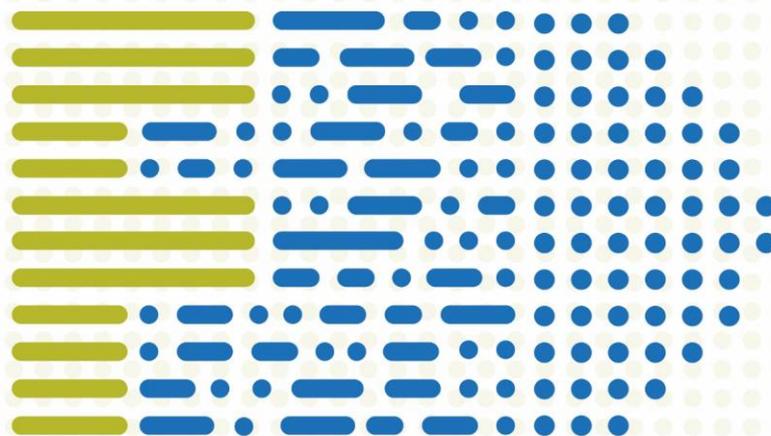


# RELATÓRIO DO 1º FÓRUM GDI



## 1º FÓRUM GESTÃO DE DADOS DE INVESTIGAÇÃO

23 setembro 2016  
FACULDADE DE PSICOLOGIA  
E DE CIÊNCIAS DA EDUCAÇÃO  
U.PORTO



organização



apoio



# VERSÃO

**Autor:** Pedro Príncipe, Universidade do Minho

**Contribuição:** Ricardo Saraiva, Vasco Vaz, João Castro, Maria João Amante (relatores dos grupos de discussão) e Filipe Furtado, Universidade do Minho.

**Versão:** 1

**Distribuição:** Pública

**Data de Criação:** 28 de Novembro de 2016

**Última Atualização:** 9 de fevereiro de 2017

## FICHA DO EVENTO:

**Data e Hora:** 23 de setembro de 2016, 09h30 – 16h00

**Local:** Faculdade de Psicologia e Ciências da Educação da Universidade do Porto

**Organização:** Ana Alves Pereira, Gabinete da Secretária de Estado da Ciência, Tecnologia e Ensino Superior; João Nuno Ferreira e João Mendes Moreira, FCT – Fundação para a Ciência e Tecnologia - FCCN – Fundação para a Computação Científica Nacional; Eloy Rodrigues e Pedro Príncipe, Universidade do Minho.

**Informações:** <http://confdados.rcaap.pt/1forumgdi>

**Página web do evento:** <http://forumgdi.rcaap.pt>

**Fotos do evento:** <https://www.flickr.com/groups/forumgdi>

**Vídeo do evento em imagens:** <https://youtu.be/FdXtlyXsiul>

# ÍNDICE

<b>VERSÃO</b> .....	<b>2</b>
<b>FICHA DO EVENTO:</b> .....	<b>2</b>
<b>INTRODUÇÃO</b> .....	<b>4</b>
<b>APRESENTAÇÃO DO FÓRUM</b> .....	<b>5</b>
<b>PROGRAMA DO 1º FÓRUM-GDI</b> .....	<b>6</b>
Agenda do evento .....	7
<b>RELATOS DOS GRUPOS DE DISCUSSÃO</b> .....	<b>9</b>
Dados abertos e repositórios de dados .....	9
Armazenamento de dados, backup e preservação digital .....	10
Organização dos dados, documentação e metadados .....	12
Formação e desenvolvimento de competências .....	13
<b>WORKSHOP</b> .....	<b>16</b>
<b>PARTICIPAÇÃO</b> .....	<b>16</b>
Tabela resumo .....	16
<b>CONSIDERAÇÕES FINAIS</b> .....	<b>17</b>

# INTRODUÇÃO

Este relatório visa reunir a informação essencial relativa ao 1º Fórum de Gestão de Dados de Investigação (1º Fórum GDI), evento que decorreu na Faculdade de Psicologia e de Ciências de Educação da Universidade do Porto, no dia 23 de setembro de 2016. Neste relatório está reproduzida a informação de apresentação do evento e o programa realizado, incluindo ainda os relatos dos grupos de discussão promovidos no âmbito do programa do Fórum e o resumo das comunicações apresentadas, bem como, as conclusões e recomendações finais dos organizadores.

Este 1º Fórum GDI foi uma iniciativa que beneficiou da realização da Conferência nacional de “Dados de Investigação e Ciência Aberta: rumo a uma estratégia nacional” que decorreu no dia 22 de setembro de 2016. Toda a informação sobre a conferência consta do site criado para o efeito e disponível em: <http://confdados.rcaap.pt>.

Decorrente do trabalho de planificação e preparação da Conferência de Dados de Investigação e Ciência Aberta (ConfDados), realizada no mês de setembro de 2016, iniciativa integrada na atividade do Ministério da Ciência, Tecnologia e Ensino Superior (MCTES) para o desenvolvimento da Política Nacional de Ciência Aberta, foi estabelecido pela organização da conferência que seria de todo adequado aproveitar esta realização nacional para promover um segundo evento de cariz mais interativo para procurar dinamizar um espaço que reunisse profissionais do sistema científico nacional envolvidos em atividades de suporte à Gestão de Dados de Investigação. Foi com este pressuposto que se propôs a realização desde primeiro fórum e se calendarizou precisamente para o dia seguinte à conferência e a decorrer no mesmo local (a informação de divulgação do Fórum GDI foi à data disponibilizada em: <http://confdados.rcaap.pt/1forumgdi>). O evento contou com a presença de 87 participantes, tendo sido apresentadas 7 comunicações, promovidos 4 grupos de discussão e a realização de um workshop.

A referida conferência - “Dados de investigação e Ciência Aberta: rumo a uma estratégia nacional” - teve como objetivo “estabelecer as linhas orientadoras da estratégia nacional para os dados de investigação, reunindo responsáveis e decisores políticos, investigadores e coordenadores de projetos, gestores de ciência e gestores de infraestruturas e de serviços de informação”. O programa da conferência pretendia contribuir para a compreensão dos diferentes componentes e desafios dos dados de investigação colocados a nível institucional e das práticas disciplinares, e procurou promover o conhecimento sobre estratégias e iniciativas de âmbito nacional e internacionais relevantes para o programa de Ciência Aberta do MCTES.

## APRESENTAÇÃO DO FÓRUM

Com a realização do 1º Fórum de Gestão de Dados de Investigação, o evento passou a designar-se por **Fórum GDI** e a definir como um espaço de debate e partilha de ideias, projetos e boas práticas de Gestão de Dados de Investigação que procura juntar gestores de repositórios digitais e data centers, técnicos de informação, bibliotecas, arquivos e curadoria de dados, especialistas de informática, investigadores, cientistas de dados e gestores de ciência de instituições de investigação e organismos de financiamento de ciência.

Com os eventos do Fórum GDI procura-se congregar e dinamizar a comunidade de profissionais e investigadores envolvidos em atividades de suporte à Gestão de Dados de Investigação com o intuito de desenvolver competências técnicas e capitalizar saberes e práticas.

A proposta de organização do Fórum surgiu de necessidades identificadas por diferentes especialistas e atores no domínio da Gestão de Dados de Investigação e resulta do desenvolvimento de projetos e infraestruturas de informação e dados científicos, no âmbito da FCT-FCCN, do projeto RCAAP e de várias instituições de investigação e ensino superior e, como já referido, no contexto da elaboração e implementação da Política Nacional de Ciência Aberta do MCTES.

O primeiro Fórum quis desde logo definir um tipo de evento com forte componente interativa e de dimensão prática, onde se procurasse escutar as necessidades dos diferentes agentes que apoiam a gestão de dados de investigação nas instituições do espaço de ciência e inovação português, procurando ainda capacitar esses profissionais na utilização de ferramentas que apoiam diferentes processos da Gestão dos Dados de Investigação.



**Fórum GDI**

**1º FÓRUM  
GESTÃO DE DADOS  
DE INVESTIGAÇÃO**

**23 setembro 2016**  
FACULDADE DE PSICOLOGIA  
E DE CIÊNCIAS DA EDUCAÇÃO  
U.PORTO

**CIÊNCIA  
ABERTA**

**ESPAÇO DE DEBATE E PARTILHA DE IDEIAS, PROJETOS  
E BOAS PRÁTICAS DE GESTÃO DE DADOS DE  
INVESTIGAÇÃO.**

- congregar e dinamizar a comunidade de profissionais e investigadores envolvidos em atividades de suporte à Gestão de Dados de Investigação.
- desenvolver competências técnicas e capitalizar saberes e práticas.

1 - Slide de abertura do evento com apresentação dos objetivos do Fórum

# PROGRAMA DO 1º FÓRUM-GDI

O programa do 1º Fórum GDI foi composto por três momentos essenciais:

1) O primeiro contou com a apresentação de comunicações do tipo (24 x 7)<sup>1</sup> com a apresentação dos detalhes essenciais de diferentes projetos, iniciativas, planos e ferramentas relacionados com competências e dinamização da Gestão de Dados de Investigação.

**MESA REDONDA (10h00-11h00)**

**Vasco Vaz**, Fundação para a Ciência e Tecnologia  
Financiadores de Ciência e políticas para os Dados Abertos

**Pedro Príncipe**, Universidade do Minho (OpenAIRE)  
OpenAIRE e Dados Abertos no Horizonte 2020

**Afonso Duarte**, ITQB-NOVA  
Infraestruturas Digitais Europeias ao serviço das Ciências da Vida – Desafios e Oportunidades. O ponto de vista do Utilizador

**João Rocha**, Faculdade de Engenharia da Universidade do Porto  
DataPublication@UPorto: multi-disciplinary data description and deposit linking the Dendro staging platform with the EUDAT European Infrastructure

**Luís Faria**, KEEP SOLUTIONS  
Curation Costs Exchange: Understanding and comparing digital curation costs to support smarter investments

**João Moreira**, FCT-FCCN (RCAAP)  
Ações de suporte à gestão de dados no projeto RCAAP

**Eloy Rodrigues**, Universidade do Minho (FOSTER)  
Recursos e objetivos de aprendizagem – resultados do projeto FOSTER

1º FÓRUM  
GESTÃO DE DADOS  
DE INVESTIGAÇÃO  
23 setembro 2016  
FACULDADE DE PSICOLOGIA  
E DE CIÊNCIAS DA EDUCAÇÃO  
U.PORTO

CIÊNCIA ABERTA  
FCT  
U.PORTO

2 - Slide com a lista das apresentações 24x7 realizadas

2) O segundo foi dedicado a debater com os participantes quatro domínios chave do desenvolvimento da Gestão de Dados de Investigação, tendo sido realizados 4 grupos de discussão com quatro diferentes temáticas.

a) *Dados abertos e repositórios de dados (questões sobre a partilha e abertura dos dados de investigação, disponibilização em repositórios de dados disciplinares ou instituições, incluindo licenciamento e citação dos dados abertos).*

b) *Armazenamento de dados, backup e preservação digital (questões sobre a gestão ativa dos dados e serviços para armazenamento, backup e segurança, incluindo os desafios e riscos da preservação a longo prazo).*

c) *Organização dos dados, documentação e metadados (questões sobre práticas de organização, gestão doméstica dos dados e reutilização, descrição e documentação dos dados, incluindo questões de proteção e direitos de acesso).*

d) *Formação e desenvolvimento de competências (questões sobre as necessidades de formação dos diferentes agentes do processo de gestão de dados de investigação, incluindo a identificação de competências e iniciativas a desenvolver).*

3) Por último, promoveu-se um workshop sobre “Planos de Gestão de Dados: ferramentas e requisitos dos financiadores”, integrado na atividades do projeto OpenAIRE.

<sup>1</sup> As apresentações “24x7” consistiram numa apresentação de 7 minutos e não mais que 24 slides. Este tipo de apresentações no Fórum visa apresentar de forma breve o essencial de projetos, iniciativas, casos de estudo ou ideias inovadoras e boas práticas relevantes na Gestão de Dados de Investigação.

## Agenda do evento

De seguida é reproduzida a agenda do evento com as respetivas ligações aos conteúdos apresentados (informação disponível em <http://confdados.rcaap.pt/1forumgdi/>).

### 09h30 – Abertura: apresentação do Fórum de Gestão de Dados de Investigação [\[Apresentação\]](#)

- **Ana Alves Pereira**, Gabinete da Secretária de Estado da Ciência, Tecnologia e Ensino Superior
- **Pedro Príncipe**, Serviços de Documentação da Universidade do Minho

### 10h00 – Mesa redonda: *projetos e iniciativas, competências e áreas temáticas de ação na Gestão de Dados de Investigação*

- **Vasco Vaz**, Fundação para a Ciência e Tecnologia  
Financiadores de Ciência e políticas para os Dados Abertos [\[Apresentação\]](#)
- **Pedro Príncipe**, Universidade do Minho (OpenAIRE)  
OpenAIRE e Dados Abertos no Horizonte 2020 [\[Apresentação\]](#)
- **Afonso Duarte**, ITQB-NOVA  
Infraestruturas Digitais Europeias ao serviço das Ciências da Vida – Desafios e Oportunidades. O ponto de vista do Utilizador [\[Apresentação\]](#)
- **João Rocha**, Faculdade de Engenharia da Universidade do Porto  
DataPublication@UPorto: multi-disciplinary data description and deposit linking the Dendro staging platform with the EUDAT European Infrastructure [\[Apresentação\]](#)
- **Luís Faria**, KEEP SOLUTIONS  
Curation Costs Exchange: Understanding and comparing digital curation costs to support smarter investments [\[Apresentação\]](#)
- **João Moreira**, FCT-FCCN (RCAAP)  
Ações de suporte à gestão de dados no projeto RCAAP [\[Apresentação\]](#)
- **Eloy Rodrigues**, Universidade do Minho (FOSTER)  
Recursos e objetivos de aprendizagem – resultados do projeto FOSTER [\[Apresentação\]](#)

### 11h00 – Grupos de discussão: *áreas temáticas de desenvolvimento da Gestão de Dados Científicos nas Instituições, Serviços e Unidades de Investigação*

**a) Dados abertos e repositórios de dados** (questões sobre a partilha e abertura dos dados de investigação, disponibilização em repositórios de dados disciplinares ou instituições, incluindo licenciamento e citação dos dados abertos).

- Moderador: Pedro Príncipe, Universidade do Minho; Relator: Ricardo Saraiva, Universidade do Minho

**b) Armazenamento de dados, backup e preservação digital** (questões sobre a gestão ativa dos dados e serviços para armazenamento, backup e segurança, incluindo os desafios e riscos da preservação a longo prazo).

- Moderador: João Nuno Ferreira, FCT-FCCN; Relator: Vasco Vaz, FCT

**c) Organização dos dados, documentação e metadados** (questões sobre práticas de organização, gestão doméstica dos dados e reutilização, descrição e documentação dos dados, incluindo questões de proteção e direitos de acesso).

- Moderador: Cristina Ribeiro, Universidade do Porto; Relator: João Castro, Universidade do Porto

**d) Formação e desenvolvimento de competências** (questões sobre as necessidades de formação dos diferentes agentes do processo de gestão de dados de investigação, incluindo a identificação de competências e iniciativas a desenvolver).

- Moderador: Eloy Rodrigues; Relator: Maria João Amante, ISCTE-IUL

Nota: Os grupos de discussão decorrem em sessões paralelas pelo que os participantes escolherão o tema e grupo do seu interesse.

12h00 – Relato e debate sobre as conclusões dos Grupos de discussão

12h30 – Almoço

**14h00 – Workshop: Planos de Gestão de Dados: ferramentas e requisitos dos financiadores**

- Pedro Príncipe, Serviços de Documentação da Universidade do Minho (OpenAIRE) [[Apresentação](#)]

**16h00 – Conclusão**



3 - Participantes no 1º Fórum GDI: foto de grupo

# RELATOS DOS GRUPOS DE DISCUSSÃO

Nos subcapítulos deste relatório seguintes apresentam-se resumidamente as principais questões apresentadas e debatidas pelos participantes nos grupos de discussão.

## Dados abertos e repositórios de dados

*Grupo de discussão sobre questões sobre a partilha e abertura dos dados de investigação, disponibilização em repositórios de dados disciplinares ou instituições, incluindo licenciamento e citação dos dados abertos)*

A introdução ao workshop foi feita com base na demonstração das características essenciais do Repositório Zenodo ([www.zenodo.org](http://www.zenodo.org) - repositório que resulta do projeto OpenAIRE e que é gerido pelo CERN), servindo de exemplo para explorar as componentes de um repositório no processo de partilha e licenciamento de conjuntos de dados. Esta demonstração foi realizada:

- elencando o passos do workflow de depósito de publicações e conjuntos de dados;
- identificando registos com conjuntos de dados em acesso aberto;
- mostrando a utilização que projetos de investigação (7ºPQ e H2020) fazem do repositório;
- demonstrando as valências das comunidades no Zenodo e a utilização feita por instituições, grupos de investigação, projetos e mesmo editores;
- destacando no processo de depósito de conjuntos de dados a associação de licenças e ligação ao projetos.

O conjunto dos participantes no grupo de discussão, composta por uma diversidade de perfis (Bibliotecários, Gestores de Ciências, Gestores de Informação e Repositórios, Responsáveis institucionais), desenvolveu um debate bastante participado com enfoque particular nas seguintes questões:

- como assegurar a privacidade dos dados e as o respeito por limitações éticas à abertura dos dados,
- quais as licenças necessárias para licenciamento de diferentes conjuntos de dados;
- quais as soluções institucionais de repositórios de dados abertos existentes no mercado;
- qual o modelo de utilização do Zenodo no contexto institucional português e como podem as bibliotecas usar a componente das comunidades no Zenodo.

No decorrer do debate foram feitas também demonstrações adicionais, nomeadamente do diretório de repositório de dados RE3data ([www.re3data.org](http://www.re3data.org)) e do Repositório internacional de ciência da vida PANGAEA (<https://pangaea.de>).

O grupo de participantes reunidos neste debate concluiu da necessidade de fomentar espaços como o do Fórum para desenvolver competências nestes domínios e mais informação das soluções existentes relativamente a repositórios de dados e licenciamento de dados abertos.

## **Armazenamento de dados, backup e preservação digital**

*Questões sobre a gestão ativa dos dados e serviços para armazenamento, backup e segurança, incluindo os desafios e riscos da preservação a longo prazo.*

O debate no seio do grupo de discussão sobre armazenamento de dados, backup e preservação digital incidiu sobre variadas questões, incluindo os suportes de armazenamento, o volume de armazenamento, os modelos de disposição das estruturas de armazenamento – centralizados, em rede, por tiers – a seleção do tipo e quantidade de dados que devem ser armazenados ou descartados, a disponibilidade e acessibilidade dos dados selecionados para preservação e o próprio conceito de preservação digital, bem como as políticas de preservação digital.

1) O início do debate pautou-se pelo consenso na visão que, atualmente, o volume de dados a armazenar não será um problema por causa dos últimos avanços tecnológicos, que permitem armazenar um petabyte de dados num simples bastidor. Além disso, no campo dos dados científicos, a maioria dos projetos gera quantidades limitadas de dados, havendo apenas uma percentagem muito pequena, que segundo estudos recentes é inferior a 10%, com capacidade de gerar terabytes.

- Manifestou-se preocupação pela enorme dependência nos suportes magnéticos, intrinsecamente frágeis, considerando-se uma alternativa promissora para o futuro o armazenamento em DNA, não sendo ainda porém plenamente viável.
- Discutiram-se as vantagens da replicação – cópia integral em disco – por oposição ao simples backup em tape, referindo-se a preferência pelo backup em tape para os casos de dados para os quais não há necessidade de reutilização num horizonte previsível ou mais frequente.
- Referiram-se as múltiplas condicionantes associadas à possibilidade de distribuição geográfica dos backups como os custos, a dimensão dos dados, que torna muito complexa a transferência dos dados após determinado volume, e a eventual necessidade de pós-processamento local, que implica considerar a configuração e performance (velocidade) dos sistemas locais. A combinação do tier storage com a distribuição geográfica do armazenamento pareceu ser uma solução atraente de backup, dada a probabilidade muito baixa de bit rot.

2) Num segundo momento, o debate centrou-se nas práticas de tratamento dos dados científicos, nomeadamente o seu armazenamento a partir do momento em que estão em condições de integrarem um acervo partilhado. Defendeu-se o planeamento e estruturação

de dados desde os primeiros passos da concepção da investigação, hábito que se considera não estar assimilado em largos setores das diferentes comunidades científicas nacionais, pelo que importa trabalhar no sentido de promover essa cultura. Inclusivamente foi citada a importância do armazenamento não só dos dados recolhidos mas dos próprios métodos e técnicas utilizados para sua produção e recolha (nalguns casos mais importantes do que os próprios dados obtidos pois estes, por vezes, podem-se facilmente reproduzir desde que a técnica esteja disponível).

- Um tópico que motivou troca acesa de argumentação foi o do próprio conceito de preservação digital. Onde é comum as empresas de engenharia informática pensarem num conceito de preservação para reutilização a 15, 10 ou menos anos, no campo das instituições culturais, históricas e de preservação do património, o horizonte temporal da disponibilidade dos objetos preservados estende-se até às centenas de anos. Importa pois distinguir entre dados em preservação e dados operacionais ou em acesso, mesmo que se esteja a aludir a um horizonte de 15 anos.
- Por não ser viável preservar todos os dados, dado o custo e exigência de recursos, deverão ser considerados dados para preservação aqueles com valor intrínseco, mesmo que esse valor e a previsão da utilidade futura dos dados seja extremamente difícil de definir.
- É fundamental a definição de políticas de preservação digital, que estabelecem critérios de preservação, definindo o que deve ou não ser preservado e a disponibilidade dos dados em acesso antes de passar para preservação. Foi sugerido o recurso a comissões disciplinares para a decisão acerca do tempo de acesso dos dados antes de passarem à condição de preservação, atendendo às necessidades específicas das diferentes áreas de investigação. As ações de seleção e preparação para preservação devem acontecer preferencialmente desde o momento da recolha. Em momento de consulta dos objetos preservados todas as ferramentas possíveis para a reutilização devem estar disponíveis, para reproduzir a experiência semântica associada ao objeto digital. É, por essa razão, fulcral disponibilizar ferramentas e proporcionar conhecimento técnico aos investigadores, pois a escolha e a preparação de dados para preservar requer, da parte do investigador (ou equipa) responsável, perspicácia para a opção por soluções agnósticas a tecnologias e por ferramentas de visualização adequadas, e precaução para evitar a utilização de software facilmente obsolecente ou tecnologias de nicho, menos interoperáveis.

3) O último tópico debatido foi a configuração e modelo nacional para as infraestruturas de dados, tendo sido sentida clara preferência por um modelo de infraestruturas em rede, evitando a criação de uma pesada, e potencialmente ineficaz, infraestruturas nacional centralizada. Uma rede nacional de infraestruturas de dados, regida por regras e orientações

de interoperabilidade quanto a interfaces, metadados e identificadores, deve obedecer aos princípios FAIR e estar ligada a outras iniciativas de dados científicos internacionais. Haveria porém espaço para uma infraestrutura nacional, não como ponto único de armazenamento e preservação de dados, mas como organismo de suporte à rede nacional e ponto de definição dos standards e de critérios de interoperabilidade comuns às infraestruturas que integram a rede. A opção preferencial seria então o modelo de infraestruturas descentralizadas em rede e o recurso a infraestruturas centrais apenas se necessário.

## **Organização dos dados, documentação e metadados**

Questões sobre práticas de organização, gestão doméstica dos dados e reutilização, descrição e documentação dos dados, incluindo questões de proteção e direitos de acesso.

A introdução que grupo de discussão esteve relacionada com o papel dos diferentes intervenientes na gestão de dados de investigação, tendo sido destacada a questão dos bibliotecários que se apresenta com dificuldade em perceber até onde podem ir no serviço que prestam aos investigadores. Sugere-se nesta questão a necessidade de envolver os vários stakeholders da gestão de dados de investigação, por exemplo através de ações de formação, mesmo para profissionais da informação.

No debate realizado no grupo destacam-se os seguintes tópicos:

Debate sobre os meios de publicação tradicionais e os novos desafios à publicação de dados.

- O número de iniciativas e projetos neste contexto é elevado o que dificulta a consolidação da gestão de dados de investigação, e o próprio posicionamento dos bibliotecários.
- Questiona-se sobre a falta de incentivos e falta de recursos para dar suporte a atividades associadas à Gestão de dados.
- Os benefícios da gestão de dados pouco são ainda pouco tangíveis. Torna-se necessário promover esses benefícios entre os intervenientes.
- Desenvolver maior conhecimento sobre a questão da publicação de dados sensíveis.

Desafios à partilha de dados na área das humanidades.

- Investigadores estão limitados na publicação em língua inglesa porque os dados perdem valor na tradução, sendo proposta a descrição parcial dos dados em língua inglesa para disseminação e promoção do trabalho.

Questões de ordem conceptual.

- Referida a dificuldade em lidar com conceitos da gestão de dados de investigação. Os metadados são um conceito pouco claro para os investigadores, apesar dos metadados até serem capturados de forma informal pelos investigadores.

- Existe um problema relacionado com a falta de descrição dos dados por parte dos investigadores. Dados que chegam ao fim do ciclo sem documentação que os contextualize; investigadores que anotam cadernos de laboratório. A sugestão é promover a descrição dos dados desde o momento em que estes são produzidos através da introdução de ferramentas para o efeito. Foi referida a este efeito a lista de vocabulários publicados pelo Research Data Alliance - o Metadata Standards Catalog.

Discussão sobre o repositório de dados ideal.

- Debate sobre a solução de publicação em repositórios disciplinares e genéricos. Concluiu-se ser difícil apontar uma solução ideal porque cada caso tem os seus requisitos. Existem pontos fortes e fracos em abordagens disciplinares e abordagens genéricas.

## Formação e desenvolvimento de competências

Questões sobre as necessidades de formação dos diferentes agentes do processo de gestão de dados de investigação, incluindo a identificação de competências e iniciativas a desenvolver.

Foram apresentados os seguintes tópicos na introdução ao tema do grupo de discussão

- Questão da formação, das necessidades de formação e das competências das pessoas que dão suporte aos processos de investigação.
- Aprendemos todos uns com os outros.
- Que competências são necessárias?
- Que iniciativas podem ser desenvolvidas?
- Reutilizar coisas que já existem?
- Criar coisas novas, por exemplo, no âmbito do Plano de Atividades do RCAAP?

De seguida apresentam-se as principais ideias apresentadas pelos participantes (todos Gestores de Ciência, com exceção de um bibliotecário) no decorrer do debate:

- A maior parte dos financiadores de Ciência começam a exigir um Plano de Gestão de Dados. Esta deveria ser uma competência a desenvolver. Saber o que é um Plano de Gestão de Dados.
- Necessidade de conhecer em detalhe os investigadores para poder dar respostas concretas. Competências enquanto investigador são igualmente necessárias.
- Colocar nas formações casos muito concretos. De uma investigação retirar dados muito concretos quanto a produtos.
- Necessidade de formação para poder elaborar um Plano de Gestão de Dados. Questão da apresentação de situações concretas de solicitações na formação. Tipificação de situações em função de financiadores/projetos/áreas.

- *Workshops hands-on*: como elaborar um Plano de Gestão de Dados para anexar a uma candidatura. *Workshop* com 5 ou 6 investigadores em que cada um traz o seu projeto, em que cada um escreve a sua parte do projeto e apresenta exemplos.
- Formação dirigida aos Centros/unidades de Investigação: identificar quem já está a trabalhar bem e pedir que dissemine a sua boa prática internamente podendo, contudo, existir também um formador externo.
- Questão da classificação dos metadados: como se classificam?
- *Data management* não pode ser feito de forma individualizada. Pelos exemplos apresentados no dia anterior (do Fórum) percebe-se que é necessária alguma coordenação. Tratamento dos dados pressupõe uma estrutura em que os dados vão ser alojados.
- Plano de Gestão de Dados está dependente das agências financiadoras (que está relacionado com os campos disciplinares).
- Ciência Aberta, Políticas de Dados, ambas são questões transversais. Questões como a Preservação são igualmente transversais.
- Plano de Gestão de Dados vem preencher uma lacuna. Mesmo que não seja exigido pelas agências financiadoras.
- Portugal tem uma grande vantagem: pegar nos bons exemplos e evitar os maus exemplos, podendo assim adotar boas práticas.
- A maior dificuldade que se coloca é conseguir que os investigadores entreguem os dados. Levanta a questão da apropriação dos dados: podemos fazer Planos mas os dados não são entregues.
- Necessidade de competência para “desmistificar a nebulosa”. Essa apropriação já acontece com o Acesso Aberto. Mas as publicações são públicas. Quanto aos dados trata-se de permitir que outros trabalhem com os meus dados.
- Questão da citação de dados: melhoraria a disposição para entregar dados.
- Necessidade de sensibilização, disseminação.
- Possibilidade de reter uma parte do financiamento até ao final do projeto, até ter a certeza que os dados estão curados.

O debate desenvolveu-se procurando os participantes responder à questão - o que podemos fazer em concreto?

- *Workshops* de introdução à gestão dos dados de investigação (MOOCs, recursos online, RCAAP, FOSTER) e *Workshops hands-on* com projetos em concreto.
- Grupo de trabalho para realizar um inquérito sobre necessidades de formação (por exemplo, quem poderia atuar como formador numa formação em *data management*).

- Referida a criação de um Grupo que pretende reunir os Gestores de Ciência do norte do País. Os Gestores de Ciência podem atuar como difusores de práticas junto dos investigadores.
- Usar a Semana Internacional do Acesso Aberto para promover a Ciência Aberta.
- MOOC sobre Ciência Aberta com vídeos disponibilizados na Semana da Ciência Aberta.
- Referido que na segunda metade do próximo anos (agora 2017) possibilidade de formação através de candidatura FOSTER (comunidade das ciências da vida; comunidade das ciências sociais; comunidade das humanidades e artes). Formação *hands-on* mas sujeita ainda (em setembro de 2016) a aprovação.
- Desenvolvimento de competências em:
  - Planos de Gestão de Dados
  - Conhecimentos especializados sobre a investigação que se está a produzir (tomar em consideração as entidades financiadoras/projetos/áreas)
  - Ciência Aberta
  - Metadados

Como considerações finais ao debate realizado no grupo foram destacadas as seguintes ações a desenvolver:

- Aproveitar plataformas *online* já existentes (como o RCAAP) para a realização de cursos que abordem aspetos genéricos.
- *Workshops hands-on* para aspetos mais práticos.
- Levantamento das necessidades de formação e também de formadores.
- Aproveitar a Semana do Acesso Aberto para a tornar, cada vez mais, a Semana da Ciência Aberta.
- Relativamente aos Planos de Gestão de Dados foi referida a necessidade de os mesmos serem tipificados tomando em consideração a entidade financiadora, os projetos e as áreas científicas.
- Necessidade de coordenação das várias iniciativas existentes no domínio da gestão dos dados científicos (existência de uma estrutura de coordenação).
- Dificuldade em conseguir que os investigadores entreguem os dados (estabelecida uma analogia com o que se passa com as publicações científicas, sendo que no caso dos dados a situação piora pois não pode ser usado o argumento da citação).

## WORKSHOP

A organização estabeleceu como objetivo do fórum procurar responder a necessidades concretas da comunidade de profissionais implicados no apoio à gestão de dados. Desde logo, a questão relacionada com os planos de gestão de dados foi considerada de grande relevância e pertinência em função dos requisitos do piloto de dados abertos no quadro do Horizonte 2020, mas igualmente das recomendações da FCT para os dados abertos no âmbito da política de Acesso Aberto.

Para esse efeito, realizou-se um workshop de hora e meia com o tema: Planos de Gestão de Dados: ferramentas e requisitos dos financiadores. Neste workshop, da responsabilidade de Pedro Príncipe (Universidade do Minho / OpenAIRE), foram desenvolvidos os seguintes tópicos:

- O que são planos de gestão de dados e para que servem
- Requisitos da Comissão Europeia no âmbito do H2020
- Ferramentas para elaboração de planos de gestão de dados
  - DMPonline
  - TemplateH2020

## PARTICIPAÇÃO

O evento decorreu das 09h30 às 16h00 verificando-se um elevado interesse e participação - 87 é o número total de participantes. Colaboraram na organização dos grupos de discussão 4 moderadores e 4 relatores, um orador na realização do workshop e sete oradores de comunicações 24x7.

### Tabela resumo

	Conf. Dados	Fórum-GDI
<b>Participantes</b>	170	87
<b>Inscritos que não compareceram</b>	71	34

## CONSIDERAÇÕES FINAIS

A realização do 1º Fórum de Gestão de Dados de Investigação foi genericamente considerado um sucesso, contribuindo para identificar as principais necessidades dos diferentes agentes envolvidos no apoio aos processos de gestão de dados de investigação. Conseguiu-se ainda providenciar as ferramentas necessárias para o desenvolvimento de planos de gestão de dados, sendo que com as apresentações de projetos e iniciativas emergentes se contribuiu já para a capacitação dos participantes.

Decorrente da avaliação deste primeiro fórum, considera-se importante prosseguir esta iniciativa, procurando assim dinamizar e criar uma comunidade de profissionais que desenvolvem atividade de suporte à gestão de dados nas instituições de investigações em Portugal. A continuidade deste tipo de iniciativa deve procurar corresponder à dimensão prática do evento, assegurando igualmente uma forte componente de interatividade.

No sentido de colmatar o relativo atraso no desenvolvimento de serviços para a gestão de dados de investigação nas instituições e o relativo distanciamento dos responsáveis instituições para a aplicação de estratégias neste domínio, propõe-se que no ano de 2017 se realizem dois eventos seguindo a mesma linha de componente prática e interativa seguida neste 1º Fórum.

Para o futuro, propõe-se que o Fórum GDI se configure como um evento anual, que dinamize uma comunidade diversa de profissionais que apoiam a gestão de dados nas instituições de investigação (gestores de repositórios digitais e data centers, técnicos de informação, bibliotecas, arquivos e curadoria de dados, especialistas de informática, investigadores, cientistas de dados e gestores de ciência de instituições de investigação e organismos de financiamento de ciência), e que estas realizações sejam espaço de promoção do debate e partilha de ideias, projetos, iniciativas emergentes e boas práticas relacionadas com a Gestão de Dados de Investigação e desenvolvimento de competências.