

FABETIZAÇÃO DE DADOS PARA PESQUISADORES DATA LITERACY FOR RESEARCHERS



Direitos reservados a Makson Reis



Apresentação

- **BIBLIOTECÁRIO E DOCUMENTALISTA-UFS**
- **GRADUANDO EM ADMINISTRAÇÃO- UNIASSELVI**
- **ME. EM GESTÃO DA INFORMAÇÃO E DO CONHECIMENTO-UFS**
- **DOUTORANDO EM CIÊNCIA DA INFORMAÇÃO -UFBA**

MAKSON DE JESUS REIS

PORQUE APRENDER DADOS ? WHY LEARN DATA?

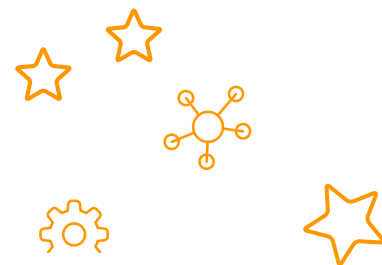
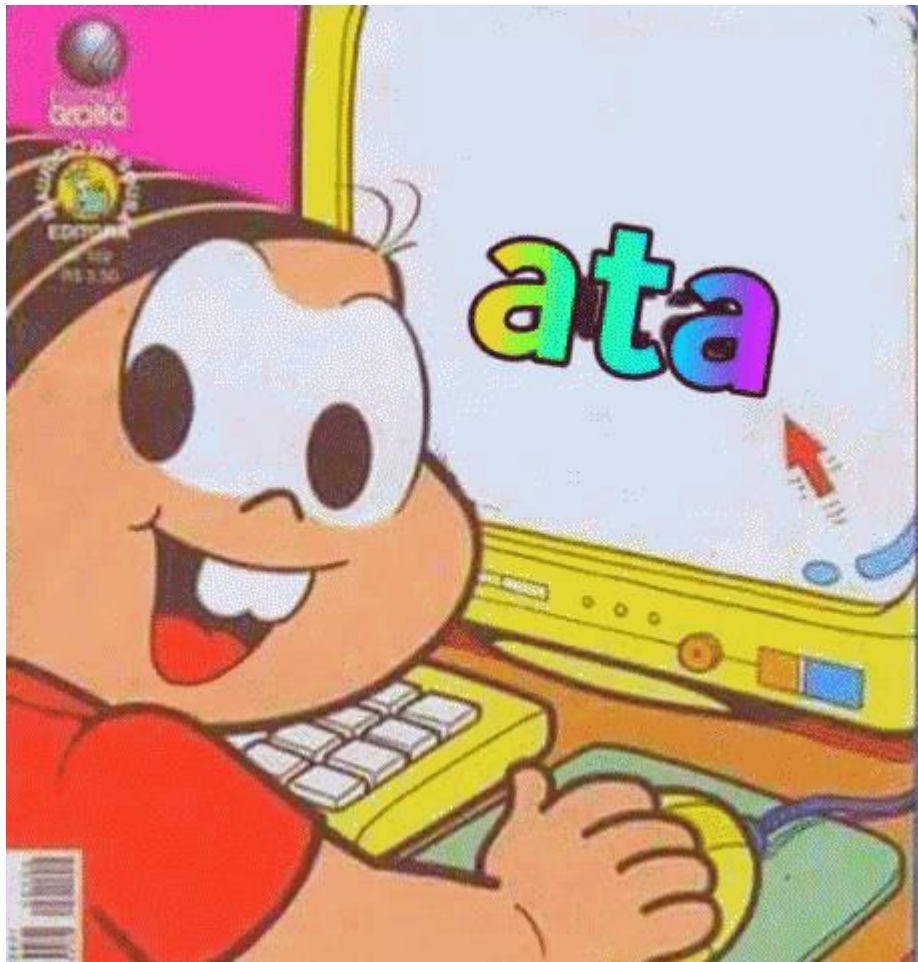




Obtenha as habilidades de dados
que sua escola não lhe deu



Entenda o que os números
realmente significam

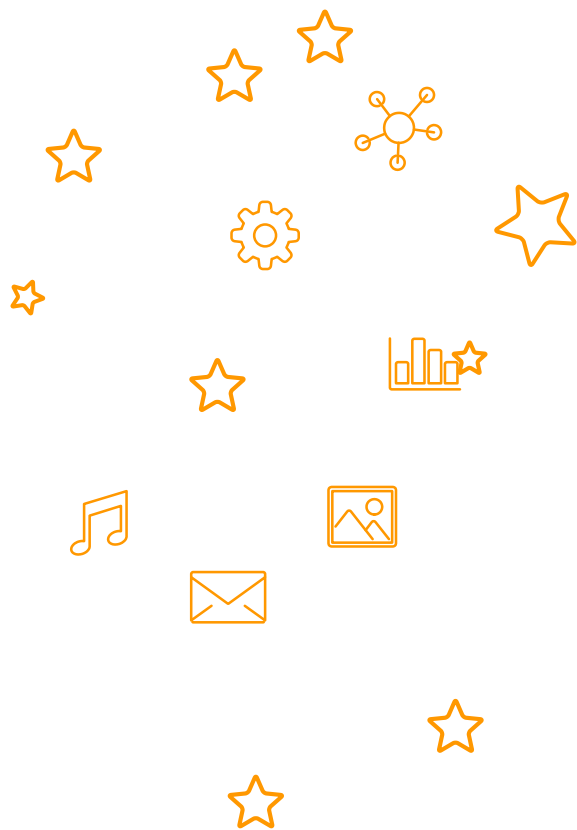


Compreender os conceitos
estatísticos subjacentes



Ajude sua equipe a colocar os
dados em bom uso

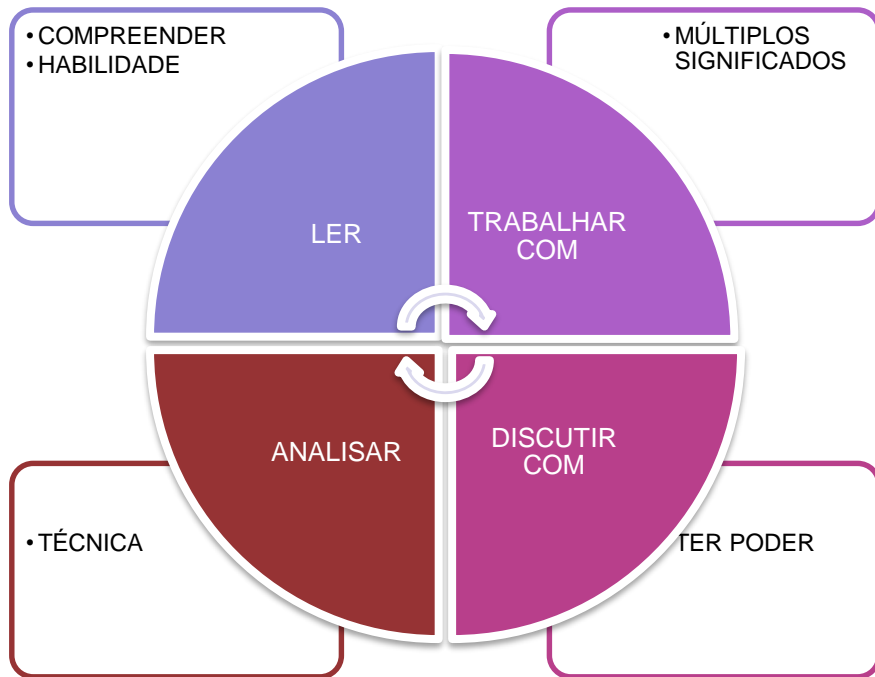
ENTENDENDO ALFABETIZAÇÃO DE DADOS ? UNDERSTANDING DATA LITERACY?





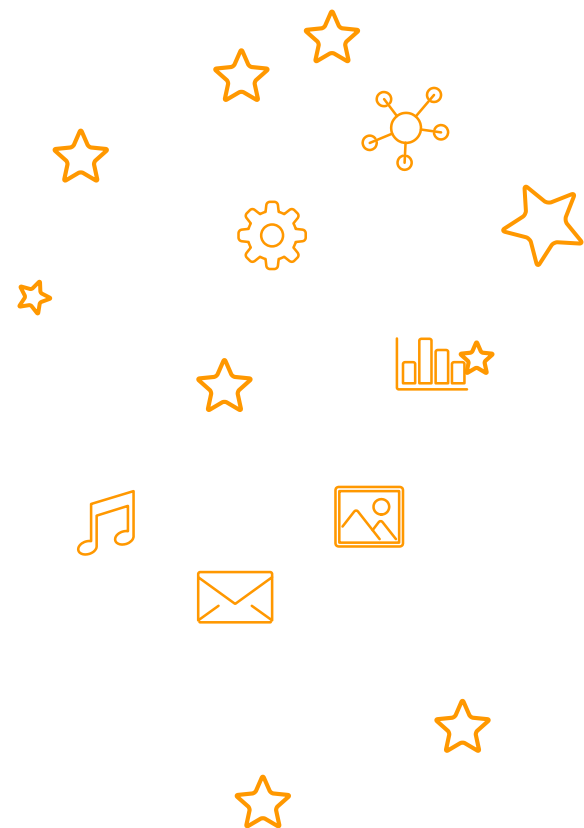
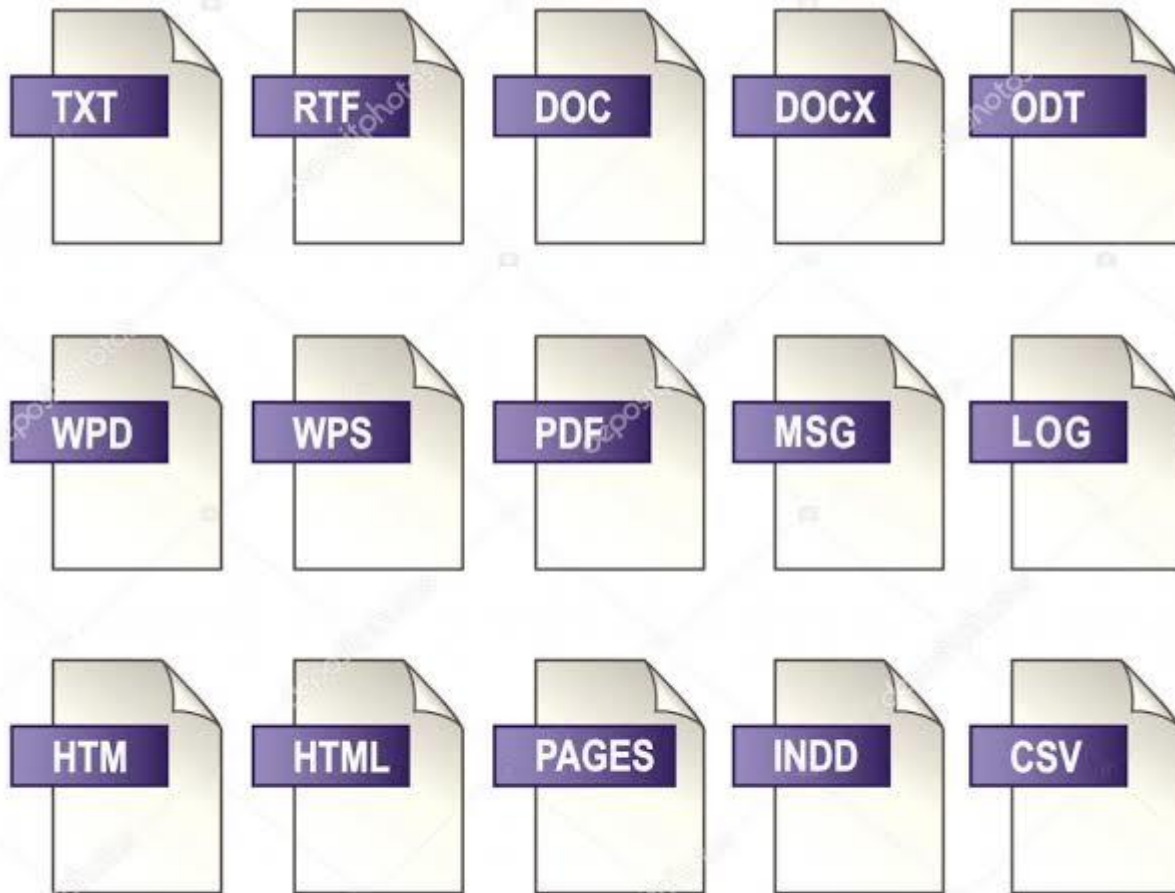
DEFINIÇÃO DE ALFABETIZAÇÃO DE DADOS

A CAPACIDADE DE LER, TRABALHAR, ANALISAR E DISCUTIR COM OS DADOS.



**Você fala a língua
dos dados?
Do you speak the
language of the
data?**





	A	C	D
1	Valor Fórmula		Resultado
2	1234,56	=TEXTO(A2;"0,00")	1234,56
3	1234,56	=TEXTO(A3;"#.##0")	1.235
4	-1234,56	=TEXTO(A4;"#.##0,00")	-1.234,56
5	1234,56	=TEXTO(A5;"R\$#.##0")	R\$ 1.235
6	1234,56	=TEXTO(A6;"R\$#.##0,00")	R\$ 1.234,56
7	-1234,56	=TEXTO(A7;"R\$#.##0,00_");(R\$#.##0,00)"	(R\$ 1.234,56)
8	1234,56	=TEXTO(A8;"R\$ * #.##0")	R\$ 1.235
9	1234,56	=TEXTO(A9;"R\$ * #.##0,00")	R\$ 1.234,56
10	-1234,56	=TEXTO(A10;"R\$ * #.##0,00_");(R\$ * #.##0,00)"	(R\$ 1.234,56)

Número

Moeda

Contábil

FORMATOS DE IMAGEM



“

Bons Hábitos de dados a partir do momento em que você começa a planejar sua pesquisa”

Good Data Habits from the moment you start planning your research. ”





CARACTERÍSTICAS DE UMA CULTURA COM CONHECIMENTO DE DADOS

CHARACTERISTICS OF A DATA KNOWLEDGE CULTURE

Figura 7 - Matriz de Competências de Alfabetização de dados

Área de Conhecimento		Competência		LEGENDA:		Competências Conceituais	Competências Essenciais	Competências Avançadas
Estrutura Conceitual	Introdução aos Dados	Conhecimento e compreensão dos dados		Conhecimento e compreensão dos usos e aplicações de dados				
	Descoberta de dados e Coleta	Realiza a exploração de dados		Identifica dados úteis		Coleta de dados		
Gerenciamento de dados	Avaliando e assegurando dados e fontes	Avalia criticamente fontes de dados para Confiabilidade		Avalia criticamente a qualidade de conjuntos de Dados para erros ou problemas				
	Organização de dados	Conhecimento da organização básica de dados métodos e ferramentas		Acesso e requisitos de organização de dados		Organizar dados		
	Manipulação de dados	Acesso e métodos para limpar dados		Identifica os valores e as anomalias		Limpa dados		
	Conversão de dados (de formato para formatar)	Conhecimento de diferentes tipos de dados e métodos de conversão		Atribui os descritores de metadados apropriados para conjuntos de dados originais				
	Criação e uso de metadados	Cria descritores de metadados		Converter dados de um formato ou tipo de arquivo para outro				
	Curadoria de Dados	Avalia os requisitos de Curadoria de dados (por exemplo, cronograma de retenção, armazenamento, acessibilidade, requisitos de partilha, etc.)		Avalia os requisitos de segurança de dados (por exemplo, acesso restrito, unidades protegidas, etc.)		Curadoria de dados		
	Preservação de dados	Avalia os requisitos de preservação		Métodos e ferramentas de acesso para preservação de dados		Preserva os dados		
	Ferramentas de dados	Conhecimento de ferramentas de análise de dados e técnicas		Seleciona a ferramenta de análise de dados apropriada ou técnica		Aplica ferramentas de análise de dados e técnica		
Avaliação de dados	Análise básica de dados	Desenvolve planos de análise		Aplica métodos e ferramentas de análise		Realiza análise exploratória	Avalia os resultados da análise	Compara os resultados da análise com outros achados
	Interpretação de dados (Noções básicas sobre dados)	Lê e compreende gráficos, tabelas		Identifica os principais pontos e integra isso com outras informações importantes		Identifica discordância dentro os dados		
	Identificando problemas usando dados	Usa dados para identificar os problemas em situações práticas (por exemplo eficiência no local de trabalho)		Usa dados para identificar problemas de nível superior (por exemplo, política, ambiente, experimentação científica, marketing, economia)				
	Visualização de dados	Cria tabelas significativas para organizar e apresentar visualmente dados		Cria representações gráficas significativas de dados		Avalia eficácia de representações gráficas	Avalia criticamente representações gráficas para precisão e deturpação de dados	
	Apresentando dados (verbalmente)	O resultado pretendido (s) para apresentar os dados		Avalia as necessidades do público e a familiaridade com sujeito (s)		Planeja a reunião apropriada ou tipo de apresentação	utiliza tabelas significativas e visualizações para comunicar dados	Apresenta argumentos e/ou resultados claros e coerentemente
	Decisões orientada a dados	Prioriza a informação obtida a partir de dados		Converter dados em informações		Pesa o mérito e os impactos de possíveis soluções/decisões	Implementa decisões/soluções	
	Pensamento Crítico	Ciente de questões de alto nível e desafios associados com dados		Pensar criticamente ao trabalhar com dados				
	Cultura de dados	Reconhece a importância dos dados		Apoia um ambiente que promove o uso crítico de dados para a aprendizagem, pesquisa e decisão				
Aplicação de dados	Ética de dados	Consciente das questões jurídicas e éticas, associado aos dados		Aplica e trabalha com dados de forma ética				
	Citação de dados	Conhecimentos de dados amplamente aceitos, métodos de citação		Cria citações corretas para conjunto de dados secundários				
	Compartilhamento de dados	Avalia métodos e plataformas para compartilhamento de dados		Compartilha dados legalmente, e eticamente				
	Avaliando decisões com base em dados	Coleta de dados de acompanhamento para avaliar a efetividade de decisões ou soluções com base em dados		Conduz a análise de dados de acompanhamento		Compara os resultados da análise com outros achados	Avalia decisões ou soluções com base em dados	Mantém as conclusões ou decisões originais, ou implementa novas decisões/soluções

Fonte: traduzido pelo autor (2019).



Web sites para alfabetização em dados

MANTRA
Research Data Management Training

MANTRA is a free online course for those who manage digital data as part of their research project.

Home About Acknowledgements DIY Training Kit for Librarians Feedback Contact Us

Learning Units: Select one to start

- Research data explained >
- File formats & transformation >
- Data protection, rights & access >
- Data management plans >
- Documentation, metadata, citation >
- Sharing, preservation & licensing >
- Organising data >
- Storage & security >
- Data handling tutorials >

EDINA | The University of Edinburgh

Mantra Privacy Notice | EDINA Cookies | Website Accessibility | Last update: May 2018

The University of Edinburgh is a charitable body, registered in Scotland, with registration number SC005336. VAT Registration Number GB652950760

FOSTER About Resources Events Courses News Search for...

The future of science is Open

START YOUR RESEARCH TRAINING NOW

Thank you for your feedback on the [tools kit](#). The period of public consultation is now closed and the courses will be amended accordingly.

USE FOSTER TO:

- Access Free Courses
- Earn Badges
- Participate

OFICINA

1. PLANEJAMENTO DE GERENCIAMENTO DE
2. ORGANIZANDO E DOCUMENTANDO DADOS
3. ARMAZENAMENTO DE DADOS E SEGURAN
4. ÉTICA E DIREITOS AUTORAIS
5. COMPARTILHAMENTO DE DADOS
6. AVALIAÇÃO DA OFICINA





OFICINA



- 1. DATA MANAGEMENT PLANNING**
- 2. ORGANIZING AND DOCUMENTING DATA**
- 3. DATA STORAGE AND SECURITY**
- 4. ETHICS AND COPYRIGHT**
- 5. DATA SHARING**
- 6. WORKSHOP EVALUATION**



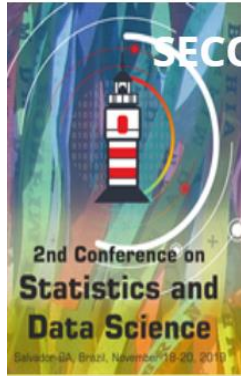
OFICINA -OBJETIVOS

- Facilidade para profissionais trabalharem com dados
- Exercícios Interativos em grupo
- Exercícios de lição de casa para reforçar aprendizagem
- Escrita Reflexiva para pensar



OFICINA -OBJETIVOS

- Ease for professionals to work with data
- Interactive group exercises
- Homework exercises to reinforce learning
- Reflective writing to think



O QUE SÃO DADOS ? WHAT ARE DATA?



O QUE SÃO DADOS?

- Dados digitais são tudo o que é criado ou manipulado em um computador:
- Correspondência de e-mail
- Arquivos de texto
- Imagens - de digitalizações digitais de objetos físicos a fotos e imagens 3D complexas
- Áudio



O QUE SÃO DADOS

- Vídeo
- Planilhas e bancos de dados - dados numéricos e textuais
- Dados da pesquisa - desde pesquisas simples sobre EDM até análises Lidar e pesquisas geofísicas
- Mensagens de texto
- Sites - até o Youtube pode ser um dado de pesquisa.
- Etc...



PROBLEMAS COM OS DADOS?



PROBLEMAS COM DADOS

- VOLUME
 - Seleção e retenção: o que você joga fora e o que guarda
- MANEJO E PRESERVAÇÃO
 - Fragilidade de dados digitais, formatos e custos
- LEGAL
 - Autenticidade dos dados digitais
 - Direitos autorais (material online por exemplo: tese eletrônica) ou dados confidenciais (dados pessoais)



O QUE SÃO DADOS

- REUSO
- Conhecendo os detalhes técnicos dos dados digitais
- Compreendendo o contexto de dados digitais

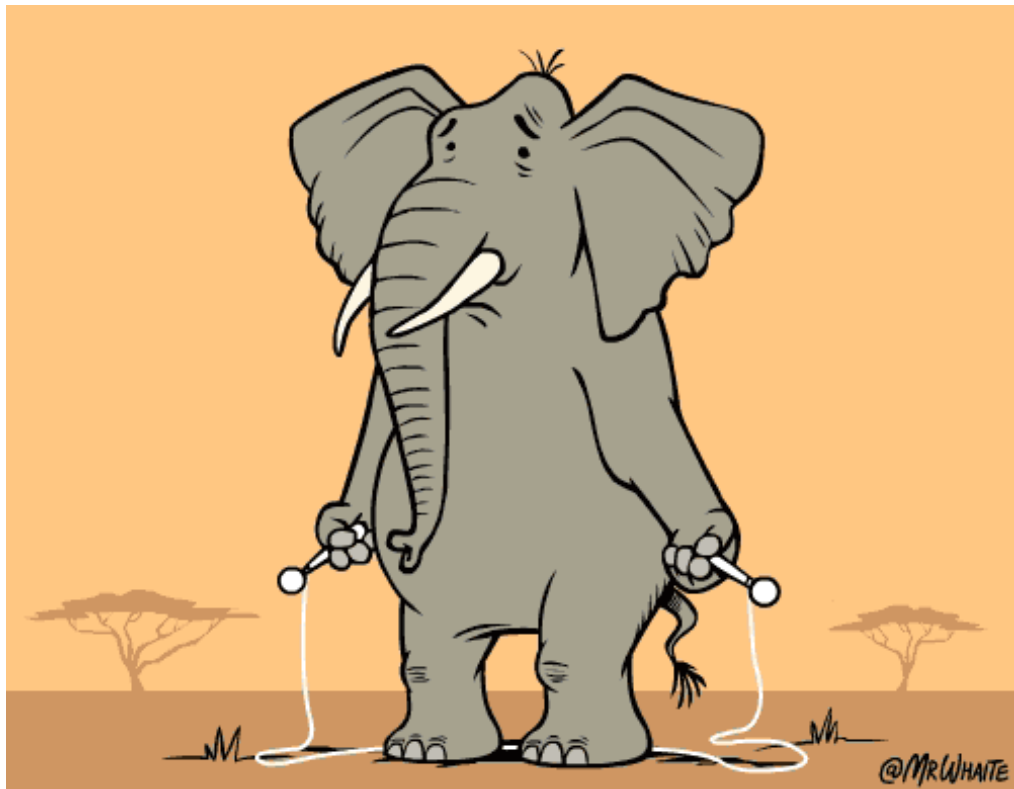


ABORDAGEM AO GERENCIAMENTO DE DADOS DATA MANAGEMENT APPROACH



ABORDAGEM AO GERENCIAMENTO DE DADOS

- Entenda os problemas maiores relacionados ao uso e arquivamento de dados digitais;
- Conhecer os requisitos e recursos de gerenciamento de dados de sua universidade;
- Estar ciente dos recursos on-line nacionais no gerenciamento de dados de pesquisa;
- Esteja preparado para o gerenciamento de dados no mundo real.



EXERCÍCIO 1

EXERCISE 1



EXERCÍCIO 1

- Qual é o conjunto de dados principais do seu projeto de pesquisa?
- O que acontecerá com os dados: dados físicos e digitais?
- Você tem autoridade para arquivar os dados digitais em que trabalha?
- Algum problema ético ?

Defining Post-Graduate Digital Research Data			
Name:			
Research Topic:			
Laboratory:			
Individual research or part of project/institution (e.g. museum)			
Research / Project Code (if any):			
Physical data: artefacts, samples, paper archives, etc.	Data origin: Where's it from?	Digitally captioned data: Scans, spreadsheets, etc.	End point: Where the data (physical and digital) goes after the project?
Digitally created data: survey data, images, analysis readouts, types of text docs, etc.			End point: Where the digital data goes after the project?
Looking after your data:	1. How do you organise your data?	2. How do you back-up your data?	3. Your most important data management issue?
1.			
2.			
3.			
Any other issues?			

- What is the core data set for your research project?
- What will happen to data: physical and digital data?
- Do you have the authority to archive the digital data you work for?
- Any ethical issues?



DISCUSSÃO DO EXERCÍCIO 1 DISCUSSION OF EXERCISE 1

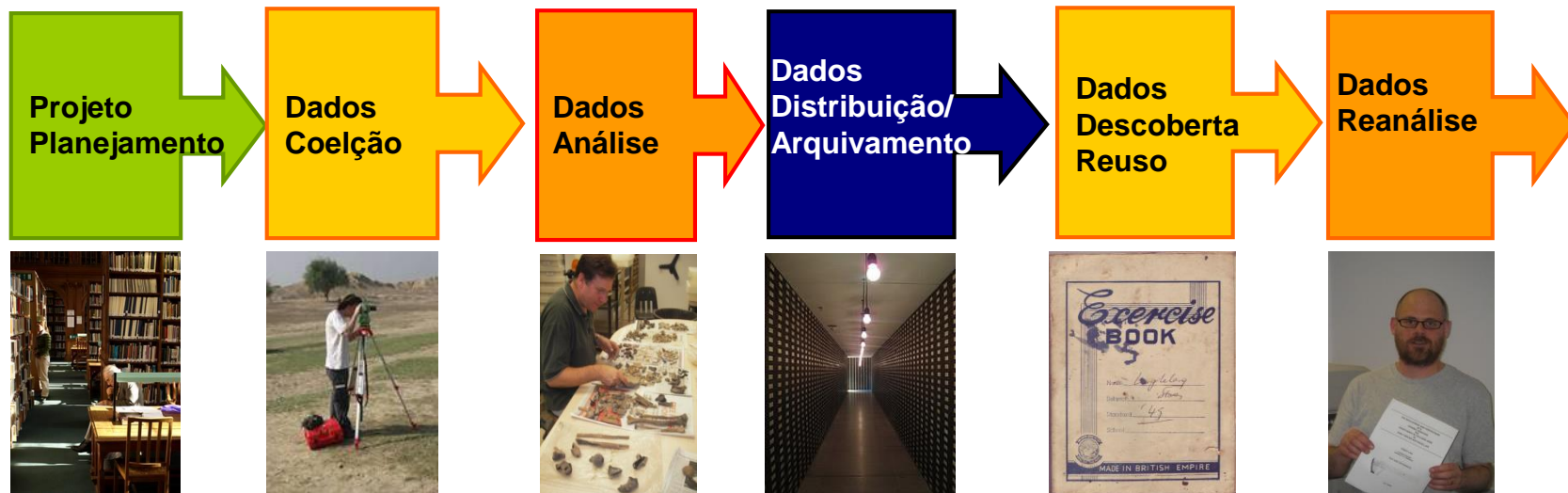


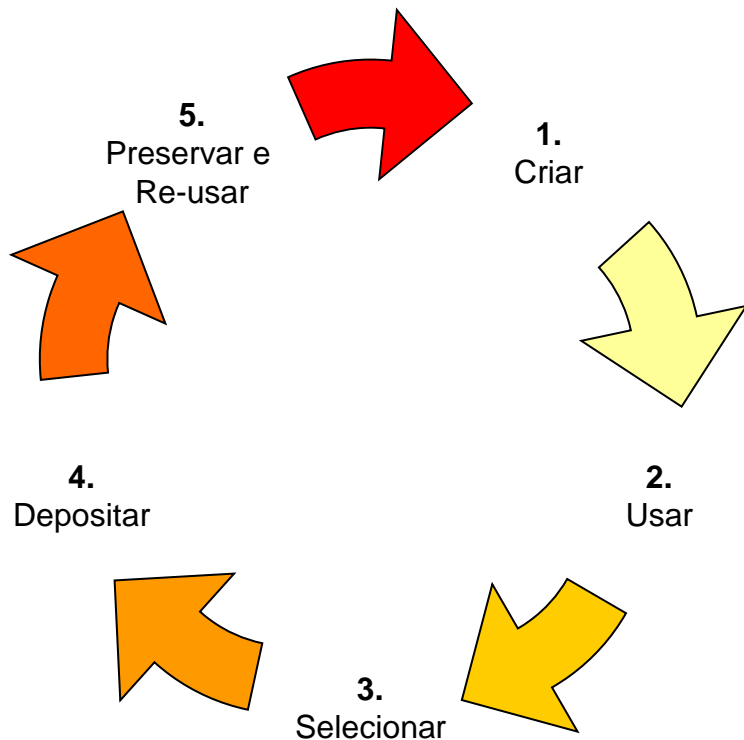
EXERCÍCIO 1

- Pontos chave
- Pense nos dados digitais no início do planejamento do projeto
- O gerenciamento de dados anda de mãos dadas com os resultados da pesquisa
- Torne seus dados de pesquisa compreensíveis para outras pessoas

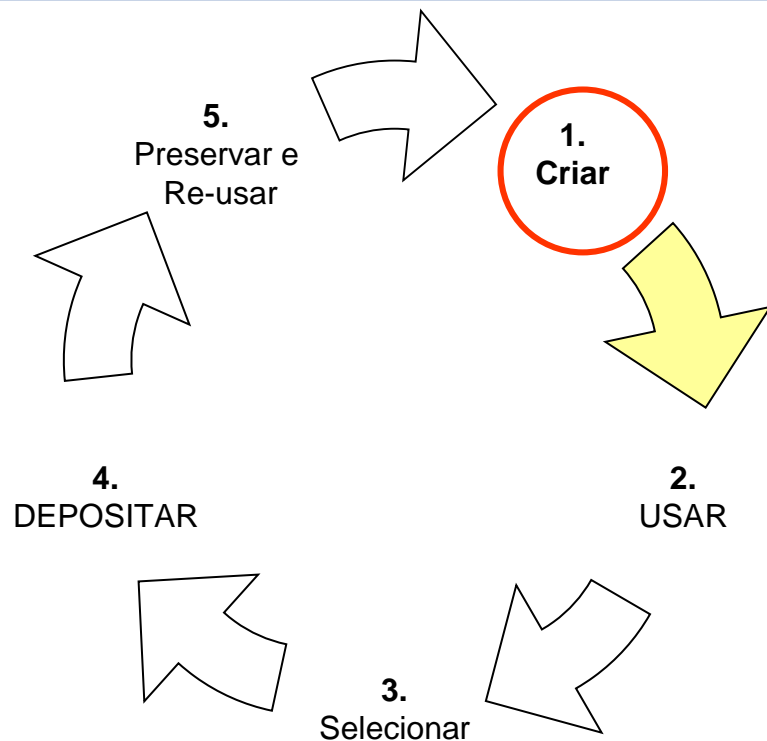


CICLO DE VIDA DOS DADOS

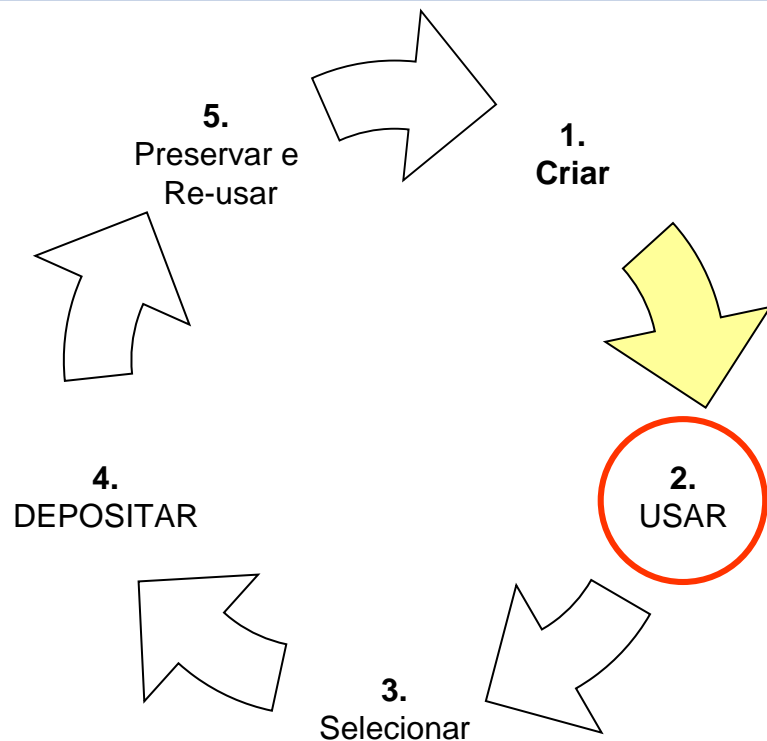




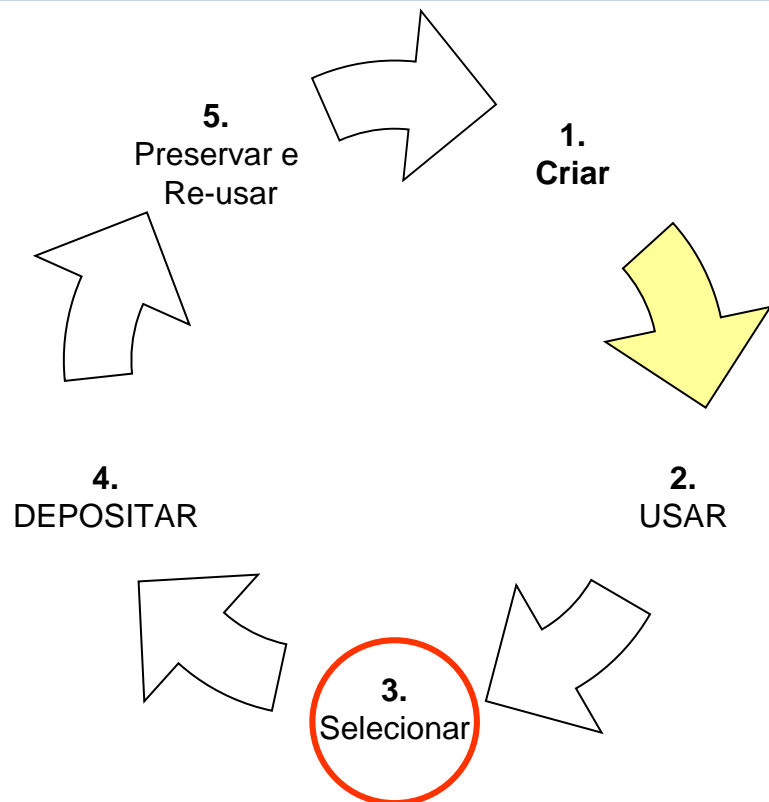
1. Quais dados vou produzir?
2. Como vou organizar os dados?
3. Meu gerenciamento de dados está funcionando bem? Quais dados vou manter?
4. Quais dados serão depositados e onde?
5. Quem estará interessado em reutilizar os dados?



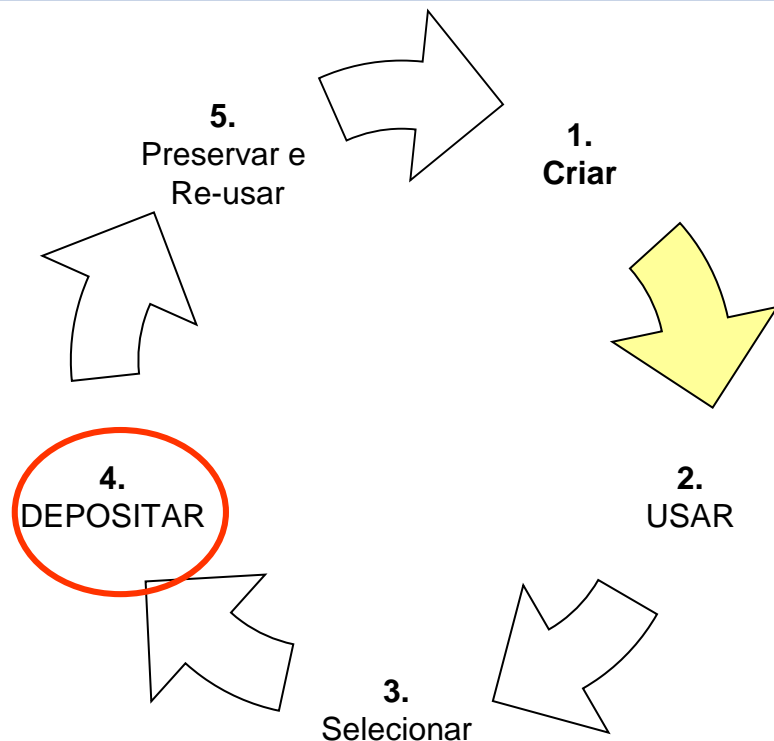
1. Quais dados vou produzir?
2. Documentos de texto
3. Análise de artefatos
4. Análise da amostra
5. Dados de pesquisa
6. Desenhos
7. Fotografias
8. Entrevistas gravadas
9. Etc ..



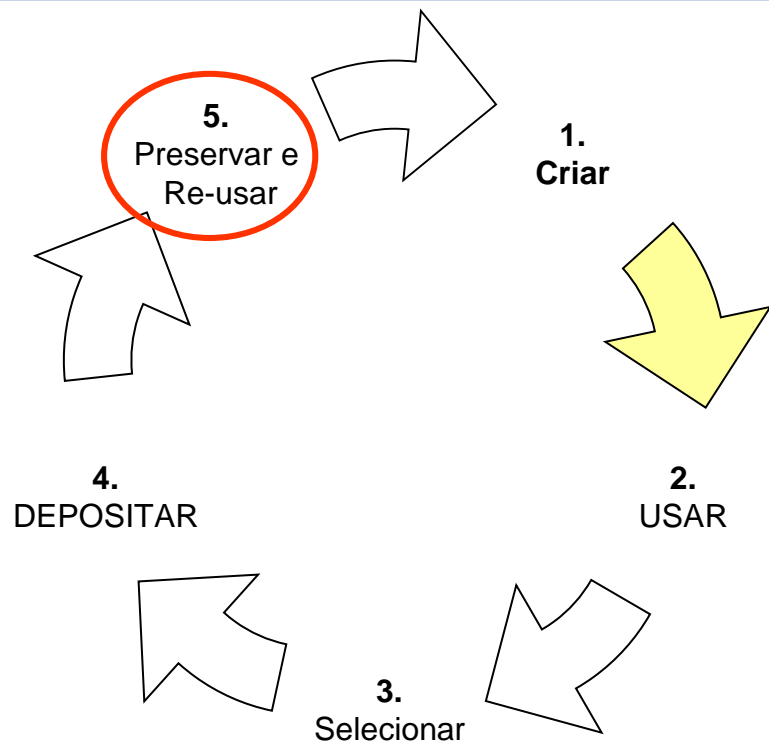
1. Estrutura de arquivo
2. Nomeação de arquivo
3. Quais formatos de arquivo vou usar?
4. Qual software vou usar?
5. Aproximadamente quantos arquivos?
6. Como vou descrever e documentar meus dados?



1. Avaliando o gerenciamento de dados:
2. A estrutura / nomeação de arquivos é compreensível para outras pessoas?
3. São necessários dados adicionais?
4. São necessários novos tipos de dados?
5. Quais dados serão mantidos?
6. Quais dados podem ser descartados?



1. Quais dados serão depositados e onde?
2. Definir o conjunto de dados principais do projeto
3. Quais dados serão incluídos na tese?
4. Quais dados são complementares?
5. Vou produzir uma tese eletrônica?
6. Onde vou depositar minha tese eletrônica?
7. Vou depositar dados suplementares?



1. Preservação e Reutilização
2. Quem estará interessado em reutilizar os dados?
3. Existem informações suficientes para permitir a fácil reutilização dos dados?



EXERCÍCIO 2

- O que você quer / tem a ver com seus dados de pesquisa após concluir seu doutorado?
- Algum dos dados é sensível?
- Quem é o proprietário dos seus dados de pesquisa e você terá autoridade para arquivá-los?



EXERCÍCIO 2

- Descreva qualquer material protegido por direitos autorais que você planeja incluir em sua tese, p. imagens.
- Você trabalhará com formatos de arquivo incomuns ou grandes?
- Descreva o conjunto de dados principal de sua pesquisa.



PLANEJAMENTO DE GERENCIAMENTO DE DADOS



OFICINA -OBJETIVOS

- Por que o RDM é importante?
- O que é o Gerenciamento de Dados de Pesquisa (RDM)?
- O que os financiadores esperam?
- O que a UEFS espera?
- Criando Planos de Gerenciamento de Dados Questões?



O QUE É GERENCIAMENTO DE DADOS DE PESQUISA?

- Um termo abrangente para descrever todos os aspectos do planejamento, organização, documentação, armazenamento e compartilhamento de dados. Também leva em conta questões como proteção de dados e confidencialidade.



O QUE É GERENCIAMENTO DE DADOS DE PESQUISA?

- Ele fornece uma estrutura que suporta pesquisadores e seus dados ao longo de suas pesquisas e além.

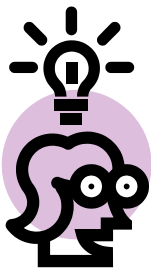


O QUE É GERENCIAMENTO DE DADOS DE PESQUISA?

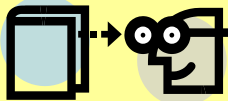
Dados
Gestão
Planejamento

Gerenciamento de dados de
pesquisa
Plataforma

Repositório Nacional
Jornal Eletrônico ou Portal da
Comunidade



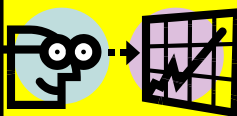
Conceber



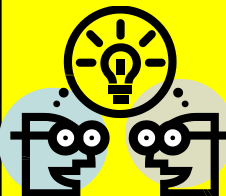
Design



Experime
ntar



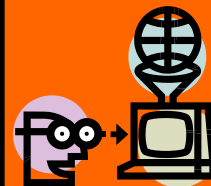
Analisar



Colaborar



Publicar



Expor



O QUE É GERENCIAMENTO DE DADOS DE PESQUISA?

O bom GD oferece vários benefícios para os pesquisadores e sua instituição, garantindo: pesquisadores atendem aos requisitos do financiador / universidade / indústria os dados são precisos, completos autênticos e confiáveis - de acordo com a boa prática de pesquisa integridade e replicação de pesquisa segurança de dados e minimizar risco de perda Maior eficiência - economizando tempo e recursos os dados estão disponíveis para uso futuro



CRIANDO PLANOS DE GERENCIAMENTO DE DADOS?

- Tipos de dados, formatos, padrões e métodos de captura
- Ética e Propriedade Intelectual
- Acesso, compartilhamento de dados e reutilização
- Armazenamento de Curto Prazo e Gerenciamento de Dados
- Depósito e Preservação a Longo Prazo
- Recursos
- Mas irá variar de acordo com as perguntas e requisitos exatos do aplicador a ser aplicado.



CRIANDO PLANOS DE GERENCIAMENTO DE DADOS?

- Tipos de dados, formatos, padrões e métodos de captura
- Quais saídas de dados sua pesquisa gerará?
- Descrever volume, tipo, conteúdo, qualidade e formato do conjunto de dados final
- Delinear os metadados, documentação ou outro tipo de material de apoio que deve acompanhar os dados para que sejam interpretados corretamente
- Quais padrões e metodologias serão utilizados para coleta e gerenciamento de dados



CRIANDO PLANOS DE GERENCIAMENTO DE DADOS?

- Ética e Propriedade Intelectual
- Demonstrar que você procurou orientação e abordou todos os problemas de direitos autorais e gerenciamento de direitos que se aplicam ao recurso
- Fazer menção explícita de consentimento, confidencialidade, anonimização e outras considerações éticas, quando apropriado



CRIANDO PLANOS DE GERENCIAMENTO DE DADOS?

- São necessárias restrições ao compartilhamento de dados - por exemplo, para proteger os participantes da pesquisa ou para obter proteção adequada à propriedade intelectual?
- Os pesquisadores devem ser capazes de apresentar um argumento forte para quaisquer restrições ao compartilhamento
- Garantir que eles tenham toda a aprovação ética necessária em vigor
- Esclarecer problemas relacionados à propriedade de dados



CICLO DE VIDA DOS DADOS





Acesso, compartilhamento de dados e reutilização

- Acesso, compartilhamento de dados e reutilização
- Quais são os usos de pesquisa pretendidos e / ou previstos para o (s) conjunto (s) de dados completo (s)?
- Como você tornará o recurso acessível ao público potencial identificado.
- Onde você disponibilizará os dados?
- Como outros pesquisadores poderão acessar os dados?



Acesso, compartilhamento de dados e reutilização

- Será necessário um acordo de compartilhamento de dados?
- Qual é o cronograma para o lançamento público dos dados?



Acesso, compartilhamento de dados e reutilização

- Afirme quaisquer dificuldades esperadas no compartilhamento de dados, juntamente com causas e possíveis medidas para superar essas dificuldades.
- Como o compartilhamento de dados fornecerá oportunidades para coordenação ou colaboração?
- Antecipar e planejar a reutilização de dados



Acesso, compartilhamento de dados e reutilização

- Tranquilizar os financiadores por serem muito claros sobre onde, quando e como os dados serão disponibilizados
- Use a infraestrutura existente, por ex. repositórios institucionais, nacionais ou internacionais



Acesso, compartilhamento de dados e reutilização

- VIDEO
- <https://www.youtube.com/watch?v=WHrrLP6v9NI&list=P L2XF5RiVI7GPXEhdyIRmJjjSXXjs0QpT7&index=2>



RECURSOS

- Recursos
- Quais recursos você precisará para entregar seu plano?
- Descrever hardware, software e conhecimentos técnicos adicionais, suporte e treinamento que provavelmente serão necessários e como serão adquiridos



Vídeo

- LINKS
- https://www.youtube.com/watch?v=PXr14Urf268&list=P_L2XF5RiVI7GOyWMmvh9Lcyi1JZ0rJ-Ldo



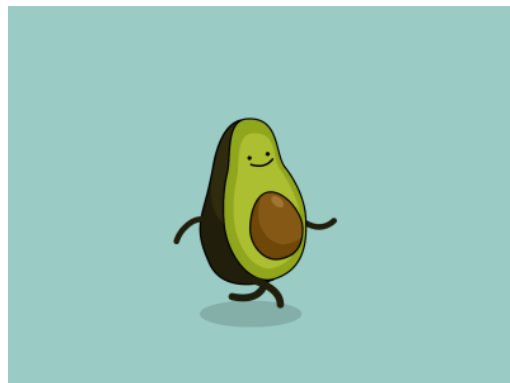
RECURSOS

- Delinear e justificar os custos
- Seja realista quanto ao tempo e esforço humanos requeridos
- Mostre que os fundos serão utilizados de forma eficiente e eficaz



Vídeo

- LINKS
- https://www.youtube.com/watch?v=FLL74eXcl8c&list=P_L2XF5RiVI7GNz2TSe3kG9uHd3UtsiV_t6



ATIVIDADES



ORGANIZANDO E DOCUMENTANDO DADOS



ORGANIZANDO SEUS DADOS

- A RDM é uma das áreas essenciais da conduta responsável da pesquisa.
- Os arquivos e pastas de dados de pesquisa precisam ser organizados de maneira sistemática para serem:
- identificável e acessível para você,
- identificáveis e acessíveis para os colegas e para futuros usuários.



ORGANIZANDO SEUS DADOS

- Por isso, é importante planejar a organização dos seus dados antes do início de um projeto de pesquisa.
- Isso evitará qualquer confusão enquanto a pesquisa estiver em andamento ou quando vários indivíduos estiverem editando e / ou analisando os dados.



CONVENÇÕES DE NOMENCLATURA DE ARQUIVOS

- Nomear conjuntos de dados de acordo com as convenções acordadas deve facilitar a nomeação de arquivos para os colegas, porque eles não precisarão "repensar" o processo a cada vez.
- Os nomes de arquivos devem fornecer contexto para o conteúdo do arquivo, tornando-o distinguível de arquivos com assuntos semelhantes ou versões diferentes do mesmo arquivo.



CONVEÇÕES DE NOMENCLATURA DE ARQUIVOS

- Muitos arquivos são usados independentemente de sua estrutura de arquivos ou diretórios, portanto forneça uma descrição suficiente no nome do arquivo.
- Estratégias sugeridas: identificar o projeto; evite caracteres especiais; use sublinhados, em vez de espaços; incluir data de criação ou modificação em um formato padrão (por exemplo, YYYY_MM_DD ou YYYYMMDD): use o número do projeto
- Ser consistente! Evite ser enigmático!



ORGANIZANDO DADOS

- Também há vários elementos comuns que devem ser considerados ao desenvolver uma estratégia de nomeação de arquivos, incluindo:
 - Número da versão
 - Data da criação
 - Nome do criador
 - Descrição do conteúdo
 - Nome da equipe / departamento de pesquisa associado aos dados
 - Data da publicação
 - Projeto número



ORGANIZANDO DADOS

- Também há vários elementos comuns que devem ser considerados ao desenvolver uma estratégia de nomeação de arquivos, incluindo:
 - Número da versão
 - Data da criação
 - Nome do criador
 - Descrição do conteúdo
 - Nome da equipe / departamento de pesquisa associado aos dados
 - Data da publicação
 - Projeto número



ORGANIZANDO DADOS

- As convenções de nomenclatura de arquivos e pastas são essenciais para manter a organização e facilitar a busca.

Um nome claro e consistente para cada arquivo ou pasta ajuda a manter a organização e facilita a busca.

Utterz naming conventions can use some work, yo :)

squisa. Na
os ou para



DICAS

- Mantenha os nomes dos arquivos curtos e relevantes - geralmente cerca de 25 caracteres têm tamanho suficiente para capturar informações descritivas suficientes para nomear um arquivo de dados
- Não use caracteres especiais em um nome de arquivo como: & *% \$ £] {! @, pois geralmente são usados para tarefas específicas em diferentes sistemas operacionais





Convenções de nomenclatura para arquivos de dados de pesquisa 2

- Use sublinhados em vez de pontos ou espaços, porque, como caracteres especiais, eles são analisados de maneira diferente em sistemas diferentes.
- O nome do arquivo deve incluir tantas informações descritivas que ajudarão na identificação, independentemente de onde estão armazenadas.
- Se incluir datas, formate-as de forma consistente.
- Não suponha que o aplicativo ou instrumento de software use dependência de caso ao nomear ou renomear arquivos no utilitário mencionado acima - suponha que TANGO, Tango e tango sejam iguais, mesmo que alguns sistemas de arquivos possam considerá-los diferentes.



Convenções de nomenclatura para arquivos de dados de pesquisa 2



- Sempre que possível, use extensões de arquivo (geralmente padrões) para refletir com precisão o ambiente de software no qual o arquivo foi criado e o formato físico do arquivo. Por exemplo, use .por para arquivos portáteis do SPSS, .xls ou .xlsx para arquivos do Excel, .ssd ou .sas7bdat conforme apropriado para arquivos SAS, .txt para arquivos de texto, etc.



ATIVIDADES



NOMENADO ARQUIVO

Escolha o nome do arquivo mais apropriado da seguinte lista:

1. 2006-03-24_Attachment
2. 24 de março de 2006 Anexo
3. 240306attch

- **Nomeando arquivos por cronologia** - se estiver usando uma data, use o formato Ano-Mês-Dia: AAAA-MM-DD ou AAAA-MM ou AAAA-AAAA. Isso manterá a ordem cronológica dos seus arquivos.
- Sim. Os arquivos que usam essa convenção de nomenclatura são fáceis de distinguir um do outro, mais fáceis de navegar e localizar cronologicamente



NOMENADO ARQUIVO

Escolha o nome do arquivo mais apropriado da seguinte lista:

1. labtox_recent_110810_old version.sps
2. 2010-08-11_bioassay_toxicity_V1.sps
3. FFTX_3776438656.sps

- **Nomeação descritiva de arquivos**
- Mantenha os nomes de arquivos curtos e relevantes - usando caracteres suficientes para capturar informações descritivas suficientes
- Sim. Data em formato uniforme e fácil de distinguir dos arquivos usando a mesma convenção de data. Nome do arquivo representa o conteúdo com mais precisão. O uso de uma convenção de número de versão também facilita a distinção de outras versões do mesmo arquivo.

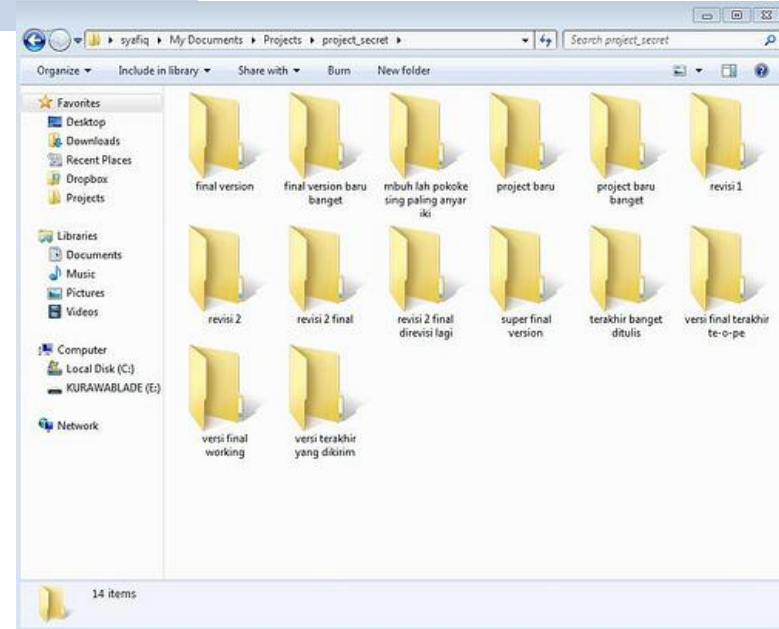


<https://www.youtube.com/watch?v=RhxVmYtKqIY>



Versionamento

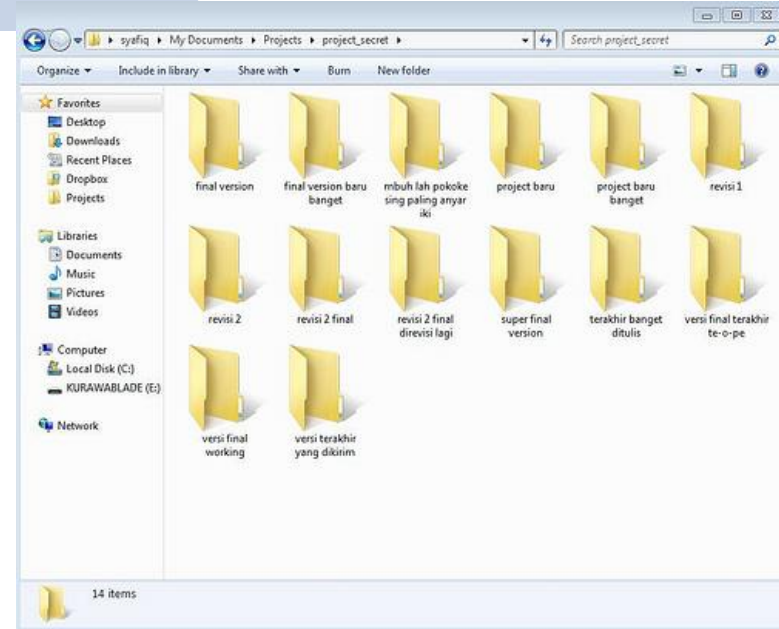
- É importante identificar e distinguir versões dos arquivos de dados de pesquisa de forma consistente. Isso garante que exista uma trilha de auditoria clara para rastrear o desenvolvimento de um arquivo de dados e identificar versões anteriores, quando necessário.
- Assim, você precisará estabelecer um método que faça sentido para você, indicando a versão dos seus arquivos de dados





Versionamento

- Uma forma comum de expressar versões de arquivos de dados é usar números ordinais (1,2,3 etc.) para alterações principais de versão e decimais para alterações menores, por exemplo, v1, v1.1, v2.6
- Cuidado ao usar rótulos confusos: revision, final, final2, definitive_copy, pois você pode achar que esses
- Registre todas as alterações, independentemente de quão pequenas elas sejam.





BENEFÍCIOS DA ROTULAGEM CONSISTENTE DE ARQUIVOS DE DADOS SÃO:

- **Arquivos de dados podem ser classificados em seqüência lógica**
- **Diferentes versões de arquivos de dados podem ser identificadas**
- **Se os arquivos de dados forem movidos para outra plataforma de armazenamento, seus nomes manterão o contexto útil**



BENEFÍCIOS DA ROTULAGEM CONSISTENTE DE ARQUIVOS DE DADOS SÃO:

- **Arquivos de dados são mais fáceis de localizar e navegar**
- **Arquivos de dados podem ser recuperados pelo criador e por outros usuários**

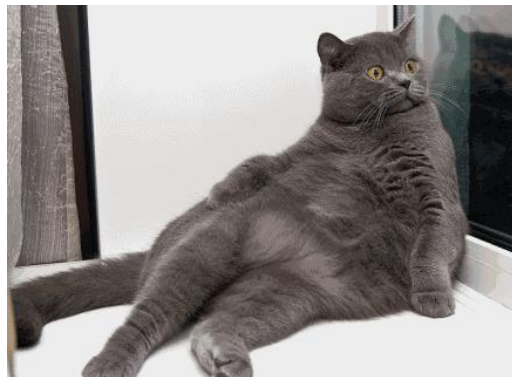


Vídeo

- LINKS
- https://www.youtube.com/watch?v=oQilZEHNf1Y&list=P_L2XF5RiVI7GNc9FkCvhfLyk46nxBggttUF



ATIVIDADES



ARMAZENAMENTO DE DADOS E SEGURANÇA- COMPARTILHAMENTO



COMPARTILHAMENTOS DE DADOS

- **O que é isso**
- **“... A prática de fazer dados usados em pesquisas acadêmicas**
- **pesquisa disponível para os outros.**
”[Wikipedia]



COMPARTILHAMENTO DE DADOS

- **Quem está envolvido**
- **o compartilhador de dados**
- **o repositório de dados**
- **o usuário de dados secundário**
- **equipe de suporte!**



Drivers para compartilhamento

- **Abertura**
- **Ciência aberta, código aberto, padrões abertos**
- **conhecimento aberto, governo aberto, arquitetura aberta, conteúdo aberto e....**



Drivers para compartilhamento

- **Dados abertos!**
- **“... Uma filosofia e prática que exigem que certos dados estejam livremente disponíveis para todos, sem restrições de direitos autorais, patentes ou outros mecanismos de controle.” [Wikipedia & other]**



Drivers para compartilhamento

- **Por que não!**
- **“... Temos as tecnologias para permitir a disponibilidade mundial e o processo distribuído de dados científicos, ampliando a colaboração e acelerando o ritmo e profundidade de descoberta...”[John Willbanks, VP de Ciência, Creative Commons - na Wikipedia]**



Drivers para compartilhamento

- **“Os cientistas preferem compartilhar sua escova de dentes do que seus dados!” [Carole Goble, discurso de abertura, Conferência EGEE (Possibilitando Grids para EscienceE) “06”.]**



Drivers para compartilhamento

- **... Research & Society**
- **Evita a duplicação de esforços e recursos.**
- **Torna disponível publicamente aquilo que foi financiado publicamente.**
- **Maximiza o retorno do investimento em pesquisa.**
- **Integridade acadêmica e científica**
- **Aumenta a transparência e a responsabilidade**



Drivers para compartilhamento

- **Facilita o escrutínio dos resultados da pesquisa**
- **Evita fraudes**
- **Amplia o alcance da pesquisa original.**
- **Facilita e promove novas linhas de pesquisa.**
- **Promove a colaboração.**
- **Permite a criação de novos conjuntos de dados**
- **Fornece conjuntos de dados para educação e treinamento**



Drivers para compartilhamento

- ... o pesquisador
- **Facilita o cumprimento dos requisitos do conselho de financiamento**
- **Autentica / valida pesquisas**
- **Melhora a reputação através de maior alcance e impacto**
- **Alcança maior visibilidade da pesquisa**
- **Assistências com armazenamento de dados a longo prazo**
- **Permite a recuperação futura dos dados pelo pesquisador e futuros pesquisadores**



Formas de compartilhar dados

Research-funder repository



<http://store.data-archive.ac.uk/store/>

Institutional repository



<http://datashare.is.ed.ac.uk/>

Discipline-specific repository



<http://figshare.com/>



Vídeo

- LINKS
- https://www.youtube.com/watch?v=n_1XFKYAcol&list=P_L2XF5RiVI7GOLsLiInXnvT8L6XHC02byD



ÉTICA E DIREITOS AUTORAIS



Ética da coleta de dados sobre seres humanos

- **O propósito e a natureza da pesquisa em si**
- **A natureza do consentimento obtido (por exemplo, participação em opt-in versus opt-out)**
- **Quais dados precisam ser salvaguardados durante a análise e destruídos após o seu uso**



Quais pesquisas estão sujeitas a revisão ética?

- **Usando participantes humanos ou animais vivos**
- **Referenciando sujeitos individuais (pessoas)**
- **Mantendo identificadores para indivíduos**



Quais pesquisas estão sujeitas a revisão ética?

- **Usando participantes humanos ou animais vivos**
- **Referenciando sujeitos individuais (pessoas)**
- **Mantendo identificadores para indivíduos**



ISENTO da revisão de ética

- **livremente disponível no domínio público, ou**
- **obtido através de estudos puramente observacionais de comportamento público:**
 - são de ação humana que ocorre em um fórum aberto ao público em geral
 - são não-invasivos
 - não requer interação com os participantes
 - não identifique participantes



- Os indivíduos são normalmente considerados como tendo um "direito à privacidade" - o estado de estar livre de intrusão ou perturbação na vida ou nos assuntos pessoais.



Privacidade e pesquisa

- **As propostas de pesquisa precisam indicar como os dados serão coletados sobre os seres humanos.**



Privacidade e pesquisa

- **Revisões éticas considerarão o quão intrusivo isso é para os sujeitos.**
- **O direito à privacidade implica a necessidade de obter consentimento dos sujeitos de pesquisa.**



Confidencialidade e pesquisa

- Muitas vezes o formulário de consentimento incluirá uma promessa ao sujeito de que o pesquisador não divulgará informações confidenciais.
- Manter esse contrato implica controles no armazenamento, no manuseio e no compartilhamento de dados pessoais.



Confidencialidade e pesquisa

- Estratégias podem incluir
- Se possível, colete os dados necessários sem usar informações de identificação pessoal.
- Se for necessária a identificação pessoal, desidentifique seus dados após a coleta ou o mais rápido possível.
- Evite transmitir eletronicamente dados pessoais não criptografados.



Confidencialidade e pesquisa

- Melhor criar um conjunto de dados de uso público que você se sinta confortável em usar, se estiver on-line.
- E não se esqueça de que a Lei de Liberdade de Informação e legislação relacionada pode obrigar você a divulgar dados, especialmente se for do interesse público!



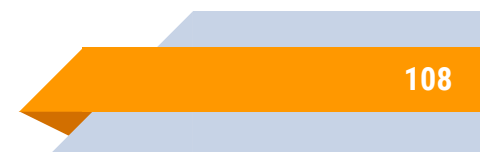
Vídeo

- LINKS
- https://www.youtube.com/watch?v=M0tRSQx-cfQ&list=PL2XF5RiVI7GMGCrHWwCYgW5Nnm_xwpvmP



Vídeo

- LINKS
- https://www.youtube.com/watch?v=FLL74eXcl8c&list=P_L2XF5RiVI7GNz2TSe3kG9uHd3UtsiV_t6





VIDEO

- <https://www.youtube.com/watch?v=pQFWZV8g3jU&list=PL2XF5RiVI7GMXjS2Q93PZo430PwNzvwht>.



AVALIAÇÃO DA OFICINA

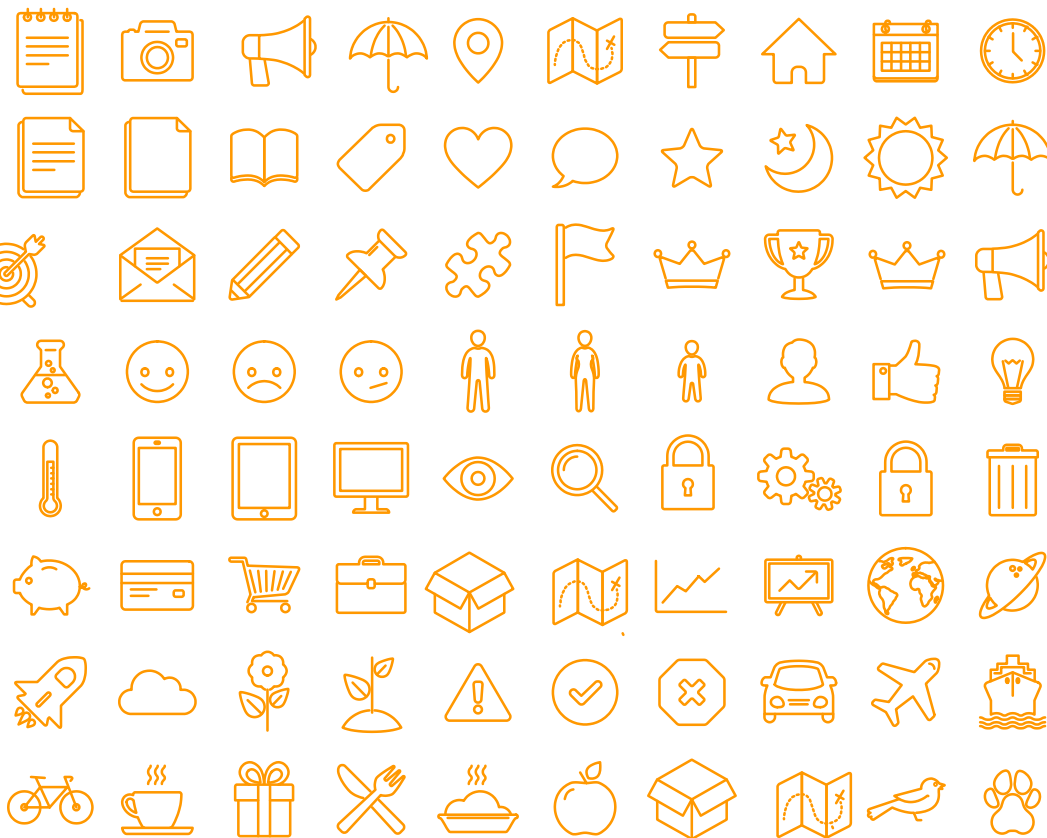
MUITO OBRIGADO! ORGANIZE SEUS DADOS!

MAKSON REIS

maksonacademico@gmail.com

 bibliotecadomakson

maksonreisblog.wordpress.com





REFERÊNCIAS

ALVARO, E., BROOKS, H., HAM, M., POEGEL, S., & ROSENCRANS, S. E-science librarianship: Field undefined. **Issues in Science and Technology Librarianship**, v. 66, 2011 Disponível em: <<http://www.istl.org/11-summer/article1.html>> Acesso em: 20 jan. 2019.

AMARAL, Fernando. **Introdução à ciência de dados**: mineração de dados e big. Rio de Janeiro: ALTA Books, 2016.

BARROS, Aidil Jesus da Silveira; LEHFELD, Neide Aparecida de Souza. **Fundamentos de metodologia científica**. 3 ed. São Paulo: Pearson Prentice Hall, 2007.

BELL, Gordon. Prefácio. In: HEY, Tony; TRANSLEY, Stewart; TOLLE, Kristin (orgs). **O quarto paradigma**: descobertas científicas na era da e-Science. São Paulo, Oficina de Textos, 2011.

BERTIN, P. R. B.; VISOLI, M. C.; DRUCKER, D. P. A gestão de dados de pesquisa no contexto da e-science: benefícios, desafios e oportunidades para organizações de p&d. **Ponto de Acesso**, v. 11, n. 2, p. 34-48, 2017. Disponível em: <<https://portalseer.ufba.br/index.php/revistaici/article/view/21449>> Acesso em: 03 jan. 2019.

BOYD, Danah; CRAWFORD, Kate. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon". In: **Information, Communication & Society** 15, no. 5: 662–79. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/1369118X.2012.678878>> Acesso em: 19 jan. 2019.



REFERÊNCIAS

BUGNION, P.; MANIVANNAN, A.; NICOLAS, P.R. *Scala: Guide for Data Science Professionals*. Birmingham: Packt Publishing, 2017, 1077 p.

CHRISTENSEN-DALSGAARD, Birte et al. Ten recommendations for libraries to get started with research data management. **Final report of the LIBER working group on E-Science/Research Data Management**. Retrieved January, v. 30, p. 2014, 2012. Disponível em: <https://libereurope.eu/wpcontent/uploads/The%20research%20data%20group%202012%20v7%20final.pdf>. Acesso em: 19 jan. 2019.

CORDEIRO, D.; BRAGHETTO, K.; GOLDMAN, A.; KON, F. Da ciência à e-ciência: paradigmas da descoberta do conhecimento. **Revista USP**, n. 97, p. 71-81, 30 maio 2013. Disponível em: < <http://www.revistas.usp.br/revusp/article/view/61867> > Acesso em: 19 de jan. 2019.

DAVENPORT, T. H. **Big Data at Work**: dispelling the myths, uncovering the opportunities. Boston: Harvard Business Review Press, 2014.

DEMO, Pedro. **Praticar ciência**: metodologias do conhecimento científico. 1. ed. São Paulo: Saraiva, 2011. 208 p.

DIEHL, Astor Antônio. **Pesquisa em ciências sociais aplicadas**: métodos e técnicas. São Paulo: Prentice Hall, 2004.

DILTHEY, Wilhelm. **Introdução às ciências humanas** – tentativa de uma fundamentação para o estudo da sociedade e da história. Trad. de Marco Antônio Casanova. Rio de Janeiro: Forense Universitária, 201

FENTANES, E.G. **A tarefa da ciência experimental**: um guia prático para pesquisar e informar resultados nas ciências naturais. Rio de Janeiro: LTC, 2014. 187p.



REFERÊNCIAS

FERREIRA, Jaider Andrade; SANTOS, Plácida L. V. A. da Costa. O modelo de dados Resource Description Framework (RDF) e o seu papel na descrição de recursos. **Informação & Sociedade: estudos**, v. 23, n. 2, 2013. Disponível em: < <http://www.brapci.inf.br/index.php/res/v/92565> > Acesso em: 21 jan. 2019.

FINZER, W. The data science education dilemma. **Technology Innovations in Statistics Education, California**, v. 7, n.2, 2013. Disponível em: <<https://escholarship.org/uc/item/7gv0q9dc> >. Acesso em: 22 dezembro. 2018

FOUREZ, Gérard. **A construção das ciências**: introdução à filosofia e à ética das ciências. São Paulo: UNESP, 1995. 319 p.

FRICKÉ, Martin. Big Data and its Epistemology. **Journal Association for Information Science and Technology**, V. 66, Issue 4. 615-661, 2014. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/asi.23212>>. Acesso em: 10 jan.2018.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. Ed. – São Paulo: Atlas, 2008.

GLEICK, J. **The information: A history, a theory, a flood**. New York, NY, US: Pantheon Books, 2011, p. 535.

GRAY, Jim. Jim Gray on eScience: A Transformed Scientific Method. Based on the transcript of a talk given by Jim Gray to the NRC-CSTB1 in Mountain View, CA, on January 11, 2007. In: HEY, Tony; TRANSLEY, Stewart; TOLLE, Kristin (orgs). **The Fourth Paradigm. Data-Intensive Scientific Discovery**. Redmond, WA: Microsoft Research, 2009. 284pp.

GUIZZO, Erico Marui. **The Essential Message**: Claude Shannon and the Making of Information Theory. [on line] Doctoral dissertation. (Master of Science in Science writing at the Massachusetts Institute of Technology), 2003. Disponível em: <<https://dspace.mit.edu/handle/1721.1/39429> >. Acesso em: 21 jan. 2019.

HJØRLAND, Birger.. “Data (with dig data and database semantics)”. **Knowledge Organization** 45, no. 8: 685-708. Also available in *ISKO Encyclopedia of Knowledge Organization*, ed. Birger Hjørland, coed. Claudio Gnol, 2018, Disponível em: <<http://www.isko.org/cyclo/data>. > Acesso em: 20 jan. 2019



REFERÊNCIAS

HOFFMANN, L. Q&A Gray's Paradigm. **Communications of the ACM**, V. v. 53, n. 10, p. 112–111, 2010. Disponível em: <<https://dl.acm.org/citation.cfm?id=1831407>>. Acesso em: 10 jan. 2019.

INMON, W. H.; LINSTEDT. **Data Architecture**: a primer for the data scientist- Big data, Data Warehouse and Data Vault. USA: Elsevier, 2015.
INTERAGENCY WORKING GROUP ON DIGITAL DATA. **Harnessing the Power of Digital Data for Science and Society**, 2009. Disponível em: <https://www.nitrd.gov/about/harnessing_power_web.pdf>. Acesso em: 19 jan. 2019

KELLEHER, John D.; TIERNEY, Brendan. **Data Science**. Cambridge, MA: The MIT Press, 2018, p.282.

KOLTAY, Tibor. Data literacy for researchers and data librarians. **Journal of Librarianship and Information Science**, v. 49, n. 1, 2017. Disponível em: <<http://journals.sagepub.com/doi/abs/10.1177/0961000615616450>>. Acesso em: 15 dez. 2018.

KUHN, T. S. **A Estrutura das Revoluções Científicas**. 9. ed. São Paulo: Perspectiva, 2005.

FEDERER, Lisa “**Defining Data Librarianship: A Survey of Competencies, Skills, and Training**”. *Journal of Medical Librarianship (JMLA)*. Vol 106, No 3 (2018)

MAATTA ,S. Placements & salaries 2013: **The emerging databrarian**. *Library Journal*. Disponível em: < <http://lj.libraryjournal.com/2013/10/placements-and-salaries/2013survey/theemerging-databrarian/#> >



REFERÊNCIAS

MARCHIONINI, Gary. Information Science Roles in the Emerging Field of Data Science. **Journal of Data and Information Science**. Vol. 1 N2, 2016 pp 1-6. Disponível em: <http://manu47.magtech.com.cn/Jwk3_jdis/Y2016/V1/I2/1> Acesso em: 15 dez. 2018.

MARTÍNEZ-URIBE, Luis, MACDONALD, Stuart Un nuevo cometido para los bibliotecarios académicos: data curation. **El profesional de la información**, vol. 17, n. 3, 2008, 273-280 p. Disponível em: <<http://www.elprofesionaldelainformacion.com/contenidos/2008/mayo/03.pdf>> Acesso em: 15 dez. 2018.

MAYER-SCHÖNBERGER Viktor, CUKIER, K. **Big Data: A Revolution That Will Transform How We Live, Work, and Think**, Boston, M: Houghton Mifflin Harcourt, 2013.

MIRANDA, Luis Felipe Sigwait de. **Introdução histórica à filosofia das ciências**. Curitiba: InterSaberes, 2016.

MONTEIRO, Elizabete Cristina de Souza de Aguiar; SANT'ANA, Ricardo Cesar Gonçalves; SANTAREM SEGUNDO, José Eduardo. e-Science Semântica: integração dos dados na comunicação científica. **Informação em Pauta**, Fortaleza, v. 1, n. 1, p. 9-29, June 2016. ISSN 2525-3468. Disponível em: <<http://periodicos.ufc.br/informacaoempauta/article/view/2942>>. Acesso em: 20 Jan. 2019.

MORAIS, Regis de. **Filosofia da ciência e da tecnologia**: introdução metodológica e crítica. Campinas, SP: Papirus, 2013.

PENTLAND, A.S. (2013). The data----driven society. *Scientific American*, 309(4), 78----83. Retrieved from Peter, Katharin and Kellam, Lynda. "Incorporating Statistical Literacy into Information Literacy Instruction." **LOEX Quarterly**, 40: 1 (2013), 2-3, 10, 2013. Disponível em: <http://www.ala.org/acrl/publications/keeping_up_with/statistical_literacy> Acesso em: 20 Jan. 2019.

POPPER. Karl Raiumund. **A Lógica da pesquisa científica**. 8 ed. São Paulo: Cultrix, 2007. 567 p.

PRADO, J., MARZAL, M. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. **Libri**, 63(2), 123–134. Disponível em: <<https://www.degruyter.com/view/j/libr.2013.63.issue-2/libri-2013-0010/libri-2013-0010.xml>> . Acesso em: 15 Jan. 2019.



REFERÊNCIAS

RICE, R.; SOUTHALL, S. **The data librarian's handbook**. London: Facet Publishing, 2016.

RIDSDALE, Chantel; ROTHWELL, James; SMIT, Mike; HOSSAM, Ali-Hassan, BLIEMEL, Michael; IRVINE Dean, KELLEY, Daniel; MATWIN ;Stan; WUETHERICK, Brad **Strategies and Best Practices for Data Literacy Education Knowledge Synthesis Report**, Dal housie University, Halifax, NS, 2015. Disponível em: <http://www.mi-kesmit.com/wp-content/papercite-data/pdf/data_literacy.pdf> Acesso em: 15 jan. 2019.

SAGIROGLU, S; SINANC, D. Big data: A review. In: Collaboration Technologies and Systems (CTS), 2013 **International Conference on. IEEE**, 2013. P. 42-47. Disponível em: <[https://www.scirp.org/\(S\(351jmbntvnsjt1aadkposzje\)\)/reference/ReferencesPapers.aspx?ReferenceID=2163642](https://www.scirp.org/(S(351jmbntvnsjt1aadkposzje))/reference/ReferencesPapers.aspx?ReferenceID=2163642)> Acesso em: 15 jan. 2019

SCHROEDER, Ralph; FRY, Jenny. Social Science Approaches to e-Science: Framing and Agenda. **Journal of Computer- Mediated Communication**, V. 12, Issue 2, p. 563-582, Jan. 2007. Disponível em: <<https://academic.oup.com/jcmc/article/12/2/563/4583047>>. Acesso em: 20 jan. 2019.

SETZER, V. W. *Dado*, informação, conhecimento e competência. **DataGramaZero**, v. 0, n. 0, p. A01, 1999. Disponível em: <<http://www.brapci.inf.br/index.php/article/download/14562>> .Acesso em: 27 jan. 2019.

SMITH, F. Jack Data Science as an academic discipline. **Data Science Journal** , v.5, (2006) p. 163–164. Disponível em: <<https://datascience.codata.org/articles/abstract/10.2481/dsj.5.163>> Acesso em: 27 jan. 2019.

SWAN, A.; BROWN, S. The skills, role and care restructure of data scientists and curators: an assessment of current practice and future needs. **Report to the Joint Information Systems Committee (JISC)**. Truro: Key Perspectives for JISC, 2008. p. 34 Disponível em: <https://eprints.soton.ac.uk/266675/1/Data_skills_report_final_draft.doc> Acesso em: 27 jan. 2019.

HEY ; Tony, HEY, Jessie **"e-Science and its implications for the library community"**, **Library Hi Tech**, Vol. 24 Issue: 4, 2016, pp.515-528. Disponível em: <<https://doi-org.ez20.periodicos.capes.gov.br/10.1108/07378830610715383>> .Acesso em: 28 dez. 2018.



REFERÊNCIAS

VAN DER AALST, W. M. P. Data scientist: the engineer of the future. In: Enterprise Interoperability VI: interoperability for agility, resilience and plasticity of collaborations. Springer: New York, 2014. Disponível em: <<https://research.tue.nl/en/publications/data-scientist-the-engineer-of-the-future>> Acesso em: 28 dez. 2018.

VELHO, Léa. Conceitos de Ciência e a Política Científica, Tecnológica e de Inovação. **Sociologias**, Porto Alegre, v. 13, n. 26, p. 128-153, 2011.

XIA, Jingfeng; WANG, Minglu. Competencies and Responsibilities of Social Science Data Librarians: An Analysis of Job Descriptions. **College & Research Libraries**, [S.l.], v. 75, n. 3, p. 362-388, may 2014. Disponível em: <<https://crl.acrl.org/index.php/crl/article/view/16367>>. Acesso em: 24 jan. 2019.

YANG X, WANG L, LASZEWSKI, G. Recent research advances in e-science. **Cluster Computing**, v.12(4), 2009, p. 353–356. Disponível em: <<https://dl.acm.org/citation.cfm?id=1666191.1666210>>. Acesso em: 10 jan. 2019.

ZHU, Y. and XIONG, Y. Towards Data Science. **Data Science Journal**, v.14, 2015 p.8. Disponível em: < <https://datascience.codata.org/article/10.5334/dsj-2015-008/>> Acesso em: 27 jan. 2019.

ZHU, Y.; XIONG, Y. Towards Data Science. **Data Science Journal**, 14, p. 8, 2015. Disponível em:<<http://doi.org/10.5334/dsj-2015-008>>. Acesso em: 06 jan. 2019.