

## São Paulo Research Foundation (fapesp.br): Digital Curation Centre (português)

### Coleta de Dados

Que dados serão coletados ou criados?

*Guidance:*

Questões a considerar:

- Qual tipo, formato e volume de dados?
- Seus formatos e programa escolhidos permitem compartilhar e ter acesso a longo prazo aos dados?
- Existem alguns dados que você possa reutilizar?

Dê uma breve descrição dos dados, incluindo quaisquer dados existentes ou fontes de terceiros que serão utilizados, em cada caso, observando seu conteúdo, tipo e cobertura. Descreva e justifique sua escolha de formato e considere as implicações do formato de dados e dos volumes de dados em termos de armazenamento, backup e acesso.

Como os dados serão coletados ou criados?

*Guidance:*

Questões a considerar:

- Quais padrões ou metodologias você usará?
- Como você vai estruturar e nomear suas pastas e arquivos?
- Como você vai lidar com versionamento?
- Quais processos de garantia de qualidade você adotará?

Descreva como os dados serão coletados / criados e quais padrões de dados da comunidade (se houver) serão usados. Considere como os dados serão organizados durante o projeto, mencionando, por exemplo, convenções de nomenclatura, controle de versão e estruturas de pastas. Explique como a consistência e a qualidade da coleta de dados serão controladas e documentadas. Isso pode incluir processos como calibração, repetição de amostras ou medições, captura ou registro de dados padronizados, validação de entrada de dados, revisão de dados por pares ou representação com vocabulários controlados.

### Documentação e Metadados

Que documentação e metadados acompanharão os dados?

*Guidance:*

Questões a considerar:

- Quais informações são necessárias para que os dados sejam lidos e interpretados no futuro?
- Como você vai capturar / criar esta documentação e metadados?
- Quais padrões de metadados você usará e por quê?

Descreva os tipos de documentação que acompanharão os dados para ajudar os usuários secundários a entender e reutilizá-los. Isso deve incluir pelo menos detalhes básicos que ajudarão as pessoas a encontrar os dados, incluindo quem criou ou contribuiu para os dados, seu título, data de criação e sob quais condições ele pode ser acessado.

A documentação também pode incluir detalhes sobre a metodologia utilizada, informações analíticas e de procedimentos, definições de variáveis, vocabulários, unidades de medida, quaisquer suposições feitas e o formato e tipo de arquivo dos dados. Considere como você irá capturar essa informação e onde ela será gravada. Sempre que possível, você deve identificar e usar os padrões existentes da comunidade.

### Ética e Conformidade Legal

Como você administrará qualquer questão ética?

*Guidance:*

Questões a considerar:

- Você obteve permissão para preservação e compartilhamento de dados?
- Como você protegerá a identidade dos participantes, se necessário? por exemplo. via anonimização
- Como os dados confidenciais serão tratados para garantir que eles sejam armazenados e transferidos com segurança?

Questões éticas afetam como você armazena dados, quem pode ver / usá-lo e quanto tempo ele é mantido. O gerenciamento de preocupações éticas pode incluir: anonimização de dados; encaminhamento para comitês de ética departamental ou institucional; e acordos formais de consentimento. Você deve mostrar que está ciente de quaisquer problemas e planejou de acordo. Se você estiver realizando uma pesquisa envolvendo participantes humanos, também deverá garantir que o consentimento seja solicitado para permitir que os dados sejam compartilhados e reutilizados.

Como você vai gerenciar os direitos autorais e os direitos de propriedade intelectual (IP / IPR)?

*Guidance:*

Questões a considerar:

- Quem possui os dados?
- Como os dados serão licenciados para reutilização?
- Existem restrições à reutilização de dados de terceiros?
- O compartilhamento de dados será adiado / restringido? Ex. para publicar ou buscar patentes?

Indique quem será o proprietário do copyright e dos Direitos de Propriedade Intelectual (IPR) de qualquer um dos dados que você irá coletar ou criar, juntamente com as licenças de uso e reuso. Para projetos de multi-parceiros, pode ser viável que os Direitos de Propriedade Intelectual (IPR) sejam descritos em um acordo de consentimento pelos donos do IPR. Considere relevante qualquer financiador, políticas de instituições, departamentos ou grupos nos direitos autorais ou nos Direitos de Propriedade Intelectual (IPR). Além disso, também considere permissões para reutilizar dados de terceiros e quaisquer restrições necessárias no compartilhamento de dados.

## **Armazenamento e Backup**

Como os dados serão armazenados e terão backup durante a pesquisa?

*Guidance:*

Questões a considerar:

- Você tem espaço de armazenamento suficiente ou precisa incluir taxas para serviços adicionais?
- Como será feito o backup dos dados?
- Quem será responsável pelo backup e recuperação?
- Como os dados serão recuperados no caso de um incidente?

Indique com que frequência os dados serão armazenados em backup e em quais locais. Quantas cópias estão sendo feitas? Armazenar dados em laptops, discos rígidos de computador ou dispositivos de armazenamento externos é muito arriscado. O uso de armazenamento robusto e gerenciado fornecido por equipes universitárias de TI é preferível. Da mesma forma, normalmente é melhor usar serviços de backup automáticos fornecidos pelos Serviços de TI do que depender de processos manuais. Se você optar por usar um serviço de terceiros, certifique-se de que isso não conflite com nenhuma política de financiador, institucional, departamental ou de grupo, por exemplo, em termos da jurisdição legal em que os dados são mantidos ou da proteção de dados confidenciais.

Como você vai gerenciar o acesso e a segurança?

*Guidance:*

Questões a considerar:

- Quais são os riscos para a segurança de dados e como eles serão gerenciados?
- Como você controlará o acesso para manter os dados seguros?
- Como você garantirá que os colaboradores possam acessar seus dados com segurança?
- Se você criar ou coletar dados no campo, como garantirá sua transferência segura para seus principais sistemas protegidos?

Se os seus dados forem confidenciais (por exemplo, dados pessoais que ainda não são de domínio público, informações confidenciais ou segredos comerciais), você deve descrever quaisquer medidas de segurança apropriadas e observar quaisquer padrões formais que você cumprirá, por exemplo. ISO 27001.

## Seleção e Preservação

Quais dados são de valor a longo prazo e devem ser mantidos, compartilhados e / ou preservados?

*Guidance:*

Questões a considerar:

- Quais dados devem ser mantidos / destruídos para fins contratuais, legais ou regulamentares?
- Como você vai decidir quais outros dados manter?
- Quais são os usos de pesquisa previsíveis para os dados?
- Por quanto tempo os dados serão mantidos e preservados?

Considere como os dados podem ser reutilizados. validar os resultados de sua pesquisa, conduzir novos estudos ou ensinar. Decida quais dados manter e por quanto tempo. Isso pode se basear em qualquer obrigação de manter certos dados, o valor potencial de reutilização, o que é economicamente viável para manter e qualquer esforço adicional necessário para preparar os dados para compartilhamento e preservação de dados. Lembre-se de considerar qualquer esforço adicional necessário para preparar os dados para compartilhamento e preservação, como a alteração de formatos de arquivo.

Qual é o plano de preservação a longo prazo do conjunto de dados?

*Guidance:*

Questões a considerar:

- Onde, por exemplo em qual repositório ou arquivo os dados serão mantidos?
- Quais os custos, se houver algum, do seu repositório de dados selecionado ou taxa de arquivamento?
- Você orçou tempo e esforço para preparar os dados para compartilhamento / preservação?

Considere como os conjuntos de dados que têm valor de longo prazo serão preservados e curados além da duração da concessão. Descreva também os planos para preparar e documentar dados para compartilhamento e arquivamento. Se você não pretende usar um repositório estabelecido, o plano de gerenciamento de dados deve demonstrar que os recursos e sistemas estarão em funcionamento para permitir que os dados sejam efetivamente curados além do tempo de vida da concessão.

## Compartilhamento de Dados

Como você vai compartilhar os dados?

*Guidance:*

Questões a considerar:

- Como os usuários em potencial podem descobrir seus dados?
- Com quem você compartilhará os dados e sob quais condições?
- Você vai compartilhar dados através de um repositório, lidar com pedidos diretamente ou usar outro mecanismo?
- Quando você disponibilizará os dados?
- Você vai buscar um identificador persistente para seus dados?

Considere onde, como e para quem os dados com valor reconhecido de longo prazo devem ser disponibilizados. Os métodos usados para compartilhar dados dependerão de vários fatores, como o tipo, tamanho, complexidade e sensibilidade dos dados. Se possível, mencione exemplos anteriores para mostrar um histórico de compartilhamento efetivo de dados. Considere como as pessoas podem reconhecer a reutilização de seus dados.

Existem restrições ao compartilhamento de dados requeridos?

*Guidance:*

Questões a considerar:

- Que medidas você tomará para superar ou minimizar as restrições?
- Por quanto tempo você precisa de uso exclusivo dos dados e por quê?
- Será necessário um contrato de compartilhamento de dados (ou equivalente)?

Descreva quaisquer dificuldades esperadas no compartilhamento de dados com valor reconhecido de longo prazo, juntamente com causas e possíveis medidas para superá-los. Restrições podem ser devidas à confidencialidade, falta de acordos de consentimento ou DPI, por exemplo. Considere se um acordo de não divulgação daria proteção suficiente para dados confidenciais.

## Responsabilidades e Recursos

Quem será responsável pelo gerenciamento de dados?

*Guidance:*

Perguntas a considerar:

- Quem é responsável pela implementação do DMP e por garantir que seja revisado e revisado?
- Quem será responsável por cada atividade de gerenciamento de dados?
- Como as responsabilidades serão divididas em sites de parceiros em projetos de pesquisa colaborativa?
- A propriedade dos dados e as responsabilidades pelo gerenciamento de dados de pesquisa serão parte de qualquer acordo de consórcio ou contrato acordado entre parceiros?

Descreva as funções e responsabilidades de todas as atividades, por exemplo captura de dados, produção de metadados, qualidade de dados, armazenamento e backup, arquivamento de dados e compartilhamento de dados. Considere quem será responsável por garantir que as políticas relevantes serão respeitadas. Os indivíduos devem ser nomeados sempre que possível.

Quais recursos você precisará para entregar seu plano?

*Guidance:*

Questões a considerar:

- É necessária especialização especializada adicional (ou treinamento para o pessoal existente)?
- Você precisa de hardware ou software que seja adicional ou excepcional à disposição institucional existente?
- As cobranças serão aplicadas pelos repositórios de dados?

Considere cuidadosamente todos os recursos necessários para entregar o plano, por ex. software, hardware, conhecimento técnico, etc. Quando são necessários recursos dedicados, estes devem ser delineados e justificados.