

MODELO GENÉRICO DO PROCESSO DE PRODUÇÃO ESTATÍSTICA DO IBGE - GSBPM

VERSÃO 5.1

Presidente da República
Jair Messias Bolsonaro

Ministro da Economia
Paulo Roberto Nunes Guedes

Secretário Especial de Fazenda
Waldery Rodrigues Junior

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE

Presidente
Susana Cordeiro Guerra

Diretora-Executiva
Marise Maria Ferreira

ÓRGÃOS ESPECÍFICOS SINGULARES

Diretoria de Pesquisas
Eduardo Luiz G. Rios Neto

Diretoria de Geociências
Claudio Stenner

Diretoria de Informática
Carlos Renato Pereira Cotovio

Centro de Documentação e Disseminação de Informações
Carmen Danielle Lins Mendes Macedo

Escola Nacional de Ciências Estatísticas
Maysa Sacramento de Magalhães

UNIDADE RESPONSÁVEL

Diretoria de Pesquisas
Coordenação de Métodos e Qualidade
Andrea Borges Paim

Ministério da Economia
Instituto Brasileiro de Geografia e Estatística - IBGE
Diretoria de Pesquisas
Coordenação de Métodos e Qualidade

Modelo Genérico do Processo de Produção Estatística do IBGE - GSBPM

Versão 5. 1



Rio de Janeiro
2021

Instituto Brasileiro de Geografia e Estatística - IBGE

Av. Franklin Roosevelt, 166 - Centro - 20021-120 - Rio de Janeiro, RJ - Brasil

© IBGE. 2021

Sumário

Introdução	4
Histórico	4
Principais Mudanças da Versão 5.1.....	4
O modelo	6
Entendendo o GSBPM.....	6
Estrutura.....	7
Aplicabilidade	8
Usando o GSBPM.....	8
Relacionamentos com outros modelos e estruturas	10
GAMSO	10
GSIM	10
Os níveis 1 e 2 do GSBPM	12
Descrições de fases e subprocessos	13
Fase Especificar Necessidades	13
Fase Planejar	15
Fase Construir.....	18
Fase Coletar.....	21
Fase Processar	24
Fase Analisar	28
Fase Disseminar	30
Fase Avaliar	32
Processos Gerais	34
Gestão da Qualidade	34
Gestão de Metadados	36
Gestão de Dados	38
Outros usos do GSBPM	39
Referências	41
Anexos	42
1 - Lista de Acrônimos.....	43

Convenções

-	Dado numérico igual a zero não resultante de arredondamento;
..	Não se aplica dado numérico;
...	Dado numérico não disponível;
x	Dado numérico omitido a fim de evitar a individualização da informação;
0; 0,0; 0,00	Dado numérico igual a zero resultante de arredondamento de um dado numérico originalmente positivo; e
-0; -0,0; -0,00	Dado numérico igual a zero resultante de arredondamento de um dado numérico originalmente negativo.

Introdução

O Modelo Genérico do Processo de Produção Estatística (GSBPM) descreve e define o conjunto de processos necessários para produzir estatísticas oficiais. Esse fornece uma estrutura padrão e terminologia harmonizada para ajudar organizações estatísticas a modernizar seus processos de produção estatística, bem como para compartilhar métodos e componentes. O GSBPM também pode ser usado para integrar dados e padrões de metadados, como um modelo para documentação de processos, para harmonizar infraestruturas de computação estatística e para fornecer uma estrutura para avaliação e melhoria da qualidade dos processos. Estas e outras finalidades para as quais o GSBPM pode ser usado são retomadas na Seção VII. Esta versão do GSBPM está alinhada com a versão 1.2 do Modelo Genérico de Informação Estatística (GSIM) e com a versão 1.2 do Modelo Genérico de Atividades para Organizações Estatísticas (GAMSO).

Histórico

O GSBPM foi desenvolvido pela primeira vez em 2008 pelo Grupo Conjunto UNECE/Eurostat/OCDE sobre Metadados Estatísticos (METIS) com base no modelo de processo de produção usado pelo Instituto de Estatística da Nova Zelândia. Após vários rascunhos e consultas públicas, a versão 4.0 do GSBPM foi lançada em abril de 2009. Posteriormente, foi amplamente adotada pela comunidade global de estatísticas oficiais, e formou uma das pedras angulares da visão e estratégia para a modernização baseada em padrões do Grupo de Alto Nível para a Modernização das Estatísticas Oficiais (HLG-MOS) (<https://statswiki.unece.org/display/hlgbas>)

A versão anterior do GSBPM (versão 5.0) foi lançada em dezembro de 2013 e uma ampla revisão foi realizada em 2018 para incorporar o feedback baseado na implementação prática do modelo e melhorar a consistência com os novos modelos HLG-MOS, o GSIM e o GAMSO. Embora a versão atual do GSBPM (versão 5.1) seja considerada como definitiva no momento de sua publicação, também se espera que futuras atualizações possam ser necessárias nos próximos anos, seja para refletir novas experiências de implementação do modelo na prática, ou devido à evolução da natureza da produção estatística. O leitor é, portanto, convidado a verificar o wiki do GSBPM (<https://statswiki.unece.org/display/GSBPM>) para ter certeza de ter a versão mais recente.

Principais Mudanças da Versão 5.1

A revisão da versão 5.0 do GSBPM tinha uma orientação do HLG-MOS de introduzir apenas as mudanças que tivessem um argumento forte e amplo apoio na comunidade estatística internacional. A justificativa para essa abordagem era que muitas organizações haviam investido pesadamente na implementação do GSBPM. Grandes mudanças

poderiam implicar custos significativos e poderiam, em última análise, ser contraproducentes em termos de garantir a adoção generalizada e o uso do modelo.

As principais mudanças no GSBPM entre as versões 5.0 e 5.1 são as seguintes:

Alguns subprocessos foram renomeados para melhorar a clareza;

A duplicação entre os processos abrangentes do GSBPM versão 5.0 e do GAMS0 foi resolvida;

As descrições das fases e dos subprocessos foi atualizada de forma a serem menos centrada em pesquisas. E as atividades referentes ao relacionamento com os fornecedores de dados não estatísticos foram adicionadas onde necessário;

Em reconhecimento à crescente importância da integração de dados estatísticos com dados geoespaciais, as descrições foram expandidas visando incluir tarefas necessárias para o uso de dados geoespaciais;

A terminologia foi modificada quando necessário, para melhorar a coerência com o GAMS0 e o GSIM;

Exemplos e descrições foram atualizados e expandidos para melhorar a clareza.

O modelo

Entendendo o GSBPM

Um processo estatístico é uma coleção de atividades e tarefas relacionadas e estruturadas para converter dados de entrada em informações estatísticas. No contexto do GSBPM, organizações ou grupos de organizações realizam processos estatísticos para criar estatísticas oficiais de forma a satisfazer as necessidades dos usuários. Os resultados do processo podem ser um conjunto misto de produtos físicos ou digitais que apresentam dados e metadados de diferentes maneiras, como publicações, mapas, serviços eletrônicos, entre outros.

O GSBPM deve ser aplicado e interpretado de forma flexível. Não é uma estrutura rígida na qual todas as etapas devem ser seguidas em uma ordem estrita, em vez disso, identifica as possíveis etapas no processo estatístico e as interdependências entre elas.

Embora a apresentação do GSBPM siga uma sequência lógica de etapas na maioria dos processos estatísticos, os elementos do modelo podem ocorrer em ordens diferentes em circunstâncias diferentes. Além disso, alguns subprocessos serão revisitados, formando loops iterativos, particularmente nas fases Processar e Analisar.

O GSBPM pode ser visto como uma lista de verificação para certificar-se de que todas as etapas necessárias foram consideradas ou como um "livro de receitas" para identificar todos os "ingredientes" de um processo estatístico.



Figura 1. Fases de trabalho em andamento e fases de mudanças do GSBPM

Em muitas organizações estatísticas, as primeiras fases são consideradas apenas quando um novo produto estatístico é criado ou quando o processo é revisado como

resultado de um processo de avaliação. Uma vez que o produto se torna parte da atividade em andamento “normal”, essas fases não são realizadas (por exemplo, não é necessário construir novas ferramentas de coleta toda vez que os dados da pesquisa da força de trabalho são coletados). A figura 1 descreve esse fenômeno.

O GSBPM deve, portanto, ser visto mais como uma matriz, através da qual existem muitos caminhos possíveis. Desta forma, o GSBPM pretende ser suficientemente genérico para ser amplamente aplicável e encorajar uma visão padrão do processo estatístico, sem se tornar nem demasiado restritivo ou demasiado abstrato e teórico.

Estrutura

O GSBPM compreende três níveis:

Nível 0, o processo estatístico;

Nível 1, as oito fases do processo estatístico;

Nível 2, os subprocessos dentro de cada fase.

Um diagrama mostrando as fases (nível 1) e subprocessos (nível 2) está incluído na Seção IV (Figura 3). Os subprocessos são descritos detalhadamente na Seção V.

O GSBPM reconhece vários processos gerais com um forte componente estatístico que se estendem ao longo das oito fases. Esses processos abrangentes estão incluídos na lista abaixo. A gestão da qualidade, a gestão de metadados e a gestão de dados são elaborados na Seção VI. Atividades que são realizadas no nível da organização para apoiar a produção estatística estão incluídas no GAMS0 (ver Seção III)

Gestão da qualidade - Inclui a avaliação da qualidade e mecanismos de controle. Reconhece a importância da avaliação e feedback durante todo o processo de produção estatística;

Gestão de metadados - Metadados são criados/reutilizados e processados dentro de cada fase, há, portanto, uma forte necessidade de um sistema de gerenciamento de metadados para garantir que os metadados apropriados mantenham as suas ligações com os dados em todo o GSBPM. Isso inclui considerações independentes de processo, como custódia e propriedade de metadados, qualidade, regras de arquivamento, preservação, retenção e descarte;

Gestão de dados - Inclui considerações independentes de processo, como segurança geral de dados, custódia e propriedade, qualidade de dados, regras de arquivamento, preservação, retenção e descarte;

Gestão de dados do processo - Inclui atividades de registro, sistematização e uso de dados sobre a implementação do processo estatístico. Os dados de processo podem

auxiliar na detecção e compreensão de padrões nos dados coletados, bem como na avaliação da execução do processo estatístico como tal;

Gestão de conhecimento - Garante que os processos estatísticos sejam replicáveis, principalmente através da manutenção da documentação do processo;

Gestão de fornecedores- Inclui o gerenciamento de sobrecarga de processos cruzados, bem como tópicos como a criação de perfis e o gerenciamento de informações de contato (e, portanto, tem vínculos particularmente estreitos com processos estatísticos que mantêm registros).

Aplicabilidade

O GSBPM destina-se a ser aplicado a todas as atividades realizadas pelos produtores de estatísticas oficiais, tanto em nível nacional como internacional, que resultam em produtos estatísticos.

O modelo é projetado para ser aplicável independentemente da fonte de dados, portanto pode ser usado para a descrição e avaliação de qualidade de processos com base em pesquisas, censos, registros administrativos e outras fontes não estatísticas ou mistas.

Enquanto processos estatísticos típicos incluem coleta e processamento de dados para gerar produtos estatísticos, o GSBPM também se aplica quando dados existentes são revisados, ou séries temporais são recalculadas, quer como resultado de fontes de dados melhorados ou uma mudança na metodologia. Nestes casos, os dados de entrada podem ser microdados originais e/ou dados adicionais, que são então processados e analisados para produzir resultados revisados. Em tais casos, é provável que vários subprocessos e possivelmente algumas fases (particularmente as primeiras) sejam omitidos. Da mesma forma, o GSBPM pode ser aplicado a processos, tais como a compilação de contas nacionais e os processos típicos em organizações estatísticas internacionais que utilizam dados secundários de países ou outras organizações.

Além de ser aplicável a processos que resultam em estatísticas, o GSBPM também pode ser aplicado ao desenvolvimento e manutenção de registros estatísticos, onde as entradas são similares àquelas para produção estatística (embora tipicamente com maior foco em dados administrativos), e as saídas são tipicamente cadastros ou outras extrações de dados, que são então usadas como entradas para outros processos.

O GSBPM é suficientemente flexível para ser aplicado em todos os cenários acima.

Usando o GSBPM

O GSBPM é um modelo de referência. Pretende-se que o GSBPM possa ser utilizado por organizações em diferentes níveis. Uma organização pode optar por implementar o GSBPM diretamente ou usá-lo como base para o desenvolvimento de uma

versão personalizada do modelo. Pode ser usado em alguns casos apenas como um modelo ao qual as organizações se referem quando se comunicam internamente ou com outras organizações para facilitar a discussão. Os vários cenários para o uso do GSBPM são todos válidos.

Quando as organizações desenvolvem adaptações específicas do GSBPM, elas podem fazer algumas especializações no modelo para caber em seu contexto organizacional. Até agora, as evidências sugerem que essas especializações não são suficientemente genéricas para serem incluídas no próprio GSBPM.

Em alguns casos, pode ser apropriado agrupar alguns dos elementos do modelo. Por exemplo, as três fases iniciais poderiam ser consideradas uma única fase de planejamento. Em outros casos, particularmente para implementações práticas, pode ser necessário adicionar um ou mais níveis detalhados à estrutura para identificar separadamente os diferentes componentes dos subprocessos.

Também pode haver a necessidade de uma aprovação formal entre as fases, em que a saída de uma fase é certificada como entrada adequada para uma próxima fase. Esta aprovação formal está implícita no modelo (exceto no subprocesso 1.6), mas pode ser implementada explicitamente de diferentes formas dependendo das necessidades organizacionais.

Relacionamentos com outros modelos e estruturas

Desde o lançamento do GSBPM, vários modelos foram desenvolvidos sob a supervisão do HLG-MOS para apoiar a modernização das estatísticas oficiais. Coletivamente, esses são chamados de modelos “ModernStats”. Os parágrafos a seguir descrevem os modelos ModernStats que têm uma forte ligação com o GSBPM.

GAMSO

O GAMSO (<https://statswiki.unece.org/display/GAMSO>) descreve e define atividades que ocorrem dentro de uma organização estatística típica. Estende e complementa o GSBPM adicionando atividades necessárias para apoiar a produção estatística (ex. As atividades nas áreas de estratégia e liderança, desenvolvimento de capacidades e suporte corporativo). No GSBPM v5.0, algumas dessas atividades foram incluídas como processos abrangentes. Atividades que não estão diretamente relacionadas com a produção de estatísticas e/ou são gerenciadas em nível corporativo ou estratégico estão agora incluídas no GAMSO (por exemplo, gestão de recursos humanos, atividades de gestão da qualidade que são realizadas em nível corporativo como o desenvolvimento de um quadro de referência de qualidade).

O GAMSO descreve atividades, isto é, o que as organizações estatísticas fazem. Inclui descrições de alto nível dessas atividades. Por outro lado, o GSBPM centra-se no processo de produção, descreve mais detalhadamente como as organizações estatísticas realizam a atividade de produção estatística.

Como o GSBPM, o GAMSO visa fornecer um vocabulário e estrutura comuns para apoiar atividades de colaboração internacional. Um valor maior será obtido do GAMSO se for aplicado em conjunto com o GSBPM.

GSIM

O GSIM (<https://statswiki.unece.org/display/GSIM>) é um quadro de referência para a informação estatística, concebido para ajudar a modernizar as estatísticas oficiais, tanto em nível nacional como internacional. Esse permite descrições genéricas da definição, gerenciamento e uso de dados e metadados em todo o processo de produção estatística. Fornece um conjunto de objetos de informações padronizados e consistentemente descritos, que são as entradas e saídas dos subprocessos do GSBPM. O GSIM ajuda a explicar as relações significativas entre as entidades envolvidas na produção estatística e pode ser usado para orientar o desenvolvimento e uso de padrões ou especificações de implementação consistentes.

Como o GSBPM, o GSIM é um dos pilares para modernizar as estatísticas oficiais e se afastar dos silos temáticos. Identifica cerca de 130 objetos de informações. Exemplos incluem conjuntos de dados, variáveis, classificações estatísticas, unidades, populações, bem como as regras e os parâmetros necessários para o funcionamento dos processos de produção (por exemplo, regras de crítica e imputação de dados).

O GSIM e o GSBPM são modelos complementares para a produção e gestão de informação estatística. Conforme mostrado na Figura 2 abaixo, o GSIM ajuda a descrever os subprocessos do GSBPM, definindo os objetos de informação que fluem entre eles, que são criados neles e que são usados por eles para produzir estatísticas oficiais. Entradas e saídas podem ser definidas em termos de objetos de informação e são formalizados no GSIM.



Figura 2. Os objetos de informação do GSIM como entrada e saída do subprocesso GSBPM

Um valor maior será obtido do GSIM se este for aplicado em conjunto com o GSBPM. Da mesma forma, maior valor será obtido do GSBPM se este for aplicado em conjunto com o GSIM. No entanto, é possível (embora não ideal) aplicar um sem o outro.

Aplicar o GSIM e o GSBPM juntos pode facilitar a construção de sistemas eficientes orientados a metadados e ajudar a harmonização de infraestruturas de computação estatística.

Os níveis 1 e 2 do GSBPM

Processos Gerais							
Especificar necessidades	Planejar	Construir	Coletar	Processar	Analisar	Disseminar	Avaliar
1.1 – Identificar as necessidades	2.1 – Desenhar os resultados	3.1 – Reutilizar ou construir instrumentos de coleta	4.1 – Criar os cadastros e selecionar as amostras	5.1 – Integrar dados	6.1 – Preparar os resultados	7.1 – Atualizar sistemas de disseminação	8.1 – Reunir os elementos para avaliação
1.2 – Consultar e confirmar as necessidades	2.2 – Definir as variáveis	3.2 – Reutilizar ou construir os componentes de processamento e análise	4.2 – Preparar a coleta	5.2 – Classificar e codificar	6.2 – Validar os resultados	7.2 – Elaborar produtos de disseminação	8.2 – Realizar a avaliação
1.3 – Estabelecer os objetivos	2.3 – Definir a metodologia de coleta	3.3 – Reutilizar ou construir os componentes da disseminação	4.3 – Coletar	5.3 – Criticar microdados	6.3 – Interpretar e explicar os resultados	7.3 – Gerenciar a disseminação	8.3 – Acordar um plano de melhorias
1.4 – Identificar os conceitos	2.4 – Especificar o cadastro e a amostragem	3.4 – Configurar fluxos de trabalho	4.4 – Finalizar a coleta	5.4 – Imputar	6.4 – Assegurar a confidencialidade	7.4 – Promover os produtos de disseminação	
1.5 – Verificar disponibilidade de dados	2.5 – Definir o processamento e a análise	3.5 – Testar os sistemas de produção		5.5 – Derivar novas variáveis e unidades	6.5 – Finalizar os resultados	7.5 – Gerenciar o suporte ao usuário	
1.6 – Preparar e submeter o plano de ação	2.6 – Desenhar fluxograma e sistema de produção	3.6 – Testar o processo estatístico de produção		5.6 – Calcular os pesos			
		3.7 – Finalizar os sistemas de produção		5.7 – Calcular resultados agregados			
				5.8 – Finalizar os arquivos de dados			

Figura 3. As fases (nível 1) e subprocessos (nível 2) do GSBPM

Descrições de fases e subprocessos

Esta seção trata de cada fase, identificando os vários subprocessos dentro dessa fase e descrevendo seu conteúdo.

Fase Especificar Necessidades

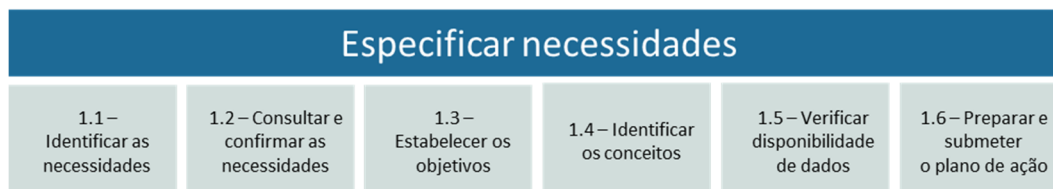


Figura 4. Fase especificar necessidades e seus subprocessos

Essa fase é acionada quando uma necessidade de novas estatísticas é identificada ou o feedback sobre as estatísticas atuais motiva uma revisão. Inclui todas as atividades associadas com o engajamento das partes interessadas de forma a identificar as suas necessidades estatísticas detalhadas (atuais ou futuras), propondo opções de solução de alto nível e preparando um plano de ação para atender a essas necessidades.

A fase "Especificar necessidades" é dividida em seis subprocessos (Figura 4), que geralmente são sequenciais, da esquerda para a direita, mas também podem ocorrer em paralelo e podem ser iterativos. Estes subprocessos são:

1.1. Identificar as necessidades

Este subprocesso inclui a investigação inicial e a identificação de quais estatísticas são necessárias e o que é necessário das estatísticas. Pode ser desencadeado por um novo pedido de informação ou por uma alteração ambiental, como um orçamento reduzido. Os planos de ação das avaliações de ocorrências anteriores do processo ou de outros processos podem fornecer uma entrada para este subprocesso. Também são consideradas as práticas de outras organizações estatísticas (nacionais e internacionais) que produzem dados semelhantes e os métodos utilizados pelas mesmas.

1.2. Consultar e confirmar as necessidades

Este subprocesso se concentra na consulta com as partes interessadas internas e externas e confirmação, em detalhes, das necessidades estatísticas. Uma boa compreensão das necessidades do usuário é essencial para que a organização estatística saiba não apenas o que se espera que ela forneça, mas também quando, como e, talvez mais importante, por quê. Para a segunda e subsequentes iterações dessa fase, o foco principal será determinar se as necessidades previamente identificadas mudaram. Essa compreensão detalhada das necessidades do usuário é a parte crítica desse subprocesso.

1.3. Estabelecer os objetivos

Este subprocesso identifica os objetivos do produto estatístico necessários para atender às necessidades do usuário identificadas no subprocesso 1.2 (Consultar e confirmar as necessidades). Inclui concordar com a adequação dos resultados propostos e suas medidas de qualidade com os usuários. As estruturas legais (por exemplo, relacionadas à confidencialidade) e os recursos disponíveis provavelmente serão restrições ao estabelecer os objetivos.

1.4. Identificar os conceitos

Este subprocesso esclarece os conceitos necessários das variáveis a serem medidas, a partir do ponto de vista do usuário. Nesta fase, os conceitos identificados podem não estar alinhados com os padrões estatísticos existentes. Este alinhamento e a escolha ou definição dos conceitos estatísticos e outros conceitos e variáveis a serem utilizados ocorrem no subprocesso 2.2 (Definir as variáveis).

1.5. Verificar disponibilidade de dados

Este subprocesso verifica se as fontes atuais de dados podem atender aos requisitos do usuário e as condições sob as quais elas estariam disponíveis, incluindo quaisquer restrições sobre seu uso. Uma avaliação de alternativas possíveis normalmente incluiria pesquisa em potenciais fontes administrativas ou outras fontes não estatísticas de dados, para:

- Determinar se seriam adequadas para utilização para fins estatísticos (por exemplo, até que ponto os conceitos administrativos correspondem aos requisitos, oportunidade e qualidade, segurança e continuidade do fornecimento de dados);
- Avaliar a divisão de responsabilidades entre fornecedores de dados e a organização estatística;
- Recursos necessários de TIC (por exemplo, armazenamento de dados, tecnologia necessária para lidar com a entrada e processamento dos dados), bem como quaisquer acordos formais com os fornecedores de dados para acessar e compartilhar os dados (por exemplo, formatos, entrega, respectivos metadados e controle de qualidade).

Quando as fontes existentes tiverem sido avaliadas, uma estratégia para preencher quaisquer lacunas remanescentes no requisito de dados deve ser preparada. Isso pode incluir a identificação de possíveis parcerias com detentores de dados. Este subprocesso inclui também uma avaliação mais geral do quadro jurídico em que os dados seriam coletados e utilizados, podendo, por conseguinte, identificar propostas de alteração da legislação existente ou a introdução de um novo quadro jurídico.

1.6. Preparar e submeter o plano de ação

Esse subprocesso documenta as saídas dos outros subprocessos dessa fase na forma de um plano de ação, para obter aprovação de implementação do processo estatístico novo ou modificado. Tal plano de ação precisa estar em conformidade com os requisitos do órgão de aprovação, normalmente inclui elementos como:

- Uma descrição do processo "As-Is" (se já existir), com informações sobre como as estatísticas atuais são produzidas, destacando quaisquer ineficiências e problemas a serem abordados;
- A solução "To-Be" proposta, detalhando como o processo estatístico será desenvolvido para produzir as estatísticas novas ou revisadas;
- Uma avaliação de custos e benefícios, bem como quaisquer restrições externas.

O plano de ação descreve opções e faz recomendações. Pode incluir os benefícios, custos, entregas, prazo, orçamento, recursos técnicos e humanos necessários, avaliação de riscos e impacto nas partes interessadas para cada opção.

Depois que o plano de ação é preparado, este é submetido à aprovação para passar para a próxima fase do processo. Neste subprocesso, uma decisão é tomada “seguir”/“não seguir”. Normalmente, o plano de ação é revisado e formalmente aprovado ou reprovado pelos patrocinadores e comitês de governança apropriados.

Fase Planejar

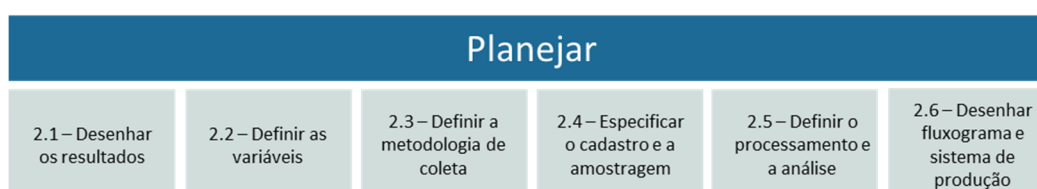


Figura 5. Fase planejar e seus subprocessos

Esta fase descreve as atividades de desenvolvimento e planejamento, e qualquer trabalho de pesquisa prática associada necessária para definir os produtos estatísticos, conceitos, metodologias, instrumentos de coleta e processos operacionais. Inclui todos os elementos de planejamento necessários para definir ou refinar os produtos estatísticos ou serviços identificados no plano de ação. Essa fase especifica todos os metadados relevantes, prontos para uso posteriormente no processo estatístico, bem como procedimentos de garantia de qualidade. Para os produtos estatísticos produzidos regularmente, essa fase geralmente ocorre para a primeira ocorrência e sempre que ações de melhoria são identificadas na fase “Avaliar” de uma ocorrência anterior.

Atividades de planejamento fazem uso substancial de padrões nacionais e internacionais, a fim de reduzir a duração e o custo do processo de planejamento, e melhorar a comparabilidade e usabilidade das saídas. As organizações são incentivadas a reutilizar ou adaptar elementos de planejamento de processos existentes e a considerar aspectos geoespaciais de dados no planejamento para aprimorar a usabilidade e o valor das informações estatísticas. Além disso, as saídas de processos de planejamento podem formar a base para futuras normas da organização nacional ou internacional.

A fase “Planejar” é dividida em seis subprocessos (Figura 5), que geralmente são sequenciais, da esquerda para a direita, mas também podem ocorrer em paralelo e podem ser iterativas. Estes subprocessos são:

2.1. Desenhar os resultados

Este subprocesso contém o projeto detalhado dos produtos estatísticos a serem gerados, incluindo o trabalho de desenvolvimento relacionado e a preparação dos sistemas e ferramentas utilizados na fase “Disseminar”. Os processos que regem o acesso a quaisquer resultados confidenciais também são projetados aqui. Os resultados devem ser projetados para seguir os padrões existentes sempre que possível, de modo que as entradas para esse processo possam incluir metadados de coletas semelhantes ou anteriores (incluindo extrações de registros e bases de dados não estatísticos, administrativos, geoespaciais e outros), padrões internacionais e informações sobre práticas em outras organizações estatísticas relacionados no subprocesso 1.1 (Identificar as necessidades). Os resultados também podem ser projetados em parceria com outros órgãos interessados, particularmente se forem considerados produtos conjuntos, ou se forem disseminados por outra organização.

2.2. Definir as variáveis

Este subprocesso define as variáveis a serem coletadas por meio do instrumento de coleta, bem como quaisquer outras variáveis que serão derivadas delas no subprocesso 5.5 (Derivar novas variáveis e unidades) e quaisquer classificações estatísticas ou geoespaciais que serão usadas. Espera-se que os padrões nacionais e internacionais existentes sejam seguidos sempre que possível. Esse subprocesso pode precisar ser executado em paralelo com o subprocesso 2.3 (Definir a metodologia de coleta), uma vez que a definição das variáveis a serem coletadas e a escolha dos instrumentos de coleta podem ser interdependentes em algum grau. A elaboração de descrições de metadados de variáveis coletadas e derivadas, classificações estatísticas e geoespaciais é uma pré-condição necessária para as fases subsequentes.

2.3. Definir a metodologia de coleta

Este subprocesso determina os instrumentos e métodos de coleta mais apropriados que podem depender do tipo de coleta de dados (censo, pesquisa amostral ou outro), do tipo de unidade de coleta (empresa, pessoa ou outro) e das fontes de dados disponíveis. As atividades reais neste subprocesso irão variar de acordo com o tipo de instrumento de coleta necessário, que pode incluir entrevistas por computador, questionários em papel, registros administrativos (por exemplo, usando interfaces de serviço existentes), métodos de transferência de dados, tecnologias de web scraping, bem como tecnologia para dados geoespaciais. O uso direto ou indireto de registros administrativos pode ser introduzido no modo de coleta de dados para controlar os dados da pesquisa ou auxiliar na captura das informações da pesquisa.

Inclui o planejamento dos instrumentos de coleta, perguntas e modelos de resposta (em conjunto com as variáveis e classificações estatísticas projetadas no subprocesso 2.2 (Definir as variáveis)) e a confirmação de quaisquer acordos formais. O subprocesso é habilitado por ferramentas como bibliotecas de perguntas (para facilitar a reutilização de perguntas e atributos relacionados), ferramentas de questionário (para permitir a rápida e fácil compilação de perguntas em formatos adequados para testes cognitivos) e modelos de convênio (para ajudar a padronizar termos e condições). Contempla, também, o planejamento de sistemas de gestão de fornecedores especificados no plano de ação em análise.

Nas situações em que as organizações de estatística não coletam dados diretamente (ou seja, uma terceira parte controla a coleta e processamento dos dados), pode incluir o desenho de mecanismos para monitorar os dados e os metadados de forma a avaliar o impacto de qualquer alteração feita por terceiros.

2.4. Especificar o cadastro e a amostragem

Este subprocesso se aplica a processos que envolvem a coleta de dados com base em amostragem. Identifica e especifica a população de interesse, define um cadastro de amostragem (e, quando necessário, as fontes de dados a partir das quais é derivado), e determina os critérios de amostragem e metodologia mais apropriados (que podem incluir enumeração completa). Fontes comuns para um cadastro de amostragem são registros administrativos e estatísticos, censos e informações de outras pesquisas por amostragem. Pode incluir dados geoespaciais e classificações. Este subprocesso descreve como essas fontes podem ser combinadas, se necessário. Deve-se analisar se o cadastro cobre a população-alvo. Um plano de amostragem deve ser feito. A amostra real é criada no subprocesso 4.1 (Criar os cadastros e selecionar as amostras), usando a metodologia especificada neste subprocesso.

2.5. Definir o processamento e a análise

Este subprocesso delinea a metodologia de processamento estatístico a ser aplicada durante as fases "Processar" e "Analisar". Isso pode incluir, entre outras, a especificação de rotinas e regras para a codificação, crítica e imputação que pode variar com base no modo de coleta e fonte de dados. Este subprocesso também inclui a definição de especificações para integração de dados a partir de múltiplas fontes, validação de dados e estimação. Os métodos estatísticos de controle de divulgação também são projetados aqui, se forem específicos para o processo em análise.

2.6. Desenhar fluxograma e sistema de produção

Este subprocesso determina o fluxo de trabalho desde a coleta de dados até a disseminação, tendo uma visão geral de todos os processos necessários em todo o processo de produção e garantindo que eles se encaixem de forma eficiente, sem lacunas ou redundâncias. Vários sistemas e bancos de dados são necessários ao longo do processo. O GSBPM pode ser usado como base para a arquitetura de negócios quando uma organização estatística possui uma arquitetura corporativa existente. O desenho pode ser ajustado para se adequar à organização. Um princípio geral é reutilizar processos e tecnologia em muitos processos estatísticos, portanto soluções de produção existentes (por exemplo, serviços, sistemas e bancos de dados) devem ser examinadas primeiro, para determinar se são adequadas para o processo em análise. Então, se lacunas forem identificadas, novas soluções devem ser projetadas. Este subprocesso também considera como a equipe irá interagir com os sistemas e quem será responsável pelo o quê e quando.

Fase Construir

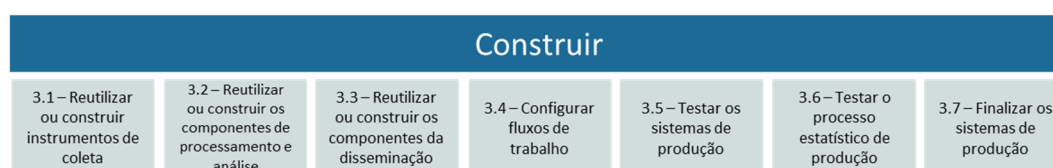


Figura 6. Fase Construir e seus subprocessos

Essa fase constrói e testa a solução de produção até o ponto em que ela está pronta para uso no ambiente "ao vivo". As saídas da fase "Planejar" são montadas e configuradas nesta fase para criar o ambiente operacional completo para executar o processo. Novos serviços são construídos por exceção, criados em resposta a lacunas no catálogo de serviços existente, originados dentro da organização e externamente. Esses novos serviços são construídos para serem amplamente reutilizáveis em alinhamento com a arquitetura de negócios da organização, sempre que possível.

Para os produtos estatísticos gerados em uma base regular, esta fase geralmente acontece para a primeira ocorrência, ou após uma revisão ou uma mudança na metodologia ou na tecnologia, e não para cada ocorrência.

A fase “Construir” é dividida em sete subprocessos (Figura 6), que geralmente são sequenciais, da esquerda para a direita, mas também podem ocorrer em paralelo e podem ser iterativos. Os primeiros três subprocessos estão preocupados com o desenvolvimento e reutilização dos sistemas para coleta, processamento, análise e divulgação de dados. Os últimos quatro subprocessos concentram-se no processo de ponta a ponta. Estes subprocessos são:

3.1. Reutilizar ou construir instrumentos de coleta

Este subprocesso descreve as atividades para construir e reutilizar os instrumentos de coleta a serem utilizados durante a fase “Coletar”. Os instrumentos de coleta são construídos com base nas especificações de planejamento criadas durante a fase “Planejar”. Uma coleta pode utilizar um ou mais meios para receber os dados (por exemplo, entrevistas pessoais ou por telefone; questionários em papel, meios eletrônicos ou pela Web; serviços da Web do SDMX). Instrumentos de coleta também podem ser rotinas de extração de dados usadas para coletar dados de registros estatísticos ou administrativos existentes (por exemplo, usando interfaces de serviço existentes). Este subprocesso inclui também a preparação e teste do conteúdo e funcionamento desses instrumentos de coleta (por exemplo, testes cognitivos das perguntas em um questionário). Recomenda-se considerar a conexão direta de instrumentos de coleta a um sistema de metadados, para que os metadados possam ser mais facilmente capturados na fase de coleta. Conectar os metadados e dados no ponto de captura pode poupar trabalho em fases posteriores. Capturar as métricas de coleta de dados (paradados) também é uma consideração importante neste subprocesso para calcular e analisar os indicadores de qualidade do processo.

3.2. Reutilizar ou construir os componentes de processamento e análise

Este subprocesso descreve as atividades para reutilizar componentes existentes ou construir novos componentes necessários para as fases “Processar” e “Analisar”, conforme projetado na fase “Planejar”. Os serviços podem incluir funções e recursos de painel (dashboard), serviços de informações, funções de transformação, serviços de dados geoespaciais, estruturas de fluxo de trabalho, serviços de gestão de fornecedores e metadados.

3.3. Reutilizar ou construir os componentes da disseminação

Este subprocesso descreve as atividades para construir novos componentes ou reutilizar componentes existentes necessários para a disseminação de produtos estatísticos, conforme projetado no subprocesso 2.1 (Desenhar os resultados). Todos os tipos de componentes de divulgação estão incluídos, desde aqueles que produzem publicações em papel tradicionais até aqueles que fornecem serviços via web, saídas de dados abertos, estatísticas geoespaciais, mapas, ou acesso aos microdados.

3.4. Configurar fluxos de trabalho

Esse subprocesso configura o fluxo de trabalho, os sistemas e as transformações usados no processo de produção, desde a coleta de dados até a disseminação. Neste subprocesso, o fluxo de trabalho é configurado com base no desenho criado no subprocesso 2.6 (Desenhar fluxograma e sistema de produção). Isso pode incluir a modificação de um fluxo de trabalho padronizado para uma finalidade específica, a montagem conjunta dos fluxos de trabalho para as diferentes fases (possivelmente com um sistema de gerenciamento de fluxo de trabalho/processo) e a configuração dos sistemas em conformidade.

3.5. Testar os sistemas de produção

Este subprocesso está relacionado ao teste de serviços montados e configurados e fluxos de trabalho relacionados. Inclui testes técnicos e aprovação de novos programas e rotinas, bem como a confirmação de que rotinas existentes de outros processos de produção são adequadas para uso neste caso. Embora parte dessa atividade referente ao teste de componentes e serviços individuais possa ser logicamente vinculada aos subprocessos 3.1, 3.2 e 3.3, esse subprocesso também inclui testes de interações entre serviços montados e configurados, garantindo que toda a solução de produção funcione de uma maneira coerente.

3.6. Testar o processo estatístico de produção

Este subprocesso descreve as atividades para gerenciar um teste de campo ou piloto do processo de produção estatística. Normalmente, inclui uma coleta de dados em pequena escala, para testar os instrumentos de coleta, seguida pelo processamento e análise dos dados coletados, para garantir que o processo estatístico funcione conforme o esperado. Após o piloto, pode ser necessário voltar ao passo anterior e fazer ajustes nos instrumentos, sistemas ou componentes de coleta. Para um processo de produção estatística importante, por exemplo, em um censo populacional, pode haver várias iterações até que o processo esteja funcionando satisfatoriamente.

3.7. Finalizar os sistemas de produção

Este subprocesso inclui as atividades para colocar em produção os processos e serviços montados e configurados, incluindo serviços modificados e recém-criados, prontos para uso. As atividades incluem:

- Produzir documentação sobre os componentes do processo, incluindo documentação técnica e manuais do usuário;
- Treinar os usuários sobre como operar o processo;
- Mover os componentes do processo para o ambiente de produção e garantir que eles funcionem conforme o esperado nesse ambiente (essa atividade também pode fazer parte do subprocesso 3.5 - Testar os sistemas de produção).

Fase Coletar

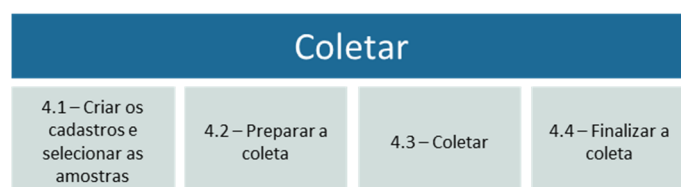


Figura 7. Fase Coletar e seus subprocessos

Esta fase coleta ou reúne toda a informação necessária (por exemplo, dados, metadados e parâmetros), utilizando diferentes modos de coleta (por exemplo, aquisição, coleta direta, extração, transferência) e os carrega no ambiente apropriado para processamento adicional. Embora possa incluir a validação de formatos de conjunto de dados, ela não inclui quaisquer transformações dos dados em si, pois todos são feitos na fase "Processar". Para produtos estatísticos gerados regularmente, essa fase acontece em cada ocorrência.

A fase "Coletar" é dividida em quatro subprocessos (Figura 7), que geralmente são sequenciais, da esquerda para a direita, mas podem ocorrer em paralelo e podem ser iterativos. Estes subprocessos são:

4.1. Criar os cadastros e selecionar as amostras

Este subprocesso estabelece o cadastro e seleciona a amostra para a coleta de uma ocorrência, tal como especificado no subprocesso 2.4 (Especificar o cadastro e a amostragem). Também inclui a coordenação de amostras entre ocorrências do mesmo processo (por exemplo, para gerir sobreposição ou rotação), e entre os diferentes processos utilizando um cadastro ou registro comum (por exemplo, para controlar a sobreposição ou distribuir a carga de resposta). A garantia da qualidade e a aprovação do

cadastro e da amostra selecionada são realizadas também neste subprocesso, ainda que a manutenção de registros subjacentes, a partir dos quais cadastros para vários processos estatísticos são obtidos, seja tratada como um processo separado. O aspecto de amostragem deste subprocesso geralmente não é relevante para processos baseados inteiramente na utilização de fontes preexistentes (por exemplo, registros administrativos, websites) uma vez que esses processos geralmente criam cadastros a partir dos dados disponíveis e, em seguida, seguem uma abordagem de censo. Variáveis oriundas de fontes administrativas e outras fontes não estatísticas de dados podem ser utilizadas como variáveis auxiliares na construção do delineamento amostral.

4.2. Preparar a coleta

Este subprocesso garante que pessoas, processos e tecnologia (por exemplo, aplicativos baseados na Web, sistema GPS) estejam prontos para coletar dados e metadados, em todos os modos, conforme projetado. Ocorre ao longo de um período, pois inclui estratégias, atividades de planejamento e treinamento na preparação para a ocorrência específica do processo de produção estatística. Quando o processo é repetido regularmente, algumas (ou todas) dessas atividades podem não ser explicitamente exigidas para cada ocorrência. Para processos únicos e novos, essas atividades podem ser demoradas. Para dados de pesquisa estatística, este subprocesso inclui:

- Preparar uma estratégia de coleta;
- Treinar a equipe de coleta;
- Sistema de treinamento usando técnicas supervisionadas de aprendizado de máquina;
- Garantir que os recursos de coleta estejam disponíveis (por exemplo, laptops, aplicativos de coleta, APIs);
- Concordar com os termos de quaisquer órgãos de coleta intermediários (por exemplo, subcontratantes para entrevistas telefônicas assistidas por computador, serviços Web);
- Configurar sistemas de coleta para solicitar e receber os dados;
- Garantir a segurança dos dados a serem coletados;
- Preparar instrumentos de coleta (por exemplo, imprimir questionários, preenchê-los previamente com os dados existentes, carregar questionários e dados nos computadores dos entrevistadores, APIs, ferramentas de web scraping);
- Fornecer informações para os entrevistados (por exemplo, redigir cartas ou brochuras explicando o objetivo da pesquisa, notificando os entrevistados quando os instrumentos de coleta on-line forem disponibilizados);

- Traduzir materiais (por exemplo, para os diferentes idiomas falados ou usados no país).

Para fontes que não são de pesquisa, este subprocesso garante que os processos, sistemas e procedimentos de confidencialidade necessários estejam em vigor, para receber ou extrair as informações necessárias da fonte. Isso inclui:

- Avaliar solicitações para adquirir os dados e registrar a solicitação em um inventário centralizado;
- Iniciar contatos com organizações que fornecem os dados e enviar um pacote introdutório com detalhes sobre o processo de aquisição dos dados;
- Verificar informações detalhadas sobre arquivos e metadados com o fornecedor de dados e receber um arquivo de teste para avaliar se os dados estão aptos para uso;
- Organizar canais seguros para a transmissão de dados.

4.3. Coletar

Este subprocesso é o momento em que a coleta é implementada. Os diferentes instrumentos de coleta são usados para coletar informações que podem incluir: microdados brutos ou agregados produzidos na origem, bem como quaisquer metadados associados; contato inicial com fornecedores e quaisquer ações subsequentes de acompanhamento ou lembrete; e entrada manual de dados no ponto de contato ou gerenciamento de trabalho de campo, dependendo da fonte e do modo de coleta.

Registra quando e como os fornecedores foram contatados e se eles responderam. Dependendo do quadro geográfico e da tecnologia utilizada, a geocodificação¹ pode ter que ser feita ao mesmo tempo que a coleta de dados usando sistemas GPS, colocando uma marca em um mapa etc. É incluído nesse subprocesso o gerenciamento dos fornecedores envolvidos na coleta atual, garantindo que o relacionamento entre a organização estatística e os fornecedores de dados seja positivo, registrando e respondendo a comentários, dúvidas e reclamações. A comunicação adequada com as unidades informantes e a minimização do número de não respostas contribuem significativamente para uma melhor qualidade dos dados coletados.

Para dados administrativos, geográficos ou outros dados não estatísticos, o fornecedor ou é contatado para enviar as informações ou as envia conforme programado. Esse processo pode ser demorado e exigir acompanhamento para garantir que os dados sejam fornecidos de acordo com os contratos. No caso em que os dados são publicados

.....
¹ Neste documento, a geocodificação é definida como "atribuir códigos relacionados a lugares geográficos" enquanto o georreferenciamento é definido de forma mais geral como "relacionando dados com o local onde o fato acontece".

sob uma Licença de Dados Abertos e existem em formato legível por máquina, eles podem ser acessados e usados livremente.

Este subprocesso também pode incluir o monitoramento, supervisão da coleta de dados e a realização de quaisquer mudanças necessárias para melhorar a qualidade dos dados. Isso inclui gerar relatórios, visualizar e ajustar o processo de aquisição para garantir que os dados estejam aptos para uso. Quando a coleta atinge seus objetivos, ela é encerrada e um relatório é produzido. Algumas verificações básicas da estrutura e integridade das informações recebidas podem ocorrer nesse subprocesso (por exemplo, verificar se os arquivos estão no formato correto e contêm os campos esperados).

4.4. Finalizar a coleta

Este subprocesso inclui o carregamento dos dados e metadados coletados em um ambiente eletrônico adequado para posterior processamento. Pode conter a captura de dados manual ou automática, por exemplo, usando uma equipe administrativa ou ferramentas de reconhecimento óptico de caracteres para extrair informações de questionários em papel ou converter os formatos de arquivos ou codificar as variáveis recebidas de outras organizações. Também pode englobar a análise dos metadados e parâmetros associados à coleta para garantir que as atividades de coleta atendam aos requisitos. Nos casos em que há um instrumento de coleta físico, como um questionário em papel, que não é necessário para processamento posterior, esse subprocesso gerencia o arquivamento desse material. Quando o instrumento de coleta usa um software como uma API ou um aplicativo, é incluído o controle de versão e o arquivamento destes.

Fase Processar

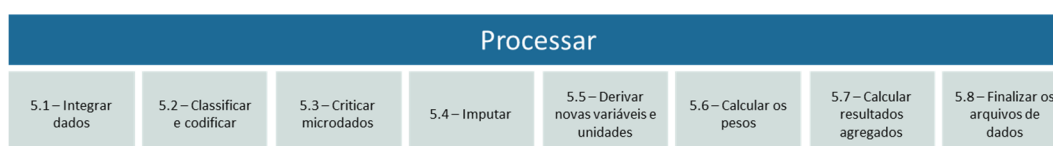


Figura 8. Fase Processar e seus subprocessos

Esta fase descreve o processamento de dados de entrada e sua preparação para análise. É composta de subprocessos que integram, classificam, verificam, limpam e transformam dados de entrada, para que possam ser analisados e divulgados como produtos estatísticos. Para produtos estatísticos gerados regularmente, essa fase acontece em cada ocorrência. Os subprocessos nessa fase podem ser aplicados a fontes estatísticas e não estatísticas (com a possível exceção do subprocesso 5.6 - Calcular pesos, que geralmente é específico dos dados de pesquisa).

As fases "Processar" e "Analisar" podem ser iterativas e paralelas. A análise pode revelar um entendimento mais amplo dos dados, o que pode tornar aparente a

necessidade de processamento adicional. Às vezes, as estimativas que estão sendo processadas podem ser dados agregados já publicados (de acordo com a Política de Revisão).

As atividades dentro das fases "Processar" e "Analisar" também podem começar antes que a fase "Coletar" seja concluída. Isso permite a compilação de resultados provisórios, nas situações em que a oportunidade é uma preocupação importante para os usuários, aumentando o tempo disponível para análise.

A fase "Processar" é dividida em oito subprocessos (Figura 8), que podem ser sequenciais, da esquerda para a direita, mas podem ocorrer em paralelo e podem ser iterativos. Estes subprocessos são:

5.1. Integrar dados

Este subprocesso integra dados de uma ou mais fontes. É onde as saídas dos subprocessos da fase "Coletar" são combinadas. Os dados de entrada podem vir de uma mistura de fontes externas ou internas e de uma variedade de instrumentos de coleta, incluindo extratos de fontes de dados administrativos e outras fontes de dados não-estatísticos. Dados administrativos ou outras fontes não estatísticas de dados podem substituir todas ou algumas das variáveis coletadas diretamente da pesquisa. Este subprocesso inclui também a harmonização ou criação de novos números que concordem entre as fontes de dados. A saída é um conjunto de dados combinados. A integração de dados pode incluir:

- Combinação de dados de múltiplas fontes, como parte da criação de estatísticas integradas, tais como contas nacionais;
- Combinação de dados geoespaciais e dados estatísticos ou outros dados não estatísticos;
- Agrupamento de dados, com o objetivo de aumentar o número efetivo de observações de alguns fenômenos;
- Rotinas de pareamento de dados, com o objetivo de ligar micro ou macrodados de diferentes fontes;
- Fusão de dados - integração seguida de redução ou substituição;
- Priorização, quando duas ou mais fontes contêm dados para a mesma variável, com valores potencialmente diferentes.

A integração de dados pode ocorrer em qualquer ponto desta fase, antes ou depois de qualquer um dos outros subprocessos. Também pode haver várias instâncias de integração de dados em um processo de produção estatística. Após a integração, dependendo dos requisitos de proteção de dados, os dados podem ser desidentificados, ou seja, os identificadores, tais como nome e endereço, são excluídos, para ajudar a proteger a confidencialidade.

5.2. Classificar e codificar

Este subprocesso classifica e codifica os dados de entrada. Por exemplo, rotinas de codificação automática (ou administrativa) podem atribuir códigos numéricos a respostas de texto de acordo com uma classificação estatística pré-determinada, para facilitar a captura e o processamento de dados. Em algumas perguntas, as categorias de resposta já estão codificadas nos questionários ou na fonte administrativa de dados, em outras são codificadas após a coleta usando um processo automatizado (que pode aplicar técnicas de aprendizado de máquina) ou um processo manual interativo.

5.3. Criticar microdados

Este subprocesso examina os dados para identificar possíveis problemas, erros e discrepâncias, tais como valores atípicos (outliers), itens sem resposta e erros de codificação. Ele também pode ser chamado de validação de dados de entrada. Pode ser executado iterativamente, validando dados usando regras de crítica predefinidas, geralmente em uma ordem especificada. Pode sinalizar dados para imputação automática ou manual. A crítica pode ser aplicada a dados de qualquer tipo de fonte, antes e depois da integração, bem como a dados imputados do subprocesso 5.4 (Imputar). Embora a crítica seja tratada como parte da fase "Processar", na prática, alguns elementos de crítica podem ocorrer juntamente com atividades de coleta, particularmente para coleta via Internet e com coletor eletrônico.

Este subprocesso cuida da crítica de microdados, ou seja, dos dados individuais, a crítica de agregados ocorre no subprocesso 6.2 (Validar os resultados), que pode ser executada em paralelo.

Enquanto este subprocesso está relacionado à detecção e localização de erros reais ou potenciais, quaisquer atividades de correção que realmente alterem os dados são feitas no subprocesso 5.4 (Imputar). Geralmente esses dois subprocessos são executados simultaneamente, as regras de crítica são definidas em conjunto com as regras de imputação.

5.4. Imputar

Quando os dados são considerados incorretos, ausentes, não confiáveis ou desatualizados, novos valores podem ser inseridos ou dados desatualizados podem ser removidos neste subprocesso. A imputação abrange uma variedade de métodos, geralmente usando uma abordagem baseada em regras. Etapas específicas geralmente incluem:

- Determinar se deve adicionar ou alterar dados;
- Selecionar o método a ser usado;
- Adicionar/alterar valores de dados;

- Escrever os novos valores de dados de volta no conjunto de dados e sinalizá-los como alterados;
- Produzir metadados no processo de imputação.

5.5. Derivar novas variáveis e unidades

Este subprocesso deriva dados para variáveis e unidades que não são explicitamente fornecidas na coleta, mas são necessárias para entregar os resultados desejados. Derivam-se novas variáveis aplicando fórmulas aritméticas a uma ou mais variáveis que já estão presentes no conjunto de dados ou aplicando diferentes hipóteses de um modelo. Essa atividade pode precisar ser iterativa, já que algumas variáveis derivadas podem ser baseadas em outras variáveis derivadas. Portanto, é importante garantir que as novas variáveis sejam derivadas na ordem correta. Novas unidades podem ser derivadas, agregando ou dividindo dados das unidades de coleta ou por vários outros métodos de estimação. Exemplos incluem famílias derivadas de unidades de coleta que são pessoas, ou unidades locais derivadas de empresas.

5.6. Calcular os pesos

Este subprocesso cria pesos para registros de dados individuais de acordo com a metodologia desenvolvida no subprocesso 2.5 (Definir o processamento e a análise). Por exemplo, os pesos podem ser usados para "aumentar os dados" para torná-los representativos da população-alvo (em pesquisas por amostragem ou em extratos de dados do scanner) ou para ajustar a não resposta em enumerações totais. Em outras situações, as variáveis podem precisar de ponderação para fins de normalização. Pode também incluir ajuste de peso por pós-estratificação ou calibração para valores conhecidos (por exemplo, totais populacionais).

5.7. Calcular resultados agregados

Este subprocesso cria dados agregados e totais populacionais de microdados ou de dados agregados de nível inferior. Inclui resumo dos dados de registros que compartilham certas características (por exemplo, agregação de dados por classificações demográficas ou geográficas), determinação de médias e medidas de dispersão e aplicação de pesos do subprocesso 5.6 (Calcular os pesos) para obter totais apropriados. No caso de produtos estatísticos provenientes de pesquisas por amostragem, os erros de amostragem correspondentes aos agregados relevantes também podem ser calculados neste subprocesso.

5.8. Finalizar os arquivos de dados

Este subprocesso reúne as saídas dos outros subprocessos desta fase em um arquivo de dados (geralmente macrodados), que é usado como entrada para a fase "Analisar". Às vezes, pode ser um arquivo intermediário, em vez de final, especialmente para processos em que existem fortes pressões de tempo e requisitos para produzir tanto estimativas preliminares quanto finais.

Fase Analisar

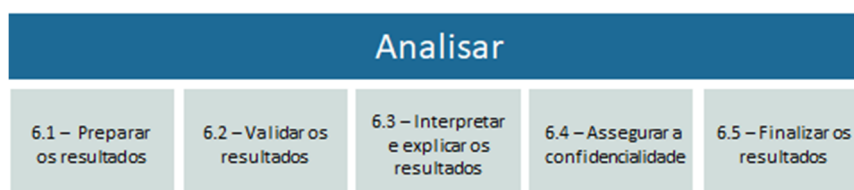


Figura 9. Fase Analisar e seus subprocessos

Nesta fase, os produtos estatísticos são gerados e examinados em detalhe. Inclui a preparação de conteúdo estatístico (incluindo comentários, notas técnicas, etc.) e a garantia de que os resultados sejam “adequados à finalidade” antes da divulgação aos usuários. Essa fase também inclui os subprocessos e atividades que permitem aos analistas estatísticos entender os dados e as estatísticas produzidas. As saídas dessa fase também podem ser usadas como entradas para outros subprocessos (por exemplo, análise de novas fontes como entrada para a fase “Planejar”). Para produtos estatísticos gerados regularmente, essa fase acontece em cada ocorrência. A fase “Analisar” e seus subprocessos são genéricos para todos os produtos estatísticos, independentemente de como os dados foram obtidos.

A fase "Analisar" é dividida em cinco subprocessos (Figura 9), que geralmente são sequenciais, da esquerda para a direita, mas também podem ocorrer em paralelo e podem ser iterativos. Estes subprocessos são:

6.1. Preparar os resultados

Este subprocesso é o momento em que os dados dos subprocessos 5.7 (Calcular resultados agregados) e 5.8 (Finalizar os arquivos de dados) são transformados em produtos estatísticos, tais como índices e estatísticas ajustadas sazonalmente, como por exemplo tendências, ciclos, componentes sazonais e irregulares, medidas de acessibilidade etc., bem como o registro de características de qualidade, como coeficientes de variação. A elaboração de mapas, saídas GIS e serviços geoestatísticos podem ser incluídos para maximizar o valor e a capacidade de analisar as informações estatísticas.

6.2. Validar os resultados

Este subprocesso é o momento em que os estatísticos validam a qualidade dos resultados produzidos, de acordo com o quadro de referência de qualidade e com as expectativas. Este subprocesso inclui atividades envolvidas com reuniões de inteligência, acumulando conhecimento sobre um domínio estatístico específico. Este conhecimento é então aplicado à coleta atual, no ambiente atual, para identificar qualquer divergência das expectativas e permitir análises embasadas. As atividades de validação podem incluir:

- Verificar se a cobertura da população e as taxas de resposta são as planejadas;
- Comparar as estatísticas com os ciclos anteriores (se aplicável);
- Verificar se os metadados, parâmetros e indicadores de qualidade associados estão presentes e alinhados com as expectativas;
- Verificar a consistência geoespacial dos dados;
- Confrontar as estatísticas com outros dados relevantes (internos e externos);
- Investigar inconsistências nas estatísticas;
- Realizar a crítica de agregados;
- Validar as estatísticas em relação às expectativas e ao conhecimento específico.

6.3. Interpretar e explicar os resultados

Este subprocesso é o momento em que uma compreensão profunda dos resultados é obtida pelos estatísticos. Eles usam esse entendimento para interpretar e explicar as estatísticas, avaliando como refletem suas expectativas iniciais, visualizando-as de todas as perspectivas com o uso de diferentes ferramentas e mídias, e realizando análises estatísticas detalhadas, como análise de séries temporais, consistência e análise de comparabilidade, análise de revisão (análise das diferenças entre estimativas preliminares e estimativas revisadas), análise de assimetrias (discrepâncias nas estatísticas espelho), etc.

6.4. Assegurar a confidencialidade

Este subprocesso garante que os dados (e metadados) a serem divulgados não violem as regras apropriadas de confidencialidade, de acordo com as políticas e regras da organização, ou com a metodologia específica do processo em análise criada no subprocesso 2.5 (Definir o processamento e a análise). Pode incluir verificações de confidencialidade primária e secundária, bem como a aplicação de supressão de dados ou de técnicas de perturbação e verificação dos resultados. O grau e o método de controle estatístico da confidencialidade podem variar para diferentes tipos de resultados. Por exemplo, a abordagem usada para conjuntos de microdados para fins de estudo e

pesquisa será diferente daquela para tabelas publicadas, produtos de estatísticas geoespaciais ou visualizações em mapas.

6.5. Finalizar os resultados

Este subprocesso garante que as estatísticas e as informações associadas sejam adequadas à finalidade e atinjam o nível de qualidade exigido e, portanto, estejam prontas para uso. Inclui:

- Completar verificações de consistência;
- Determinar o nível de disseminação e aplicar advertências;
- Reunir informações de apoio, incluindo interpretação, comentários, notas técnicas, instruções, medidas de incerteza e quaisquer outros metadados necessários;
- Produzir os documentos internos de suporte;
- Realizar a discussão de pre-release com especialistas internos apropriados;
- Traduzir os produtos estatísticos em países com disseminação multilíngue;
- Aprovar o conteúdo estatístico para disseminação.

Fase Disseminar

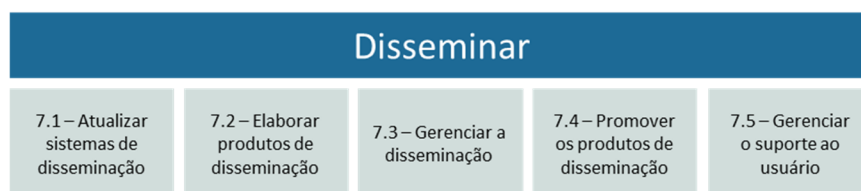


Figura 10. Fase Disseminar e seus subprocessos

Esta fase gerencia a disseminação dos produtos estatísticos aos usuários. Inclui todas as atividades associadas à montagem e disseminação de uma série de produtos estáticos e dinâmicos por meio de diversos canais. Essas atividades dão suporte aos usuários para acessar e usar os produtos disseminados pela organização estatística. Para produtos estatísticos gerados regularmente, essa fase acontece em cada ocorrência.

A fase “Disseminar” é dividida em cinco subprocessos (Figura 10), que geralmente são sequenciais, da esquerda para a direita, mas também podem ocorrer em paralelo e podem ser iterativos. Estes subprocessos são:

7.1. Atualizar sistemas de disseminação

Este subprocesso gerencia a atualização de sistemas (por exemplo, bancos de dados), onde dados e metadados são armazenados prontos para fins de disseminação, incluindo:

- Formatação de dados e metadados prontos para serem colocados em sistemas de disseminação;
- Carregamento de dados e metadados em sistemas de disseminação;
- Garantir que os dados estejam vinculados aos metadados relevantes.

A formatação, o carregamento e a vinculação de metadados devem, preferencialmente, ocorrer em fases anteriores, mas esse subprocesso inclui uma verificação final de que todos os metadados necessários estão prontos para a disseminação.

7.2. Elaborar os produtos de disseminação

Este subprocesso elabora os produtos de disseminação, conforme previamente projetado no subprocesso 2.1 (Desenhar os resultados), para atender às necessidades dos usuários. Os produtos de disseminação podem incluir publicações impressas, comunicados de imprensa e sites. Podem também assumir várias formas, incluindo gráficos interativos, tabelas, mapas, conjuntos de microdados de uso público, dados abertos vinculados e arquivos para download. Etapas típicas incluem:

- Preparar os componentes do produto (textos explicativos, tabelas, gráficos, mapas, declarações de qualidade, etc.);
- Montar os componentes em produtos;
- Editar os produtos e verificar se atendem aos padrões de publicação.

7.3. Gerenciar a disseminação

Este subprocesso garante que todos os elementos da disseminação estejam em vigor, incluindo o gerenciamento do tempo da disseminação. Inclui briefings para grupos específicos, como a imprensa ou ministros, bem como os arranjos para quaisquer embargos de pre-release. Também inclui o fornecimento de produtos aos assinantes e o gerenciamento do acesso a dados confidenciais por grupos de usuários autorizados, tais como pesquisadores. Às vezes, uma organização pode precisar retirar um produto, por exemplo, se um erro for descoberto. Isso também está incluído neste subprocesso.

7.4. Promover os produtos de disseminação

Embora o marketing em geral possa ser considerado um processo abrangente, este subprocesso diz respeito à promoção ativa dos produtos de disseminação elaborados num processo estatístico específico, para ajudá-los a atingir o maior público possível. Inclui o uso de ferramentas de gestão de relacionamento com o usuário para melhor direcionar possíveis usuários dos produtos, bem como o uso de sites, wikis e blogs, para facilitar o processo de comunicação de informações estatísticas aos usuários.

7.5. Gerenciar o suporte ao usuário

Este subprocesso garante que as consultas e solicitações de usuários, como o acesso a microdados, sejam registradas e que as respostas sejam fornecidas dentro dos prazos acordados. Essas consultas e solicitações devem ser revisadas regularmente para fornecer contribuições para o processo de gestão de qualidade global, pois elas podem indicar necessidades novas dos usuários. As respostas às solicitações do usuário também podem ser usadas para preencher um banco de dados de conhecimento ou uma página de “Perguntas frequentes”, que é disponibilizada publicamente, reduzindo assim a carga de resposta a solicitações repetidas e/ou similares de usuários externos. Este subprocesso também inclui a gestão de suporte a qualquer organização parceira envolvida na disseminação dos produtos.

Fase Avaliar

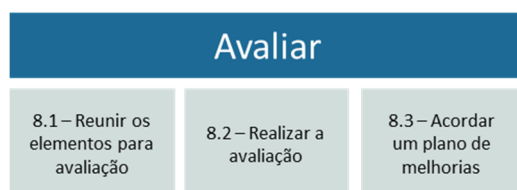


Figura 11. Fase Avaliar e seus subprocessos

Esta fase gerencia a avaliação de uma ocorrência específica de um processo estatístico, em oposição ao processo mais geral de gestão da qualidade descrito na Seção VI (Processos Gerais). Pode ocorrer no final do processo, mas também pode ser feito de forma contínua durante o processo de produção estatística. Baseia-se em insumos coletados ao longo das diferentes fases. Inclui a avaliação do sucesso de uma ocorrência específica do processo estatístico, com base em uma série de dados quantitativos e qualitativos, e na identificação e priorização de possíveis melhorias.

Para os produtos estatísticos gerados regularmente, a avaliação deve, pelo menos teoricamente, acontecer para cada ocorrência, determinando se deve haver ocorrências futuras e, em caso afirmativo, se melhorias devem ser implementadas. No entanto, em alguns casos, particularmente para processos estatísticos regulares e bem estabelecidos,

a avaliação pode não ser formalmente executada para cada ocorrência. Nesses casos, essa fase pode ser vista como uma decisão sobre se a próxima ocorrência deve começar a partir da fase “Especificar Necessidades” ou de alguma fase posterior (geralmente a fase “Coletar”).

A fase “Avaliar” é dividida em três subprocessos (Figura 11), que geralmente são sequenciais, da esquerda para a direita, mas também podem ocorrer em paralelo e podem ser iterativos. Estes subprocessos são:

8.1. Reunir os elementos para avaliação

O material de avaliação pode ser produzido em qualquer outra fase ou subprocesso. Pode ter várias formas, incluindo feedback dos usuários, metadados do processo (paradados), métricas do sistema e sugestões da equipe. Os relatórios de progresso em relação a um plano de ação acordado durante uma ocorrência anterior também podem formar uma entrada para avaliações de ocorrências subsequentes. Este subprocesso reúne todas essas informações, compila os indicadores de qualidade e os disponibiliza para a pessoa ou equipe que produz a avaliação. A coleta de alguns desses materiais de avaliação pode ser automatizada e ocorrer de forma contínua ao longo de todo o processo, conforme definido pela estrutura de qualidade (consulte Gestão da Qualidade na Seção VI). Por outro lado, para a avaliação de certos processos, pode ser necessário realizar atividades específicas, como pequenas pesquisas (por exemplo, pesquisas pós-enumeração, estudos de re-entrevista, pesquisa sobre a eficácia da disseminação).

8.2. Realizar a avaliação

Este subprocesso analisa as informações de avaliação, compara-os com os resultados esperados/alvo (quando disponíveis) e os sintetiza em um relatório de avaliação ou painel de controle. A avaliação pode ocorrer no final de todo o processo para atividades selecionadas, durante a sua execução de forma contínua, ou durante todo o processo, permitindo assim soluções rápidas ou melhoria contínua. O relatório resultante deve anotar quaisquer problemas de qualidade específicos a essa ocorrência do processo estatístico, bem como destacar qualquer desvio das métricas de desempenho dos valores esperados, e deve fazer recomendações para alterações, se apropriado. Essas recomendações podem abranger alterações em qualquer fase ou subprocesso para futuras ocorrências ou podem sugerir que o processo não seja repetido.

8.3. Estabelecer um plano de melhoria

Este subprocesso reúne o poder de decisão necessário para estabelecer um plano de melhoria com base no relatório de avaliação. Também deve incluir a consideração de um mecanismo de monitoramento do impacto dessas ações, que, por sua vez, pode fornecer um insumo para as avaliações de futuras ocorrências do processo.

Processos Gerais

O GSBPM reconhece vários processos abrangentes que se aplicam em todas as fases do processo de produção estatística. Alguns desses processos abrangentes estão listados na Seção II. Os processos de gestão da qualidade, gestão de metadados e gestão de dados são detalhados nesta seção.

Gestão da Qualidade

Gestão da Qualidade

A qualidade diz respeito a organizações, produtos, fontes e processos. Na estrutura atual, o processo abrangente de gestão da qualidade refere-se à qualidade do produto e do processo. Qualidade no nível institucional (por exemplo, adoção de uma Política de Qualidade ou Estrutura de Garantia de Qualidade) é considerada no GAMSQ.

O principal objetivo da gestão da qualidade dentro do processo estatístico é entender e gerenciar a qualidade das fontes, processos e produtos estatísticos. Existe um consenso entre as organizações estatísticas de que a qualidade deve ser definida de acordo com a norma ISO 9000-2015: “O grau em que um conjunto de características inerentes a um objeto preenche os requisitos.” Assim, qualidade é um conceito complexo e multifacetado, geralmente definido em termos de várias dimensões de qualidade. As dimensões de qualidade que são consideradas mais importantes dependem das perspectivas, necessidades e prioridades do usuário, que variam entre processos e entre grupos de usuários.

Para melhorar a qualidade, a gestão da qualidade deve estar presente em todo o processo de produção estatística. Está intimamente ligado à fase “Avaliar”, no entanto, a gestão da qualidade tem um escopo mais amplo e profundo. Além de avaliar ocorrências de um processo, também é necessário avaliar fases e subprocessos separados, idealmente a cada vez que são aplicados, mas pelo menos de acordo com um cronograma acordado. Metadados gerados pelos diferentes subprocessos também são de interesse como uma entrada para a gestão da qualidade do processo. Essas avaliações podem ser aplicadas em um processo específico ou em vários processos que usam componentes comuns. Além disso, um papel fundamental na gestão da qualidade é desempenhado pelo conjunto de ações de controle de qualidade que devem ser implementados nos subprocessos para prevenir e monitorar erros e fontes de riscos. Estes devem ser documentados e podem ser usados para relatórios de qualidade.

Dentro de uma organização, a gestão da qualidade geralmente se refere a uma estrutura de qualidade específica e, portanto, pode assumir diferentes formas e fornecer resultados diferentes dentro de diferentes organizações. A multiplicidade atual de estruturas de qualidade aumenta a importância das abordagens de benchmarking e revisão por pares para avaliação, e embora essas abordagens provavelmente não sejam

viáveis para cada iteração de cada parte de cada processo estatístico, elas devem ser usadas de maneira sistemática de acordo com um cronograma pré-determinado que permite a revisão de todas as partes principais do processo dentro de um período especificado.

Ampliando o campo de aplicação do processo geral de gestão da qualidade, a avaliação de grupos de processos estatísticos também pode ser considerada, a fim de identificar possíveis duplicações ou lacunas.

Todas as avaliações resultam em feedback, que deve ser usado para melhorar o processo, fase ou subprocesso relevante, criando um ciclo de qualidade que reforça a abordagem de melhorias contínuas e aprendizado organizacional.

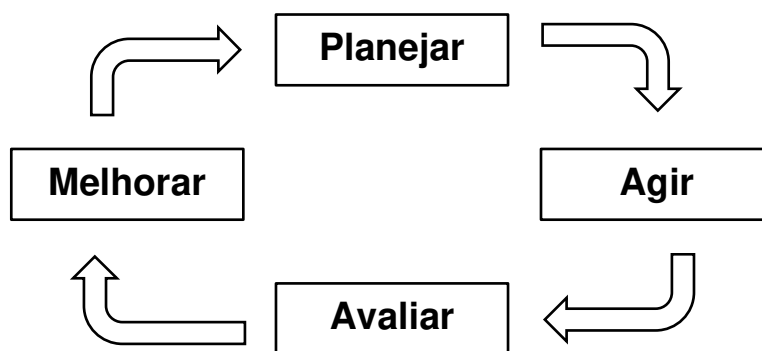


Figura 12. Ciclo de qualidade

Exemplos de atividades de gestão da qualidade incluem:

- Avaliar riscos e implementar tratamentos de risco para garantir a qualidade adequada ao objetivo;
- Definir critérios de qualidade a serem utilizados no processo;
- Definir metas de qualidade do processo e monitoramento de conformidade;
- Buscar e analisar o feedback do usuário;
- Revisar as operações e documentar as lições aprendidas;
- Analisar metadados de processos e indicadores de qualidade;
- Auditoria interna ou externa no processo.

Os indicadores de qualidade apoiam uma gestão da qualidade orientada para o processo. Uma lista sugerida de indicadores de qualidade para fases e subprocessos do GSBPM, bem como para os processos gerais de qualidade e gestão de metadados pode ser encontrada nos Indicadores de Qualidade para o GSBPM - para Estatísticas derivadas de Pesquisas e Fontes de Dados Administrativos. Entre outros, eles podem ser usados como uma lista de verificação para identificar lacunas e/ou duplicação de trabalho na organização.

Gestão de Metadados

Gestão de Metadados

Os metadados têm um papel importante e devem ser gerenciados em um nível operacional dentro do processo de produção estatística. Quando aspectos da gestão de metadados são considerados no nível corporativo ou estratégico (por exemplo, há sistemas de metadados que afetam grandes partes do sistema de produção), isso deve ser considerado no contexto do GAMSO.

A boa gestão de metadados é essencial para a operação eficiente dos processos estatísticos. Os metadados estão presentes em todas as fases, criados, atualizados ou transportados de uma fase anterior ou reutilizados de outro processo estatístico. No contexto desse modelo, a ênfase do processo geral de gestão de metadados está na criação/revisão, atualização, uso e arquivamento de metadados estatísticos, embora os metadados dos diferentes subprocessos também sejam de interesse, inclusive como um insumo para a gestão da qualidade. O principal desafio é garantir que esses metadados sejam capturados o mais cedo possível, armazenados e transferidos de fase para fase, ao lado dos dados aos quais eles se referem. A estratégia e os sistemas de gestão de metadados são, portanto, vitais para a operação deste modelo, e estes podem ser facilitados pelo GSIM.

O GSIM é uma estrutura de referência de objetos de informação, que permite descrições genéricas da definição, gestão e uso de dados e metadados ao longo do processo de produção estatística. O GSIM suporta uma abordagem consistente aos metadados, facilitando a função primária dos metadados, isto é, que os metadados devem definir de forma única e formal o conteúdo e os links entre objetos e processos de informação no sistema de informação estatística.

A Estrutura Comum de Metadados do METIS identifica os seguintes dezesseis princípios básicos para a gestão de metadados, os quais devem ser cobertos no processo de gestão de metadados geral e levados em consideração ao projetar e implementar um sistema de metadados estatísticos. Os princípios são apresentados em quatro grupos:

1- Tratamento de metadados:

- Modelo de processos estatísticos: gerenciar metadados com foco no modelo geral de processos estatísticos;
- Ativo, não passivo: tornar os metadados ativos o máximo possível. Metadados ativos são metadados que impulsionam outros processos e ações. Tratar metadados dessa maneira garantirá que eles sejam precisos e atualizados;
- Reutilizar: reutilizar metadados sempre que possível para fins de integração estatística e de eficiência;
- Versões: preservar o histórico (versões antigas) dos metadados.

2- Autoridade dos metadados:

- - Registro: Assegurar-se de que o processo de registro (fluxo de trabalho) associado a cada elemento de metadados esteja bem documentado, para que haja clara identificação de propriedade, status de aprovação, data de operação etc.;
- - Fonte única: verificar se existe uma única fonte autorizada (“autoridade de registro”) para cada elemento de metadados;
- - Uma entrada/atualização: Minimizar erros inserindo uma vez e atualizando em um lugar;
- - Variações de padrões: garantir que as variações dos padrões sejam bem gerenciadas/aprovadas, documentadas e visíveis.

3- Relação com o Ciclo Estatístico/ Processos:

- - Integridade: Tornar o trabalho relacionado a metadados uma parte integrante dos processos em toda a organização;
- - Metadados correspondentes: verificar se os metadados apresentados aos usuários finais correspondem aos metadados que direcionaram o processo ou foram criados durante o processo;
- - Descrever o fluxo: Descrever o fluxo de metadados com os processos (juntamente com o fluxo de dados e a lógica de negócios);
- - Capturar na origem: Capturar metadados em sua origem, preferencialmente automaticamente como um subproduto de outros processos;
- - Troca e uso: trocar metadados e usá-los para informar processos baseados em computador e interpretação humana. A infraestrutura para troca de dados e metadados associados deve ser baseada em componentes fracamente acoplados, com uma escolha de linguagens padrão de troca, como XML.

4- Usuários:

- - Identificar usuários: garantir que os usuários sejam claramente identificados para todos os processos de metadados e que toda a captura de metadados crie valor para eles;
- - Diferentes formatos: A diversidade de metadados deve ser reconhecida e existem diferentes visões correspondentes aos diferentes usos dos dados. Usuários diferentes exigem diferentes níveis de detalhes. Os metadados devem aparecer em diferentes formatos, dependendo dos processos e objetivos para os quais são produzidos e usados;
- - Disponibilidade: verificar se os metadados estão prontamente disponíveis e podem ser usados no contexto das necessidades de informações dos usuários (seja um usuário interno ou externo).

Gestão de Dados

Gestão de Dados

A gestão de dados é essencial, pois os dados são produzidos em muitas das atividades do processo estatístico e são os principais resultados. O principal objetivo da gestão de dados é garantir que os dados sejam usados adequadamente em todo o seu ciclo de vida. A gestão de dados em todo o seu ciclo de vida abrange atividades como planejamento e avaliação de processos de gestão de dados, além de estabelecer e implementar processos relacionados à coleta, organização, uso, proteção, preservação e descarte dos dados.

A forma como os dados são gerenciados estará intimamente ligado ao uso dos dados, que por sua vez está vinculado ao processo estatístico onde os dados são criados. Tanto os dados quanto os processos nos quais eles são criados devem estar bem definidos para garantir uma gestão de dados adequada.

Exemplos de atividades de gestão de dados incluem:

- Estabelecer uma estrutura de governança e atribuir responsabilidades de administração de dados;
- Projetar estruturas de dados e conjuntos de dados associados e o fluxo de dados pelo processo estatístico;
- Identificar o banco de dados (repositórios) para armazenar os dados e administrá-lo;
- Documentar os dados (por exemplo, registrar e inventariar dados, classificar os dados de acordo com o conteúdo, retenção ou outra classificação exigida);
- Determinar períodos de retenção de dados;
- Proteger dados contra acesso e uso não autorizados;
- Proteger dados contra mudanças tecnológicas, degradação da mídia física, corrupção de dados;
- Executar verificações de integridade de dados (por exemplo, verificações periódicas que garantem a precisão e consistência dos dados durante todo o seu ciclo de vida);
- Executar atividades de descarte quando o período de retenção dos dados expirar.

Outros usos do GSBPM

O objetivo original do GSBPM era fornecer uma base para as organizações estatísticas concordarem com a terminologia padrão para discussões sobre o desenvolvimento de sistemas e processos de metadados estatísticos. No entanto, à medida que o modelo se desenvolveu, tornou-se cada vez mais utilizado para muitos outros fins, em particular, relacionados com a modernização das estatísticas oficiais. Vários artigos descrevendo usos reais e potenciais do GSBPM estão disponíveis na plataforma wiki da UNECE. A lista abaixo tem como objetivo destacar alguns usos atuais e inspirar ideias adicionais sobre como o GSBPM pode ser usado na prática.

- Fornecer uma estrutura para documentação de processos estatísticos - O GSBPM pode fornecer uma estrutura para organizar e armazenar documentação dentro de uma organização, promovendo a padronização e a identificação de boas práticas;
- Facilitar o compartilhamento de métodos e softwares estatísticos - O GSBPM define os componentes dos processos estatísticos de uma forma que não apenas incentiva o compartilhamento de métodos e ferramentas de software entre processos estatísticos, mas também facilita o compartilhamento entre diferentes organizações estatísticas que aplicam o modelo;
- Descrever quais padrões são ou poderiam ser usados para diferentes fases do processo de produção estatística. Por exemplo, o Anexo 2 do Guia do Usuário do SDMX 2.1 explora como o SDMX se aplica ao trabalho estatístico no contexto de um modelo de processo de negócios;
- Fornecer uma estrutura para avaliação e melhoria da qualidade do processo – Para que uma abordagem de benchmarking para a avaliação da qualidade do processo seja bem-sucedida, é necessário padronizar os processos tanto quanto possível. O GSBPM fornece um mecanismo para facilitar isso;
- Melhor integração do trabalho em metadados estatísticos e qualidade - Relacionado com o ponto anterior, o quadro comum fornecido pelo GSBPM pode ajudar a integrar o trabalho internacional sobre metadados estatísticos com o da qualidade dos dados, fornecendo um quadro e uma terminologia comuns para descrever o processo estatístico;
- Fornecer modelo subjacente para estruturas de padrões metodológicos - Padrões metodológicos podem ser vinculados à(s) fase(s) ou subprocesso(s) com os quais se relacionam e podem ser classificados e armazenados em uma estrutura baseada no GSBPM;
- Desenvolver um repositório de modelos de processos estatísticos para armazenar saídas de modelagem de processos e permitir que elas sejam vinculadas ao modelo estatístico de processos;

- Fornecer um modelo subjacente para conceber uma estrutura de treinamento e um conjunto de terminologia padrão para descrever as habilidades e conhecimentos necessários no processo de produção estatística;
- Medição de custos operacionais - O GSBPM pode ser usado como base para medir os custos de diferentes partes do processo estatístico. Isso ajuda a direcionar as atividades de modernização para melhorar a eficiência das partes do processo que são mais caras;
- Desempenho do sistema de medição - Relacionado ao ponto acima dos custos, o GSBPM também pode ser usado para identificar componentes que não estão funcionando de maneira eficiente, que estão se duplicando desnecessariamente ou que requerem substituição. Da mesma forma, pode identificar lacunas para as quais novos componentes devem ser desenvolvidos;
- Fornecer uma ferramenta para alinhar os processos de fornecedores de dados não-estatísticos (por exemplo, dados administrativos, dados geoespaciais), facilitando a comunicação entre estatísticos e especialistas de outros domínios e harmonizando a terminologia relacionada;
- Fornecer uma ferramenta para construir capacidades e construir conhecimento técnico metodicamente, referindo-se aos detalhes de cada fase;
- Fornecer uma ferramenta para o desenvolvimento/revisão de classificações estatísticas.

Referências

UNECE Statistics Wikis - HLG-MOS (<https://statswiki.unece.org/display/hlgbas>)

UNECE Statistics Wikis - GSBPM (<https://statswiki.unece.org/display/GSBPM>)

UNECE Applying the Generic Statistical Business Process Model to business register maintenance (<https://www.unece.org/index.php?id=18256>)

UNECE Statistics Wikis - GAMSO (<https://statswiki.unece.org/display/GAMSO>)

UNECE Statistics Wikis - GSIM (<https://statswiki.unece.org/display/GSIM>)

UNECE Statistics Wikis - Generic Statistical Data Editing Models (<https://statswiki.unece.org/display/sde/GSDEMs>)

ISO 9000:2015, Quality management systems - Fundamentals and vocabulary. Organização Internacional para Padronização

National Quality Assurance Framework (<http://unstats.un.org/unsd/dnss/QualityNQAF/nqaf.aspx>)

UNECE Statistics Wikis - Quality Indicators for the GSBPM (<https://statswiki.unece.org/display/GSBPM/Quality+Indicators>)

UNECE Statistics Wikis - Uses of the GSBPM (<https://statswiki.unece.org/display/GSBPM/Uses+of+GSBPM>)

SDMX community (http://sdmx.org/index.php?page_id=38)

Anexos

1. Lista de Acrônimos

1 - Lista de Acrônimos

API – Application Programming Interface

GAMSO – Generic Activity Model for Statistical Organisations: Uma estrutura de referência que descreve e define as atividades que ocorrem dentro de uma organização estatística típica

GIS – Geographic Information System

GPS – Global Positioning System

GSBPM – Generic Statistical Business Process Model: Modelo flexível para descrever e definir o conjunto de processos necessários para produzir estatísticas oficiais

GSIM – Generic Statistical Information Model: Uma estrutura de referência de objetos de informação, que permite descrições genéricas da definição, gestão e uso de dados e metadados ao longo do processo de produção estatística

HLG-MOS – High-Level Group for the Modernisation of Official Statistics

ICT – Information and Communications Technology

METIS – A "marca" para o grupo de trabalho em Statistical Metadata under the Conference of European Statisticians

OECD – Organisation for Economic Cooperation and Development

SDMX - Statistical Data and Metadata eXchange: Um conjunto de normas técnicas e diretrizes orientadas a conteúdo, juntamente com uma arquitetura e ferramentas de TI, a serem usadas para a troca e o compartilhamento eficientes de dados estatísticos e metadados

UNECE – United Nations Economic Commission for Europe

XML – eXtensible Mark-up Language: Uma linguagem que define um conjunto de regras para codificar documentos em um formato legível por máquina

Equipe técnica

Diretoria de Pesquisas

Coordenação de Métodos e Qualidade

Andrea Borges Paim

Gerência de Qualidade Estatística

Raquel Rose Silva Correia

Maria Emilia Freitas Haussmann

Gerência de Metodologia Estatística

Marcos Paulo Soares de Freitas

Se o assunto é **Brasil**,
procure o **IBGE**.



/ibgecomunica



/ibgeoficial



/ibgeoficial



/ibgeoficial

www.ibge.gov.br 0800 721 8181

