

AVALIAÇÃO DE FORMATOS DE PUBLICAÇÃO DE DADOS ABERTOS GOVERNAMENTAIS ATRAVÉS DE INDICADORES DE USABILIDADE

Jorge Luiz de Faria Fernandes

Douglas Farias Cordeiro

Resumo: A exploração de tecnologias e conceitos relacionados à Gestão de Dados Abertos Governamentais é um campo de grande relevância, devido aos seus impactos sociais e políticos. Entretanto, a literatura correlata é consideravelmente restrita, principalmente em aspectos voltados à parte de tecnologia. Neste sentido, este trabalho se propõe a apresentar uma análise, com base em conceitos de usabilidade, sobre os formatos de publicação de dados abertos, a fim de auxiliar o processo de gestão e implementação das fases inerentes ao ciclo de abertura de dados.

Palavras-chave: Dados Abertos Governamentais. Formato de dados. Usabilidade.

1 INTRODUÇÃO

De maneira geral, um dado pode ser descrito como qualquer elemento que trate da representação de algo, originalmente em um formato “bruto”, que de forma aparente não possui um valor específico, mas que através de manipulação e processamento pode ser transformado em uma informação, a qual, por outro lado, gerará conhecimento e agregará valor aos objetivos considerados, e que em um contexto organizacional, pode ser considerada como um dos principais ativos (CHOO, 2003). Embora em uma visão inicial, a exploração de conceitos e ferramentas relacionados à dados estejam diretamente ligada a fatores tecnológicos, este campo também possui considerável influência acerca das ciências sociais aplicadas, principalmente no que se refere a obtenção de informação, alteração de perspectivas sociais, e construção e melhoria de políticas públicas.

Embora no âmbito de várias organizações existam tipos específicos de dados que necessitam ser resguardados de divulgação e acesso não autorizado, existem também aqueles passíveis de publicação e reutilização para diferentes fins, os quais são conhecidos como dados

abertos. Este tipo de dado ganha uma especial atenção quando coletados, mantidos e utilizados pelo Governo, tanto pelo aspecto da centralidade, quanto também pelo direito de acesso, garantido constitucionalmente. De acordo com *Open Knowledge International* (2011):

“Dado aberto é aquele que pode ser livremente utilizado, reutilizado e redistribuído por qualquer pessoa, e deve, no máximo, atribuir à fonte original e/ou fazer o compartilhamento desses dados utilizando a mesma licença em que as informações foram apresentadas”.

Neste sentido, o avanço de estudos e tecnologias relacionadas à gestão de dados abertos é de fundamental importância para o debate e crescimento social e político. Entretanto, a exploração desta área ainda é consideravelmente restrita, sendo que, a literatura aponta, basicamente, para direções voltadas a explicação do que são os dados abertos e alguns estudos de caso de experiências. A partir disso, e considerando questões de usabilidade, este trabalho se propõe a realizar uma análise avaliativa acerca das características e propriedades relacionadas aos principais formatos de publicação de dados abertos, a fim de permitir uma maior eficiência em projetos de abertura de dados governamentais.

2 DADOS ABERTOS GOVERNAMENTAIS

Dados Abertos Governamentais são informações do setor público que estão disponíveis para serem usadas, reutilizadas e redistribuídas por qualquer pessoa, sendo que no máximo se fará necessário informar a sua autoria e compartilhar com a mesma licença utilizada pelo órgão que disponibilizou a mesma. O objetivo é que essas informações sejam publicadas e difundidas pela Web, compartilhadas em formatos diversos e livre, de modo que sejam compreensíveis às máquinas para poderem ser reutilizadas em *API's* desenvolvidas pela sociedade. Eles são pautados em três leis e oito princípios que serão apresentados a seguir.

As três leis dos Dados Abertos Governamentais propostas por EAVES (2009), e que se aplicam também a Dados Abertos, são:

1. Se o dado não pode ser encontrado e indexado na Web, ele não existe.
2. Se o dado não estiver em formato aberto e compreensível por máquina, ele não pode ser reaproveitado.

3. Se algum dispositivo legal não permitir sua reaplicação, ele não é útil.

Os oito princípios, definidos em 2007 por um grupo de trabalho de 30 pessoas na Califórnia (Estados Unidos da América), que também se aplicam a Dados Abertos, exceto o primeiro, são os seguintes:

1. **Completo** - Todos os dados públicos são disponibilizados. Dados são informações eletronicamente gravadas, incluindo, mas não se limitando a, documentos, bancos de dados, transcrições e gravações audiovisuais. Dados públicos são dados que não estão sujeitos a limitações válidas de privacidade, segurança ou controle de acesso, reguladas por estatutos.
2. **Primários** - Os dados são publicados na forma coletada na fonte, com a mais fina granularidade possível, e não de forma agregada ou transformada.
3. **Atuais** - Os dados são disponibilizados o quanto rapidamente seja necessário para preservar o seu valor.
4. **Acessíveis** - Os dados são disponibilizados para o público mais amplo possível e para os propósitos mais variados possíveis.
5. **Compreensíveis por máquina** - Os dados são razoavelmente estruturados para possibilitar o seu processamento automatizado.
6. **Não discriminatórios** - Os dados estão disponíveis a todos, sem que seja necessária identificação ou registro.
7. **Não proprietários** - Os dados estão disponíveis em um formato sobre o qual nenhum ente tenha controle exclusivo.
8. **Livres de licenças** - Os dados não estão sujeitos a regulações de direitos autorais, marcas, patentes ou segredo industrial. Restrições razoáveis de privacidade, segurança e controle de acesso podem ser permitidas na forma regulada por estatutos.

O Tribunal de Contas da União do Brasil elaborou um documento propondo 5 razões para que os órgãos públicos pensem e invistam em iniciativas de abertura dos dados governamentais (Tribunal de Contas da União, 2015), que são as seguintes:

- Transparência na gestão pública;
- Contribuição da sociedade com serviços inovadores ao cidadão;

- Aprimoramento na qualidade dos dados governamentais;
- Viabilização de novos negócios;
- Obrigatoriedade por lei.

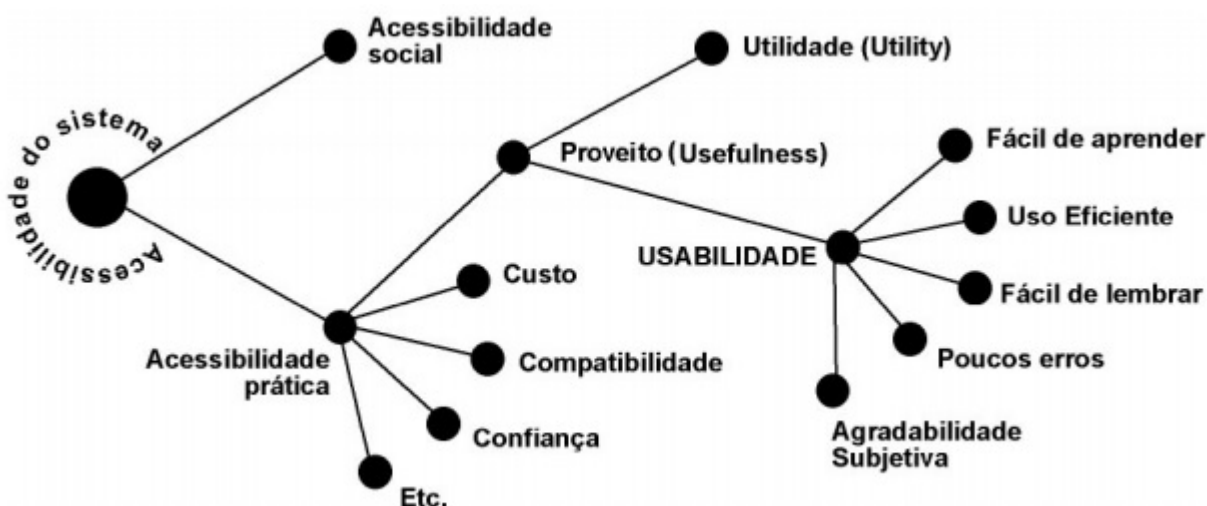
Pensando no Âmbito social, os Dados Abertos Governamentais podem auxiliar das mais diversas maneiras, uma vez que uma gama muito expressiva de informações podem ser processadas por máquinas e podem resultar em serviços que cruzam essas informações, que podem auxiliar a sociedade em atividades diárias, fiscalização sobre os gastos públicos, dentre várias outras aplicações.

Segundo o *Open Knowledge International* (2014), quando um Órgão Governamental torna pública uma informação que nunca foi exibida previamente, deve ser escolhido um formato que permita o melhor equilíbrio entre custo e adequação para fins. O foco desse artigo é nos formatos que os dados serão disponibilizados pelos órgãos públicos, para que possam ser reaproveitados da melhor maneira possível. Na seção a seguir serão apresentados três formatos, que são o XML, o CSV e o JSON, os mesmos serão apresentados e para cada um deles serão levantados pontos que devem ser levados em consideração para a utilização dos mesmos.

3 USABILIDADE

Nos primórdios da era computacional um sistema era avaliado com o termo *user friendly* (amigável). Segundo NIELSEN (1993) tal termo não se aplica por duas questões, os usuários não precisam de máquinas que sejam amigáveis, mas sim de máquinas que não dificultam a operação, e a outra questão é que tal declaração implica que as necessidades do usuário podem ser descritas ao longo de um único plano por sistemas pouco o muito amigáveis. Os usuários são seres humanos diferentes e que possuem necessidades diferentes, sendo assim, um sistema pode ser ótimo para um usuário e extremamente chato para outro.

Figura 01 - Modelo de atributos de acessibilidade do sistema de NIELSEN (1993).



Fonte: NIELSEN (1993).

Como pode ser visto na Figura 01, o autor destaca a Usabilidade e a utilidade como atributos de qualidade, igualmente importantes dentro do atributo maior de Proveito (*Usefulness*) do sistema. A utilidade então responderia a questão que se refere à funcionalidade: o sistema faz o que o usuário necessita? NIELSEN (1993 e 2003) resume o gráfico afirmando que pouco importa se um sistema é fácil de usar, se ele não é o que o usuário quer. Por outro lado, de nada adianta se é o que o usuário quer, mas a interface é extremamente complicada.

Para MORAES e MONT' ALVÃO, a "usabilidade implica que o sistema deve oferecer sua funcionalidade de tal maneira que o usuário, para o qual foi planejado, seja capaz de controlá-lo e utilizá-lo sem constrangimentos demasiados sobre suas capacidades e habilidades" (...) (MORAES e MONT' ALVÃO, 2000). "Como conceito, a usabilidade trata da adequação entre o produto e as tarefas a cujo desempenho ele se destina, da adequação entre o produto e as tarefas a cujo desempenho ele se destina, da adequação com o usuário que o utilizará, e da adequação ao contexto em que será usado."

Para PREECE (2002), a usabilidade está normalmente relacionada a dar garantias de que um produto iterativo é fácil de aprender, possui o uso efetivo e agradável do ponto de vista dos usuários. Isso envolve otimizar a interação que as pessoas desempenham com produtos

interativos que permitam a elas realizar suas atividades no trabalho, no estudo e na vida diária.

Para a autora a usabilidade é quebrada nos seguintes objetivos:

- Efetividade: deve possuir uso efetivo;
- Eficiência: deve ser eficiente;
- Segurança: deve ser seguro de usar;
- Utilidade: deve possuir boa utilidade;
- Fácil Aprendizagem: deve ser fácil de aprender;
- Fácil Lembrança: deve ser fácil lembrar como usar.

Para medir a usabilidade de um sistema podemos utilizar indicadores, que são a maneira de avaliar a performance de um sistema ao longo do tempo, para verificar se os objetivos de projeto são alcançados.

Esse artigo tem como base os 8 indicadores selecionados pela cartilha Indicadores de Métricas para avaliação de e-Serviços, elaborado pelo Ministério do Planejamento, Orçamento de Gestão, que são:

1. Maturidade do serviço prestado: Traduz em que medida o serviço é prestado de maneira totalmente desmaterializada e completa. O indicador é medido na forma de escala, sendo o serviço pontuado de acordo com o nível de maturidade alcançado.
2. Comunicabilidade: Busca valorizar a receptividade e abertura do Estado à influência e participação dos cidadãos, assim como a capacidade de oferecer ajuda para facilitar o acesso às informações prestadas.
3. Confiabilidade: A confiabilidade traduz a qualidade do sistema que permite ao usuário/cidadão confiar, justificadamente, no serviço prestado. Ela valoriza a capacidade do governo de oferecer segurança às informações e canais relativos ao serviço prestado.
4. Multiplicidade de acesso: A Multiplicidade de Acesso traduz a abrangência que cada iniciativa de Governo Eletrônico tem de disponibilizar serviços em outros canais eletrônicos além da Internet.
5. Acessibilidade: A Acessibilidade diz respeito à participação das pessoas portadoras de deficiência no uso de produtos, serviços e informações de forma irrestrita.

6. Disponibilidade: Verifica quanto o serviço prestado está disponível para acesso do cidadão/usuário.
7. Facilidade de uso: A facilidade de uso ou usabilidade trata da qualidade de interação e navegação da interface para acesso ao serviço.
8. Transparência: A Transparência é a responsabilidade pela eficiente gerência de recursos públicos (*accountability*), permitindo o maior controle da sociedade ao andamento do serviço solicitado. Essa forma de controle torna a alocação de recursos simultaneamente mais eficiente e mais democrática.

Na secção seguinte serão apresentados os formatos de arquivos para publicação de dados abertos que serão abordados por esse artigo e em seguida serão aplicados sobre os mesmos indicadores baseados nos que foram apresentados acima.

4 FORMATOS DE PUBLICAÇÃO DE DADOS ABERTOS

O trabalho tem como finalidade apresentar três formatos de representação dos Dados Abertos Governamentais que estão presentes em manuais (Governo do Estado de São Paulo e Governo do Reino Unido, 2015) e que são muito utilizados como maneira de extração desses dados para uma futura utilização, que são o XML, o CSV e o JSON. Os três formatos foram escolhidos por serem populares, livres e por serem editáveis, ou seja, podem ser facilmente manipulados e processados por máquinas. Cada um deles será apresentado e serão indicadas as melhores utilidades para cada um.

4.1 XML

O *XML (Extensible Markup Language)* é um formato de arquivo amplamente usado para troca de informações por permitir diversas maneiras de manter a estrutura da informação e a forma que o arquivo é construído, além de permitir aos desenvolvedores escrever partes da documentação dentro, com a informação e sem interferir o processo de leitura da mesma. Trata-se de uma linguagem de marcação (assim como o HTML, utilizado para construir páginas Web) mantida pela W3C (*World Wide Web Consortium*) e ele tem como objetivo a simplicidade, generalidade e usabilidade por toda a internet. Ele enfatiza a geração de documentos, porém

também é muito utilizado para representar estruturas de dados arbitrárias para integrar diferentes sistemas de computadores.

Figura 02 - Arquivo em formato XML.

```
1 <xml version="1.0" encoding="UTF-8">
2 <Exemplo>
3   <Relatório número="1">
4     <Objeto>Cadeiras</Objeto>
5     <Quantidade>500</Quantidade>
6     <Valor>21.957,00</Valor>
7   </Relatório>
8   <Relatório número="2">
9     <Objeto>Livros</Objeto>
10    <Quantidade>300</Quantidade>
11    <Valor>13.965,00</Valor>
12  </Relatório>
13  <Relatório número="3">
14    <Objeto>Kit de material de limpeza</Objeto>
15    <Quantidade>800</Quantidade>
16    <Valor>33.890,00</Valor>
17  </Relatório>
18  <Relatório número="4">
19    <Objeto>Kit de utensílios de cozinha</Objeto>
20    <Quantidade>650</Quantidade>
21    <Valor>25.987,00</Valor>
22  </Relatório>
23 </Exemplo>
```

Fonte: Elaborado pelos autores

Conforme podemos visualizar na figura 02, um arquivo XML possui marcadores e conteúdos, objeto e quantidade são marcadores e cadeiras e 500 são conteúdos, por exemplo. Existem marcadores do tipo *tags*, que iniciam e finalizam com "<" e ">" e podem ser do tipo de início (<Objeto>), de fim (</Objeto>) ou de elemento vazia (<line break/>, quebra de linha). Existem elementos que são componentes do XML que começam com uma *tag* de início e terminam com uma *tag* de fim, ou possuem apenas uma *tag* de elemento vazia, caso houver caracteres entre as tags, são o conteúdo do elemento (<Valor>21.957,00</Valor>). Os atributos são pares de "valor/nome" dentro de uma tag de início ou de elemento vazia, no exemplo acima a *tag* Relatório possui um atributo número e um valor correspondente (<Relatório número="1">). Um documento xml começa com a sua declaração (<xml version="1.0" encoding="UTF-8">), na qual informa sobre si mesmo.

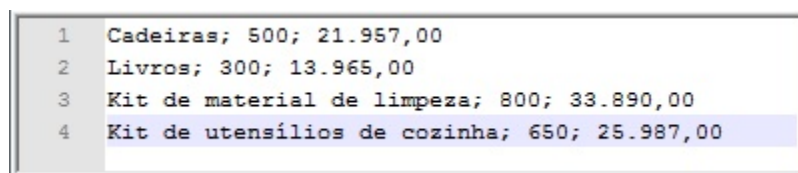
Os arquivos no formato XML são muito úteis quando é necessário manter uma estrutura dos dados, gerar documentação dentro dos dados sem atrapalhar na leitura dos mesmos e quando deseja-se transferir dados entre diferentes sistemas computacionais.

4.2 CSV

O CSV (*Character Separated Values*, valores separados por um delimitador) é um formato de arquivo que pode ser muito útil porque é compacto, e portanto, se adequa bem para transferências de grandes conjuntos de dados com a mesma estrutura. Entretanto, é um formato tão rigoroso que o arquivo pode se tornar inútil caso não possua uma documentação explicando como os dados são separados. Por isso, é importante que os arquivos do formato CSV tenham documentações precisas dos campos individuais. Além disso, é essencial que a estrutura dos arquivos CSV seja respeitada, pois uma simples omissão de um campo pode distorcer a leitura de todos os dados restantes no arquivo sem qualquer possibilidade de retificação, uma vez que não será possível determinar a forma como os dados restantes devem ser interpretados.

Os arquivos CSV, geralmente possuem informações em linhas e para cada linha existe a delimitação de colunas, feitas por caracteres como vírgula (","), ponto e vírgula (";"), tabulação, dentre outros. As linhas podem ser interpretadas como sendo os registros e cada informação antes de um delimitador é uma coluna, The Internet Society (2005).

Figura 03 - Arquivo em formato CSV.



```
1 Cadeiras; 500; 21.957,00
2 Livros; 300; 13.965,00
3 Kit de material de limpeza; 800; 33.890,00
4 Kit de utensílios de cozinha; 650; 25.987,00
```

fonte: Elaborado pelos autores.

Conforme Pode ser visualizado na figura 03, o delimitador utilizado foi o ponto e vírgula (";") e cada linha é referente a um registro, que é separado em três colunas que no caso do exemplo são objeto, quantidade e preço total. Dessa forma seria como pensar que a informação

é referente a uma tabela com três colunas e para cada linha do arquivo existe um novo registro da tabela.

Os arquivos no formato CSV são excelentes para armazenar uma quantidade muito elevada de informações tabuladas, uma vez que a compactação obtida com esse formato é muito satisfatória e também é bastante útil quando utilizado por um usuário final, através de processadores de planilhas, pois é possível carregá-los nesses editores com facilidade e obter arquivos já delimitados por linhas e colunas, sem os delimitadores.

4.3 JSON

O JSON (*JavaScript Object Notation*) é um formato de arquivo aberto e simples que pode ser facilmente lido por qualquer linguagem de programação. Sua simplicidade significa que geralmente ele é mais fácil de ser processado do que outros tipos de arquivos, XML, por exemplo. É uma alternativa ao XML para transferência de dados estruturados entre um servidor Web e uma aplicação Web. Sua lógica de organização tem semelhanças com o XML, mas possui notação diferente. É bastante difundido em serviços Web (e-commerce, serviços de *email*, dentre outros), pois consegue transferir uma taxa elevada de informação entre o cliente e o servidor utilizando uma quantidade menor de caracteres, pois seus marcadores são simples, não possuem fechamento e por possibilitar a serialização de estrutura de objetos complexos, como listas e sub propriedades.

Figura 04 - Arquivo em formato JSON.

```
1 {
2   "relatório 1": {
3     "Objeto": "Cadeiras",
4     "Quantidade": 500,
5     "Valor": 21.957,00
6   },
7   "relatório 2": {
8     "Objeto": "Livros",
9     "Quantidade": 300,
10    "Valor": 13.965,00
11  },
12  "relatório 3": {
13    "Objeto": "Kit de material de limpeza",
14    "Quantidade": 800,
15    "Valor": 33.890,00
16  },
17  "relatório 4": {
18    "Objeto": "Kit de utensílios de cozinha",
19    "Quantidade": 650,
20    "Valor": 25.987,00
21  }
22 }
```

Fonte: elaborado pelos autores.

Conforme figura 04, um arquivo JSON também utiliza pares de atributos e valores, porém não utiliza marcador como no XML, utiliza delimitador em cadeias: {}, [], e "". O delimitador { marca o início de uma seção e o } marca o fim. Os pares de valor e atributo são separados por : e seus valores, quando texto, ficam entre aspas, número não recebem aspas. "relatório 3" é um atributo que recebe uma série de valores (Objeto, Quantidade e Valor). A lógica de encadeamento de conjunto de pares pode ser repetida inúmeras vezes e podem ser criados diversos níveis para a estrutura de dados que se deseja, *ECMA International* (2013).

Os arquivos no formato JSON estão se tornando o padrão mais utilizado para integração de dados entre repositórios e frameworks, também está se tornando o padrão nativo de armazenamento em alguns bancos de dados modernos, *SLTI/MP* (2012).

5 ANÁLISE DA USABILIDADE DE DADOS ABERTOS PUBLICADOS

Nessa seção serão aplicados indicadores sobre os 3 formatos de dados abordados na seção anterior. Os indicadores que serão utilizados não são especificamente os mesmos propostos na cartilha Indicadores de Métricas para avaliação de e-Serviços, apresentados na seção 3, porém serão baseados nos mesmos, de modo a garantir que a informação seja extraída com sucesso, seja fácil de ser reutilizada, compreendida pelo receptor, com uma taxa adequada de processamento, de modo estruturado ou não e ainda permitindo que haja interoperabilidade entre diferentes sistemas, eles serão classificados em “sim”, “não” ou por ordem de classificação para o indicador.

Os indicadores selecionados foram:

- **Facilidade de utilização:** capacidade da informação ser utilizada no formato em que está.
- **Compreensão por humanos:** capacidade de um arquivo em um determinado formato ser compreendido por um ser humano.
- **Velocidade de processamento:** taxa de processamento do arquivo em um determinado formato.
- **Estrutura dos dados:** modo em que os dados são armazenados no arquivo em um formato, sendo que eles podem ser armazenados isoladamente ou podem ser estruturados por seções.
- **Interoperabilidade entre os sistemas:** capacidade da informação extraída de um sistema ser lida e inserida em outro.

Foi realizada uma pesquisa e experimentos para fundamentar dois pontos dos 5 indicadores acima apresentados, que foram a “compreensão humanos” com um formulário apresentando arquivos nos 3 formatos para verificar a capacidade de pessoas de diferentes áreas do conhecimento compreender os arquivos e a “velocidade de processamento” com um serviço que consumiu arquivos nos 3 formatos.

Velocidade de processamento:

Foram utilizados três arquivos para cada formato, de aproximadamente 6.7 MB, 23MB e 113MB e os mesmos foram consumidos 5 vezes, utilizando blocos de 4096 KB, 4 threads simultâneas e o tempo médio foi calculado.

1) Arquivo pequeno de aproximadamente 6.7 MB

XML - tempo médio de 1.935 segundos.

JSON - tempo médio de 1.399 segundos.

CSV - tempo médio de 1.433 segundos.

2) Arquivo médio de aproximadamente 23 MB

XML - tempo médio de 4.017 segundos.

JSON - tempo médio de 3.851 segundos.

CSV - tempo médio de 3.995 segundos.

3) Arquivo grande de aproximadamente 113 MB

XML - tempo médio de 17.795 segundos.

JSON - tempo médio de 17.213 segundos.

CSV - tempo médio de 17.666 segundos.

Com base nos dados acima, podemos verificar que a velocidade de processamento do JSON se mostrou um pouco superior, posteriormente a do CSV e por fim a do XML.

Compreensão por humanos:

Foi criado um formulário contendo uma pergunta sobre cada um dos três formatos de arquivos. Segue abaixo as perguntas e as respostas obtidas de 32 usuários distintos, sendo que os mesmos não são somente da área de Tecnologia. A maioria das respostas foram obtidas de pessoas de outras áreas.

1) Arquivo no formato CSV

A pergunta foi a seguinte:

O que você entende ao ler o trecho abaixo?

João Carlos Teixeira; 26; Goiânia

Pedro Paulo Silva; 15; Anápolis

Carlos Henrique Pereira; 53; Aparecida de Goiânia

As respostas foram as seguintes:

Figura 05 - Respostas do questionário sobre o formato de arquivos CSV.



Fonte: elaborado pelos autores.

2) Arquivo no formato XML

A pergunta foi a seguinte:

O que você entende ao ler o trecho abaixo?

```
<?xml version="1.0" encoding="UTF-8">
  <usuario>
    <código=1>
      <nome>João Carlos</nome>
      <idade>26</idade>
      <cidade>Goiânia</cidade>
    </código>
  </usuário>
  <usuario>
    <código=2>
      <nome>Pedro Paulo Silva</nome>
      <idade>15</idade>
      <cidade>Anápolis</cidade>
```

```

        </código>
    </usuário>
    <usuario>
        <código=3>
            <nome>Carlos Henrique Pereira</nome>
            <idade>53</idade>
            <cidade>Aparecida de Goiânia</cidade>
        </código>
    </usuário>

```

As respostas foram as seguintes:

Figura 06 - Respostas do questionário sobre o formato de arquivos XML.



Fonte: elaborado pelos autores

3) Arquivo no formato JSON

A pergunta foi a seguinte:

O que você entende ao ler o trecho abaixo?

```

[
  {
    "usuário 1": {
      "nome": "João Carlos",

```

```
    "idade": "26",
    "cidade": "Goiânia"}
}
{
  "usuário 2": {
    "nome": "Pedro Paulo Silva",
    "idade": "15",
    "cidade": "Anápolis"}
}
{
  "usuário 3": {
    "nome": "Carlos Henrique Pereira",
    "idade": "53",
    "cidade": "Aparecida de Goiânia"}
}
]
```

As respostas foram as seguintes:

Figura 07 - Respostas do questionário sobre o formato de arquivos JSON.



Fonte: Elaborado pelos autores.

Com base nos dados dos três gráficos acima, podemos concluir que para o usuário final, os arquivos CSV são mais fáceis de ser compreendidos, posteriormente os arquivos XML e por último os arquivos JSON.

Abaixo segue o quadro de aplicação dos indicadores aos formatos de arquivos XML, CSV e JSON, sendo que para os indicadores “Compreensão por humanos” e “Velocidade de processamento” foi aplicado uma metodologia diferente, de acordo com a classificação obtida do experimento e da pesquisa que foram realizados e descritos acima..

Quadro 01 - Aplicação dos indicadores aos formatos de arquivos.

	XML	CSV	JSON
Facilidade de utilização	Sim	Sim	Sim
Compreensão por humanos	2º lugar	1º lugar	3º lugar
Velocidade de processamento	3º lugar	2º lugar	1º lugar
Estrutura dos dados	Sim	Não	Sim
Interoperabilidade entre os sistemas	Sim	Sim	Sim

Fonte: Elaborado pelos autores.

Conforme o quadro 01 e a seção 4, nota-se que os padrões XML e JSON possuem características e aplicações semelhantes, sendo que o XML é um padrão mais estruturado e a melhor opção quando trata-se de transferir arquivos entre sistemas diferentes, para serem processados à parte e preservados, enquanto os arquivos no formato JSON se aplicam mais quando o importante é que a informação seja processada diretamente pela aplicação sem a necessidade de arquivos externos e com um maior desempenho. Já os arquivos CSV são mais interessantes quando deseja-se extrair uma quantidade de informações muito grande, devido a sua capacidade de compactação, quando a estruturação dos dados não é interessante e também quando os dados serão lidos por programas de edição de planilhas eletrônicas.

6 CONCLUSÃO

Os Dados Abertos Governamentais são de grande valia para a sociedade, porém conforme foi tratado nesse artigo, a maneira como os mesmos são extraídos é bastante importante, pois pode impactar diretamente na maneira que serão processados ou reutilizados posteriormente.

No artigo foram tratados 3 formatos de dados, XML, CSV e JSON, sendo que ambos podem ser utilizados para a extração de dados em qualquer projeto, foram propostos alguns indicadores de usabilidade e aplicados sobre esses formatos. Conforme foi mostrado, quando o assunto é um volume extremamente elevado de informações e não se deseja a estruturação dos mesmos, o formato CSV é mais recomendado. Para a utilização de dados estruturados, e integração entre diferentes sistemas ou frameworks XML e JSON são a melhor opção, sendo que o segundo é uma opção ao primeiro. A maior diferença entre os dois é que o JSON tem uma maneira mais simples de utilizar os marcadores e conseqüentemente reduz a quantidade de caracteres a serem processados e quando a opção é um rápido processamento pela máquina sem a necessidade de arquivos externos, ele é uma opção melhor do que o XML.

7 REFERÊNCIAS

Brasil. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Logística e Tecnologia da Informação. Departamento de Governo Eletrônico. **Indicadores e métricas para avaliação de e-Serviços/Departamento de Governo Eletrônico**. Brasília: MP, 2007. 44 p.

CHOO, C. W. **A organização do conhecimento: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões**. São Paulo: Senac São Paulo, 2003. 425p.

EAVES, David. **The three laws of open government data**. September 2009. Disponível em: <<http://eaves.ca/2009/09/30/three-law-of-open-government-data/>> . Acesso em 18 de jun. 2016.

ECMA International. **The JSON Data Interchange Format**. 2013. 14p. Disponível em: <<http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-404.pdf>>. Acesso em 13 maio. 2016.

MORAES, Anamaria de e MONT' ALVÃO, Claudia. **Ergonomia: conceitos e aplicações**. 2a Ed. Rio de Janeiro (Rio de Janeiro), 2AB - Série Oficina, 2000. 1332p.

Governo do Estado de São Paulo e Governo do Reino Unido. **Guia de dados abertos - Projeto de cooperação entre o Governo do Estado de São Paulo e o Reino Unido**. 2015. Disponível em: <<http://pt.slideshare.net/colaborativismo/guia-de-dados-abertos>>. Acesso em 10/06/2016.

NIELSEN, Jakob. **Usability Engineering**. San Francisco (California), Morgan Kaufmann, 1993, 362p.

NIELSEN, Jakob. **Usability 101**. August 2003. Useit.com. Disponível em: <<http://www.useit.com/alertbox/20030414.html>>. Acesso em 25 de maio. 2016.

Open Knowledge International. **The Open data Handbook Documentation**. 2014. 27p. Disponível em: <<http://opendatahandbook.org/>> . Acesso em 02 de abr. 2016.

Open Knowledge International. **What is Open?** 2011. Disponível em: <<https://okfn.org/opendata/>>. Acesso em 02 de maio. 2016.

PREECE, J.; Rogers, Y.; Sharp, E. **Interaction Design: Beyond Human-computer Interaction**. New York, NY: John Wiley & Sons. 2002.

Secretaria de Logística e Tecnologia da Informação (SLTI) e Ministério do Planejamento Orçamento e Gestão (MP). **Cartilha Técnica para Publicação de Dados Abertos no Brasil v1.0**. 2012. Disponível em: <http://www.governoaberto.sp.gov.br/wp-content/uploads/2016/02/Cartilha_Tecnica.pdf> . Acesso em 11 de maio. 2016.

The Internet Society. **Common Format and MIME Type for Comma-Separated Values (CSV) Files**. 2005. Disponível em: <https://tools.ietf.org/html/rfc4180>. Acesso em 13/05/2016.

Tribunal de Contas da União. **5 Motivos para a abertura de dados na Administração Pública**. Brasília, 2015. 23p. Disponível em: <<http://portal3.tcu.gov.br/portal/pls/portal/docs/2689107.PDF>> . Acesso em: 09 de maio. 2016.

W3C. **XML**. Disponível em: <<http://www.w3c.org/XML>>. Acesso em 13 de maio. 2016.