

Recomendações e desafios técnicos PNCP: API & Acesso aos dados

Junho/2024

Versão 1.0



Este trabalho está sob a licença [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/). Mediante atribuição de crédito à organização autora, pode ser copiado e redistribuído em qualquer suporte ou formato; remixado e adaptado para qualquer fim, inclusive comercial (nestes casos, as alterações feitas devem ser indicadas).

FICHA TÉCNICA

junho/2024

DIREÇÃO EXECUTIVA

Juliana Sakai

SUPERVISÃO E EDIÇÃO

Talita Lôbo

REDAÇÃO

Jessé Oliveira

Joellen Silva

Lívia Almeida

Talita Lôbo

Colin Maudry

REVISÃO

Mariana Lopez

Raul Durlo

Yohanna Lisnichuk

PARCERIA

**OPEN
CONTRACTING
PARTNERSHIP**

1. Resumo executivo

Criado para centralizar os processos de contratação de todo país, o Portal Nacional de Contratações Públicas (PNCP) entrou em uso em janeiro de 2024. Na intenção de identificar compras de medicamentos e analisar o preço dos itens dessas compras, a Transparência Brasil utilizou o portal e suas ferramentas (APIs, manuais etc.) para acessar dados de contratação publicamente disponibilizados pelo governo federal, estados e municípios.

Em sua primeira versão, este relatório descreve os principais entraves no processo de busca, download, pré-processamento e uso de dados.

Além da impossibilidade de download de dados em massa, também verificou-se limitações nas APIs disponibilizadas, oportunidades de melhoria no esquema de dados e, em destaque, a pouca acessibilidade no acompanhamento do processo de contratação e suas possíveis atualizações. Destaca-se, também, a pulverização dos dados de uma contratação em múltiplos arquivos ou endpoints que contém parte(s) do processo.

Para além da complexidade do tema, os desafios levantados implicam na necessidade de letramento técnico avançado para cruzamento, filtragem e entendimento dos arquivos de dados. Esse processo representa, por si, um empecilho relevante para o acesso e fiscalização do gasto público pela sociedade civil, academia e, muitas vezes, até mesmo pelo poder público.

Tendo em vista o potencial estratégico do PNCP na transparência das contratações públicas e economia de recursos, foram feitas recomendações acerca da infraestrutura tecnológica, esquema e qualidade de dados. Dentre as sugestões está a adoção e implementação de um padrão de dados aberto, comum e extensível, o Open Contracting Data Standard (OCDS). A implementação do OCDS resultaria na solução dos problemas aqui apresentados, resultando na:

- Interoperabilidade, em matéria de dados, entre os sistemas utilizados pelos diversos órgãos dos mais de 5 mil municípios brasileiros, órgãos estaduais e da esfera federal através da adoção de um esquema comum;

- Publicação de todas as contratações disponibilizadas pelos entes nacionais e subnacionais de um único conjunto de dados, em um formato homogêneo - resultando em transparência efetiva;
- Publicação de cada contratação em arquivo único e legível por máquina, assim como as atualizações das contratações - minimizando a necessidade de pré-processamento e ocorrência de erros de cruzamento;
- Garantia do uso de boas práticas em arquitetura e desenvolvimento de software - facilitando o acompanhamento e manutenção do sistema a médio e longo prazo;
- Descrição apropriada de produtos e serviços através do uso ou proposição de extensões - possibilitando a filtragem e comparação apropriada entre itens de compra;
- Consistência semântica entre campos de dados e comparabilidade entre itens similares de contratações distintas a nível nacional ou internacional - proporcionada pela normalização do dado desde o momento em que é inserido no sistema.

Considerando a escala de abrangência e volume de dados estimado para o portal, é importante salientar que qualquer recomendação ou solução que não envolva a adoção de um modelo de dados comum — em esquema e conteúdo — representa um paliativo, estando fadada a dar continuidade às questões aqui listadas; e que hoje representam ônus à transparência, acessibilidade e uso efetivo dos dados públicos para geração de inteligência.

2. Sumário

1. Resumo executivo.....	3
2. Sumário.....	5
3. Visão geral.....	6
3.1 Objetivo.....	6
4. Padronização de dados.....	8
4.1 Open Contracting Data Standard.....	9
5. Arquivos de dados abertos.....	11
5.1 Ausência de acessibilidade aos dados abertos.....	11
5.2 Adequação na extensão do dado disponibilizado.....	13
6. APIs.....	13
6.1 Consulta.....	15
6.1.1 Limitações gerais.....	15
6.2 Busca.....	17
6.2.1 Limitação do número de registros no endpoint de listagem de contratos do portal.....	18
6.3 Integração e serviços.....	19
6.3.1 Disponibilização efetiva de dados via API de integração.....	19
6.3.2 Endpoints que retornam um único valor.....	20
6.3.3 Campos idênticos com nomes distintos.....	20
6.3.4 Inconformidade entre API de integração e sua documentação (manual versão 2.2.8).....	21
7. Esquema e qualidade de dados.....	21
7.1 Insuficiência de campos para descrição e filtragem de medicamentos.....	21
7.2 Ausência de normalização dos dados.....	22
7.3 Preenchimento do campo ano no endpoint pode ser incoerente.....	23
7.4 Campos com preenchimento nulo / descontinuados.....	23
7.5 Não é possível acompanhar o contrato de itens individualmente.....	24
7.6 Não há filtragem de tipo de bem adquirido.....	25
8. Mapeamento de erros na coleta.....	25

3. Visão geral

O PNCP¹, exibido na Figura 1, é uma plataforma oficial do governo brasileiro criada a partir dos atos exigidos pela Lei nº 14.133, de 2021. De uso obrigatório de diferentes poderes, e esferas a partir de janeiro de 2024, o portal visa centralizar documentos e dados de contratações, planos de contratações anuais, catálogos e outros arquivos. Essa ferramenta deixa à disposição da sociedade a possibilidade de acompanhar, de maneira única, contratações públicas de todo o país.

Considerando o volume de dados previsto, os benefícios potenciais do uso desses dados para realização de análises de preço, análise de comportamento de compra e até mesmo produção de estudos acerca da efetividade de políticas públicas são, de partida, credores do investimento feito pelo governo federal na construção da ferramenta. Entretanto, um grande desafio desse tipo de sistema é apresentar os dados de maneira a gerar informações e *insights* relevantes ao próprio poder público e à sociedade civil. Isto é, o investimento no PNCP tem o potencial de gerar economia de recursos nas contratações, fomentando melhor aproveitamento dos recursos públicos.

Em face do cenário apresentado, este relatório tem o objetivo de elencar os principais entraves no acesso aos dados e documentos públicos sobre processos de contratação, trazendo uma visão de acessibilidade de dados na perspectiva de profissionais técnicos e da sociedade civil.

3.1 Objetivo

Possibilitar a consulta e comparação de preços e itens similares é uma das funcionalidades previstas na lei que rege o PNCP. Tendo em vista o volume de dados esperado considerando a abrangência do portal, essa é uma das funcionalidades mais importantes para o acompanhamento do uso dos recursos públicos e realização de contratações mais vantajosas.

"§ 3º O PNCP deverá, entre outras funcionalidades, oferecer:

¹ <https://pncp.gov.br/>

II - painel para consulta de preços, banco de preços em saúde e acesso à base nacional de notas fiscais eletrônicas; (...)"

Lei 14.133, art. 174

Assim, a partir da necessidade da Transparência Brasil de analisar dados dos itens contratados para comparar preços de medicamentos, fez-se preciso, então, realizar o download de dados presentes no portal, em especial dados de itens de contratos.



Figura 1 - Captura de tela do site do PNCP

Fonte - PNCP

Este relatório está fundamentado no uso do PNCP e suas ferramentas acessórias (APIs, manuais etc.) para acesso e análise de dados sobre itens adquiridos em contratações, a fim de sugerir melhorias no processo de disponibilização e possível economia de recursos através das contratações.

A seguir, são descritos processos de busca e download de dados, os problemas encontrados ao longo desse caminho, sugestões de solução ou melhoria baseadas nos padrões de disponibilização de dados e boas práticas sugeridas pela W3C² e Open Contracting Partnership (OCP).

² [W3C](#)

4. Padronização de dados

Segundo o Tribunal de Contas da União, a abertura de dados públicos, além de fomentar a participação e empoderamento social, promove melhorias também na qualidade dos dados gerados, do serviço oferecido e nas políticas públicas propostas³. Entretanto, apesar de ser uma importante ferramenta, a disponibilização de parte dos dados ou ferramentas consultivas por si só não garantem o uso pleno e efetivo do dado para compreensão do processo da compra pública, **sendo de suma importância assegurar a qualidade, padronização, acessibilidade e disponibilidade desse dado.**

Hoje, as administrações públicas federal, estadual e municipal contam com diversos sistemas para coleta, armazenamento, manipulação e visualização dos dados de contratações públicas. Dois grandes desafios ocasionados por essa diversidade de fontes são: i) a interoperabilidade de dados/sistemas, e ii) a ausência de normalização entre os dados compartilhados. Por meio da interoperabilidade garante-se que diversos sistemas compartilhem dados entre si, de maneira segura, com mínima interferência humana. Já a normalização garante que dados sejam comparáveis entre si de modo a possibilitar a produção de inteligência e geração de valor. Em ambos os casos, protocolos, processos e padrões são os mecanismos que sustentam essas propriedades.

Neste relatório é adotado como referência um padrão internacional utilizado por mais de 50 países e cidades ao redor do mundo, o Open Contracting Data Standard (OCDS)⁴ - ou Padrão de Dados de Contratação Aberta em tradução livre, descrito a seguir.

Desde já destacamos que a contratação deveria ser o grão mínimo referenciado no processo de abertura de dados, possibilitando acompanhamento completo do processo com mínimo esforço, sem a necessidade de cruzamentos de dados. Por **contratação**, entende-se, aqui, como o processo de compra disponível no PNCP: desde a publicação dos planos anuais de contratação e instrumentos convocatórios, passando pela homologação da compra, até a publicação de documentos vinculantes (contrato ou nota de empenho, por exemplo).

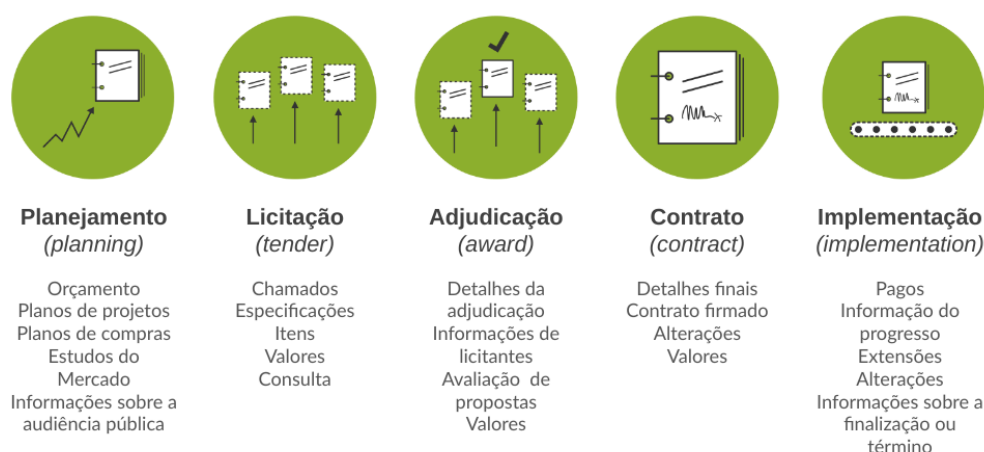
³ "Dados abertos: uma estratégia para o aumento da transparência e modernização da gestão pública." *Revista do TCU* 131 (2014): 22-29

⁴ [Open Contracting Data Standard](#) - OCDS

4.1 Open Contracting Data Standard

O OCDS é um padrão que permite a descrição e divulgação de todos os estágios da contratação pública por meio de um esquema de dados comum que permita a comparação de contratações independente de onde tenham sido geradas. O uso de um padrão assegura a qualidade dos dados, possibilita inspeções mais assertivas e eficazes, tanto pelo governo quanto pela sociedade civil organizada, além de democratizar o acesso ao dado.

A contratação é o elemento fundamental do OCDS, em que cada fase do processo de contratação é conectada por um identificador (*ocid - open contracting identifier*) único. O publicador dos dados, cada item, cada licitante e outras entidades envolvidas também possuem identificadores únicos. Dessa maneira, é garantida a comparabilidade entre processos de compra.



Identificador único do processo de contratação

Open Contracting ID (OCID)

Figura 2 - Etapas do processo de contratação

Fonte: [OCDS](#) - Tradução livre

Para facilitar a publicação de dados de contratação para atender a uma variedade de necessidades, o OCDS fornece:

- Um modelo de dados estruturado comum, incluindo um esquema, listas de códigos e regras e definições comuns para campos de dados e processos de contratação;

- Orientações e ferramentas para apoiar a implementação e utilização de dados;
- Mecanismo de extensão para adicionar campos complementares aos já existentes no OCDS;
- Suporte global gratuito da equipe de suporte de dados⁵.

O modelo de dados estruturado comum inclui como modelar:

- metadados de lançamento - informações contextuais sobre cada lançamento de dados;
- partes - informações sobre as organizações e participantes do processo de contratação;
- planejamento - informações sobre as metas, orçamentos e projetos aos quais se refere um processo de contratação;
- licitação - informações sobre como uma licitação será realizada ou ocorreu;
- adjudicações - informações sobre adjudicações realizadas no âmbito de um processo de contratação;
- contrato - informações sobre contratos assinados como parte de um processo de contratação;
- implementação - informações sobre o progresso de cada contrato até a conclusão.

A interoperabilidade é garantida na prática por meio da implementação deste modelo de dados comum, oferecendo bases e convenções para a inserção de dados naturalmente normalizados. Cada contratação, com *ocid* único, é disponibilizada em um arquivo JSON que contém os detalhes de cada fase do processo. Múltiplas contratações podem ser disponibilizadas em um mesmo arquivo, assim como também múltiplas atualizações de uma mesma contratação.

A adoção e manutenção de um modelo de dados comum para contratações, órgão-agnóstico, possibilita e sedimenta a capacidade do poder público, sociedade civil e outros atores interessados de acessar, analisar, visualizar e compartilhar dados; viabiliza o intercâmbio de ferramentas entre

⁵ [OCP: What is the OCDS and why use it?](#)

órgãos e remove impeditivos técnicos relacionados aos dados (cruzamentos, normalização, limpeza etc.)⁶. Com isso, abre-se espaço para engajar esforços na atividade efetiva de fiscalização da compra pública.

A partir das premissas do *quê* publicar e *como* publicar, partindo de um processo de contratação do PNCP e utilizando as diretrizes do OCDS, realizamos o mapeamento entre os campos disponíveis no portal em relação aos campos listados no padrão. Nesta planilha⁷ é possível visualizar esse esquema.

5. Arquivos de dados abertos

5.1 Ausência de acessibilidade aos dados abertos

Para acessar os dados de contratações, o primeiro passo foi buscar pelos dados abertos do PNCP. Tipicamente, sites governamentais possuem link ou referência para uma página onde é possível realizar o download de arquivos de dados com o conteúdo exibido. Verificou-se que até abril de 2024 ainda **não estavam disponíveis aparatos que oferecessem arquivos de dados abertos no portal**.

No cenário encontrado em 10 de janeiro de 2024, apenas usuários programadores conseguiriam acessar os dados abertos do PNCP de maneira massiva -- uma vez que estavam até então disponíveis apenas via APIs ou por meio da raspagem das páginas utilizando robôs. Outros possíveis interessados no tema de contratações e compras públicas, tais como: jornalistas, pesquisadores e o próprio cidadão comum, por exemplo, não conseguiriam ter acesso ao download completo da base, ou até mesmo pacotes de dados com filtros específicos para manipulação em seu próprio computador pessoal.

Se a motivação por trás da abertura dos dados é, essencialmente, "aumentar o número de pessoas, ferramentas e aplicações que podem acessar e fazer uso efetivo de um conjunto de dados", possibilitar o download massivo dos dados é uma maneira de contemplá-la.

Recomendação: permitir download massivo de dados de contratações em formato aberto

⁶ [OCP: Why implement the OCDS](#)

⁷ [TB-PNCP-OCDS - Field Level Mapping Template for OCDS 1.1 \(Template Version 0...](#)

Atualmente, **os dados de contratação são fragmentados**. Faz-se necessário acessar diversas páginas para acompanhar uma compra desde o lançamento do seu edital, a divulgação de atas de registro de preços (quando aplicável), até a contratação das empresas para fornecimento dos itens homologados. **Em um cenário ótimo, os dados de uma ou mais contratações deveriam estar disponíveis para download através de uma única requisição ou clique.**

Em 30 de abril de 2024⁸ foi lançado o painel [PNCP em números](#), que fornece uma visão geral das contratações publicadas e homologadas, bem como visões de compras realizadas por estados, municípios ou órgãos e unidades da administração federal direta, autárquica e fundacional. A ferramenta provê, além das visualizações, a possibilidade de baixar os dados apresentados em tela, após filtragens.

O uso do próprio painel e filtros disponíveis (esfera, poder, UF, cidade, ano, mês etc.) para realizar a seleção e download de arquivos de dados de todo o processo de contratação, até a etapa em que esteja disponível (certames em aberto ou homologados), é uma solução que contemplaria parcialmente a descarga massiva de dados. Arquivos complementares à contratação ou gerados durante o processo (editais ou contratos, por exemplo) podem ser referenciados aos links originais, não onerando tempo de download ou necessitando de espaço físico adicional em disco.

A depender do volume de dados gerado pelo(s) filtro(s) escolhido(s), o tamanho do arquivo gerado pode ser um empecilho ao download (dado o consumo de banda) ou uso efetivo do dado (quanto maior o arquivo, maior o poder computacional necessário para processá-lo), o OCDS sugere boas práticas⁹ acerca desse problema, propondo diversos caminhos para solução: limitação de tamanho do arquivo, número de contratações presentes em cada arquivo, estratégias de compressão, paginação, e até mesmo a disponibilização de um servidor de arquivos para atender a esse requisito.

⁸ [Governo federal: Gestão lança painel para visualização de gastos públicos com compras e contratação de serviços](#)

⁹ [OCP: Boas práticas - Download de arquivos em massa](#)

5.2 Adequação na extensão do dado disponibilizado

Em complemento à sugestão anterior, modificar o formato do arquivo disponibilizado ou possibilitar a escolha do mesmo é fundamental para atender aos requisitos de abertura de dados, possibilitando o livre acesso independente do software escolhido para realização de análises e manipulação.

Formatos comumente utilizados são: CSV, JSON, XML, por exemplo.

A partir da implementação dessa estratégia, mitiga-se a necessidade de realizar pré-processamentos ou cruzamentos de dados que possam comprometer sua integridade, e assegurado o pleno acesso ao gasto público. Assim, qualquer usuário que deseje acompanhar as compras de seu estado, município ou qualquer outro órgão poderá fazê-lo de maneira independente e sem custos adicionais através do arquivo baixado.

6. APIs

Hoje, o PNCP conta com três Interfaces de Programação de Aplicação (APIs): Integração e Serviços, Consulta e Busca.

API	Documentação
Integração e serviços	Link da API rest e manual técnico (Versão 2.2.9)
Consulta	Link da API rest e manual técnico (Versão 1.0)
Busca	Não encontrada em transparência ativa

Tabela 1 - APIs disponíveis

A API de Integração e Serviços¹⁰ tem a finalidade de possibilitar o acesso, inclusão, modificação e deleção de dados no portal pelos entes nacionais e subnacionais, cujos sistemas individuais devem ser interoperáveis com as interfaces disponibilizadas pela equipe técnica do portal.

¹⁰ [PNCP: API de Integração e Serviços](#)

Atualmente, a *API de Consulta*¹¹ oferece uma visão limitada, não abrangendo o processo de contratação em sua totalidade. Estão disponíveis dados sobre contratos por data de publicação, consulta a contratações específicas, contratações com edital aberto, atas de registro de preço e planos de contratação.

Ainda, há uma *API de Busca*¹² (ou um conjunto de endpoints de busca) que é utilizada para alimentar a interface navegável do próprio portal. Nas imagens abaixo, vemos um exemplo da apresentação dos dados na interface e do retorno de um endpoint dessa API, respectivamente. Não foi encontrada documentação desta API no site do PNCP.

A partir das aplicações e serviços disponíveis e do Manual de Integração¹³ foram listados abaixo problemas relacionados à completude e qualidade dos dados, uso das APIs e também sugestões de melhoria baseadas em boas práticas de desenvolvimento de software e padrões reconhecidos internacionalmente.



The screenshot shows the 'Portal Nacional de Contratações Públicas' interface. At the top, there is a search bar with the text 'Buscar no PNCP' and an 'Entrar' button. Below the search bar, the main navigation includes 'Editais e Avisos de Contratações', 'Atas de Registro de Preços', and 'Contratos'. The 'Editais e Avisos de Contratações' section is active, displaying a search filter interface. This interface includes a 'Palavra-chave' search box, a 'Status' section with radio buttons for 'A Receber/Recebendo Proposta' (selected), 'Encerradas', 'Em Julgamento/Propostas Encerradas', and 'Todos'. There are 'Limpar' and 'Pesquisar' buttons. Below the search filters, the page title is 'Editais e Avisos de Contratações A Receber/Recebendo Propostas'. A 'FILTROS' sidebar is visible on the left with dropdowns for 'Modalidades da Contratação' and 'Órgãos'. The main content area shows a contract listing for 'Aviso de Contratação Direta nº 90027/2024' with details on modality, update date, organ, and object.

Figura 3 - Captura de tela do site do PNCP

Fonte: Screenshot tomado a partir do PNCP

¹¹ [PNCP: API de Consulta](#)

¹² [PNCP: API de Busca](#)

¹³ [PNCP: Manual de Integração PNCP Versão 2.2.6](#)



```

{
  "items": [
    {
      "id": "85de771c1c34d97d70afbd875ce009c5",
      "index": "catalog2",
      "doc_type": "_doc",
      "title": "Aviso de Contratação Direta nº 90027/2024",
      "description": "O objeto da presente dispensa é a escolha da proposta mais vantajosa para a Aquisição, por dispensa de licitação de Relógio de Parede, analógico, quartz, com ponteiros de horas, minutos e segundos, conforme condições, quantidades e exigências estabelecidas neste Aviso de Contratação Direta e seus anexos.",
      "item_url": "/compras/00508903000188/2024/721",
      "document_type": "edital",
      "createdAt": "2024-04-25T16:03:24.388323",
      "numero": null,
      "ano": "2024",
      "numero_sequencial": "721",
      "numero_sequencial_compra_ata": null,
      "numero_controle_pncp": "00508903000188-1-000721/2024",
      "orgao_id": "87693",
      "orgao_cnpj": "00508903000188",
      "orgao_nome": "JUSTICA FEDERAL DE PRIMEIRA INSTANCIA",
      "orgao_subrogado_id": null,
      "orgao_subrogado_nome": null,
      "unidade_id": "4114",
      "unidade_codigo": "090031",
      "unidade_nome": "TRIBUNAL REGIONAL FEDERAL DA 5A.REGIAO",
      "esfera_id": "F",
      "esfera_nome": "Federal",
    }
  ]
}

```

Figura 4 - Captura de tela do retorno de endpoint de busca
 Fonte: Screenshot tomado a partir do retorno [de endpoint de busca](#)

6.1 Consulta

Apesar de disponibilizar alguns endpoints, a API de consulta não atende ao propósito de acompanhamento completo de uma contratação.

6.1.1 Limitações gerais

Os campos disponíveis no retorno dos endpoints *não permitem a associação do instrumento convocatório (edital, aviso de contratação e compra direta etc.) com os instrumentos de contratos*. Além do detalhamento de itens contratados não estar disponível em nenhuma das entidades citadas. Apenas a partir do parâmetro `numero_controle_pncp` é possível recuperar os documentos de contrato de uma contratação, porém é na API de integração que esse dado pode ser acessado.

Ademais, as contratações podem ser pesquisadas dentro de um intervalo de 365 dias estabelecido pelo usuário, permitindo o usuário limitar resultados para um determinado órgão contratante ou usuário que publicou os dados¹⁴, mas não há busca pelos identificadores especificados no próprio manual técnico da API¹⁵.

Dada a maturidade da API de consulta, as recomendações ordenadas abaixo seguem com o objetivo de auxiliar em sua evolução e abertura efetiva do processo de contratação.

Recomendações

A partir do OCDS¹⁶, o desenho dos endpoints deve ser realizado a partir da contratação e de suas versões/histórico levando em consideração:

- um endpoint com paginação para recuperação de conjuntos de contratações sem a necessidade de prover parâmetros (identificador da contratação ou identificador da versão da contratação);
- um endpoint com paginação para recuperação de todas as versões (caso hajam) de uma mesma contratação para múltiplas contratações sem a necessidade de prover parâmetros (identificador da contratação ou identificador da versão da contratação);
- um endpoint para recuperação de uma contratação específica através de seu identificador e OCID;
- um endpoint para recuperação de todas as versões de uma contratação específica (*record*) através de seu identificador e OCID.

Para os dois primeiros endpoints, de maneira a não sobrecarregar a infraestrutura e facilitar o acesso, opções de filtragem e ordenação podem ser oferecidas, considerando filtros de datas e/ou ordenação cronológica reversa - isso irá garantir que usuários acessem dados mais recentes prioritariamente. Em caso de não haver possibilidade de paginação, **é preciso fornecer uma lista legível por máquina dos identificadores de todas as contratações**. Assim é

¹⁴ Os portais integrados e credenciados ao ambiente de produção do PNCP possuem função de publicação dos dados via API de integração e são denominados usuários. A lista de sistemas integrados ao PNCP pode ser consultada [neste link](#).

¹⁵ Os números de controle são os identificadores únicos para contratações, contratos, planos de contratação anual (pca) e atas de registro de preços ([Manual Técnico V.1.0, p. 4](#)).

¹⁶ [OCP: OCDS - Endpoints](#)

possível automatizar o acesso a todo dado. Mais detalhes podem ser vistos [nesta lista de recomendações](#).

Paliativo: Disponibilização de endpoint com múltiplos objetivos

Para a listagem de contratações a partir dos dados que já são disponibilizados pelo PNCP, é necessária a criação de um endpoint (ou múltiplos endpoints) que possibilite múltiplos resultados, desde mais específicos até mais abrangentes. Por exemplo:

```
pncp.gov.br/api/consulta/id_orgao/ano/id_contratacao
```

Parâmetros:

- `id_orgao`: identificador do órgão responsável pela contratação;
- `ano`: limitador temporal onde foi realizada a contratação;
- `id_contratacao`: identificador único da contratação (opcional).

Dessa maneira, ao utilizar todos os parâmetros seria possível acessar uma contratação específica, e ao remover o especificador da contratação seria retornado um conjunto de contratações. Analogamente à recomendação acima, parâmetros de filtragem, ordenação e paginação podem ser incorporados. Também, em caso de impossibilidade de paginação, o fornecimento da lista de identificadores das contratações é fundamental para permitir a automatização do acesso a toda massa de dados.

6.2 Busca

A API de Busca (ou um conjunto de endpoints de busca) é utilizada para alimentar a interface navegável do próprio portal, e apesar de não dispor de documentação em transparência ativa, foi útil para que os dados de contratações fossem coletados pela Transparência Brasil.

6.2.1 Limitação do número de registros no endpoint de listagem de contratos do portal

O [endpoint](#)¹⁷ responsável por alimentar a página inicial do PNCP, que fornece acesso à listagem de contratos existentes no portal, apesar de possuir parâmetros de consulta (como paginação, por exemplo) tem seu número máximo de contratos exibidos limitado a 10 mil por resposta.

Considerando que, no momento da coleta, o portal possuía mais de 200 mil contratos registrados, essa limitação afeta diretamente a disponibilidade e manipulação dos dados por meio da API. Por exemplo, ao definir `tamanho_pagina = 100` (para retornar 100 itens de consulta por página) e consultar a `página = 200`, a seguinte mensagem é retornada:

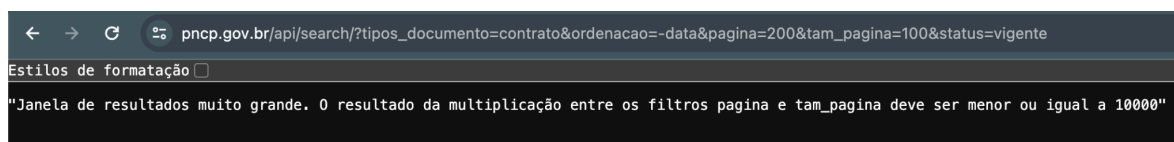


Figura 4 - Captura de tela do retorno de endpoint de busca

Fonte: Screenshot tomado a partir do retorno de [endpoint de busca](#)

Paliativo

Para acessar os dados em sua totalidade por meio deste endpoint, foi necessária a filtragem dos dados por estado, município e órgão. Vale destacar que, no momento da primeira coleta de dados, a quantidade total de contratos por estado ainda estava abaixo de 10 mil, o que possibilitou a solução paliativa.

Recomendação

Aparentemente, a limitação de 10 mil resultados retornados a partir do endpoint parece ser devido à configuração padrão da ferramenta utilizada para consulta de dados, Elasticsearch. Desse modo, a solução seria ajustar o parâmetro `index.max_result_window`¹⁸ para um valor que contemple a totalidade dos registros disponíveis no PNCP.

¹⁷https://pncp.gov.br/api/search/?tipos_documento=contrato&ordenacao=-data&pagina=1&tam_pagina=10&status=vigente

¹⁸ [Elasticsearch - parâmetro index.max_result_window](#)

Esta problemática foi encontrada na *API de Busca*, considerando a sua finalidade e limitações de negócio, direcionamos as demais sugestões à [API de Consulta](#).

6.3 Integração e serviços

Em teoria, essa API tem o objetivo de prover serviços e integração do PNCP com sistemas individuais. Porém, na prática, ela também possui função consultiva e esse "acúmulo de funções" pode onerar o host servidor responsável pela API ou mesmo dificultar a acessibilidade ao dado.

Na busca pelos dados abertos de contratação, essa API tornou possível o acesso ao detalhamento de editais, itens a serem adquiridos, itens efetivamente contratados, histórico da contratação etc. Contudo, no processo de aquisição do dado, expostas as possibilidades atuais, houveram dificuldades para extração. A seguir, foram listadas as mais relevantes e que merecem especial atenção para assegurar a acessibilidade e uso efetivo dos dados públicos.

6.3.1 Disponibilização efetiva de dados via API de integração

O propósito da API – integração e provimento de serviços – justifica sua arquitetura atual e, em partes, alguns dos desafios/problemas enfrentados no acesso ao dado completo, uma vez que a API de Consulta não atende ao propósito de disponibilizar dados do processo de contratação como um todo.

Recomendação

Para possibilitar a acessibilidade aos dados de contratação da maneira adequada, é essencial a implementação das melhorias mencionadas no Tópico [6.1 Consulta](#) para criação de endpoints úteis e complementares.

Ressaltamos que, como sugerido no OCDS, a unidade mínima de disponibilização de dados deve ser a contratação, o processo por inteiro, sem a necessidade de cruzamentos de dados. Essa prática mitiga a possibilidade de que usuários de dados gerem arquivos inconsistentes ou com ruído, o que pode levar a análises equivocadas.

6.3.2 Endpoints que retornam um único valor

Atualmente, nos endpoints de consulta a quantidade de arquivos em uma contratação ([exemplo](#)), quantidade de itens de histórico ([exemplo](#)) e quantidade de itens adquiridos ([exemplo](#)), é necessário consultar o número de itens retornados. Fazer isso para cada um dos cerca de 200 mil contratos disponíveis no momento da extração é demorado e pouco eficiente — tomando quase 34 horas para cada um dos endpoints citados.

Paliativo

Ajustar os endpoints principais de forma que informações adicionais e úteis para páginas subsequentes, como a quantidade de itens ou de arquivos da contratação, por exemplo, possam ser retornados em conjunto. Isso mitiga a necessidade de múltiplas requisições e garante uma maior eficiência no processo de obtenção de dados.

Recomendação

Implementação do padrão OCDS, no qual uma única requisição forneceria um JSON completo contendo todos os dados referentes à contratação, mitigando a necessidade de endpoints intermediários, simplificando a obtenção de dados e melhorando a acessibilidade aos dados e gastos públicos. Como já exposto anteriormente, a aplicação de paginação, filtros e operações de ordenação são úteis nesse contexto.

6.3.3 Campos idênticos com nomes distintos

Em todos os endpoints, alguns campos são idênticos mas têm nomes diferentes. Por idênticos, queremos dizer que apontam para a mesma peça de informação. Dar-lhes nomes diferentes é confuso e pode levar a pensar que contêm informações diferentes.

Exemplos:

- `listagemDeContratosVigentes/description` e `consultarDadosDoContrato/objetoContrato`

- `listagemDeContratosVigentes/unidade_codigo` e `consultarDadosDoContrato/unidadeOrgao.codigoUnidade`

Paliativo

Fazer uma lista de todos os campos que são afetados por este problema, agrupados por campos idênticos. Renomear os campos de dados para que os campos idênticos tenham o mesmo nome em todos os endpoints.

Recomendação

A adoção do OCDS mitiga a possibilidade de campos sintaticamente idênticos mas que contêm informações distintas.

6.3.4 Inconformidade entre API de integração e sua documentação (manual versão 2.2.8)

A API, versão 1.0, retorna campos com nomenclatura diferente do informado em sua documentação (ver. 2.2.8). Por exemplo, `anoCompra` (campo na API) diferente de `ano` (campo no manual).

Recomendação

Atualização do nome dos campos.

7. Esquema e qualidade de dados

7.1 Insuficiência de campos para descrição e filtragem de medicamentos

Na busca por medicamentos no PNCP, ao filtrar pelo portal ou pelos dados brutos, informações relevantes na diferenciação de compras de medicamentos não estão estruturadas ou com descrição padrão. Esse problema de normalização de dados impossibilita duas funcionalidades fundamentais do portal: a busca por medicamentos específicos e a comparação de preços de itens similares.

Paliativo

Limitar o cadastro no portal a partir de medicamentos cadastrados em catálogos de referência (por exemplo, o CATMAT - catálogo de materiais) ou bancos de dados oficiais, utilizando códigos identificadores.

Recomendação

Implementação do OCDS para descrição e caracterização apropriada de medicamentos, utilizando a extensão proposta para esse tipo de compra¹⁹. Isso implica na criação de campos descritivos para especificação de dosagem, apresentação, unidade de fornecimento, medicamentos compostos, entre outras informações relevantes. É fundamental a limitação do formato dos campos, uma vez que utilizar texto livre pode amplificar o problema.

7.2 Ausência de normalização dos dados

Os itens adquiridos são escritos em texto livre, sem um padrão para descrição do princípio ativo, unidade ou apresentação.

Paliativo

Convencionar o preenchimento de concentrações e dosagens na menor unidade possível. Alguns estados já adotam essa prática.

Por exemplo:

- Álcool Etílico 1L > Álcool Etílico 1000 ml

Recomendação

Idêntica à recomendação do tópico 7.1.

¹⁹ [Extensão para dados de medicamentos - OCDS](#)

7.3 Preenchimento do campo ano no [endpoint](#) pode ser incoerente

No endpoint de [histórico de contratos](#), o parâmetro `contratoAno`, que se refere ao ano da compra, possui preenchimentos inválidos, até mesmo superiores ao ano corrente. Por exemplo: 4761.

```
JavaScript
[
  {
    "contratoAno": 4761,
    "justificativa": null,
    "tipoLogManutencao": 0,
    "categoriaLogManutencao": 3,
    "logManutencaoDataInclusao": "2023-08-22T15:51:42",
    "usuarioNome": "Compras.gov.br",
    "sequencialTermoContrato": null,
    "numeroTermoContrato": null,
    "tipoLogManutencaoNome": "Inclusão",
    "categoriaLogManutencaoNome": "Contrato",
    "contratoOrgaoCnpj": "00394502000144",
    "contratoSequencial": 1,
    "sequencialDocumentoContrato": null,
    "tituloDocumentoContrato": null,
    "sequencialDocumentoTermoContrato": null,
    "tituloDocumentoTermoContrato": null
  }
]
```

Recomendação

Realizar validação do campo *ano* para que seja consistente.

7.4 Campos com preenchimento nulo / descontinuados

A versão 1.0 da API de integração, para determinados endpoints, retorna campos que, segundo o manual (ver. 2.2.8), foram descontinuados e são preenchidos com `NULL`. Como é o caso dos campos `orgaoSubRogado` e `unidadeSubRogada` no [endpoint de contratos](#).

Recomendação

Uma vez que as APIs estão em constante mudança, destaca-se a necessidade de documentação pública nesse processo através de, por exemplo, da disponibilização de um log de mudanças (*change log*) público ou da adição do motivo de descontinuação no próprio versionamento do manual técnico. Além disso, tendo em vista o espaço em disco consumido por campos descontinuados, estabelecer um prazo para que os sistemas integrados sejam atualizados para atender a versões mais recentes da API é fundamental.

7.5 Não é possível acompanhar o contrato de itens individualmente

O endpoint sobre o instrumento convocatório retorna, para cada item, a sua quantidade e preço estimados. Após homologação, o portal também retorna, para cada item, sua quantidade e preço homologados. No entanto, o PNCP não disponibiliza de forma estruturada nos endpoints de contrato as quantidades que estão sendo efetivamente contratadas. É possível saber apenas o valor global de um contrato. Para saber as quantidades e preços realizados no contrato é preciso abrir os documentos de contrato, que muitas vezes são arquivos PDF de notas de empenho, para recuperar tal informação. Isso impossibilita que haja uma avaliação de sobrepreços efetiva, indo contra o propósito do PNCP de garantir transparência em todo o processo de contratação.

Paliativo

Disponibilizar, em dado estruturado, no endpoint que exibe a informação dos contratos, a quantidade sendo contratada e o preço sendo pago **para cada item individualmente**.

Recomendação

Como exibido na [Figura 2](#), uma vez que o OCDS contempla todas as fases da contratação, este impedimento seria sanado considerando a adoção do padrão.

7.6 Não há filtragem de tipo de bem adquirido

Incluir filtro por tipo de compra (bens, contratações e serviços etc.). Está disponível via API, mas não no front para o usuário.

Recomendação

Inclusão de um filtro de bens e serviços no PNCP.

8. Mapeamento de erros na coleta

[Nesta planilha](#) estão listados os arquivos/endpoints coletados que tiveram erro durante a coleta, assim o time técnico responsável pode mapear possíveis indisponibilidades do serviço. [Nesta planilha](#) estão listados endpoints específicos que falharam mas que, possivelmente, possuem dados.

É importante ressaltar que durante o processo de coleta houve pausa entre requisições de modo a não sobrecarregar o serviço e causar sua indisponibilidade. Acreditamos que os casos listados são exceções, mas merecem atenção tendo em vista a indisponibilidade do dado.