAIR 原则的词汇表?	组级信息至少包含 SNOMED 和疾病本体标识符。 参与者级信息可能使用以下内容进行标记： Neurobagel，	I
22	该存储库是否提供 机器可读的登录页面?	由于 Q10 的答案为“是”，因此未到达此处。该存储库 本身具有等效的机器可读 JSON 记录和 XML 导出。	A, C
23	着陆页是否支持数据 引用?	是：特定版本的 DOI 和引文元数据 已显示，并且机器可读元数据已 导出。	C
24	该存储库是否采用公认的 基本元数据社区标准?	是：BIDS、JSON Schema、Crossref DOI 元数据、 OpenAIRE CERIF 和 re3data 表示均 使用。	R
25	能否访问数据或元数据 通过符合标准的 API?	是：可通过我们的 API 获取数据。所有元数据 均可通过 HTTPS/标准 git 命令行界面以开放 JSON 和 XML 导出的形式获取。 HTTPS/标准 git 命令行界面。	A

Q	评估问题	PublicnEUro 答案和证据	P
26	元数据是否使用正式、可访问的 知识表示共享语言 ？	是：由 JSON Schema 管理的 JSON，以及 符合模式的 XML 表示是公开的。	I
27	该存储库是否提供基于 API 的 搜索以及对数据和 元数据的访问？	没有搜索 API；提供静态机器可读记录以及 命令行数据检索； 可使用 Neurobagel API 查询数据集。	F, A
28	存储库治理是否透明？	是：治理、角色、政策、生命周期 承诺和机构责任均 公开且经过版本控制。	T
29	存储库是否由利益相关者治理？	部分符合：监督和管理结构已 记录在案，但用户和存储者参与 尚未正式纳入管理。	T
30	运行该数据基础设施的代码是否 受某种 开源许可证许可？	部分组件：用户和数据数据库、数据目录以及 元数据组件使用开源代码，但 完整的身份验证和机构服务 栈并未公开。	T
31	该存储库是否获得了数据认可印章 、CoreTrustSeal 或 等效认证？	没有。PublicnEUro 已独立注册，并且 已在 re3data (r3d100014932)、FAIRsharing (10.25504/FAIRsharing.4c07be) 和研究资源标识符 (SCR_027770) 中有所描述，但尚未获得 CoreTrustSeal 或等效认证。	T

表 1. 使用包含 31 个问题的 OFCT评估工具 对 PublicnEUro 进行评估。P 表示原则：O，开放；F，可发现；A，可访问；I，可互操作；R，可复用；C，可引用；T，可信。

4. 讨论

PublicnEUro 为在欧洲法律与机构背景下实现敏感神经影像数据的 FAIR数据发现和受控复用提供了一个实用模式，也可应用于与数据保护或知识产权相关的其他问题。虽然该平台旨在为欧盟研究人员提供帮助，但任何国家的数据集都可以托管在 PublicnEUro 上。FAIR 评估说明，公开元数据可以与受保护的参与者级数据并存。FAIR 原则 A1.2 明确允许在必要时进行身份验证和授权，而 A2 则要求即使数据不可访问，元数据仍应保持可访问（Wilkinson 等，2016）。另一个知名的神经影像数据共享基础设施 OpenNeuro（Markiewicz 等，2021）要求数据以 CC0许可证 发布，因此其已发布的数据集可直接下载，无需创建账户。
PublicnEUro 对 公开 一词的理解有所不同：任何人都可以发现并查看数据集记录、标识符、可用性和访问条件，但获取受限数据则需要授权。不过，PublicnEUro 也支持开放访问的数据集。

这三项评估考察了 PublicnEUro 的不同方面。FAIR 评估已发布数据和元数据的属性：数据集版本能否被找到、其访问条件是否清晰易懂，以及其内容能否被解读和再利用。TRUST 评估负责管理这些对象的存储库，包括其治理、托管、对用户的响应能力、可持续性和技术。包含 31 个问题的 OFCT评估工具则从操作层面考察可观察到的数据仓储实践，尤其是开放性和可引用性。它并未取代 TRUST 评估；后者的五项原则

并未明确纳入原始评估工具。应用这三项评估也揭示了 OFCT 对比的一项局限。Murphy 等人开发并测试该评估工具时，所用的数据存储库中的数据无需经过审批即可访问；其问题并未涉及需要受控访问的敏感的人类受试者数据。因此，即使公开元数据、持久标识符、有据可查的使用条件和授权检索均支持 FAIR再利用，PublicnEUro 在 OFCT 的开放相关问题上仍表现不同。结果应按维度解读：无限制下载、FAIR可访问性和可靠的数据仓储管理是彼此不同的属性，不应将基于数据特点而设的访问限制视为另外两项未能实现的证据。

数据共享依赖于基础设施，而在产生数据集的研究资助结束后，这些基础设施仍必须得到维护。存储只是其中一项费用；数据存储库还需要持续开展数据整理、保存、安全保障、元数据维护和用户支持（经济合作与发展组织，2017；Palaiologk等人，2012）。fMRI数据中心说明了长期可持续性所面临的风险：资金变化导致全面数据整理和新研究接收工作的终止，尽管其现有档案库仍以缩减后的形式继续运行（Van Horn和Gazzaniga，2013）。虽然PublicnEUro的初期开发得到诺和诺德基金会的支持，但其持续运营通过数据集级服务协议向数据提供者收回成本，同时继续向用户免费提供发现和访问请求服务。根据数据集的不同，协议可能涵盖安装、发布、数据整理、定期管理和存储，从而支持在长期内保持数据可发现和可访问所需的服务。这引出了数据共享领域一个更广泛的问题：当研究人员在项目结束后才寻求存储库时，谁来支付其持续托管的费用？对存入者而言，免费托管之所以免费，只是因为资助方、机构或其他贡献者承担了这些成本。

PublicnEUro目前正致力于强化内部流程，依据2026–2028年要求，努力获得CoreTrustSeal认证（Dillo & Leeuw，2018），并提升数据可发现性，具体包括努力与EBRAINS知识图谱（Amunts等人，2022）集成，以及促进元数据知识交流。

致谢

诺和诺德基金会（NNF20OC0063277）为C. Pernet、M. Ganz以及PublicNeuro知识库的创建提供了支持。

参考文献

Amunts, K.、DeFelipe, J.、Pennartz, C.、Destexhe, A.、Migliore, M.、Ryvlin, P.、Furber, S.、Knoll, A.、Bitsch, L.、Bjaalie, J. G.、Ioannidis, Y.、Lippert, T.、Sanchez-Vives, M. V.、Goebel, R. 与 Jirsa, V.（2022）。通过计算连接脑结构、活动与认知功能。*eNeuro*，*9*(2)。https://doi.org/10.1523/ENEURO.0316-21.2022 Botvinik-Nezer, R.、Holzmeister, F.、Camerer, C. F.、Dreber, A.、Huber, J.、Johannesson, M.、Kirchler, M.、Iwanir, R.、Mumford, J. A.、Adcock, A.、Avesani, P.、Baczkowski, B.、Bajracharya, A.、Bakst, L.、Ball, S.、Barilari, M.、Bault, N.、Beaton, D.、Beitner, J.、…

Schonberg, T.（2019）。多个团队对单个神经影像数据集进行分析时的变异性。https://doi.org/10.1101/843193 Dillo, I. 与 Leeuw, L. de（2018）。CoreTrustSeal。奥地利图书馆员与图书馆协会通讯，*71*(1)，162–170。https://doi.org/10.31263/voebm.v71i1.1981 DLA Piper Intelligence。（2026）。全球数据保护法律。https://www.dlapiperdataprotection.com/ Ganz, M.、Norgaard, M.、Pernet, C.、Matheson, G. J.、Galassi, A.、Ceballos, E.、Bilgel, M.、Eierud, C.、Gonzalez-Escamilla, G.、Wighton, P.、Markiewicz, C. J.、Hardcastle, N.、Buckholtz, J.、Blair, R.、Greve, D. N.、Thomas, A. G.、Poldrack, R. A.、Calhoun, V.、Innis, R. B. 与 Knudsen, G. M.（2026）。为分子神经影像构建开放基础设施：OpenNeuroPET 倡议的标准与工具。《脑血流与代谢杂志》，0271678X261490245。https://doi.org/10.1177/0271678X261490245 Gorgolewski, K. J.、Auer, T.、Calhoun, V. D.、Craddock, R. C.、Das, S.、Duff, E. P.、Flandin, G.、Ghosh, S. S.、Glatard, T.、Halchenko, Y. O.、Handwerker, D. A.、Hanke, M.、Keator, D.、Li, X.、Michael, Z.、Maumet, C.、Nichols, B. N.、Nichols, T. E.、Pellman, J.、… Poldrack, R. A.（2016）。脑成像数据结构：用于组织和描述神经影像实验输出的格式。《科学数据》，*3*，160044。Halchenko, Y. O.、Meyer, K.、Poldrack, B.、Solanky, D. S.、Wagner, A. S.、Gors, J.、MacFarlane, D.、Pustina, D.、Sochat, V.、Ghosh, S. S.、Mönch, C.、Markiewicz, C. J.、Waite, L.、Shlyakhter, I.、Vega, A. de la、Hayashi, S.、Häusler, C. O.、Poline, J.-B.、Kadelka, T.、… Hanke, M.（2021）。DataLad：用于联合管理代码、数据及其关系的分布式系统。《开放源代码软件杂志》，*6*(63)，3262。https://doi.org/10.21105/joss.03262 Heunis, S. 与 Pernet, C. R.（2024）。*PublicnEUro* 数据目录。神经生物学研究单元。https://doi.org/https://doi.org/10.70883/cuot3926 Jeanson, F.、Gibson, S. J.、Alper, P.、Bernier, A.、Woolley, J. P.、Mietchen, D.、Strug, A.、Becker, R.、Kamerling, P.、Sanchez Gonzalez, M. del C.、Mah, N.、Novakowski, A.、Wilkinson, M. D.、Benhamed, O. M.、Landi, A.、Krog, G. P.、Müller, H.、Riaz, U.、Veal, C.、… Brookes, A. J.（2024）。让您的 DUC 排成一列——规范化数字使用条件的表示。《科学数据》，*11*(1)，464。https://doi.org/10.1038/s41597-024-03280-6 Kibsgaard, E.、Jwa, A. S.、Markiewicz, C. J.、Gonzalez, D. R.、Pardo, J. S.、Poldrack, R. A. 与 Pernet, C. R.（2025）。评估神经影像学中的元数据隐私（arXiv:2509.15278）。arXiv。https://doi.org/10.48550/arXiv.2509.15278 Knudsen, G. M.、Ganz, M.、Appelhoff, S.、Boellaard, R.、Bormans, G.、Carson, R. E.、Catana, C.、Doudet, D.、Gee, A. D.、Greve, D. N.、Gunn, R. N.、Halldin, C.、Herscovitch, P.、Huang, H.、Keller, S. H.、Lammertsma, A. A.、Lanzenberger, R.、Liow, J.-S.、Lohith, T. G.、… Innis, R. B.（2020）。出版物和档案库中 PET 脑数据的内容与格式指南：共识论文。《脑血流与代谢杂志》，*40*(8)，1576–1585。https://doi.org/10.1177/0271678X20905433 Lefort-Besnard, J.、Pron, A.、Clement, P.、Barker, G.、Erez, Y.、Fernández-Seara, M. A.、Feyzioglu, S. B.、Hartman, S.、Iglesias-Rey, R.、Keil, V.、Mitrović, T.、Nechifor, R. E.、Ozturk-Isik, E.、Passarinho, C.、Petrušić, I.、Trbalić, A. Š.、Sollmann, N.、Thust, S. C.、Turhan, G.、… Maumet, C.（2025）。我可以使用哪个基础设施来分享人类神经影像数据？关于面向欧盟研究人员的当前解决方案的调查与文献综述。*Aperture Neuro*，*5*(SI 2)。https://doi.org/10.52294/001c.144839 Lin, D.、Crabtree, J.、Dillo, I.、Downs, R. R.、Edmunds, R.、Giaretta, D.、De Giusti, M.、L'Hours, H.、Hugo, W.、Jenkyns, R.、Khodiyar, V.、Martone, M. E.、Mokrane, M.、Navale, V.、Petters, J.、Sierman, B.、Sokolova, D. V.、Stockhause, M. 与 Westbrook, J.（2020）。数字存储库的 TRUST 原则。《科学数据》，*7*(1)，144。https://doi.org/10.1038/s41597-020-0486-7

Machanavajjhala, A.、Kifer, D.、Gehrke, J. 和 Venkitasubramaniam, M.（2007）。L-多样性：超越 k-匿名的隐私保护。*ACM*知识发现与数据汇刊，*1*（1），3-es。https://doi.org/10.1145/1217299.1217302 Markiewicz, C. J.、Gorgolewski, K. J.、Feingold, F.、Blair, R.、Halchenko, Y. O.、Miller, E.、Hardcastle, N.、Wexler, J.、Esteban, O.、Goncavles, M.、Jwa, A. 和 Poldrack, R.（2021）。用于共享神经科学数据的 OpenNeuro 资源。*eLife*，*10*，e71774。https://doi.org/10.7554/eLife.71774 Milham, M. P.、Craddock, R. C.、Son, J. J.、Fleischmann, M.、Clucas, J.、Xu, H.、Koo, B.、Krishnakumar, A.、Biswal, B. B.、Castellanos, F. X.、Colcombe, S.、Di Martino, A.、Zuo, X.-N. 和 Klein, A.（2018）。共享脑成像数据对科学文献影响的评估。自然通讯，*9*（1），文章 1。https://doi.org/10.1038/s41467-018-04976-1 Murphy, F.、Bar-Sinai, M. 和 Martone, M. E.（2021）。用于评估生物医学数据存储库与开放、FAIR、引用和可信原则一致性的工具。*PLOS ONE*，*16*（7），e0253538。https://doi.org/10.1371/journal.pone.0253538 OECD。（2017）。可持续研究数据存储库的商业模式。经合组织科学、技术与工业政策文件。https://doi.org/10.1787/302b12bb-en Palaiologk, A. S.、Economides, A. A.、Tjalsma, H. D. 和 Sesink, L. B.（2012）。用于长期保存和传播数字研究数据的作业成本法模型：以 DANS 为例。《国际数字图书馆杂志》，*12*（4），195–214。https://doi.org/10.1007/s00799-012-0092-1 Poldrack, R. A.、Markiewicz, C. J.、Appelhoff, S.、Ashar, Y. K.、Auer, T.、Baillet, S.、Bansal, S.、Beltrachini, L.、Benar, C. G.、Bertazzoli, G.、Bhogawar, S.、Blair, R. W.、Bortoletto, M.、Boudreau, M.、Brooks, T. L.、Calhoun, V. D.、Castelli, F. M.、Clement, P.、Cohen, A. L.、..... Gorgolewski, K. J.（2024）。脑成像数据结构（BIDS）的过去、现在与未来。《成像神经科学》，*2*，imag–2–00103。https://doi.org/10.1162/imag_a_00103 Poline, J.-B.、Breeze, J. L.、Ghosh, S.、Gorgolewski, K.、Halchenko, Y. O.、Hanke, M.、Haselgrove, C.、Helmer, K. G.、Keator, D. B.、Marcus, D. S.、Poldrack, R. A.、Schwartz, Y.、Ashburner, J. 和 Kennedy, D. N.（2012）。神经影像学研究中的数据共享。《神经信息学前沿》，*6*。https://doi.org/10.3389/fninf.2012.00009 re3data.org。（2026）。*PublicnEUro*（p. 24 个数据集）。re3data.org——研究数据存储库注册库。https://doi.org/10.17616/R31NJO0K Schwarz, C. G.、Kremers, W. K.、Therneau, T. M.、Sharp, R. R.、Gunter, J. L.、Vemuri, P.、Arani, A.、Spychalla, A. J.、Kantarci, K.、Knopman, D. S.、Petersen, R. C. 和 Jack, C. R.（2019）。利用面部识别软件识别匿名 MRI 研究参与者。《新英格兰医学杂志》，*381*（17），1684–1686。https://doi.org/10.1056/NEJMc1908881 Urchs, S.、Dai, A.、Jahanpour, A.、Wang, M.、Dugré, M.、Bhagwat, N.、McPherson, B.、Pfarr, J.-K.、van Heese, E. 和 Laansma, M.（2026）。*Neurobagel*：构建分布式数据发现的国际网络。https://osf.io/x64tn_v1 Van Horn, J. D. 和 Gazzaniga, M. S.（2013）。为什么要共享数据？从 fMRIDC 中汲取的经验。《神经影像》，*82*，677–682。https://doi.org/10.1016/j.neuroimage.2012.11.010 Wilkinson, M. D.、Du montier, M.、Aalbersberg, Ij. J.、Appleton, G.、Axton, M.、Baak, A.、Blomberg, N.、Boiten, J.-W.、da Silva Santos, L. B.、Bourne, P. E.、Bouwman, J.、Brookes, A. J.、Clark, T.、Crosas, M.、Dillo, I.、Dumon, O.、Edmunds, S.、Evelo, C. T.、Finkers, R.、.....Mons, B.（2016）。科学数据管理与托管的 FAIR指导原则。《科学数据》，*3*（1），160018。https://doi.org/10.1038/sdata.2016.18