



## **Proposta para Grupo de Trabalho Fase 2 2020**

GT-RecMEM - Recomendação de Mídias Educacionais

Reinaldo Matushima

Jairo Francisco de Souza

25 de abril de 2020

### **1. Título**

GT-RECMEM - Recomendação de Mídias Educacionais

### **2. Coordenador Geral**

## **Coordenador**

Prof. Dr. Jairo Francisco de Souza  
Departamento de Ciência da Computação - DCC  
Universidade Federal de Juiz de Fora – UFJF  
E-mail: jairo.souza@ice.ufjf.br  
Tel: +55-32-21023311 ramal 4037  
Lattes: <http://lattes.cnpq.br/4516605108233899>

## **3. Assistente de Inovação**

Reinaldo Matushima  
TV On Demand Consultoria e Soluções LTDA - BRAINYIT  
E-mail: reinaldo@tvod.com.br  
Tel: +55-11-2389-4589  
Lattes: <http://lattes.cnpq.br/6861921652323765>  
Linkedin: <https://www.linkedin.com/in/reinaldo-matushima-9a1a134/>

## **4. Resumo**

O desafio para todos que querem disponibilizar conteúdo educativo é como facilitar e ampliar o processo de busca em seus acervos. Este GT objetiva disponibilizar uma solução para indexação de repositórios de conteúdos educativos que permite atuar em duas frentes: (i) auxiliar o professor através de um sistema de recomendação de conteúdos indexados de diversos repositórios abertos; e (ii) fornecer dados comparativos sobre o processo de ensino de diferentes escolas para secretarias de ensino. Os principais diferenciais da solução são: (a) a indexação de conteúdos privados de LMS; (b) a anotação semântica desses conteúdos (textos, vídeos) para auxiliar na busca; (c) a associação dos recursos educacionais com conteúdos e habilidades da Base Nacional Comum Curricular; (d) uma plataforma para recomendação de conteúdo para professores e alunos; (e) a coleta e disponibilização de dados de uso da plataforma para gestores tomarem conhecimento sobre práticas de ensino em diferentes escolas ou turmas.

## **5. Abstract**

How to facilitate and expand the search process in their collections is the challenge for educational content creators. This WG aims to provide a solution for indexing educational content repositories that allows it to act on two fronts: (i) assist the teacher through a system for recommending indexed content from several open repositories; and (ii) provide comparative data on the teaching process of different schools for decision makers. The main contributions of the solution are (a) the indexing of private LMS content; (b) the semantic annotation of these contents (texts, videos) to improve the search; (c) the association of educational resources with the contents and skills of the BNCC; (d) a platform for recommending content for teachers and students; (e) the collection and availability of data on the use of the platform for decision makers to learn about teaching practices in different schools or courses.

## 6. Link do vídeo-pitch

[https://youtu.be/yU\\_y15\\_8nEY](https://youtu.be/yU_y15_8nEY)

## 7. Parcerias e respectivas contrapartidas

Este projeto é uma parceria entre a Universidade Federal de Juiz de Fora e a BrainyIT - TV on Demand Consultoria e Soluções SS LTDA.

A Universidade Federal de Juiz de Fora será responsável pela pesquisa aplicada em sistemas de recomendação, o desenvolvimento da solução e o treinamento dos recursos humanos para manutenção do serviço após entrada em produção. O projeto será desenvolvido na sua maior parte na Universidade Federal de Juiz de Fora (UFJF), pelos laboratórios de pesquisa:

- Laboratório de Aplicações e Inovação em Computação – LapIC (Multimídia, IPTV e videoaulas)
- Núcleo de Engenharia do Conhecimento – NEnC (Recuperação de Informação, Resolução de Identidade e Representação do Conhecimento)

BrainyIT - TV on Demand Consultoria e Soluções SS LTDA dará o apoio necessário para adequação do produto para posicionamento no mercado e no rol de serviços da RNP. A empresa irá entrar também como contrapartida com horas de trabalho de seus colaboradores, no que se relaciona a questões técnicas e desenvolvimento associadas à integração ao serviço nas Nuvens e outras demandas associadas à produção do MVP. A princípio é previsto de 40 a 80 horas mensais de envolvimento pela empresa, conforme necessidades que sejam verificadas. Além disso a empresa irá entrar com apoio financeiro de até R\$ 10.000,00 para viabilização das atividades.

## 8. Descrição da evolução do MVP com destaque para a entrada na loja virtual da RNP

Ao longo do desenvolvimento do RECMEM, foi realizado um conjunto de entrevistas com diferentes públicos, como professores de escolas públicas e privadas, gestores de educação, empresários, abrangendo setores como educação a distância, treinamentos, entre outros. Por conta dessas entrevistas, o MVP entregue ao final da fase 1 se difere consideravelmente da proposta inicial do GT, estando mais alinhado às necessidades do mercado, com uma solução validada que apresenta maior valor aos potenciais clientes.

As entrevistas finais do projeto com as secretarias de ensino evidenciaram o interesse dos estados em fornecer uma solução de fácil uso para professores e alunos. Algo que possibilite liberdade de uso e planejamento pelos professores e, que auxilie as secretarias no cruzamento de informações de seus dados de acompanhando da rede de ensino, como frequência e desempenho, com as práticas pedagógicas de cada escola. Com isso eles conseguem enriquecer suas análises, permitindo identificar boas práticas dentro da rede de ensino e mitigar com celeridade possíveis indicadores de abandono e baixo rendimento.

Percebemos que o cenário vivido no Brasil com a COVID-19 a partir de março/2020 jogou o foco dos gestores estaduais para soluções que apoiem professores a escolher bons materiais e que permita um melhor acompanhamento das secretarias das práticas fora de sala de aula. Neste sentido, entendemos que o cenário é propício para o produto do GT.

O GT pretende trabalhar em pontos que tornam inviável o reuso e a recomendação automática de mídias nos repositórios existentes. O principal aspecto é a falta de metadados adequados nos conteúdos dos repositórios de Objetos de Aprendizagem (OA). Apesar do grande potencial de reuso gerado pelo correto preenchimento dos metadados, os repositórios existentes apresentam grandes deficiências de catalogação, devido a diversos aspectos: falta de conhecimento para preenchimento dos metadados técnicos e/ou pedagógicos, falta de tempo, o fato do OA não ser inserido na plataforma por seus idealizadores (e sim por terceiros) e até mesmo o desinteresse em fornecer esses metadados (por desconhecimento da importância de um cadastro adequado).

Para isso, diferentes resultados de pesquisa do LApIC são empregadas na solução. O RECMEM utiliza o BAVi (Gomes Jr et al, 2019a; Souza et al, 2017; Dias et al, 2017), o qual é aplicado para extrair metadados de materiais didáticos automaticamente. Ainda, esta abordagem possui como principal diferencial o processamento do grafo de conhecimento do DBpedia para atribuir categorias (tópicos) aos vídeos. Diferente de outras abordagens presentes na literatura e mercado (Kim et al, 2019; Dias & Wives; 2019; Neves et al, 2017; Ribeiro et al, 2013; Ferreira & Raabe, 2010), o RECMEM considera o professor como ator principal do processo de ensino e sua atuação como importante direcionador do processo de recomendação.

Assim, embora o processo de recomendação se dê de forma automática com base na similaridade entre recursos didáticos e com o conteúdo apresentado no curso, a recomendação se dá, primeiro, para que o professor possa escolher o conteúdo que mais se adequa ao curso, agindo como facilitador da busca em grandes repositórios, na mesma linha do projeto desenvolvido pelos autores da presente proposta (Barrére et al. 2018), cujo foco era a recomendação de vídeos, mas ampliando para mais de um tipo de mídia e mais de um repositório.

Todo o processo de extração de dados de repositórios é feito de forma transparente para o usuário, permitindo que novos coletores sejam facilmente integrados na plataforma. Além da integração com o Moodle, o qual aperfeiçoa as pesquisas anteriores do laboratório (Barrére et al. 2018), o RECMEM permite coletar vídeos, documentos e jogos interativos de plataformas abertas. O processo de definição dos coletores e processamento do conteúdo aplica resultados em *web scraping* do LApIC (Gomes Jr et al, 2019b).

Como resultado dessa mudança de trajetória do GT, o escopo do sistema se ampliou e o esforço de desenvolvimento teve que ser readequado, priorizando as funcionalidades principais que permitisse a demonstração do produto para possíveis clientes. Foi decidido focar em três públicos (professor, aluno e secretaria de ensino). Para cada um destes usuários, foi considerado o suporte aos seguintes cenários (casos de uso):

Alunos:

- 1) Acessar cadastro;
- 2) Realizar login;
- 3) Visualizar as recomendações gerais do sistema (\*);

- 4) Visualizar as recomendações com base nas preferências do aluno (\*);
- 5) Visualizar as recomendações para suas turmas;
- 6) Visualizar as recomendações diretas do professor;
- 7) Buscar um material selecionando conceitos e tipos de mídia;
- 8) Ver anotações do professor;
- 9) Avaliar e comentar o material (\*);
- 10) Acessar página para entrar em uma turma.

#### Professores:

- 1) Acessar cadastro;
- 2) Realizar login;
- 3) Visualizar as recomendações gerais do sistema (\*);
- 4) Visualizar as recomendações com base nas preferências do professor (\*);
- 5) Visualizar as recomendações com base nas metas das turmas (\*);
- 6) Visualizar as recomendações de outros professores de turmas similares;
- 7) Buscar um material selecionando conceitos e tipos de mídia;
- 8) Criar turma e definir suas metas;
- 9) Acessar página de turma: alunos, metas e descrições;
- 10) Gerar código para acesso às turmas;
- 11) Visualizar estatísticas de turmas;
- 12) Acompanhar progresso dos alunos das turmas;
- 13) Recomendar materiais para alunos ou turmas; 14) Inserir tags em materiais;
- 15) Gerenciar biblioteca de materiais.

#### Secretarias de ensino

- 1) Cadastrar escolas e professores;
- 2) Configurar instâncias de moodles para ser indexado pelo RECMEM;
- 3) Visualizar dados de escolas, gerando relatórios de acesso, conteúdos assistidos;
- 4) Comparar escolas, analisando práticas de ensino e dados sobre motivação; 5) Integrar dados com base de dados das secretarias estaduais (\*).

Através desses cenários, permitimos demonstrar que a solução para indexação de repositórios de conteúdos educativos permite atuar em duas frentes:

- (i) Auxiliar o professor através de um sistema de recomendação de conteúdos indexados de diversos repositórios;
- (ii) Fornecer dados comparativos sobre o processo de ensino de diferentes escolas para secretarias de ensino.

Neste ponto, destacamos os diferenciais da solução como:

- (a) A indexação de conteúdos privados de LMS e de repositórios de mídias externos;
- (b) A anotação semântica desses conteúdos (textos, vídeos) para auxiliar na busca;
- (c) A associação dos recursos educacionais com conteúdos e habilidades da Base Nacional Comum Curricular;
- (d) Uma plataforma para recomendação de conteúdo para professores e alunos;

- (e) A coleta e disponibilização de dados de uso da plataforma para gestores tomarem conhecimento sobre práticas de ensino em diferentes escolas ou turmas.

Das funcionalidades destacadas nos roteiros, as destacadas com (\*) são foco do trabalho nesta segunda fase. Resumidamente o projeto previsto para a fase 1 cresceu muito e o objetivo para a continuidade do projeto, é voltar a atenção para pontos que foram tratados de forma superficial no MVP. O objetivo assim, é amadurecer bem a essência para o qual o projeto se propõe a atender, realizando eventuais adequações que sejam verificadas ao longo do processo de maturação do modelo de negócios junto à RNP.

Ao longo da primeira fase, foram levantados diversos modelos de negócio e estratégias distintas de monetização que poderiam ser explorados. Por exemplo, verificou-se que os mecanismos de indexação e busca poderiam ser explorados em diversas áreas. Optou-se por seguir focando na área de educação, alvo inicial da proposta. Com relação a monetização do trabalho desenvolvido, foram identificadas diversas oportunidades. A seguir são apresentadas as principais identificadas até então:

1. Mensalidade por instância de Moodle indexado: cada instituição que desejar obter dentro de sua instância de Moodle as facilidades de indexação e recomendação de conteúdos do serviço, pagaria uma licença de uso. No moodle o professor teria recomendações de outros conteúdos da própria instância do Moodle e de acervos externos que sejam indexados;
2. Cobrança de relatórios customizados para secretarias/gestores de ensino, cruzando dados de sistemas da secretaria/instituição com dados da plataforma, permitindo análises e tomadas de decisão ricas: objetivo é atender uma das grandes dificuldades das secretarias, acompanhar o uso das tecnologias de apoio por parte dos alunos e conseguir extrair informações que permitam detectar e mitigar problemas como abandono e baixo rendimento, além de permitir identificar áreas estratégicas para investimento;
3. Licenciamento do serviço para outras plataformas de ensino. O acesso das facilidades da plataforma (como indexação e recomendação) mediante mensalidades proporcionais a número de alunos/conteúdos indexados: o objetivo é levar os benefícios para outras plataformas, gerando valor para investimentos que a instituição já tenha realizado (em outras plataformas que não sejam o Moodle);
4. Parcerias com acervos privados. Acervos privados pagar para que parte de seus conteúdos sejam indexados/recomendados, como uma forma de divulgação de seus materiais: o serviço passaria a ser um canal de divulgação adicional de acervos privados de mídias educacionais;
5. Pacotes voltados para o usuário final que incluiria acesso a conteúdos de acervos privados: estratégia complementar à parceria com acervos privados. Com este modelo, os acervos privados conseguem potencializar sua monetização;
6. Parceria com empresas que atuem com o Moodle: o serviço poderia agregar valor aos projetos de Moodle oferecidos por estas empresas.

O modelo final de monetização pode ser uma das estratégias identificadas ou uma combinação das mesmas. O modelo será amadurecido ao longo da segunda fase, provavelmente focando no modelo que tenha maior potencial e que possa ser mais

facilmente explorado inicialmente dentro do sistema RNP, evoluindo com o decorrer do serviço.

Resumindo a proposta, a fase dois visa amadurecer o MVP, que se transformou em um meta agregador de conteúdos educacionais, que indexa conteúdos de plataformas de ensino (inicialmente Moodle, mas com potencial de integração com outros LMS) e de repositórios de mídias diversas (de acervos públicos como o EduPlay a acervos privados quaisquer) e possibilita via sistema de recomendação, auxiliar o professor a fazer reuso e escolha adequada de conteúdos, além de disponibilizar relatórios de comportamento de acesso aos materiais que permitam às secretarias de ensino extrair dados importantes para o acompanhamento e direcionamento de suas ações.

O esforço será, no viés técnico, focar em questões a qual não foi possível tratar adequadamente dada a ampliação das atividades do projeto na fase 1 em relação ao escopo inicial previsto, tal como a recomendação de conteúdos, que precisa de um grande esforço para incorporar todos os requisitos funcionais levantados como importantes pelos professores. Outro exemplo de aspecto técnico a ser trabalhado, é nos relatórios disponibilizados às secretarias/gestores de forma a envidar esforços em relatórios que atendam as principais dores do cliente. Sob o viés de negócio, o objetivo será amadurecer o modelo de monetização mais adequado a curto prazo considerando as instituições usuárias da RNP. **9. Cronograma de marcos**

Marcos	Meses											
	1	2	3	4	5	6	7	8	9	10	11	12
Disponibilização de nova versão do repositório com novo conjunto de mídias coletas de repositórios de terceiros	x	x	x									
Disponibilização de novo serviço de indexação de mídia com uso de solução de speech2text proprietário		x	x	x								
Disponibilização de melhoria no módulo de recomendação com atualização do modelo de recomendação para professores e alunos			x	x	x	x						
Disponibilização do serviço no NasNuvens, após Levantamento de requisitos e adequação do serviço.				x	x	x						
Disponibilização de melhorias nos classificadores de habilidades e conceitos da BNCC							x	x	x			
Liberação do serviço para potenciais clientes							x					
Levantamento de novos requisitos levantados pelos clientes e atendimento a novas demandas										x	x	x

Além das atividades definidas acima, salienta-se as atividades de frente de negócios já estabelecidas no edital da Fase 2 pela RNP.

## 10. Recursos financeiros

### 10.2. Infraestrutura

#### 10.2.1. Equipamentos

Descrição	Instituição de Destino	de Qtd.	Valor em R\$ estimado
Desktop Modelo i7 (Core i7 - 16GB - SSD 256GB) + Kit teclado e mouse com fio  Incluso: Garantia 3 anos ProSupport, troca de peças onsite	UFJF	3	R\$10.807,92
Monitor 24"	UFJF	3	R\$3.149,37
Display 42" LED 4K com suporte de parede para monitoramento de datacenter	UFJF	1	R\$2.500,00
Display 21" LED para PC	UFJF	1	R\$ 900,00
No-breaks 2KVA para redundância de alimentação para switches de rede	UFJF	2	R\$2.400,00
Bandeja móvel para rack 19"	UFJF	3	R\$ 900,00
Disco Rígido para servidor Dell SAS 300Gb 12Gbps 15k RPM	UFJF	1	R\$ 3.300,00
Memória Kingston 8GB, 1600MHz, DDR3	UFJF	4	R\$ 1.000,00
<b>Subtotal</b>			R\$24.957,29

## 11. Referências

- Barrère, E., Souza, J., Vitor, M.A. & Almeida, M.A. (2018) Recomendação automática de videoaulas no Moodle. In: XXIX Simpósio Brasileiro de Informática na Educação (Brazilian Symposium on Computers in Education).
- Dias, L., Barrere, E., Siqueira, J., Souza, J.F. (2017). Uma Abordagem para Identificação de Similaridade Entre Recursos Educacionais Utilizando Bases de Conhecimento Externas. *Revista Brasileira de Informática na Educação*. 25 (2).
- Dias AD, Wives LK. Recommender system for learning objects based on the fusion of social signals, interests, and preferences of learner users in ubiquitous e-learning systems. *Personal and Ubiquitous Computing*. 2019 Apr 1;23(2):249-68.
- Ferreira, V. H. and Raabe, A. L. A. (2009). Lorsys: Um sistema de recomendação de objetos de aprendizagem SCORM para o Moodle. XX Simpósio Brasileiro de Informática na Educação.
- GOMES JR, J.; DIAS, L. L. ; SOARES, E. R. ; BARRÉRE, EDUARDO ; SOUZA, J. F. A framework for knowledge discovery in educational video repositories. *Computing and Informatics*, v. 38, p. 1375-1402, 2019a.
- GOMES JR, JORÃO ; FERRANTI, NICOLAS ; DE SOUZA, JAIRO FRANCISCO . Semantic Enrichment of Web Data for the Provision of an Unified Data Repository of Brazilian Missing Persons. In: the XV Brazilian Symposium, 2019, Aracaju. Proceedings of the XV Brazilian Symposium on Information Systems - SBSI'19. New York: ACM Press, 2019b. p. 1.
- Kim Y, Jung S, Ji S, Hwang E, Rho S. IoT-based personalized NIE content recommendation system. *Multimedia Tools and Applications*. 2019 Feb 1;78(3):3009-43.
- Neves, D. , Brandão, W. C., & Ishitani, L. (2017). Automatic content recommendation and aggregation according to SCORM. *Informatics in Education*, 16(2).
- Ribeiro, F. A. A., Fonseca, L. C. C., and de Sousa Freitas, M. (2013). Recomendando objetos de aprendizagem a partir das hashtags postadas no moodle. In *Simpósio Brasileiro de Informática na Educação*, volume 24, page 82.
- SOUZA, J. F.; GOMES JR, J. ; BARRERE, E. . Comparativo entre fontes de dados para anotação automática de videoaulas. In: *Simpósio Brasileiro de Informática na Educação*, 2017, Recife. *Anais do Simpósio Brasileiro de Informática na Educação*, 2017. p. 1127-1136.