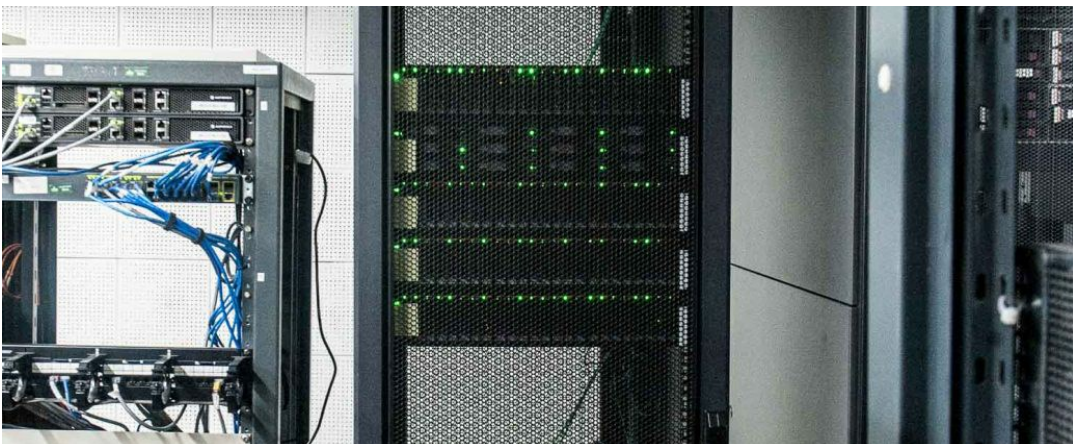


## Para uma história de sites de projetos de pesquisa da UFRGS

Jornal da Universidade / 24 de novembro de 2022 / Artigo



### Artigo | Maria José B. Finatto, docente do PPG Letras, promove o diálogo na Universidade sobre como manter a memória digital de bancos de dados e ferramentas hospedados em servidores da instituição

\*Por: Maria José B. Finatto

\*Foto: Flávio Dutra/JU

Sou linguista. Desde 2001, investigo fenômenos da comunicação técnico-científica, vocabulários e modos de comunicar ciência para pessoas com escolaridade limitada e pouca experiência de leitura. Com apoio de órgãos de fomento à pesquisa, da Pró-reitora de Pesquisa e também da Secretaria de Educação a Distância (SEAD-UFRGS), desenvolvemos uma série de sites, bases de dados e ferramentas online para análise linguística.

A maioria dos recursos visa apoiar atividades de aula, graduação e pós-graduação, formação de professores, de tradutores e de estudiosos da linguagem. São vários produtos, todos hospedados e colocados em funcionamento em servidores da UFRGS. Servidores são grandes computadores que hospedam dados, interligam sistemas e guardam materiais de sites e softwares. Nesses servidores – hoje chamados nuvens –, ferramentas e recursos funcionam, vários sem que você precise baixar nada no seu computador.

Realizar coisas desse tipo envolveu e ainda envolve alunos de iniciação científica, mestrandos e doutorandos. São recursos públicos para auxiliar diferentes pessoas no Brasil e em diferentes partes do mundo. É o caso, entre outros, do [Dicionário colaborativo de português para estrangeiros](#), um protótipo modesto, realizado parcialmente em 2012. Está conectado a uma versão antiga, em formato mais estático, entrelaçado com um banco de textos de jornais populares.

*Quando pouco se falava em ensino remoto ou suporte à EAD, já produzíamos materiais de acesso público de uso online. Um exemplo é o [VIRTUALERN](#), ambiente para o estudo e ensino da tradução da língua alemã – construído em 2007, em meio a uma tese – que segue ativo.*

Muito do que criamos foi sendo hospedado no servidor Colossus, do Centro de Processamento de Dados (CPD-UFRGS), e em outros servidores, como o Nuvem UFRGS. O Colossus hospeda uma série de módulos e de bases de dados dos nossos recursos. Vale explicar: temos funcionalidades conectando diferentes servidores entre si. Isto é, uma base de dados A, hospedada no servidor 1, pode funcionar “chamando dados” de uma base de dados B, hospedada no servidor 2.

Esse é o caso da nossa [Ferramenta MedSimples](#) – um sistema que ajuda a simplificar um texto sobre temas de Saúde para pessoas de escolaridade limitada. MedSimples foi reconhecida pela empresa Google em 2019, o que nos rendeu um prêmio internacional. Outros projetos antigos, sobre linguagem e terminologias médicas, conectam-se e deram suporte à MedSimples: caso do [TextPneumo](#) – sobre Pneumopatias Ocupacionais. A imagem de um de seus módulos pode ajudar a entender o seu foco linguístico:

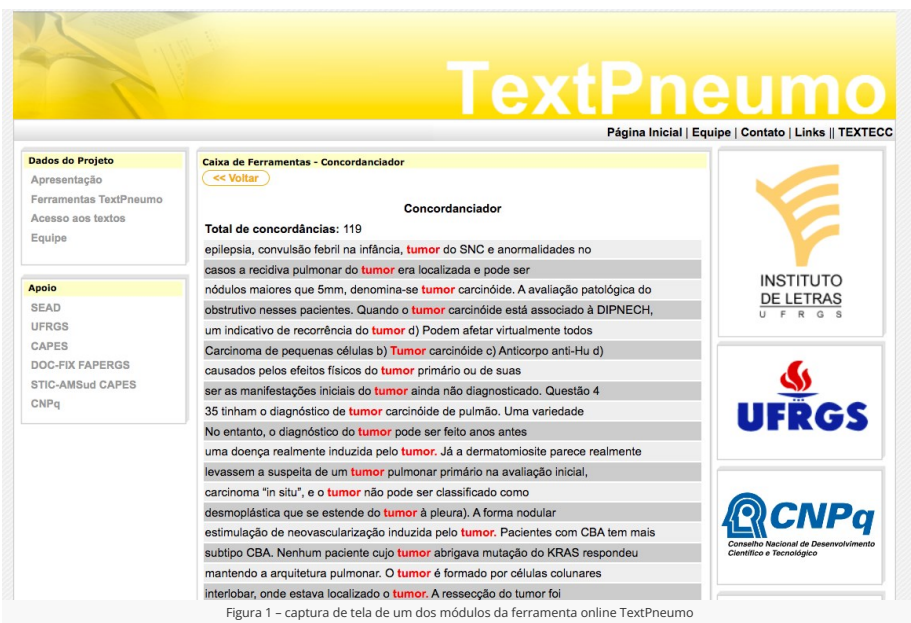


Figura 1 – captura de tela de um dos módulos da ferramenta online TextPneumo

A desativação de um servidor X significa, na prática, que será preciso mover dados, sites e todos os seus arquivos, bases e ferramentas para um novo servidor Y. No nosso caso, vale frisar, uma nova máquina já nos foi disponibilizada em forma de espaço de dados e infraestrutura, pelo CPD, pelo que muito agradecemos.

*Mas, em “mudanças” desse tipo, o trabalho técnico de migração, em geral, não é trivial, sobretudo quando já temos algumas tecnologias “velhas”.*

Esse serviço o nosso CPD não tem condições de oferecer, o que posso entender. Além disso, há centenas de casos de sites e de recursos como os nossos. Por isso, resolvi escrever este texto.

Após o aviso do CPD sobre a descontinuidade do Colossus, preocupados, buscamos apoios para preservar tal massa de dados e essa memória digital. Felizmente, o Programa de Pós-graduação em Letras está nos ajudando e a migração sendo feita. Como se pode imaginar, exige-se um verdadeiro trabalho de equipe, uma certa engenharia computacional e recursos para pagar o trabalho.

Por fim, menciono o nosso sistema [CORPOP](#) – base de referência para apontar padrões de um português escrito simples. São *corpora* e ferramentas – no site [PORLEXBRAS](#). Esse recurso, como tantos outros de vários colegas da UFRGS, tem sido bastante utilizado, reconhecido e citado. Seu uso mais recente é iniciativa do Tribunal de Justiça do RS (TJ-RS), que publicou um guia para a [Linguagem Jurídica Simples](#). O TJ-RS valoriza e recomenda a pesquisa linguística na UFRGS.

Meu leitor, sendo também pesquisador, caso tenha produzido materiais e os hospedado em plataformas da UFRGS, pode perguntar: e se isso acontecer comigo?

*Como podemos preservar nossas histórias de trabalho, sem o risco de ter dados e produtos simplesmente apagados dos registros da nossa Universidade? Como cuidaremos da nossa memória digital?*

Muitas vezes, tem-se notícia de um site ou recurso importante oferecido em uma pesquisa na UFRGS. Mas, muitas vezes, vemos, frustrados, a mensagem: site fora do ar/inexistente. E a leitura é que tal recurso teria sido descontinuado, envelhecido. Ou seria um caso de um projeto acabado?

Será apenas responsabilidade do pesquisador arcar com isso? E em caso de quem se desligue da UFRGS ou que tenha falecido sem deixar “herdeiros de pesquisa”? Quem cuidará de salvaguardar dados e produtos oferecidos online?

Por isso, convido você a nos ajudar a pensar sobre essa situação, que pode se repetir, a despeito de “nuvens mágicas” que parecem eternas. Como construir uma memória digital com o que fazemos na UFRGS com recursos públicos? Temos, sim, a base LUME, os recursos para guarda de objetos de ensino do NAPEAD e da SEAD. Temos o Moodle Histórico. Mas vale perguntar: como poderíamos contar toda essa história online, que certamente é feita e foi feita nas áreas das Exatas, Saúde e Humanidades?

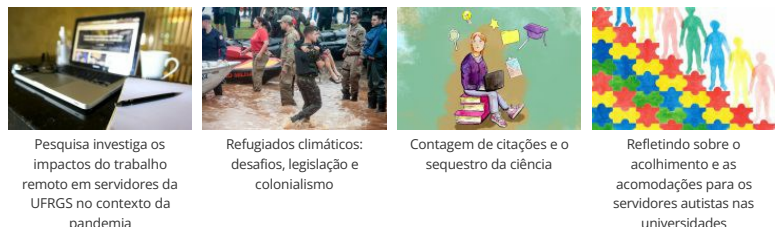
A maioria dos sistemas desse gênero foram previstos para funcionar de modo gratuito, sem necessidade de baixar nada no seu PC. Não são “simples sites”. Um exemplo importante de um colega meu, fora da área de Letras, é o [SOBEK](#) – uma ferramenta de mineração de textos. O [SOBEK online](#) foi feito para apoiar aplicações educacionais. Não há quem não fique encantado ao encontrar esse recurso gratuito, funcionando. É algo que tenho mostrado em cursos e citado em publicações.

Todos esses recursos reconhecidos e em uso – que nos orgulham como um grupo em diferentes unidades da UFRGS – foram sendo construídos ao longo de anos, sempre com apoio de verbas e de bolsas, especialmente da SEAD e de outros órgãos. Há muito a preservar. A questão é: como fazer isso? Considerando a necessidade de salvaguarda desse patrimônio, convido colegas pesquisadores e a comunidade a refletir sobre isso. Eu não tenho uma resposta.

Maria José B. Finatto é professora titular, docente colaboradora convidada (aposentada) do Programa de Pós-graduação em Letras da UFRGS.

*“As manifestações expressas neste veículo não representam obrigatoriamente o posicionamento da UFRGS como um todo.”*

### Posts relacionados



Pesquisa investiga os impactos do trabalho remoto em servidores da UFRGS no contexto da pandemia

Refugiados climáticos: desafios, legislação e colonialismo

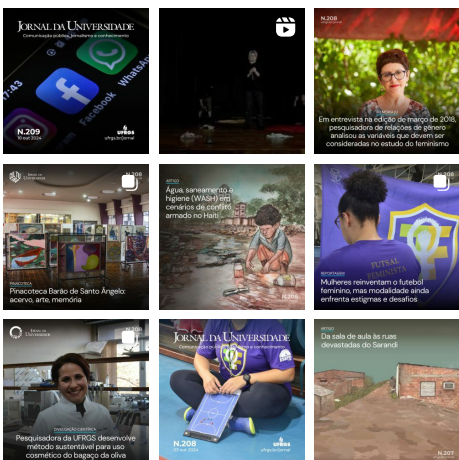
Contagem de citações e o sequestro de ciência

Refletindo sobre o acolhimento e as acomodações para os servidores autistas nas universidades

### INSTAGRAM

Jornal da Universidade UFRGS  
@jornaluniversidadeufrgs

Follow



View on Instagram

### REALIZAÇÃO

JORNAL DA UNIVERSIDADE

UFRGS  
SECOM

UFRGS

### CONTATO

Jornal da Universidade  
Secretaria de Comunicação Social/UFRGS

Av. Paulo Gama, 110 | Reitoria – 8. andar | Câmpus Centro |  
Bairro Farrópilha | Porto Alegre | Rio Grande do Sul | CEP:  
90040-060

(51) 3308.3368

jornal@ufrgs.br