



Gestão de Dados: Tudo em Todos os Lugares, Ao Mesmo

Tempo

Bruno Direito^a, João Pereira^b, Alexandre Sayal^c *

^aCentro de informática e Sistemas da Universidade de Coimbra, bruno.direito@uc.pt

^bCentro de Imagem Biomédica e Investigação Translacional (CIBIT), Universidade de Coimbra, jpereira@uc.pt

^cCentro de Imagem Biomédica e Investigação Translacional (CIBIT), Universidade de Coimbra, alexandre.campos@uc.pt

Resumo

A correta gestão de dados de investigação é fundamental para a concretização da ciência aberta. Esta gestão assegura a integridade dos processos científicos, permitindo que os resultados sejam verificados e reproduzidos por outros investigadores, promovendo a transparência, investigação responsável e a reprodutibilidade. Estas práticas também facilitam o acesso a dados de qualidade, contribuindo de forma significativa para o progresso da ciência.

Neste trabalho, apresentamos um caso de uso que destaca a relevância de ferramentas de controlo de versões para acompanhar a evolução de dados, código e resultados ao longo do ciclo de vida de um projeto de investigação. O objetivo deste projeto é prever valência e excitação (*arousal*) nos participantes enquanto ouvem diferentes trechos musicais, com os sinais funcionais adquiridos por espectroscopia funcional de infravermelho próximo (fNIRS). Durante o processo, tarefas de gestão e processamento de dados envolvem a organização de acordo com standards estabelecidos (*Brain Imaging Data Standard*, BIDS, e *shared near-Infrared spectroscopy format*, SNIRF), normalização, processamento, desenvolvimento de modelos computacionais de classificação e apresentação de resultados.

A utilização destas ferramentas permite não só o rastreamento das alterações em cada

*todos os autores contribuíram de igual forma para o trabalho.

fase, mas também a reversibilidade das mesmas, garantindo maior rigor e controlo nas modificações.

A ferramenta DataLad permite integrar de forma eficiente a gestão de versões de dados e código, facilitando o trabalho internamente e a partilha de repositórios entre múltiplos investigadores e instituições. Assim, fomenta-se uma colaboração eficiente, assegurando ao mesmo tempo a consistência e a reprodutibilidade dos resultados. Nesta apresentação, mostramos que esta ferramenta permite um histórico detalhado e integrado de alterações, sendo uma mais valia no contexto colaborativo deste projecto e proporcionando uma plataforma robusta para a ciência aberta, elevando a transparência da investigação científica.

Palavras-chave: gestão de dados, controlo de versões, neuroimagem

Designação do projeto/infraestrutura/iniciativa

FCT exploratory project Brainplayback (EXPL/PSI-GER/0948/2021), CIBIT (UIDB/04950/2020, UIDP/04950/2020, LA/P/0104/2020), and CISUC (UIDB/00326/2020), AS is funded by Siemens Healthineers Portugal and the FCT PhD fellowship (2023.04365.BD). BD is funded by FCT (CEECINST/00117/2021/CP2784/CT0002).

Público-alvo

gestores de repositórios e data centers, gestores de ciência, curadores de dados, especialistas de informática, investigadores

Ligações web úteis

<https://handbook.datalad.org/>