

VIII ENARQUIFES

Encontro Nacional dos Arquivistas e Técnicos
das Instituições Federais de Ensino Superior

Preservando o legado: Integrando o passado com novas tendências arquivísticas para garantir um futuro seguro

6 a 10 de outubro de 2025

Campus Pampulha da UFMG

DIARQ
DIRETORIA DE
ARQUIVOS
INSTITUCIONAIS

UFMG

A CLASSIFICAÇÃO DE DOCUMENTOS NO FALA.BR POR MEIO DO USO DO CHATGPT

Autor: Zenóbio dos Santos Júnior

Arquivista da Universidade Federal de Ouro Preto

Doutor em Sistemas de Informação e Gestão do Conhecimento

INTRODUÇÃO

A LAI garante ao cidadão o acesso à informação de duas formas: **transparência ativa e passiva**.

- i. Os órgãos públicos disponibilizam um rol de informações em seus **sítios eletrônicos**;
- ii. Disponibiliza um canal de atendimento para o cidadão, denominado **Fala.BR**.

INTRODUÇÃO

Problema de pesquisa:

- Grande concentração em **um único assunto** (campo obrigatório);
- **Ausência de registro** de alguns campos nos pedidos de acesso à informação como “Subassunto” e “Tags” (campo opcional).

Objetivo: Propor a classificação de documentos nos pedidos de acesso à informação do Fala.BR por meio do uso do ChatGPT.

DESENVOLVIMENTO

Figura 1 – Fluxo do pedido de acesso à informação



Fonte: Santos Júnior; Dutra; Luz (2025)

DESENVOLVIMENTO

Figura 2 – Estruturação de um esquema de classificação do Fala.BR e CCDA¹

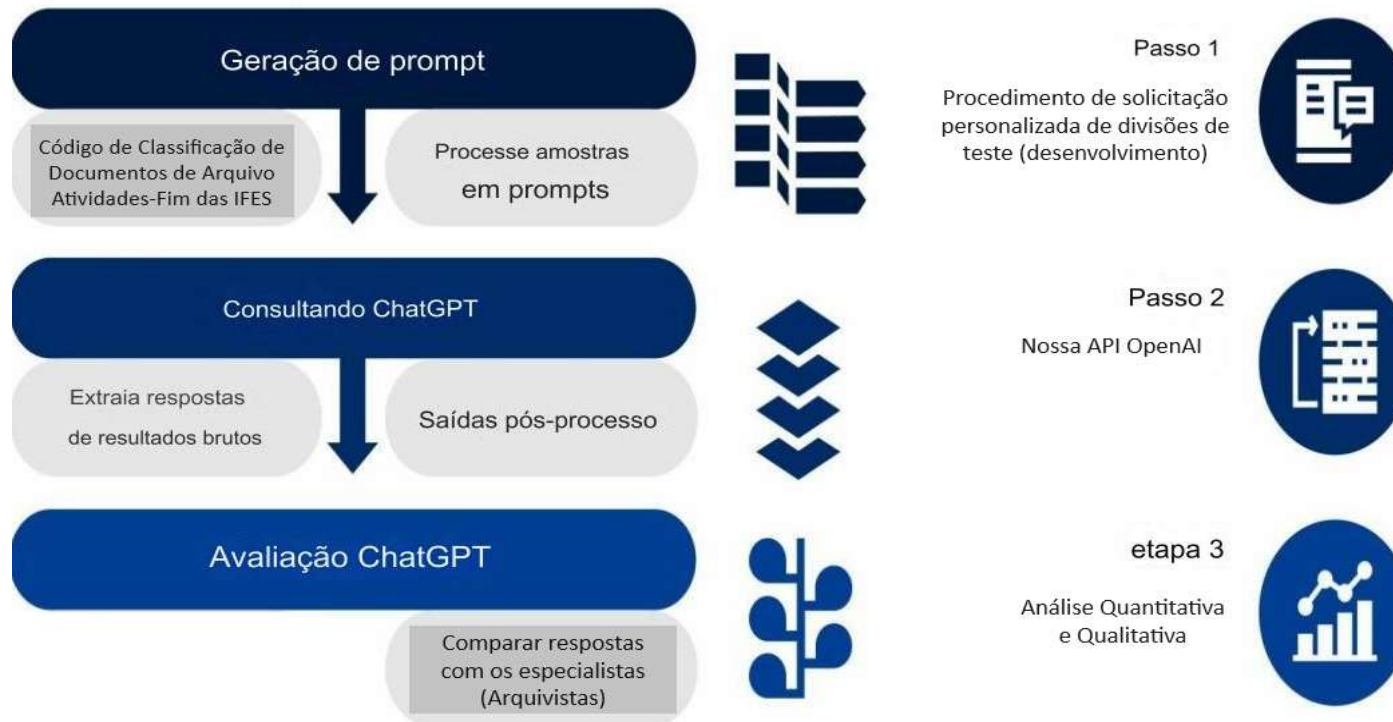


Fonte: Santos Júnior; Corrêa; De Faria (2023)

¹ Código de Classificação de Documentos de Arquivo

DESENVOLVIMENTO

Figura 3 – Estágios do processamento de dados do ChatGPT



Fonte: Adaptado de Kocón et al. (2023)

METODOLOGIA

Pesquisa qualitativa, exploratória e bibliográfica

1. Levantamento das informações dos **pedidos de acesso à informação do Fala.BR** cadastrados no ano de 2023;
2. Selecionados pedidos que contenham a palavra **Educação**;
3. Envio de 50 textos da **busca de pedidos e respostas/Fala.BR** para 10 arquivistas (5 texto para cada);
4. Os arquivistas **classificaram os textos** dos pedidos de acesso à informação do Fala.BR, utilizando, o Código de Classificação de Documentos de Arquivo para atividades-fim da IFES;
5. Comparar os resultados dos materiais produzidos pelos **arquivistas com o do ChatGPT**, a fim de avaliar o grau de eficiência da ferramenta.

RESULTADOS

Tabela 1 – Assuntos, Subassuntos e *Tags* das Universidades Federais em 2023

Descrição	ASSUNTOS	SUBASSUNTOS	TAGS
Cadastro	Obrigatório	Opcional	Opcional
Quantidade de registros	127 registrados	715 registrados	203 registrados
Quantidade de Universidades Federais	64 Universidades preencheram o campo	35 Universidades que preencheram o campo	33 Universidades que preencheram o campo
Termo mais registrado	Acesso à Informação (45,37%)	Transparência/ Institucional (0,69%)	SIC (1,30%)
Percentual de preenchimento	100%	18,85%	18,49%
Percentual de campo vazio	0%	81,15%	81,51%

Fonte: Dados da pesquisa



RESULTADOS

Tabela 2 – Assuntos das Universidades Federais em 2023

Ordem	Assunto	Quantidade	%
1	Acesso à informação	5.418	45,37
2	Universidades e Institutos	691	5,79
3	Educação Superior	642	5,38
4	Recursos Humanos	544	4,56
5	Concursos	438	3,67
6	Ações Afirmativas	405	3,39
7	Outros em Administração	341	2,86
8	Transparência	281	2,35
9	Outros em Educação	244	2,04
10	Planejamento e Gestão	224	1,88

Fonte: Elaborado pelo autor.

RESULTADOS

Tabela 3 – Subassuntos das Universidades Federais em 2023

Ordem	Tipo de Subassunto	Quantidade	%
1	VAZIO	9.690	81,15
2	Transparência Institucional	82	0,69
3	Pesquisa Acadêmica	58	0,49
4	Documentos	48	0,40
5	Graduação	46	0,39
6	Informações administrativas	37	0,31
7	Ações Afirmativas	32	0,27
8	Assistência Estudantil	31	0,26
9	Informações e dados acadêmicos	30	0,25
10	Informações Institucionais	28	0,23

Fonte: Elaborado pelo autor.

RESULTADOS

Tabela 4 – Tags das Universidades Federais em 2023

+

Ordem	Tag	Quantidade	%
1	VAZIO	9.733	81,51
2	SIC	155	1,30
3	Pesquisa	146	1,22
4	Pró-Reitoria de Gestão de Pessoas	122	1,02
5	Dados e Informações	79	0,66
6	Pró-Reitoria de Graduação	78	0,65
7	Ouvidoria	51	0,43
8	PROGEP	48	0,40
9	Superintendência de Gestão de Pessoas	48	0,40
10	Pró-Reitoria de Administração	33	0,28

Fonte: Dados da pesquisa.

RESULTADOS

Tabela 5 – Resultados da análise de conteúdo: Homem x Máquina

Principais Temas	Análise de Concordância/ Similaridade de cosseno	Resultados	Média de concordância
Assunto	Acurácia Kappa de Cohen	38% 0.1066	Concordância Ligeira
Subassunto	<i>BERTimbau e Word Embedding</i>	0.5945	Concordância Moderada
Tag	<i>BERTimbau e Word Embedding</i>	0.7702	Concordância Relativamente Alta

Fonte: Dados da pesquisa

- O coeficiente de Kappa apontou um valor de 0.1066, das classificações feitas pelo ChatGPT coincidiram exatamente com as do especialista humano, indicando uma concordância ligeira para o Assunto.
- A similaridade dos valores de cosseno (BERTimbau) dos Subassuntos apresentaram, aproximadamente, 0.5945, o que sugere uma concordância moderada.
- E para as Tags, cosseno (BERTimbau), apresentou um valor em 0.7702, o que sugere uma concordância relativamente alta.

RESULTADOS

Quadro 1 – Resultado dos principais temas indexados

TEXTO 12	Participação em pesquisa para Pós-graduação em ensino de matemática Pedido 23546079412202313 - 08/09/2023 - Acesso Concedido		
Detalhamento Solicitacao	Prezados, Meu nome é [REDACTED], sou doutoranda do curso de Pós-graduação em Educação Matemática e Tecnológica da Universidade Federal de Pernambuco - EDUMATEC - UFPE. Estou desenvolvendo um projeto de pesquisa sobre questões étnico-raciais em cursos de pós-graduação em Educação Matemática. Um dos objetivos do meu trabalho é coletar informações acerca das ações desenvolvidas pelas instituições de Ensino Superior para promoção do enfrentamento ao racismo, em cursos de pós-graduação stricto sensu na área de ensino de matemática. Assim, solicito, gentilmente, apoio para minha pesquisa, respondendo às questões no corpo desse e-mail, pois sem essa participação não será possível identificar os dados iniciais para o desenvolvimento do meu estudo. Seguem as perguntas: 1. O curso dispõe de políticas públicas de ações-afirmativas para estudantes/professores negros? Quais? 2. No caso de acesso por cotas raciais, quantos estudantes acessaram o curso através da política de cotas em 2023? 3.O curso desenvolve grupos/linhas de pesquisa, disciplinas, eventos, formações, pesquisas que contemplem temáticas que promovam a equidade racial? Se sim, quais? Desde já sou imensamente grata pela participação. Atenciosamente,		
TEMA	USUÁRIO FALA.BR (Pedido Original)	CHATGPT (Análise Digital)	ARQUIVISTA (Análise profissional)
ASSUNTO	Acesso à informação	Ensino Superior	Pesquisa
SUBASSUNTO	-	Cursos de pós-graduação stricto sensu (inclusive na modalidade a distância)	Outros assuntos relacionados à pesquisa
TAG 1	-	Políticas públicas de ações-afirmativas	Pedido de informações
TAG 2	-	Cotas raciais	Ações afirmativas
TAG 3	-	Equidade racial	Cotas raciais
TAG 4	-	Grupos de pesquisa	Equidade racial
TAG 5	-	Eventos acadêmicos	Educação matemática

Exemplo de texto da Busca de pedidos e respostas avaliado pelo Fala.BR, ChatGPT e Arquivistas.

Fonte: Dados da pesquisa

CONSIDERAÇÕES FINAIS

A pesquisa apontou que um único assunto “acesso à informação” **possui quase a metade dos registros** no preenchimento dos Subassuntos e Tags, além de uma **grande ausência no preenchimento**, superando os 80%;

Pode-se observar a inversão de valores no Fala.BR, pois **o que era negativo** (ausência de preenchimento em 80%) **passou a ser positivo com uso do ChatGPT e dos Arquivistas** (59% para os Subassuntos e 77% das Tags preenchidas);

O ChatGPT pode vir a apresentar as opções de preenchimento para os usuários, inclusive, os Subassuntos e Tags; ou sugerir a **inclusão do código de classificação de documentos de arquivo das atividades-meio**.

REFERÊNCIAS

Brasil. (2011). **Lei nº 12.527**, de 18 de novembro de 2011.

Brasil. (2023). Controladoria-Geral da União. Secretaria de Transparência e Prevenção da Corrupção. Acesso à Informação. **Busca de Pedidos e Respostas**. *Download de Dados*.

Brasil. (2013). Ministério da Gestão e da Inovação em Serviços Públicos. Conselho Nacional de Arquivos. Portaria MEC Nº 1.224, de 18 de dezembro de 2013. **Código de Classificação de Documentos de Arquivo Relativos às Atividades-fim das Instituições Federais de Ensino Superior e na Tabela de Temporalidade e Destinação de Documentos de Arquivo Relativos às Atividades-Fim das Instituições Federais de Ensino Superior**. Rio de Janeiro: Arquivo Nacional.

Brasil. (2020). Ministério da Gestão e da Inovação em Serviços Públicos. Conselho Nacional de Arquivos. **Código de classificação e tabela de temporalidade e destinação de documentos relativos às atividades-meio do Poder Executivo Federal**. Rio de Janeiro: Arquivo Nacional.

Kocoń, Jan et al. (2023). **ChatGPT: Jack of all trades, master of none**. Information Fusion, 99, 101861.

Santos Júnior, Zenóbio; Corrêa, Fábio; De Faria, Vinícius F. (2023). **Classificação de documentos de arquivo: principais temas dos pedidos de acesso à informação nas universidades federais mais demandadas pela LAI**. *Arquivos, democracia e justiça social [livro eletrônico]* / organização Mariana Lousada Pinha, Marcia Cristina de Carvalho Pazin Vitoriano, Paulo Roberto Elian dos Santos. -- 1. ed. -- São Paulo: ARQ-SP.

Santos Júnior, Zenóbio; Dutra, Frederico Giffoni de Carvalho; Luz, Eduardo José da Silva (2025). **O uso do ChatGPT na indexação dos pedidos de Acesso à Informação no Fala.BR**. Advances in Knowledge Representation, [S. l.], v. 5, n. 1.

CONTATO

Zenóbio dos Santos Júnior

zenojr19@gmail.com

arquivocentral@ufop.edu.br