



O curso Research Data Management da Nova Escola Doutoral

Antónia Correia^a, Isabel Andrade^b, Sean Story^c, Susana Lopes^d

^a *Universidade do Minho, Portugal, antonia.correia@sdum.uminho.pt*

^b *Escola Nacional de Saúde Pública, Universidade Nova de Lisboa, Portugal, isabel.andrade@ensp.unl.pt*

^c *Nova School of Business and Economics, Universidade Nova de Lisboa, Portugal, sean.story@novasbe.pt*

^d *Nova School of Business and Economics, Universidade Nova de Lisboa, Portugal, susana.lopes@novasbe.pt*

Resumo

Pretende-se com esta apresentação divulgar e caracterizar o curso *Research Data Management* da Escola Doutoral da Universidade Nova de Lisboa, bem como o resultado da avaliação feita por meio de questionário aos participantes deste curso. Este curso teve a sua primeira edição em 2015 e conta já com dezenas de participantes que na sua grande maioria são, pela primeira vez, confrontados com as questões relacionadas com o acesso aberto às publicações e dados de investigação.

O objetivo, para além de divulgar esta iniciativa pioneira, é avaliar o trabalho realizado ao longo das várias edições realizadas, apresentar os resultados e identificar melhorias a introduzir futuramente. A avaliação é feita por todos os estudantes de doutoramento que participaram no curso de *Research Data Management*, através de questionário entregue após o teste de avaliação de conhecimentos e tem como objetivo aferir a qualidade e adequação dos conteúdos lecionados, e também do impacto que estes podem ter no seu trabalho de investigação.

Palavras-chave: Gestão de dados de investigação; Literacia da informação; NOVA Escola Doutoral

Apresentação

A NOVA Escola Doutoral – Escola Doutoral da Universidade Nova de Lisboa – tem como objetivo promover a qualidade, a interdisciplinaridade e a internacionalização dos programas doutorais de toda a Universidade, oferecendo formação de carácter complementar à oferta formativa das várias Unidades Orgânicas (UO's), contribuindo para a criação de redes e potenciando a colaboração entre estudantes e docentes.

Os cursos são gratuitos, de inscrição voluntária, realizados em língua inglesa e sujeitos a avaliação. Os ECTS obtidos são creditados no programa doutoral ou em suplemento ao diploma, e as turmas geralmente são constituídas por estudantes provenientes de várias Faculdades ou Escolas da Universidade, possuindo diferentes *backgrounds*. Esta troca de experiências enriquece e reflete-se na dinâmica de cada curso lecionado.

O curso *Research Data Management* – Gestão de Dados de Investigação – foi proposto em outubro de 2014 à Escola Doutoral da Universidade Nova de Lisboa como um curso de iniciação. A par do curso de Literacia de Informação, o seu corpo docente é constituído na

totalidade por bibliotecários.

Objetivos

A competência gestão de dados de investigação assume cada vez maior importância dado que as principais entidades financiadoras, como a Fundação para a Ciência e Tecnologia, vêm emitindo recomendações para o livre acesso aos dados de investigação incentivando uma política de dados abertos e indicando que os "dados de pesquisa científica devem ser facilmente detetáveis, acessíveis, avaliáveis, inteligíveis, utilizáveis, e sempre que possível interoperáveis" (Universidade Nova de Lisboa, 2018).

No âmbito da Comissão Europeia, nomeadamente no quadro do programa Horizonte 2020, e tendo em conta as devidas exceções, os dados de investigação deverão estar em acesso aberto por defeito uma vez que estes constituem a base para os resultados de investigação baseados em medições, observações ou pesquisas e a sua adequada gestão se revela essencial para a sua utilização, reutilização e para a validação dos respetivos resultados.

A gestão e publicação de dados de investigação, como forma de democratização do conhecimento científico e de rentabilização do investimento feito na produção desses dados, assegura a transparência do processo científico havendo que considerar como relevante para esse efeito não só o período após publicação dos manuscritos como também o próprio processo de revisão por pares.

O curso de *Research Data Management* da Escola Doutoral da Universidade Nova de Lisboa pretende ajudar os investigadores e estudantes de doutoramento a gerir os seus dados de investigação em conformidade com as diretrizes das agências de financiamento, e a torná-los visíveis e citáveis.

Programa do curso

Este curso está organizado em quatro módulos distintos que visam aumentar a consciência do movimento de ciência aberta.

- No primeiro módulo faz-se uma introdução aos conceitos-chave relacionados com o progresso e iniciativas de acesso aberto, publicação aberta, modelos de divulgação científica e de publicação e políticas de ciência aberta o que é fundamental no sentido de incentivar uma mudança comportamental dinamizando uma verdadeira “cultura de gestão de dados” (Carvalho, Furtado e Príncipe, 2017) ao nível da comunidade científica no seu todo. De facto, no que diz respeito aos dados de investigação, constata-se uma quase total inexistência de estratégias ou de políticas sobre a partilha e gestão de dados de investigação resultantes das atividades desenvolvidas ou financiadas pelas instituições nacionais (Portugal. Ministério da Ciência, Tecnologia e Ensino Superior. Grupo de Trabalho para a Política Nacional de Ciência Aberta, 2016).
- No segundo módulo, destacam-se algumas iniciativas de dados abertos e apresentam-se os principais diretórios e repositórios de dados - Zenodo, Figshare, Github, Arxiv.org, re3data.org. Este módulo introduz uma dimensão prática, com a realização de diversos exercícios sendo dada relevância aos cuidados a ter para o acesso, partilha e reutilização de dados e apontadas as três principais “regras” a ter em linha de conta: planear e antecipar a reutilização de dados; especificar os detalhes para o acesso e usar uma

infraestrutura existente (Carvalho, Furtado e Príncipe, 2017). Acentua-se, ainda, que a principal justificação para tornar os dados de investigação disponíveis não se centra nas métricas de avaliação e benefícios da citação mas sim no facto de que dar conta de um processo experimental ou investigação em curso ser um princípio da ciência e de que, quando financiada por dinheiros públicos, a ciência é um recurso público (Piwowar e Todd, 2013).

Em termos de aprendizagem, espera-se que os estudantes, no final do módulo, reconheçam a importância de uma adequada gestão de dados de investigação e se familiarizem com alguns dos mais importantes e credíveis repositórios abertos de dados de investigação: Zenodo, Figshare, Github, Arxiv.org, re3data.org, etc.

- No módulo seguinte é continuada a dimensão prática abordando-se o ciclo de vida dos dados e os planos de gestão de dados, procedendo-se à utilização de *checklists* para aferição de dimensões, ao registo na ferramenta gratuita DMPonline e ao início de uma proposta de plano de gestão de dados (University of Edinburgh, 2018).

Em termos de aprendizagem, espera-se que, no final do módulo, os estudantes compreendam o ciclo de vida dos dados; reconheçam a necessidade da gestão dos dados de investigação; conheçam os requisitos das agências de financiamento no que se refere aos dados de investigação e ser capaz de elaborar o seu próprio plano de gestão de dados.

- No último módulo pretende-se dar consistência aos elementos apresentados anteriormente, fazendo uma revisão e explicando a *big picture* – o alinhamento entre políticas nacionais e internacionais e a definição de *workflows* possíveis pela interoperabilidade entre sistemas, terminando com a apresentação do agregador/ sistema de monitorização da Comissão Europeia OpenAire (European Commission, 2018b).

Em termos de aprendizagem, espera-se que os estudantes, no final do módulo, saibam como atribuir uma licença a resultados de investigação; compreender as vantagens e desafios associados à partilha de dados de investigação e saber como tornar os dados visíveis e citáveis.

Material e métodos

Trata-se de aulas expositivas complementadas por uma componente prática que envolve os estudantes através de discussões e trabalho de grupo existindo um acompanhamento constante relativo aos exercícios propostos.

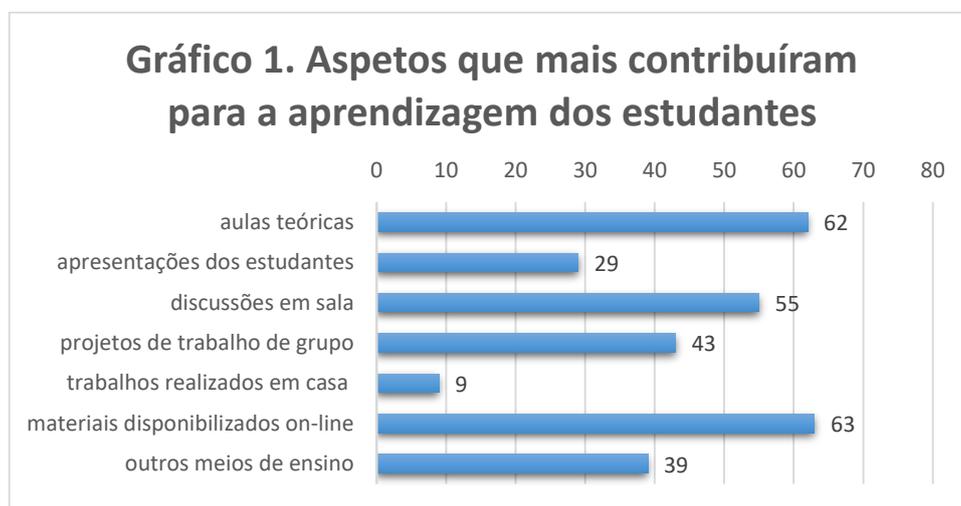
A avaliação final é feita com base no teste de escolha múltipla, na participação individual e coletiva em contexto de sala de aula e em exercícios práticos resolvidos em pequenos grupos e cujas soluções são posteriormente apresentadas oralmente por um porta-voz que o grupo escolhe e discutidas pelo coletivo de alunos e formadores.

Avaliação do ensino

No final de cada edição do curso, os estudantes fazem a avaliação da qualidade dos conteúdos,

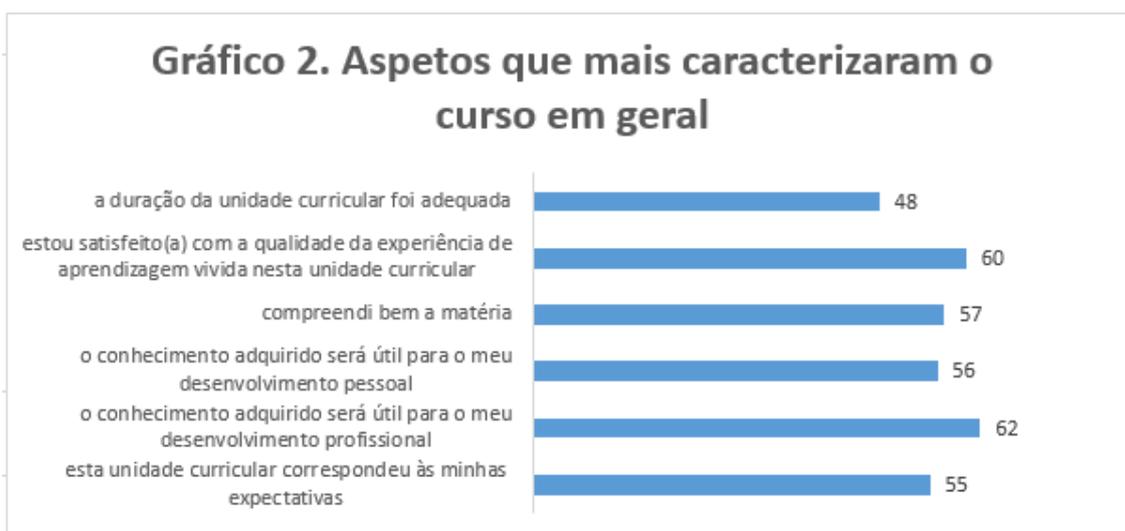
materiais distribuídos, exercícios e o impacto gerado pela aquisição de novos conhecimentos.

Tomando como elemento de análise os dados referentes às respostas *Concordo* e *Concordo Totalmente* dadas pela totalidade dos 77 estudantes às várias questões colocadas no questionário de avaliação do ensino distribuído, pretendeu-se aferir quais os vários aspetos que na opinião daqueles mais contribuíram para a aprendizagem (Gráfico 1) e mais caracterizaram o curso em geral (Gráfico 2).



Da análise das respostas obtidas é possível verificar que a maioria dos participantes considerou que os aspetos que mais contribuíram para a aprendizagem foram a disponibilização dos materiais *online*, as aulas teóricas ministradas e a discussão em sala.

Em termos de características do curso em geral pode-se constatar que o impacto dos novos conhecimentos adquiridos, quer ao nível pessoal quer profissional, é aquele que é valorizado por um maior número de participantes.



De salientar que, ao logo das várias edições realizadas, todos os parâmetros foram avaliados positivamente pela maioria dos inquiridos.

Considerações finais

Ao apresentar diferentes iniciativas políticas e institucionais ao nível nacional e internacional para a gestão dos dados de investigação e a promoção do seu acesso aberto, sobretudo no que se refere aos que resultam de investigação financiada com dinheiros públicos; ao salientar o potencial benefício da partilha de dados para o progresso científico; ao apontar razões para as reservas e desconhecimento dos investigadores quanto à forma como podem proceder a essa disponibilização em termos práticos; ao apresentar alguns repositórios que cumprem com as políticas de gestão de dados e ao apelar ao auto-arquivo dos dados de investigação, o curso de “Research data Management” da NOVA Escola Doutoral tem vindo a constituir importante marco para a disseminação de boas práticas de Ciência Aberta ao nível dos estudantes do 3º ciclo de estudos.

Nesta fase, reconhece-se que é importante criar algum distanciamento das edições já realizadas, registar a sua avaliação e repensar os conteúdos para fazer face aos desafios impostos pela atualização das políticas das entidades financiadoras.

Face aos resultados da avaliação do ensino revela-se essencial investir na componente prática e em ferramentas associadas bem como nas questões relacionadas com a visibilidade dos dados de investigação e seu potencial de reutilização.

As sugestões de melhoria apresentadas pelos participantes: i) incrementar a componente prática; ii) clarificar alguns tópicos relativos às unidades curriculares e iii) eliminar alguma informação repetida ao longo dos diferentes módulos, serão tidas em linha de conta em edições futuras.

Aos formadores caberá, ainda, procurar os melhores recursos e técnicas de formação dirigidas àqueles que precisam de desenvolver estratégias e aptidões que lhes permitam implementar as práticas de Acesso Aberto e de Ciência Aberta nos seus fluxos de trabalho. A plataforma do Projecto FOSTER, financiado pela EU, revela-se um recurso essencial para esse fim (European Commission, 2018a).

Referências bibliográficas

Carvalho, J.; Furtado, F.; Príncipe, P. – Kit sobre dados de investigação. [Braga]: Repositório Científico de Acesso Aberto de Portugal – RCAAP. Universidade do Minho, 2017.

European Commission. FOSTER Plus. The future of science is Open: start your research training now. [Brussels]: FOSTER Plus (Fostering the Practical implementation of Open Science in Horizon2020 and beyond), 2018a. Disponível em <https://www.fosteropenscience.eu/>

European Commission – OpenAIRE Portal: Science. Set Free. [Em linha]. [Brussels]: European Commission, 2018b. Disponível em <https://www.openaire.eu/>

Heather A. Piwowar, H.A.; Vision, T.J.V. – Data reuse and the open data citation advantage. PeerJ. 1 (2013) e-175. DOI 10.7717/peerj.175.

Portugal. Ministério da Ciência, Tecnologia e Ensino Superior. Grupo de Trabalho para a Política Nacional de Ciência Aberta – Primeiro Relatório da Comissão Executiva do Grupo de Trabalho para a Política Nacional de Ciência Aberta: Comissão Executiva GT-PNCA. Lisboa: Comissão Executiva GT-PNCA, 2016.

Universidade Nova de Lisboa. Nova Escola Doutoral – Curso de «Research Data Management». [Em linha]. Lisboa: Nova Escola Doutoral, 2018. Disponível em <https://www.unl.pt/curso/ensino/escola-doutoral/curso-de-research-data-management>

University of Edinburgh. Digital Curation Centre. DMPonline. [Em linha]. Edinburgh: Digital Curation Centre, 2018. Disponível em <https://dmponline.dcc.ac.uk/>