



UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI
Programa de Pós-Graduação em Educação
Henrique Carlos Fonte Boa Carvalho

Categorização de vídeos educacionais do Youtube por meio de comentários

Diamantina
2021

Henrique Carlos Fonte Boa Carvalho

Categorização de vídeos educacionais do Youtube por meio de comentários

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Educação, como parte dos requisitos exigidos para a obtenção do título de Mestre em Educação.

Orientador: Cristiano Grijó Pitangui

Coorientador: Alessandro Vivas Andrade

Diamantina

2021

Catálogo na fonte - Sisbi/UFVJM

F682 Fonte Boa Carvalho, Henrique Carlos
2021 Categorização de vídeos educacionais do Youtube por meio de
comentários [manuscrito] / Henrique Carlos Fonte Boa Carvalho. --
Diamantina, 2021.
69 p. : il.

Orientador: Prof. Cristiano Grijó Pitangui.
Coorientador: Prof. Alessandro Vivas Andrade.

Dissertação (Mestrado Profissional em Educação) -- Universidade
Federal dos Vales do Jequitinhonha e Mucuri, Programa de Pós-
Graduação em Educação, Diamantina, 2021.

1. Categorização. 2. Educação. 3. Mineração de Texto. 4.
Objetos de Aprendizagem. 5. Youtube. I. Pitangui, Cristiano Grijó.
II. Andrade, Alessandro Vivas. III. Universidade Federal dos Vales
do Jequitinhonha e Mucuri. IV. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFVJM
com os dados fornecidos pelo(a) autor(a).
Bibliotecário Rodrigo Martins Cruz / CRB6-2886
Técnico em T.I. Thales Francisco Mota Carvalho



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI

HENRIQUE CARLOS FONTE BOA CARVALHO

Categorização de Vídeos Educacionais do Youtube por meio de Comentários

Dissertação apresentada ao Programa de Pós-Graduação em Educação da Universidade Federal dos Vales do Jequitinhonha e Mucuri, **nível de Mestrado**, como requisito parcial para obtenção do título de **Mestre em Educação**.

Orientador: Prof. **Cristiano Grijó Pitangui**

Data de aprovação 16/07/2021.

Prof. Cristiano Grijó Pitangui - (UFVJM)

Prof. Alessandro Vivas Andrade - (UFVJM)

Prof. Cristiano Maciel da Silva - (UFSJ)

Prof. Fabiano Azevedo Dorça - (UFU)

Profa Luciana Pereira Assis - (UFVJM)



Documento assinado eletronicamente por **Luciana Pereira de Assis, Servidor**, em 03/08/2021, às 17:34, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Alessandro Vivas Andrade, Servidor**, em 03/08/2021, às 17:36, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Cristiano Grijó Pitangui, Usuário Externo**, em 03/08/2021, às 18:04, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Fabiano Azevedo Dorça, Usuário Externo**, em 05/08/2021, às 14:43, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **CRISTIANO MACIEL DA SILVA, Usuário Externo**, em 31/08/2021, às 22:36, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufvjm.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0419918** e o código CRC **136F78F2**.

Referência: Processo nº 23086.005099/2021-71

SEI nº 0419918

Dedico esse trabalho a minha família, em especial a minha esposa e minha filha, Aline e Sabrina, pelo apoio, pelo amor, pela compreensão, pelo incentivo e pela inspiração.

AGRADECIMENTOS

Ao meu orientador, professor Cristiano Grijó Pitangui, pela dedicação, pela disponibilidade, pela paciência, pela confiança, por todo conhecimento compartilhado e pela excelente orientação.

Aos membros da banca, pelas críticas construtivas visando o enriquecimento do trabalho.

Aos amigos por todo apoio e confiança.

Aos colegas de trabalho pelo incentivo à ingressar no programa.

Aos colegas de mestrado pelas experiências compartilhadas, pela união e pelo incentivo.

À UFVJM por prestar um ensino superior de excelência.

I am a great believer in luck. The harder I work, the more of it I seem to have. (COX, 1922)

RESUMO

O *Youtube* foi lançado em 2005 com o objetivo de realizar compartilhamento de vídeos de diversos tipos, sendo possível encontrar: trailers; shows; filmes; entrevistas; educativos; entre outros, e desde o seu lançamento possui uma interface simples e intuitiva. A plataforma está em constante expansão sendo acessada mensalmente por cerca de 2 bilhões de usuários, recebendo, aproximadamente, 400 horas de vídeo por minuto. Do ponto de vista educacional, a plataforma disponibiliza Objetos de Aprendizagem de diversas áreas do conhecimento, por isso é acessada em larga escala por estudantes interessados nessas áreas e por outros públicos, totalizando 9 a cada 10 usuários no Brasil, com o objetivo de aprender algo novo. O *Youtube* também auxilia os aprendizes através da Recomendação de Objetos de Aprendizagem, fornecendo uma relação de vídeos retornados da expressão de busca utilizada. A plataforma realiza a categorização de seus vídeos, porém não é transparente quanto a isso e não possibilita a busca de vídeos por estas categorias. O presente trabalho foi realizado em etapas. Inicialmente, foi desenvolvido o Educavídeos, um sistema capaz de realizar buscas de vídeos através das categorias do *Youtube*. Resultados preliminares apontaram uma melhoria ao identificar vídeos educacionais por meio da categoria *education*. Posteriormente, foi realizada uma análise comparativa de buscas da ferramenta desenvolvida e do *Youtube*. Resultados demonstraram que apesar dos vídeos possuírem uma categoria, algumas foram erroneamente atribuídas. Neste sentido, tais resultados apontaram para a necessidade de melhor categorização dos vídeos da plataforma. Por fim, foram utilizadas técnicas de Mineração de Texto com a finalidade de identificar se os comentários postados em vídeos educacionais e não educacionais podem ser utilizados como parâmetro para categorizá-los. Resultados iniciais demonstram que existem diferenças significativas nos vocabulários das duas classes de vídeos, o que sugere que os comentários postados possuem potencial para categorização de vídeos educacionais.

Palavras-chave: Categorização. Educação. Mineração de Texto. Objetos de Aprendizagem. Youtube.

ABSTRACT

Youtube was launched in 2005 in order to share videos of different types, as trailers; shows; movies; interviews; educational; among others and since its launch it has a simple and intuitive interface. The platform is in constant expansion, being monthly accessed by approximately 2 billion users and receiving about 400 hours of video per minute. From an educational point of view, the platform provides Learning Objects of different knowledge areas, reason why it is accessed on a large scale by students and by other people, counting 9 out of 10 users in Brazil, aiming at learning something new. The Youtube also supports learners through the Recommendation of Learning Objects, providing a list of videos returned from the search expression used. The platform categorizes its videos, but it is not transparent about it and it does not allow you to search for videos through these categories either. The present work was done in stages. Initially, Educavídeos was developed, a system to search videos through Youtube categories. Preliminary results pointed to an improvement in identifying educational videos through the Education category. Afterwards a comparative analysis of searches of the developed tool and Youtube was performed. Results showed that although videos have a category, some of them have been wrongly assigned to videos. Therefore, such results showed the need to better categorize the platform's videos. Finally, Text Mining techniques were used to identify whether comments posted on educational and non-educational videos could be used as a parameter to categorize them. Initial results demonstrate that there are significant differences in the vocabularies of the two classes of videos, which suggests that the comments posted on the videos may be used for educational videos' categorization.

Keywords: Categorization. Education. Learning Objects. Text Mining. Youtube.

LISTA DE ILUSTRAÇÕES

Figura 1 – Categoria <i>Education</i>	33
Figura 2 – Funcionamento básico do sistema EducaVídeo	34
Figura 3 – Tela inicial do EducaVídeo	35
Figura 4 – Tela de Pesquisa do EducaVídeo	35
Figura 5 – Retorno da API do Youtube para a categoria <i>Education</i>	47
Figura 6 – Funcionamento básico do aplicativo de busca por API.	47
Figura 7 – Metodologia utilizada.	61

LISTA DE TABELAS

Tabela 1	– Pesquisa em modo padrão.	36
Tabela 2	– Pesquisa utilizando a categoria Education.	37
Tabela 3	– Resultados da busca pelo <i>Google Chrome</i> (modo anônimo).	48
Tabela 4	– Resultados da busca pelo aplicativo.	49
Tabela 5	– Resultados da busca pelo aplicativo e <i>categoryID = Education</i>	50
Tabela 6	– Termos mais frequentes nos comentários dos vídeos.	63
Tabela 7	– Vocábulo mais frequentes nos comentários dos vídeos.	63

SUMÁRIO

1	INTRODUÇÃO	21
	REFERÊNCIAS	25
2	EDUCAVÍDEOS: UM SISTEMA DE RECOMENDAÇÃO DE OBJETOS DE APRENDIZAGEM DE VÍDEOS EDUCACIONAIS DO YOUTUBE	27
2.1	Introdução	29
2.2	Desenvolvimento	30
2.2.1	<i>Trabalhos Relacionados</i>	30
2.2.2	<i>Metodologia</i>	32
2.2.3	<i>Resultados Preliminares</i>	34
2.3	Considerações Finais e Trabalhos Futuros	35
	REFERÊNCIAS	39
3	LEARNING OBJECTS AND YOUTUBE: AN ANALYSIS OF VIDEOS AND THEIR CATEGORIES	41
3.1	Introdução	43
3.2	Trabalhos Relacionados	44
3.3	Metodologia	46
3.4	Resultados	47
3.5	Considerações Finais e Trabalhos Futuros	51
	REFERÊNCIAS	53
4	CATEGORIZAÇÃO DE VÍDEOS EDUCACIONAIS DO YOUTUBE POR MEIO DE COMENTÁRIOS	55
4.1	Introdução	57
4.2	Trabalhos Relacionados	58
4.3	Metodologia	61
4.4	Resultados	62
4.5	Considerações Finais e Trabalhos Futuros	64
	REFERÊNCIAS	65
5	CONSIDERAÇÕES FINAIS	67

1 INTRODUÇÃO

Em 2005 Chad Hurley, Steve Chen e Jawed Karim desenvolveram um sistema focado em realizar o compartilhamento de vídeos, o sistema foi denominado como *Youtube*. Desde o seu lançamento o sistema se apresentava amigável para ser utilizados por todos os tipos de usuários, possibilitando que pessoas sem elevado conhecimento técnico pudessem fazer *upload* e assistir vídeos.

Atualmente, o *Youtube* é uma das maiores plataformas de vídeo da Internet, possuindo grande acervo em expansão, que recebe mais de 400 horas de vídeos a cada minuto (HEILBRON, 2019). Além disso, a plataforma é acessada mensalmente por mais de 2 bilhões de usuários (YOUTUBE, 2019b).

No âmbito brasileiro a plataforma é altamente utilizada, sendo acessada pelo menos uma vez ao mês por 95% da população conectada a Internet (YOUTUBE, 2017). Ademais, 9 em cada 10 usuários do sistema o acessam com o objetivo de adquirir conhecimento sobre algo novo e mais da metade acredita que tudo que desejam ver e aprender será encontrado no sistema (YOUTUBE, 2019a).

Seu acervo é composto por vídeos de diversos tipos sendo possível encontrar: trailers; shows; filmes; entrevistas; educativos; entre outros. Pode ser entendido como um repositório de Objetos de Aprendizagem (OA) quando o analisamos sob o ponto de vista educacional. Wiley (2000) define os OAs como “qualquer recurso digital que possa ser reutilizado com o intuito de auxiliar na educação, desde que possa ser entregue pela rede, como imagens, vídeos, animações, textos e outros”.

A oferta de vídeos do *Youtube* é realizada por meio do seu motor de busca, busca-se por um conteúdo e a plataforma apresenta os resultados adequados ao termo utilizado. O fornecimento desses resultados, na área da educação, faz parte do campo de estudo que trabalha com a Recomendação de Objetos de Aprendizagem. A Recomendação de Objetos de Aprendizagem é utilizada para identificar usuários, armazenar suas preferências e recomendar itens que podem ser produtos, serviços e/ou conteúdos, de acordo com suas necessidades e interesses (PONTES *et al.*, 2014).

Apesar de possuir um extenso e diversificado acervo, identifica-se alguns problemas relacionados ao seu mecanismo de busca. Três problemas relacionados ao seu motor de busca foram percebidos.

O primeiro problema refere-se aos resultados fornecidos pela plataforma a partir de uma busca, pois, em diversos casos, a quantidade de resultados é alta e muitas vezes uma grande quantidade é de baixa qualidade e/ou não muito relacionados a busca realizada.

A segunda dificuldade está relacionada pela metodologia utilizada para executar a pesquisa. Basicamente, a busca é realizada através da comparação dos termos da pesquisa com os dados do vídeo, a saber: título, descrição e *tags*. Assim sendo, os resultados apresentados serão os vídeos que contenham o(s) termo(s) (da expressão de pesquisa) nesses seus dados. No caso da expressão buscada conter mais de um termo, a plataforma retorna os vídeos que con-

templem, em seus dados, todos os termos da expressão. Por exemplo, os vídeos apresentados ao se utilizar a expressão de pesquisa “Futebol”, serão aqueles que, em seu título, descrição, ou *tags*, incluam o termo “Futebol”. De maneira similar, caso a busca seja realizada pela expressão “Futebol educação”, serão retornados vídeos que contenham ambos termos em seus dados.

O motor de busca do *Youtube* pode ser considerado bastante restritivo, quando a expressão de pesquisa possuir mais de um termo, quando é capaz de desconsiderar diversos vídeos que podem ser de interesse ao usuário e que possuam, apenas um ou parte do(s) termo(s) da busca, nos seus dados.

Por fim, outro problema encontrado está relacionado a categorização dos vídeos do *Youtube*. A plataforma não possibilita realizar buscas por vídeos segundo as suas categorias. De fato, o *Youtube* categoriza seus vídeos, porém, não proporciona métodos de realizar buscas por elas, e falta transparência quando a categorização, pois, a maioria dos usuários não tem conhecimento sobre essa categorização. Além disso, apresenta um elevado erro quando a categorização de seus vídeos, principalmente os educacionais.

Este trabalho realiza uma análise dos OAs do *Youtube*, demonstrando os problemas apresentados anteriormente e, adicionalmente, propõe maneiras de sobrepujá-los através da utilização dos comentários para a identificação. O presente trabalho se organiza como segue.

O capítulo 2 apresenta o artigo “EducaVídeo: Sistema de Recomendação de Objetos de Aprendizagem dos vídeos educacionais do *Youtube*” publicado no XVII Congresso Brasileiro de Ensino Superior a Distância (ESUD 2020) nele é introduzido o sistema Educavídeos, um sistema capaz de realizar buscas pelas categorias do *Youtube*, objetivando que o sistema possua melhores resultados para a busca de OA da plataforma através do uso da categoria “education”. Em seus resultados preliminares é demonstrado que a busca pelo sistema Educavídeos através do uso da categoria “education” fornece mais vídeos de conteúdo educacional do que a busca padrão (CARVALHO *et al.*, 2020a).

O capítulo 3 apresenta o artigo “Learning Objects and Youtube: an analysis of videos and their categories” publicado no XV Latin American Conference on Learning Technologies (LACLO 2020), é realizado a análise do mecanismo de busca dos vídeos de educação da plataforma, buscando também verificar se a categorização de vídeos da própria plataforma é satisfatória para identificação de vídeos. Observa-se que a plataforma apesar de possuir o metadado categoria nos vídeos, diversos vídeos estão categorizados erroneamente, demonstrando, desta maneira, que as categorias fornecidas pela plataforma não devem ser consideradas como determinantes para a categorização dos vídeos (CARVALHO *et al.*, 2020c).

O capítulo 4 apresenta o artigo “Categorização de vídeos educacionais do *Youtube* por meio de comentários” publicado na Revista Novas Tecnologias na Educação (RENTE). Neste trabalho é realizada a análise e coleta dos comentários de de 200 vídeos, sendo 100 educacionais e 100 não educacionais. Seus resultados mostram diferenças significativas entre os vocabulários presentes nos comentários dos vídeos educacionais e não educacionais, demonstrando

que os comentários possuem potencial para serem utilizados como método para categorização de vídeos educacionais da plataforma *Youtube* (CARVALHO *et al.*, 2020b).

Por fim, o capítulo 5 apresenta as considerações finais e apontamentos a trabalhos futuros.

REFERÊNCIAS

- ABU-EL-HAIJA, S.; KOTHARI, N.; LEE, J.; NATSEV, P.; TODERICI, G.; VARADARAJAN, B.; VIJAYANARASIMHAN, S. Youtube-8m: A large-scale video classification benchmark. **arXiv preprint arXiv:1609.08675**, 2016.
- ADERMANN, N.; BOGGIANO, J. Composer: Dependency manager for php. **URL: <https://getcomposer.org/>**, 2020. Último acesso em 16 de junho de 2020.
- BRAGA, J.; MENEZES, L. **Objetos de aprendizagem, volume 1: introdução e fundamentos**. UFABC, 2014. v. 1. 145 p. Disponível em: <<http://proec.ufabc.edu.br/uab/metdesOA2/2014-BRAGA-livro-oa-v1.pdf>>.
- BURGESS, J.; GREEN, J. **YouTube e a revolução digital**. [S.l.: s.n.], 2009. 24 p.
- CARVALHO, H. C. F. B.; PITANGUI, C. G.; ASSIS, L.; VIVAS, A. Educavídeos: Um sistema de recomendação de objetos de aprendizagem de vídeos educacionais do youtube. In: CONGRESSO BRASILEIRO DE ENSINO SUPERIOR A DISTÂNCIA, XVII., 2020, Goiânia. **2020 XVII Congresso Brasileiro de Ensino Superior a Distância**. Goiânia, 2020.
- CARVALHO, H. C. F. B.; PITANGUI, C. G.; TRINDADE, E. A. C.; ASSIS, L. P. de; ANDRADE, A. V.; SOUZA, D. P. B. de. Categorização de vídeos educacionais do youtube por meio de comentários. **RENOTE**, v. 18, n. 2, p. 621–629, 2020.
- CARVALHO, H. F. B.; PITANGUI, C. G.; TRINDADE, E. C.; ASSIS, L. P. de; ANDRADE, A. V. Learning objects and youtube: an analysis of videos and their categories. In: IEEE. **2020 XV Conferencia Latinoamericana de Tecnologias de Aprendizaje (LACLO)**. [S.l.], 2020. p. 1–6.
- CAZELLA, S. C.; BEHAR, P.; SCHNEIDER, D.; SILVA, K. K. da; FREITAS, R. Desenvolvendo um sistema de recomendação de objetos de aprendizagem baseado em competências para a educação: relato de experiências. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. [S.l.: s.n.], 2012. v. 23, n. 1.
- COX, C. **Listen to this**. [S.l.]: C. Cox Publishing Company, 1922.
- GOMES, L. Vídeos didáticos: uma proposta de critérios para análise. **Revista Brasileira de Estudos Pedagógicos**, v. 89, n. 223, 2008.
- HEILBRON, F. D. C. **Understanding Human Activities at Large Scale**. Tese (Doutorado), 2019.
- IEEE. Ieee learning technology standards committee (ltsc): Ieee p1484.12.1 ieee standard for learning object metadata. 2002. Acessado em: 27 de março de 2019. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1032843>>.
- JÚNIOR, C. B.; DORÇA, F. **Uma Abordagem para a Criação e Recomendação de Objetos de Aprendizagem usando um Algoritmo Genético, Tecnologias da Web Semântica e uma Ontologia**. 2018. 1533 p.
- LERDORF, R. Php: O que é o php? **URL: https://www.php.net/manual/pt_BR/intro-what-is.php**, 2020. Último acesso em 16 de junho de 2020.

MENOLLI, A.; MALUCELLI, A.; REINEHR, S. Criação semi-automática de objetos de aprendizagem a partir de conteúdos da wiki. In: BRAZILIAN SYMPOSIUM ON COMPUTERS IN EDUCATION (SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO-SBIE), XXII., 2011, Aracaju. **Anais**. Aracaju, 2011.

MIRANDA, R. M. d. **GROA: um gerenciador de repositórios de objetos de aprendizagem**. Dissertação (mathesis), 2004.

MORAIS, E. A. M.; AMBRÓSIO, A. P. L. Mineração de textos. **Relatório Técnico–Instituto de Informática (UFG)**, 2007.

NASCIMENTO, P. do; BARRETO, R.; PRIMO, T.; GUSMÃO, T.; OLIVEIRA, E. **Recomendação de objetos de aprendizagem baseada em modelos de estilos de aprendizagem: Uma revisão sistemática da literatura**. 2017. 213 p.

OTTO, M.; THORNTON, J. Bootstrap. URL: <https://getbootstrap.com/>, 2020. Último acesso em 16 de junho de 2020.

PINHEIRO, R. R. A. *et al.* Sistema de recomendação de vídeos educacionais: um estudo de caso no you tube. Universidade Federal de Alagoas, 2018.

PONTES, W. L.; FRANÇA, R. M.; COSTA, A. P. M.; BEHAR, P. **Filtragens de Recomendação de Objetos de Aprendizagem: uma revisão sistemática do CBIE**. 2014. 549 p.

RESIG, J. Jquery. URL: <https://jquery.com/>, 2020. Último acesso em 16 de junho de 2020.

SORGALLA, J. Lity: lightweight, accessible and responsive lightbox. URL: <https://sorgalla.com/lity/>, 2020. Último acesso em 16 de junho de 2020.

THELWALL, M. Social media analytics for youtube comments: Potential and limitations. **International Journal of Social Research Methodology**, Taylor & Francis, v. 21, n. 3, p. 303–316, 2018.

VIEIRA, F. J. R.; NUNES, M. A. S. N. Dica: Sistema de recomendação de objetos de aprendizagem baseado em conteúdo. **Scientia Plena**, v. 8, n. 5, 2012.

WILEY, D. A. **Learning object design and sequencing theory**. Tese (Doutorado) — Brigham Young University, 2000.

YOUTUBE. **API Reference**. 2015. Acesso em: 08/05/2019. Disponível em: <<https://developers.google.com/youtube/v3/docs/?hl=pt-br>>.

YOUTUBE. **Youtube Insights 2017**. 2017. Acesso em: 16 de Abril de 2019. Disponível em: <<https://www.thinkwithgoogle.com/intl/pt-br/youtubeinsights/2017/introducao/>>.

YOUTUBE. **YouTube Insights**. 2019. Acesso em: 17 de Abril de 2019. Disponível em: <<https://youtubeinsights.withgoogle.com>>.

YOUTUBE. **Youtube para imprensa**. 2019. Acesso em: 16 de Abril de 2019. Disponível em: <<https://www.youtube.com/intl/pt-BR/yt/about/press/>>.

2 EDUCAVÍDEOS: UM SISTEMA DE RECOMENDAÇÃO DE OBJETOS DE APRENDIZAGEM DE VÍDEOS EDUCACIONAIS DO YOUTUBE

RESUMO

O Youtube é uma das maiores plataformas de vídeos da Internet - a cada minuto são enviadas 400 horas de vídeos, e 1,5 bilhão de pessoas acessam a plataforma mensalmente. Sob o ponto de vista educacional, a plataforma oferta diversos tipos de Objetos de Aprendizagem sobre o mais variados temas. De forma a organizar este vasto conteúdo, o Youtube utiliza categorias para a classificação de seus vídeos. Entretanto, essa categorização não pode ser utilizada pelo usuário para a realização de buscas mais específicas. O presente trabalho introduz o EducaVídeos, um sistema capaz de realizar buscas por vídeos no Youtube utilizando categorias. Resultados preliminares utilizando-se 3 expressões de busca e a categoria Education apontam um melhoria na qualidade de retorno da busca quando se utiliza a categoria do vídeo.

Palavras-chave: Objetos de Aprendizagem. Recomendação. Youtube. Categorização.

ABSTRACT

Youtube is one of the largest online video-sharing platforms on the Internet – at each minute 400 hours of videos are uploaded, and 1.5 billion people access the platform monthly. From an educational point of view, the platform offers several types of Learning Objects over different topics. In order to organize this vast content, Youtube classifies its' videos in some categories. However, this categorization cannot be used by the user to perform specific searches. The present work introduces EducaVideos, a system capable of searching Youtube videos using categories. Preliminary results using 3 search expressions and the Education category point to an improvement in the quality of search result when using the video category.

Keywords: Learning Objects. Recommendation. Youtube. Categorization.

2.1 Introdução

No final de Junho de 2005 foi lançada a plataforma *Youtube*, desenvolvida por Chad Hurley, Steve Chen e Jawed Karim, com o objetivo de compartilhar vídeos na Web. Burgess e Green (2009) afirmam que desde o seu lançamento, o sistema se fez simples e possuía uma interface intuitiva e simplificada. A plataforma possibilita que seus usuários possam assistirem e publicarem vídeos sem a necessidade de elevado conhecimento técnico.

A plataforma está voltada para vídeos e possui conteúdos sobre diversas áreas do conhecimento. Segundo Heilbron (2019), a cada minuto são enviadas 400 horas de vídeos, e mais de 1,9 bilhões de usuários acessam a plataforma mensalmente (YOUTUBE, 2019b). Segundo Youtube (2017), 1,5 bilhão de pessoas conectadas pelo mundo acessam o *Youtube* mensalmente e 95% da população brasileira conectada acessa a plataforma ao menos uma vez por mês. No Brasil, 9 em cada 10 usuários do *Youtube* acessam a plataforma com o intuito de aprender algo novo e mais da metade acredita que é o lugar onde encontra-se tudo o que deseja-se ver e aprender (YOUTUBE, 2019a).

O *Youtube*, sob o ponto de vista educacional, pode ser entendido como um repositório de Objetos de Aprendizagem (OA). De forma geral, os OAs podem ser definidos como “qualquer recurso digital que possa ser reutilizado com o intuito de auxiliar na educação, desde que possa ser entregue pela rede, como imagens, vídeos, animações, textos e outros” (WILEY, 2000).

Neste sentido, a oferta de OA é realizada pela plataforma por meio do seu mecanismo de busca, i.e., busca-se por um conteúdo e a plataforma apresenta os OA condizentes à expressão de busca utilizada. Na área da Educação, essa “assistência” faz parte da área de estudos que aborda a Recomendação de Objetos de Aprendizagem. Segundo Pontes *et al.* (2014), a Recomendação de Objetos de Aprendizagem é utilizada para identificar usuários, armazenar suas preferências e recomendar itens que podem ser produtos, serviços e/ou conteúdos, de acordo com suas necessidades e interesses.

Apesar de disponibilizar conteúdos sobre as mais diversas áreas e possuir um extenso acervo de vídeos, é possível identificar problemas relacionados ao mecanismo de busca da plataforma. De forma geral, identificaram-se dois problemas relacionados ao dispositivo de busca do *Youtube*.

O primeiro problema se relaciona aos resultados retornados pela plataforma ao se realizar uma busca, pois, em muitos casos, a quantidade de resultados retornados é muito alta, sendo muitos destes de baixa qualidade e/ou não muito relacionados à busca realizada. Outra dificuldade encontrada é a impossibilidade de se realizar buscas por vídeos segundo suas categorias. De fato, o *YouTube* categoriza seus vídeos, porém, não é possível realizar pesquisas por elas, e a maioria dos usuários não tem conhecimento sobre essa categorização.

Este trabalho introduz o sistema denominado EducaVídeos que busca maneiras de sobrepujar os problemas de busca do *Youtube* apontados anteriormente. O sistema foi desenvolvido para auxiliar a etapa de escolha de OAs, identificando e retornando vídeos na categoria

Education. Neste sentido, para esta classe específica, objetiva-se que o EducaVÍdeos apresente melhores resultados que o mecanismo de busca padrão do Youtube.

Resultados experimentais, realizados utilizando os termos de busca “Herança”, “Futebol”, e “Gastronomia”, apontam que o EducaVÍdeos apresenta melhores resultados quanto a quantidade e qualidade de vídeos educacionais quando comparado a pesquisa padrão da plataforma Youtube, i.e., o sistema proposto retorna um número maior de vídeos educacionais, mais condizentes com a expressão de busca, que a busca padrão do Youtube.

O presente trabalho se organiza como segue. A seção 2 apresenta, primeiramente, os principais trabalhos relacionados a esta pesquisa. Em seguida, apresenta-se o funcionamento geral do sistema EducaVÍdeos, bem como uma breve explicação de como se deu o seu desenvolvimento, e tecnologias utilizadas. Posteriormente apresentam-se os resultados preliminares obtidos. Por fim, a seção 3 apresenta as considerações finais e apontamentos a trabalhos futuros.

2.2 Desenvolvimento

2.2.1 Trabalhos Relacionados

Esta seção aborda os trabalhos relacionados a esta pesquisa. Destaca-se, entretanto, que poucos trabalhos foram identificados referentes a categorização e recomendação de vídeos educacionais no *Youtube*.

Menolli, Malucelli e Reinehr (2011) objetivam gerar OAs com a Web 2.0, por meio da *Wikipedia*, utilizando tecnologias semânticas e o padrão *Learning Object Metadata* (LOM). Em sua proposta, acessam os conteúdos inseridos na plataforma, e realizam a mineração de textos a fim de extrair e classificar os conteúdos de acordo com o padrão LOM. Com a utilização desse padrão, encontram os atributos e metadados da página, gerando um *XML-schema* com os metadados da página em questão. Concluem que essa abordagem é necessária, pois facilita a utilização dos conteúdos devido as ferramentas wikis não considerarem como o conteúdo será utilizado.

Cazella *et al.* (2012) apresentam um relato de experiência de desenvolvimento e validação de um Sistema de Recomendação de OAs baseado em competências. Apresentam o modelo denominado RECoaComp (REComendador de OAs baseado em Competências). O sistema funciona em três etapas, a saber: 1) o professor informa quais OAs ele utilizará em uma disciplina; 2) o aluno interage com o sistema, viabilizando as recomendações; 3) o sistema realiza uma filtragem de informação, recomendando os OAs. Concluem que o sistema realiza as recomendações que melhor atendem às necessidades individuais, e destacam que o sistema se apresenta robusto, eficiente, e de fácil utilização.

Vieira e Nunes (2012) desenvolveram um Sistema de Recomendação de OAs denominado DICA, que utiliza o Banco Internacional de Objetos Educacionais (BIOE). O desenvolvimento do sistema ocorreu em 3 etapas, a saber: 1) coleta de informações - desenvolveu-se o modelo dos dados da plataforma; 2) processamento de dados - utilizando técnica de agrupamento (*K-means*), e de vizinhos mais próximos (*KNN*), encontraram um valor para distancia dos

temas, possibilitando agrupamento de OAs de temas diferentes; 3) recomendação - desenvolvimento do Sistema de Recomendação. Concluíram que o sistema foi bem aceito, sendo muitas vezes preferido ao BIOE, pois apresenta a possibilidade de facilitar as pesquisas por temas.

Abu-El-Haija *et al.* (2016) abordam a classificação de vídeos do *Youtube* com o objetivo de desenvolver um sistema de multi-classificação de vídeos. A base de dados utilizada possui aproximadamente 8 milhões de vídeos englobando um total de 1,9 bilhão de quadros de vídeos, e 500 mil horas de vídeos categorizados. A pesquisa é realizada em duas etapas, a saber: 1) os rótulos dos vídeos são obtidos por meio do *Knowledge Graph entities*; 2) os vídeos são processados *frame a frame*, e categorizados por uma Rede Neural Convolutiva pré-treinada no ImageNet. O ImageNet é um banco de dados visual com diversos objetos/entidades já classificados. Por meio do processamento de mais de 50 anos de vídeos, provendo 2 bilhões de *frames*, e mais de 8 milhões de vídeos que podem ser modelados rapidamente em uma única máquina, o trabalho aponta sua contribuição no sentido de auxiliar o desenvolvimento de pesquisas sobre compreensão de vídeos.

Apesar da categorização em diversas classes, não foi encontrada uma categoria específica para vídeos educacionais. O trabalho cita a categoria “*Jobs & Education*” e nela estão universidades, salas de aulas, palestras, etc. Desta forma, um vídeo que possua imagens de um *campus* universitário, por exemplo, será enquadrado nessa categoria, e não necessariamente é um vídeo educacional.

Júnior e Dorça (2018) apresentam uma abordagem para criação e recomendação de OAs por meio da plataforma de textos *Wikipédia*. A abordagem é definida por três etapas, a saber: 1) enriquecimento da ontologia através dos metadados das seções wiki; 2) a recomendação dos OAs - são utilizadas técnicas de Problema de Cobertura de Conjuntos com a utilização de Algoritmo Genético; 3) uso de operações CRUD (*Creat, Read, Update, Delete*). O trabalho conclui que a abordagem adotada resolve o problema da recomendação de OAs, retornando soluções de elevada qualidade.

Pinheiro *et al.* (2018) apresentam o *Easy Youtube*, um Sistema de Recomendação de OAs baseado no Youtube. O funcionamento do sistema se faz em seis etapas, a saber: 1) enriquecimento de consultas - definição de temas pré-definidos, cadastrados por especialistas; 2) extração de vídeos - busca de vídeos, que pode ser realizada por meio de pesquisa ou de temas pré-definidos; 3) pré-processamento - tratamento dos textos (em português), com remoção de pontuação, espaços, etc.; 4) classificação - utilização de um algoritmo para classificar os vídeos considerados como educacionais e de qualidade; 5) engenho de recomendação - o sistema recebe os vídeos considerados “bons” e classifica-os; 6) coletor de *feedback* - o usuário avalia, por meio de notas (de 1 a 7 estrelas), a recomendação fornecida pelo sistema. O trabalho aponta suas principais contribuições nos pontos: 1) o Sistema de Recomendação desenvolvido pode ser utilizado como solução para vários domínios de aplicação; 2) o sistema serviu como prova de conceito para melhorar as recomendações, através de características do Youtube, como a avaliações dos usuários, e linguagem nativa do vídeo.

O trabalho apresentado não detalha partes importantes de pesquisa. Por exemplo, para a classificação de vídeos considerados de qualidade, afirma-se que se utilizou um conjunto de treinamento de 100 vídeos, contendo avaliações realizadas por especialistas e alunos que já tiveram em sua grade curricular o assunto “Orientação a Objetos/Herança”. Porém, não se explica de que maneira esta análise foi realizada, e quais características dos vídeos foram consideradas. Outro ponto que causa confusão é o apontamento de que, devido ao prazo para a realização da pesquisa, o trabalho focou em “algumas características para o experimento”. Tais características não foram descritas.

Thelwall (2018) analisa os comentários de vídeos do *Youtube* relacionados a estilos de dança. A base de dados utilizada contém 36.702 vídeos. O trabalho objetiva identificar, por meio dos comentários postados nos vídeos da plataforma, os tipos de dança, relações quanto aos gêneros (masculino e feminino), sentimentos expressados, e discussões referente aos estilos. Utiliza-se, para tanto, o método denominado *Comment Term Frequency Comparison* (CTFC) na tentativa de identificação de subtópicos/subtemas das discussões sobre um tópico nos comentários do *Youtube*, questões de gênero, sentimentos, e relacionamento entre tópicos. O método utilizado define com sucesso diversas atitudes predominantes em mulheres ou homens. Os 10 termos homem-associados foram: *shit, fuck, shuffle, man, fucking, cripp, dude, bro, shuffling, hardstyle*. Por sua vez, os 10 termos mulher-associados foram: *she, amazing, her, beautiful, cute, omg, belly, ballet, really, workout*. A análise de sentimentos forneceu ideias plausíveis dos motivos pelo qual as danças eram apreciadas. Os 10 termos positivos mais utilizados foram: *please, nice, wow, beautiful, loved, job (e.g. nice/great/good job), pretty, hope, perfect, keep (going/up the good work/it up)*. Por sua vez, os 10 termos negativos mais utilizados foram: *shit, fuck, killed, stupid, wtf, hate, idiot, dislike, die, dead*.

2.2.2 Metodologia

Inicialmente realizou-se um estudo da Interface de Programação de Aplicativos(API) da plataforma *Youtube* para verificar as funcionalidades disponíveis e para a identificação dos vídeos e seus dados. A API permite que funções da própria plataforma possam ser incorporadas e utilizadas em outro site ou aplicativo. Entretanto, alguns requisitos são necessário para a sua utilização (YOUTUBE, 2015). São eles:

- Cada solicitação deverá especificar uma chave de API (com o parâmetro *key*);
- Deve-se enviar um código de autorização para cada solicitação de inserção, atualização e exclusão.

A API trabalha por meio de cotas que servem para garantir que os desenvolvedores utilizem o serviço de maneira correta sem acarretar diminuição da qualidade do serviço e sem limitar o acesso a outros usuários. O usuário possui um limite de 10 mil cotas diárias. Nota-se, entretanto, que uma requisição pode utilizar mais de uma cota. Por exemplo, uma requisição para listar vídeos pode gastar 21 unidades de cota (YOUTUBE, 2015).

A API apresenta um limite de 50 valores retornados por requisição do usuário. Por meio da API, é possível identificar alguns metadados de cada vídeo. Os metadados fornecidos pela API que podem ser utilizados pelo usuário são:

- Título: título do vídeo;
- ID: identificação de um vídeo da plataforma. É encontrada no final do endereço, como: `https://www.youtube.com/watch?v=ID`;
- Categoria: definida como **categoryId**, apresenta em quais categorias o vídeo classifica-se. Algumas categorias de vídeos do *Youtube* são: *Education, People & Blogs, Sports, Music*, dentre outras;
- Visualizações: número de vezes que o vídeo já foi visto;
- *Likes*: número de pessoas que clicaram no botão “Gostei”, sinal de aprovação do vídeo;
- *Dislikes*: número de pessoas que clicaram no botão “Não Gostei”, sinal de reprovação do vídeo;
- Número de comentários: quantidade de comentários do vídeo.

A figura 1 apresenta o retorno da API para a categoria *Education*.

Figura 1 – Categoria *Education*

```
{
  "kind": "youtube#videoCategory",
  "etag": "hHUhloYhyMMqVkJQ4LH9sr3aRQoU",
  "id": "27",
  "snippet": {
    "title": "Education",
    "assignable": true,
    "channelId": "UCBR8-60-B28hp2BmDPdntcQ"
  }
},
```

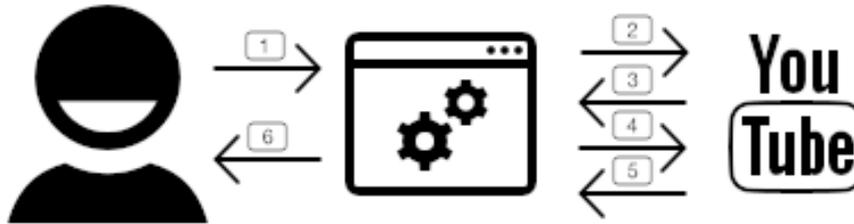
Fonte: Elaborado pelos autores

A figura 2 apresenta o funcionamento preliminar do EducaVídeo. O sistema funciona basicamente em 6 etapas, a saber:

1. O usuário acessa o sistema e preenche os dados para pesquisa;
2. O sistema envia a solicitação para a API;
3. A API responde a solicitação com os dados: Título e ID;
4. O sistema realiza nova busca na API para obter as demais informações dos vídeos (categoria, visualizações, *likes*, *dislikes* e número de comentários);
5. A API retorna os dados solicitados;
6. O sistema processa os dados recebidos e envia para o usuário

O EducaVídeo, foi desenvolvido com PHP, CSS, HTML e *Javascript*, com a utilização das bibliotecas *Bootstrap, JQuery, Lity* e o gerenciador de pacotes *Composer*.

Figura 2 – Funcionamento básico do sistema EducaVídeo



Fonte: Elaborado pelos autores

O PHP é uma linguagem aberta (*open source*), utilizada principalmente para o desenvolvimento Web, pois pode ser embutida dentro do HTML. O PHP é simples e oferece recursos avançados tanto para programadores iniciantes como profissionais (LERDORF, 2020).

O *Composer* é um gerenciador de pacotes para a linguagem PHP. Ele fornece um formato padrão para gerenciar bibliotecas e dependências para um projeto. Foi desenvolvido por Nils Adermann e Jordi Boggiano, que continuam a gerenciar o projeto (ADERMANN; BOGGIANO, 2020).

O *Bootstrap* é um *framework Web* com código-fonte aberto para desenvolvimento de componentes de interface e *front-end* para sites e aplicações web usando HTML, CSS e *JavaScript* (OTTO; THORNTON, 2020).

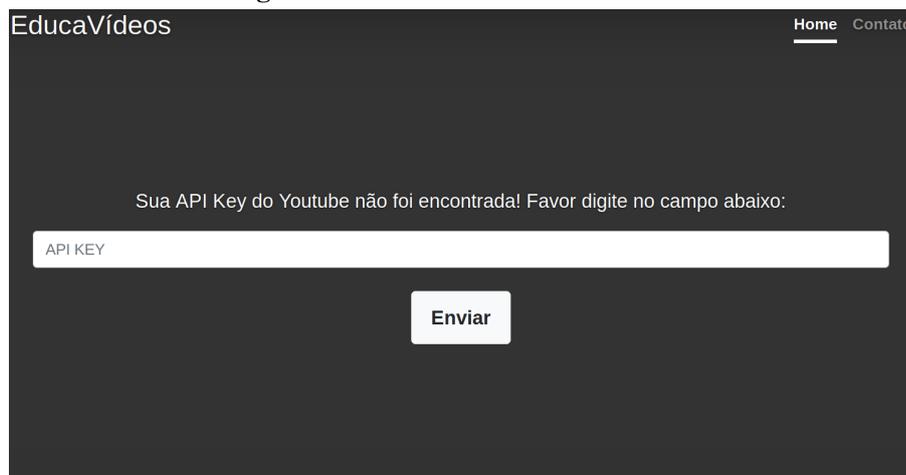
A utilização do *Composer* possibilitou a instalação do *JQuery*, uma biblioteca de funções *JavaScript* que interage com o HTML, desenvolvida para simplificar os *scripts* interpretados no navegador do cliente. Utilizada por cerca de 74.4% dos 10 mil sites mais visitados do mundo, *JQuery* é a mais popular das bibliotecas *JavaScript* (RESIG, 2020). Utilizou-se o *Lity* por ser um um *plug-in lightbox* ultraleve, acessível e responsivo que suporta imagens, iframes e conteúdo embutido (SORGALLA, 2020).

2.2.3 Resultados Preliminares

A figura 3 apresenta a tela inicial do EducaVídeo onde é necessário que o usuário forneça a sua chave (API Key) do *Youtube*. Após fornecida a chave, é permitido que o usuário realize as pesquisas. Caso a chave esteja incorreta, o sistema retorna para a tela inicial e solicita novamente a chave.

A figura 4 apresenta a tela de pesquisa. Nessa etapa o usuário: 1) digita a expressão de busca; 2) escolhe a número de vídeos a serem retornados pela busca; 3) clica no botão pesquisar.

Realizaram-se 3 experimentos com expressões de buscas distintas, a saber: “Herança”, “Futebol” e “Gastronomia”. Cada experimento foi realizado de duas formas: 1) utilizado-se a busca padrão do Youtube; 2) utilizado-se a busca do Youtube por vídeos classificados na categoria Education. A Tabela 1 apresenta os 5 primeiros resultados da busca, respectivamente, para as expressões “Herança”, “Futebol” e “Gastronomia”, quando a busca padrão é utilizada, enquanto a Tabela 2 apresenta os 5 primeiros resultados para cada expressão de busca com a utilização da categoria *Education*.

Figura 3 – Tela inicial do EducaVídeo

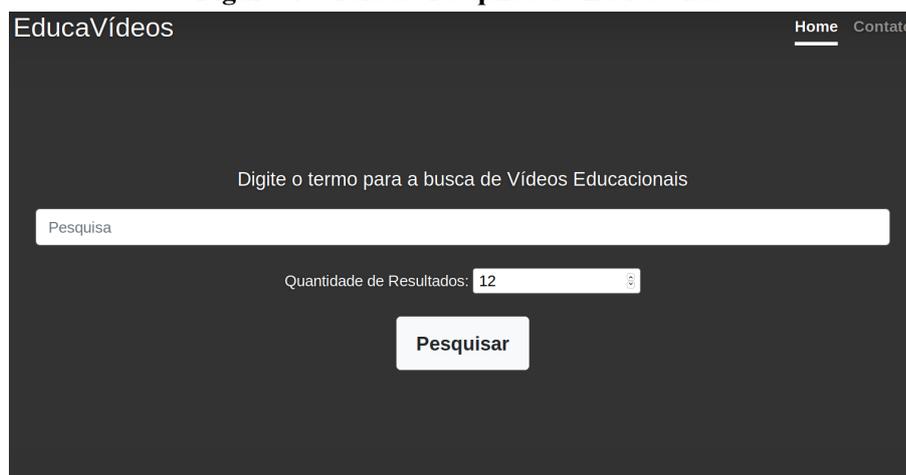
EducaVídeos Home Contato

Sua API Key do Youtube não foi encontrada! Favor digite no campo abaixo:

API KEY

Enviar

Fonte: Elaborado pelos autores

Figura 4 – Tela de Pesquisa do EducaVídeo

EducaVídeos Home Contato

Digite o termo para a busca de Vídeos Educacionais

Pesquisa

Quantidade de Resultados: 12

Pesquisar

Fonte: Elaborado pelos autores

De maneira geral, observa-se que o sistema apresenta melhores resultados quando a categoria Education é utilizada. Para a pesquisa pelo termo “Herança”, utilizando-se a categoria Education, o sistema retorna 4 vídeos educacionais e 1 vídeo de música, enquanto que a busca padrão retorna 4 vídeos de música e 1 vídeo educacional (apesar deste de estar classificado como *News Politics*). Para o termo “Futebol”, utilizando-se a categoria Education, o sistema retorna 1 vídeo educacional e 4 vídeos não educacionais, enquanto que a busca padrão retorna 5 vídeos não educacionais. Finalmente, para o termo “Gastronomia”, utilizando-se a categoria Education, o sistema retorna 4 vídeos educacionais e 1 vídeo não educacional, enquanto que a busca padrão retorna 2 vídeos educacionais e 3 vídeos não educacionais.

2.3 Considerações Finais e Trabalhos Futuros

O Youtube é uma plataforma muito popular, com um crescente acervo de vídeos sobre os mais diversos temas. Porém, apesar do crescimento constante de conteúdos disponíveis, nem sempre os conteúdos sugeridos atendem a necessidade do usuário, demandando diversas

Tabela 1 – Pesquisa em modo padrão.

Título	idVideo	id - Categoria
Jozyanne- Herança (legendado)	OQGMGWv3_Z4	10 - Music
Drik Barbosa - Herança part. Anna Tréa (Álbum visual)	CtguPlt4_7s	10 - Music
HERANÇA Música para apresentação de Bebês	YPVyEaWI2fA	27 - Education
Tudo o que você deveria saber sobre herança... mas certamente não sabe	ITfdWxHc2tA	25 - News & Politics
Herança	MQFE3ecqq5w	10 - Music
Dribles Mais Humilhantes Da História Do Futebol	CWWOW7T82Hw	17 - Sports
GAROTADA VIRA JOGO NA RAÇA E CONQUISTA O TETRA MUNDIAL PARA NOSSA SELEÇÃO HD 17/11/2019	piJycdsGDEI	10 - Music
DESAFIO DOS GOLS DE FALTA IMPOSSÍVEIS NO FUTEBOL	aaIZk0Bh150	17 - Sports
🎵 LEI DO RETORNO (Me Dibrrou, Vou Te Dibrar) Paródia de Futebol - MC Don Juan e MC Hariel	4ExGMuMMnVk	17 - Sports
DESAFIO DOS GOLAÇOS DE CURVA NO FUTEBOL	iBytohoQG2c	17 - Sports
TIRANDO DÚVIDAS SOBRE O CURSO DE GASTRONOMIA- TUDO SOBRE MEU CURSO	uHRCa61wOWw	22 - People & Blogs
Como é a faculdade de Gastronomia? Stephanie Ferreira	PJQ4ehx4Oqk	22 - People & Blogs
1º Aula de Gastronomia Utensílios de Cozinha, Tomates e Molhos / De Volta aos 30 - 19/09/2016	XkmCN7oBorw	27 - Education
COZINHEIRO 02 de 04 - Gastronomia - Cursos completos Grátis Agora	kfpaeBQY43M	27 - Education
AULA DE GASTRONOMIA	yNNsyXzZIEM	24 - Entertainment

Fonte: Elaborado pelo autor, com base na pesquisa realizada

visualizações e/ou extensivas buscas. Devido a esses problemas, um estudo específico sobre o mecanismo de busca desta plataforma torna-se necessário.

Por meio do estudo realizado, verificou-se que a plataforma Youtube não permite que sejam realizadas buscas por vídeos segundo suas categorias. De fato, o Youtube categoriza seus vídeos, porém, não é possível realizar pesquisas por categorias, e a maioria dos usuários não tem conhecimento sobre essa categorização.

O presente trabalho introduziu o EducaVídeos, um sistema capaz de realizar buscas no por vídeos no Youtube utilizando categorias. Resultados preliminares utilizando-se 3 expressões de busca e a categoria Education apontam um melhoria na qualidade de retorno da busca quando se utiliza a categoria do vídeo.

Para a continuidade desta pesquisa, analisar-se-á a qualidade das recomendações do EducaVídeos, assim como as recomendações do Youtube. Estudos sobre a categorização em outras classes e buscas pelas mesmas estão em desenvolvimento.

Tabela 2 – Pesquisa utilizando a categoria Education.

Título	idVideo	id - Categoria
HERANÇA Música para apresentação de Bebês	YPVyEaWI2fA	27 - Education
HERANÇA LIGADA AO SEXO Parte 01 - HERANÇA SEXUAL - Genética - Prof. Kennedy Ramos	1o7weCgWFCE	27 - Education
Curso POO Teoria #10a - Herança (Parte 1)	_PZldwo0vVo	27 - Education
Herança Sexual (parte 1) - Aula 15 - Módulo II: Genética Prof. Gui	iujcgOnooVY	27 - Education
Noções Gerais do Direito das Sucessões - Herança e sua administração	pK4Cd4xIcUY	27 - Education
Turma da Mônica FUTEBOL NOVO !! Cebolinha x Cascaõ TiaCris TURMADAMÔNICA TurmadaMonica	7ZUx6OvMrwo	27 - Education
Tutorial de Banner para Futebol no Photoshop □ o futebol que nao foi	urr6pnKfLI0	27 - Education
FUTEBOL A FARSA MUNDIAL □	PHigOIqh5SY	27 - Education
Futebol de Pano	ISyb5-nUGVI	27 - Education
1º Aula de Gastronomia Utensílios de Cozinha, Tomates e Molhos / De Volta aos 30 - 19/09/2016	YJpqnf9qlkU	27 - Education
COZINHEIRO 02 de 04 - Gastronomia - Cursos completos Grátis Agora	XkmCN7oBorw	27 - Education
2 Aula Pratica Gastronomia Basica Corte Brunoise Técnica de corte da Cebola, Cenoura e Abobrinha	kfpaeBQY43M	27 - Education
Especial Sobre a Gastronomia Italiana Parte I	qYuKpK4XqbI	27 - Education
Aula prática cortes básicos de legumes - Parte 1	JeoWhXNPnc8	27 - Education
	u3YPNWJvdo0	27 - Education

Fonte: Elaborado pelo autor, com base na pesquisa realizada

REFERÊNCIAS

- ABU-EL-HAIJA, S.; KOTHARI, N.; LEE, J.; NATSEV, P.; TODERICI, G.; VARADARAJAN, B.; VIJAYANARASIMHAN, S. Youtube-8m: A large-scale video classification benchmark. **arXiv preprint arXiv:1609.08675**, 2016.
- ADERMANN, N.; BOGGIANO, J. Composer: Dependency manager for php. **URL: <https://getcomposer.org/>**, 2020. Último acesso em 16 de junho de 2020.
- BRAGA, J.; MENEZES, L. **Objetos de aprendizagem, volume 1: introdução e fundamentos**. UFABC, 2014. v. 1. 145 p. Disponível em: <<http://proec.ufabc.edu.br/uab/metdesOA2/2014-BRAGA-livro-oa-v1.pdf>>.
- BURGESS, J.; GREEN, J. **YouTube e a revolução digital**. [S.l.: s.n.], 2009. 24 p.
- CARVALHO, H. C. F. B.; PITANGUI, C. G.; ASSIS, L.; VIVAS, A. Educavídeos: Um sistema de recomendação de objetos de aprendizagem de vídeos educacionais do youtube. In: CONGRESSO BRASILEIRO DE ENSINO SUPERIOR A DISTÂNCIA, XVII., 2020, Goiânia. **2020 XVII Congresso Brasileiro de Ensino Superior a Distância**. Goiânia, 2020.
- CARVALHO, H. C. F. B.; PITANGUI, C. G.; TRINDADE, E. A. C.; ASSIS, L. P. de; ANDRADE, A. V.; SOUZA, D. P. B. de. Categorização de vídeos educacionais do youtube por meio de comentários. **RENOTE**, v. 18, n. 2, p. 621–629, 2020.
- CARVALHO, H. F. B.; PITANGUI, C. G.; TRINDADE, E. C.; ASSIS, L. P. de; ANDRADE, A. V. Learning objects and youtube: an analysis of videos and their categories. In: IEEE. **2020 XV Conferencia Latinoamericana de Tecnologias de Aprendizaje (LACLO)**. [S.l.], 2020. p. 1–6.
- CAZELLA, S. C.; BEHAR, P.; SCHNEIDER, D.; SILVA, K. K. da; FREITAS, R. Desenvolvendo um sistema de recomendação de objetos de aprendizagem baseado em competências para a educação: relato de experiências. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. [S.l.: s.n.], 2012. v. 23, n. 1.
- COX, C. **Listen to this**. [S.l.]: C. Cox Publishing Company, 1922.
- GOMES, L. Vídeos didáticos: uma proposta de critérios para análise. **Revista Brasileira de Estudos Pedagógicos**, v. 89, n. 223, 2008.
- HEILBRON, F. D. C. **Understanding Human Activities at Large Scale**. Tese (Doutorado), 2019.
- IEEE. Ieee learning technology standards committee (ltsc): Ieee p1484.12.1 ieee standard for learning object metadata. 2002. Acessado em: 27 de março de 2019. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1032843>>.
- JÚNIOR, C. B.; DORÇA, F. **Uma Abordagem para a Criação e Recomendação de Objetos de Aprendizagem usando um Algoritmo Genético, Tecnologias da Web Semântica e uma Ontologia**. 2018. 1533 p.
- LERDORF, R. Php: O que é o php? **URL: https://www.php.net/manual/pt_BR/intro-what-is.php**, 2020. Último acesso em 16 de junho de 2020.

MENOLLI, A.; MALUCELLI, A.; REINEHR, S. Criação semi-automática de objetos de aprendizagem a partir de conteúdos da wiki. In: BRAZILIAN SYMPOSIUM ON COMPUTERS IN EDUCATION (SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO-SBIE), XXII., 2011, Aracaju. **Anais**. Aracaju, 2011.

MIRANDA, R. M. d. **GROA: um gerenciador de repositórios de objetos de aprendizagem**. Dissertação (mathesis), 2004.

MORAIS, E. A. M.; AMBRÓSIO, A. P. L. Mineração de textos. **Relatório Técnico–Instituto de Informática (UFG)**, 2007.

NASCIMENTO, P. do; BARRETO, R.; PRIMO, T.; GUSMÃO, T.; OLIVEIRA, E. **Recomendação de objetos de aprendizagem baseada em modelos de estilos de aprendizagem: Uma revisão sistemática da literatura**. 2017. 213 p.

OTTO, M.; THORNTON, J. Bootstrap. URL: <https://getbootstrap.com/>, 2020. Último acesso em 16 de junho de 2020.

PINHEIRO, R. R. A. *et al.* Sistema de recomendação de vídeos educacionais: um estudo de caso no you tube. Universidade Federal de Alagoas, 2018.

PONTES, W. L.; FRANÇA, R. M.; COSTA, A. P. M.; BEHAR, P. **Filtragens de Recomendação de Objetos de Aprendizagem: uma revisão sistemática do CBIE**. 2014. 549 p.

RESIG, J. Jquery. URL: <https://jquery.com/>, 2020. Último acesso em 16 de junho de 2020.

SORGALLA, J. Lity: lightweight, accessible and responsive lightbox. URL: <https://sorgalla.com/lity/>, 2020. Último acesso em 16 de junho de 2020.

THELWALL, M. Social media analytics for youtube comments: Potential and limitations. **International Journal of Social Research Methodology**, Taylor & Francis, v. 21, n. 3, p. 303–316, 2018.

VIEIRA, F. J. R.; NUNES, M. A. S. N. Dica: Sistema de recomendação de objetos de aprendizagem baseado em conteúdo. **Scientia Plena**, v. 8, n. 5, 2012.

WILEY, D. A. **Learning object design and sequencing theory**. Tese (Doutorado) — Brigham Young University, 2000.

YOUTUBE. **API Reference**. 2015. Acesso em: 08/05/2019. Disponível em: <<https://developers.google.com/youtube/v3/docs/?hl=pt-br>>.

YOUTUBE. **Youtube Insights 2017**. 2017. Acesso em: 16 de Abril de 2019. Disponível em: <<https://www.thinkwithgoogle.com/intl/pt-br/youtubeinsights/2017/introducao/>>.

YOUTUBE. **YouTube Insights**. 2019. Acesso em: 17 de Abril de 2019. Disponível em: <<https://youtubeinsights.withgoogle.com>>.

YOUTUBE. **Youtube para imprensa**. 2019. Acesso em: 16 de Abril de 2019. Disponível em: <<https://www.youtube.com/intl/pt-BR/yt/about/press/>>.

3 LEARNING OBJECTS AND YOUTUBE: AN ANALYSIS OF VIDEOS AND THEIR CATEGORIES

RESUMO

A tecnologia tem contribuído para a área Educacional ao disponibilizar inúmeros conteúdos digitais. Porém, o excesso de materiais disponíveis torna sua seleção muito cansativa e frustrante. Para superar essa dificuldade, Objetos de Aprendizagem são usados durante o aprendizado de tecnologia. A plataforma de compartilhamento de vídeos online YouTube oferece diversos Objetos de Aprendizagem e possui mais de 2 bilhões de acessos mensais, sendo amplamente utilizada para fins de aprendizagem. Este trabalho apresenta uma análise da ferramenta de busca de vídeos de conteúdo educacional do YouTube. Os resultados apontam que, para conteúdos educacionais, a categorização fornecida pela plataforma geralmente não é assertiva.

Palavras-chave: educação, objetos de aprendizagem, motor de busca, youtube.

ABSTRACT

Technology has contributed to the Educational area by providing numerous digital content. However, the excess of available materials makes its selection very tiring and frustrating. In order to overcome this difficulty Learning Objects are used during technology learning. The online video-sharing platform YouTube offers several Learning Objects and has more than 2 billion monthly accesses, being widely used for learning purposes. This work presents an analysis of the search engine for educational content videos from YouTube. Results point out that, for educational content, the categorization provided by the platform is generally not assertive.

Keywords: education, learning objects, search engine, youtube.

3.1 Introdução

A evolução tecnológica tem possibilitado uma sociedade cada vez mais conectada, com acesso a diversas informações e com enorme interação social (NASCIMENTO *et al.*, 2017). Diariamente, diversas novas páginas Web são criadas, textos são escritos e vídeos são compartilhados. Essa expansão tecnológica propiciou um enorme avanço em diversas áreas, incluindo a Educação.

Um vasto acervo de arquivos, como os disponíveis atualmente na Internet, também trazem limitações. No contexto da Educação, por exemplo, novos conteúdos são produzidos e disponibilizados, porém, o acesso a determinados assuntos são frustrantes e/ou exaustivos (MIRANDA, 2004; BRAGA; MENEZES, 2014). Um dos motivos causadores de frustração advém do fato que ao se realizar a pesquisa, uma grande quantidade de materiais é retornada, o que causa confusão e dificulta a seleção dos conteúdos conforme o interesse e necessidade do usuário (BRAGA; MENEZES, 2014, p. 20).

No âmbito educacional, percebe-se que a grande quantidade de material disponibilizado na Internet causa confusão e dificuldade de seleção, atrapalha docentes e discentes, fazendo com que estes utilizem a grande parte do tempo em busca dos conteúdos ideais. Diante disto, uma nova maneira de planejar o desenvolvimento, a descrição, a busca, e a reutilização de materiais educacionais emergiu, dando origem ao conceito de Objetos de Aprendizagem (OAs) (BRAGA; MENEZES, 2014). Os OAs podem ser definidos como “qualquer entidade, digital ou não digital, que pode ser usada, reutilizada ou referenciada durante a aprendizagem apoiada por tecnologia” (IEEE, 2002, p. 1).

Diversos OAs podem ser encontrados na plataforma *YouTube*, bem como vídeos de diferentes contextos. O *YouTube* é uma plataforma voltada para vídeos e possui conteúdos sobre as mais diversas áreas do conhecimento e de diferentes tipos como: *vlogs*, desenhos, música, clipes, trailers, entre outros. O *YouTube* é uma grande plataformas de vídeos *online* e possui mais de 2 bilhões de acessos mensais. Os usuários desta plataforma geram bilhões de visualizações e assistem mais de um bilhão de horas de vídeos diariamente (YOUTUBE, 2019b).

Apesar de disponibilizar conteúdos sobre as mais diversas áreas, é possível identificar problemas quanto à qualidade da busca dos vídeos pesquisados. De forma geral, identificaram-se dois problemas relacionados ao mecanismo de busca do *YouTube*.

A primeira dificuldade se relaciona ao retorno dos resultados da pesquisa através da plataforma. De forma geral, a busca é realizada pela comparação entre os termos da expressão da pesquisa, com os dados dos vídeos, a saber: título, descrição e *tags*. Assim, os resultados da pesquisa são vídeos que contém o(s) termo(s) (da expressão de pesquisa) em seus dados. No caso da expressão de pesquisa conter mais de um termo, a busca retorna os vídeos que contenham, em seus dados, todos os termos da expressão. Por exemplo, ao se utilizar a expressão de pesquisa “Herança”, serão retornados vídeos que contenham o termo “Herança” em seu título, descrição, ou *tags*. Analogamente, caso a busca seja realizada pela expressão “Herança educação”, serão retornados vídeos que contenham ambos os termos em seus dados.

Como apontado, no caso de expressões de pesquisa com múltiplos termos, o mecanismo de busca da plataforma pode ser considerado bastante restritivo, já que, neste caso, é capaz de desconsiderar vídeos úteis e de interesse ao usuário que possuam, em seus dados, apenas alguns dos termos da expressão.

Outra dificuldade encontrada é a impossibilidade de se realizar buscas por vídeos segundo suas categorias. De fato, o *YouTube* categoriza seus vídeos, porém, não é possível realizar pesquisas por categorias, e a maioria dos usuários não tem conhecimento sobre essa categorização.

Este trabalho tem como objetivo realizar uma análise do mecanismo de busca de vídeos de conteúdo educacional da plataforma *YouTube*. Pretende-se, adicionalmente, verificar se a categorização de vídeos da própria plataforma é satisfatória para identificação de vídeos educacionais. Para estes fins, desenvolveu-se um aplicativo que acessa os dados do *YouTube*, apresenta as informações dos vídeos, e realiza buscas utilizando a categoria “*Education*”.

O presente trabalho se organiza como segue. A seção 2 apresenta os principais trabalhos relacionados à presente pesquisa. A metodologia utilizada para a realização da análise proposta é apresentada na seção 3. A seção 4 apresenta os resultados obtidos. Por fim, a seção 5 apresenta as considerações finais e apontamentos a trabalhos futuros.

3.2 Trabalhos Relacionados

Esta seção aborda os trabalhos relacionados a esta pesquisa. Destaca-se que poucos trabalhos foram identificados referentes à categorização de vídeos educacionais no *YouTube*. Neste sentido, esta seção relaciona os trabalhos que abordam o tema de Recomendação de OAs, por se tratarem de pesquisas que mais se aproximam do problema abordado.

Menolli, Malucelli e Reinehr (2011) objetivam gerar OAs com a Web 2.0, através da *Wikipedia*, utilizando tecnologias semânticas e o padrão *Learning Object Metadata* (LOM). Em sua proposta, acessam os conteúdos inseridos na plataforma, e realizam a mineração de textos a fim de extrair e classificar os conteúdos de acordo com o padrão LOM. Com a utilização desse padrão, encontram os atributos e metadados da página, gerando um *XML-schema* com os metadados da página em questão. Concluem que essa abordagem é necessária, pois facilita a utilização dos conteúdos devido as ferramentas wikis não considerarem como o conteúdo será utilizado.

Cazella *et al.* (2012) apresentam um relato de experiência de desenvolvimento e validação de um Sistema de Recomendação de OAs baseado em competências. Apresentam o modelo denominado RECoaComp (REComemendador de OAs baseado em Competências). O sistema funciona em três etapas, a saber: 1) o professor informa quais OAs ele utilizará em uma disciplina; 2) o aluno interage com o sistema, viabilizando as recomendações; 3) o sistema realiza uma filtragem de informação, recomendando os OAs. Os autores concluem que o sistema realiza as recomendações que melhor atendem às necessidades individuais, e destacam que o sistema se apresenta robusto, eficiente, e de fácil utilização.

Vieira e Nunes (2012) desenvolveram um Sistema de Recomendação de OAs denominado DICA, que utiliza o Banco Internacional de Objetos Educacionais (BIOE). O desenvolvimento do sistema ocorreu em 3 etapas, a saber: 1) coleta de informações - desenvolveu-se o modelo dos dados da plataforma; 2) processamento de dados - utilizando técnica de agrupamento (*K-means*), e de vizinhos mais próximos (*KNN*), encontraram um valor para a distância dos temas, possibilitando agrupamento de OAs de temas diferentes; 3) recomendação - desenvolvimento do Sistema de Recomendação. Os autores concluíram que o sistema foi bem aceito, sendo muitas vezes preferido ao BIOE, pois apresenta a possibilidade de facilitar as pesquisas por temas.

Abu-El-Haija *et al.* (2016) abordam a classificação de vídeos do *YouTube* com o objetivo de desenvolver um sistema de multi-classificação de vídeos. A base de dados utilizada possui aproximadamente 8 milhões de vídeos englobando um total de 1,9 bilhão de quadros de vídeos, e 500 mil horas de vídeos categorizados. A pesquisa é realizada em duas etapas, a saber: 1) os rótulos dos vídeos são obtidos por meio do *Knowledge Graph entities*; 2) os vídeos são processados *frame a frame*, e categorizados por uma Rede Neural Convolutiva pré-treinada no ImageNet. O ImageNet é um banco de dados visual com diversos objetos/entidades já classificados. Por meio do processamento de mais de 50 anos de vídeos, provendo 2 bilhões de *frames*, e mais de 8 milhões de vídeos que podem ser modelados rapidamente em uma única máquina, o trabalho aponta sua contribuição no sentido de auxiliar o desenvolvimento de pesquisas sobre compreensão de vídeos.

Apesar da categorização em diversas classes, não foi encontrada uma categoria específica para vídeos educacionais. O trabalho cita a categoria “*Jobs & Education*” e nela estão universidades, salas de aulas, palestras, etc. Desta forma, um vídeo que possua imagens de um *campus* universitário, por exemplo, será enquadrado nessa categoria, e não necessariamente é um vídeo educacional.

Júnior e Dorça (2018) apresentam uma abordagem para criação e recomendação de OAs por meio da plataforma de textos *Wikipédia*. A abordagem é definida por três etapas, a saber: 1) enriquecimento da ontologia através dos metadados das seções wiki; 2) a recomendação dos OAs - são utilizadas técnicas de Problema de Cobertura de Conjuntos com a utilização de Algoritmo Genético; 3) uso de operações CRUD (*Creat, Read, Update, Delete*). O trabalho conclui que a abordagem adotada resolve o problema da recomendação de OAs, retornando soluções de elevada qualidade.

Pinheiro *et al.* (2018) apresentam o *Easy YouTube*, um Sistema de Recomendação de OAs baseado no *YouTube*. O funcionamento do sistema se faz em seis etapas, a saber: 1) enriquecimento de consultas - definição de temas pré-definidos, cadastrados por especialistas; 2) extração de vídeos - busca de vídeos, que pode ser realizada por meio de pesquisa ou de temas pré-definidos; 3) pré-processamento - tratamento dos textos (em português), com remoção de pontuação, espaços, etc.; 4) classificação - utilização de um algoritmo para classificar os vídeos considerados como educacionais e de qualidade; 5) engenho de recomendação - o sistema

recebe os vídeos considerados “bons” e classifica-os; 6) coletor de *feedback* - o usuário avalia a recomendação fornecida pelo sistema através de notas, que variam de 1 a 7 estrelas. O trabalho aponta suas principais contribuições nos pontos: 1) o Sistema de Recomendação desenvolvido pode ser utilizado como solução para vários domínios de aplicação; 2) o sistema serviu como prova de conceito para melhorar as recomendações, através de características do *YouTube*, como a avaliações dos usuários, e linguagem nativa do vídeo.

O trabalho apresentado não detalha partes importantes de pesquisa. Por exemplo, para a classificação de vídeos considerados de qualidade, afirma-se que utilizou-se um conjunto de treinamento de 100 vídeos, contendo avaliações realizadas por especialistas e alunos que já tiveram em sua grade curricular o assunto “Orientação a Objetos/Herança”. Porém, não explica-se de que maneira esta análise foi realizada, e quais características dos vídeos foram consideradas. Outro ponto que causa confusão é o apontamento de que, devido ao prazo para a realização da pesquisa, o trabalho focou em “algumas características para o experimento”. Tais características não foram descritas.

3.3 Metodologia

Primeiramente, realizou-se um estudo da API da plataforma *YouTube* para verificar as funcionalidades disponíveis e para identificar os vídeos e seus dados. Em seguida, desenvolveu-se um aplicativo que realiza a busca e retorna o nome, *link*, e categoria dos vídeos do *YouTube*. O aplicativo realiza a busca por nome, como na plataforma, e também pela categoria do vídeo. De posse do aplicativo desenvolvido, realizou-se a coleta de dados para posterior análise dos mesmos.

Por meio da API do próprio *YouTube*, é possível identificar alguns metadados de cada vídeo. Dentre os metadados disponíveis, foi identificada a característica **categoryId**, que apresenta em quais categorias os vídeos se classificam. Algumas categorias de vídeos do *YouTube* são: **Education**, **People & Blogs**, **Sports**, **Music**, dentre outras. A Figura 5 apresenta o retorno da API para a categoria *Education*.

Posteriormente, desenvolveu-se um aplicativo para a coleta dos dados da plataforma. De forma geral, o aplicativo possui 4 etapas principais, a saber: 1) o usuário determina a expressão de pesquisa segundo o “tipo” de busca a ser realizada (normal ou pela **categoryID = Education**); 2) o aplicativo envia a solicitação de pesquisa para a API; 3) a API responde a solicitação de pesquisa com os dados; 4) O aplicativo retorna o resultado da pesquisa para o usuário. A figura 6 apresenta o funcionamento básico do aplicativo desenvolvido.

Realizaram-se três experimentos para a obtenção dos dados. Cada experimento diferencia-se pela maneira em que a busca dos vídeos foi realizada. A três formas de busca adotadas foram:

1. Por meio do navegador *Google Chrome* executado de maneira anônima;
2. Por meio do aplicativo desenvolvido em Python;
3. Por meio do aplicativo desenvolvido em Python e com o código da *categoryId = Education*.

Figura 5 – Retorno da API do Youtube para a categoria *Education*.

```
{
  "kind": "youtube#videoCategory",
  "etag": "hHUhloYhyMMqVkQ4LH9sr3aRQoU",
  "id": "27",
  "snippet": {
    "title": "Education",
    "assignable": true,
    "channelId": "UCBR8-60-B28hp2BmDPdntcQ"
  }
},
```

Fonte: Elaborado pelos autores

Figura 6 – Funcionamento básico do aplicativo de busca por API.



Fonte: Elaborado pelos autores

Cada um dos experimentos utilizou oito expressões de pesquisa, a saber: 1) “Medicina”, 2) “Inteligência Artificial”, 3) “Futebol”, 4) “Cerveja”, 5) “Herança”, 6) “Linguagem de Programação”, 7) “Java”, e 8) “Estrutura de Repetição”. Para cada pesquisa realizada, armazenaram-se os 12 primeiros resultados de vídeos retornados. Os três primeiros resultados de cada experimento, para cada uma das oito expressões de busca, são apresentados nas Tabelas 3, 4, e 5.

A coluna denominada #Exp., nas Tabelas 3, 4, e 5, identifica o experimento (tipo de busca realizada, de 1 a 3), bem como a expressão de busca utilizada, de 1 a 8. Neste sentido, o termo “Exp.1-1” representa os três primeiros resultados da busca para o tipo de busca 1), i.e., por meio do navegador *Google Chrome* executado de maneira anônima, para expressão de busca 1), i.e., “Medicina”. De maneira análoga, o termo “Exp.2-2” representa os três primeiros resultados da busca para o tipo de busca 2), i.e., por meio do aplicativo desenvolvido em Python, para expressão de busca 2), i.e., “Inteligência Artificial”. Todas as outras entradas da coluna #Exp. devem ser interpretadas desta mesma maneira.

3.4 Resultados

A Tabela 3 apresenta os resultados do experimento 1. Observa-se que alguns vídeos não estão categorizados de maneira precisa, principalmente quando os termos são mais genéricos. Alguns exemplos deste caso são os resultados: Exp.1-1 código: 7NKlihwokyk; Exp.1-3 código: 7ukt4fveex; Exp.1-4 código: CsPrHiDPr0Y. Nota-se, adicionalmente, que a plataforma apresenta dificuldade na classificação de vídeos de Educação voltados para a área de tecnologia, uma vez que não é capaz de identificar se um vídeo se classifica na categoria “*Education*” ou na categoria “*Science & Technology*”, como por exemplo os vídeos: Exp.1-2 código: W95YIM5-iPk; Exp.1-6 códigos: Uh-GNH-t89w e 2_FJrmft3uQ; Exp.1-8 código: 6BLB0fBqzlg.

Tabela 3 – Resultados da busca pelo *Google Chrome* (modo anônimo).

#Exp	Título	idVideo	id - Categoria
Exp.1-1	Anitta - Medicina (Official Music Video)	UhurAzTl5gs	10 - Music
	QUERO MEDICINA, E AGORA? - Débora Aladim	7NKlihwokyk	27 - Education
Exp.1-2	Minha despedida da Medicina	SGnALLiVpeY	22 - People & Blogs
	Discovery Brasil Inteligência Artificial - IBM	W95YIM5-iPk	2 - Autos & Vehicles
	A história da Inteligência Artificial - TecMundo	Lhu8bdmkMCM	28 - Science & Technology
Exp.1-3	Inteligência Artificial ZERANDO o jogo mais difícil do mundo!	QD-gHp81G4M	28 - Science & Technology
	DESAFIO DOS GOLS DE FALTA IMPOSSÍVEIS NO FUTEBOL	aalZk0Bhl50	17 - Sports
	AS PIORES BRIGAS, FALTAS E CONFUSÕES DO FUTEBOL Ft. Neymar, Messi, Cristiano Ronaldo, Mbappé [2018]	mfp_5FcvTD0	17 - Sports
	TEVE GOL E EXPULSÃO DO GABISHOW Grêmio 0 x 1 Flamengo Melhores Momentos HD 17/11/2019	7ukt4fveexM	10 - Music
Exp.1-4	Atitude 67 - Cerveja De Garrafa (Fumaça Que Eu Faço)	cPjhBow3dLw	10 - Music
	Leandro & Leonardo - Cerveja - Videoclipe Oficial	-0zjb5SZck	10 - Music
Exp.1-5	Cervejas Puro Malte	CsPrHiDPr0Y	22 - People & Blogs
	Jozyanne- Herança (legendado)	OQGMGWv3_Z4	10 - Music
	Drik Barbosa - Herança part. Anna Tréa (Álbum visual)	CtguPlt4_7s	10 - Music
Exp.1-6	Herança - Os Arrais	vHjHh8CTPUU	10 - Music
	Curso online de programação para iniciantes - AULA GRÁTIS: O que é uma linguagem de programação?	PmefpISZ7Ew	27 - Education
	Qual é a melhor linguagem de programação? #HiptersPontoTube	Uh-GNH-t89w	28 - Science & Technology
Exp.1-7	MINHA OPINIÃO: Qual a melhor linguagem para começar a programar?	2_FJrmft3uQ	28 - Science & Technology
	Curso de Java #4 - Primeiro Programa em Java - Gustavo Guanabara	ijXXaIYsVx4	27 - Education
	Parto Gak Tahu Anwar & Marshel Datang ke Rumahnya OPERA VAN JAVA (06/01/20) Part 1 (en)	bp6CdeULMPY	24 - Entertainment
	Java // Dicionário do Programador	sZAxLRMxEUo	28 - Science & Technology
Exp.1-8	Estruturas de Repetição 1 - Curso de Algoritmos #9 - Gustavo Guanabara	U5PnCt58Q68	27 - Education
	13 - Lógica de Programação - Estruturas de Repetição (Loop) - ENQUANTO	6BLB0fBqzlg	28 - Science & Technology
	Curso Python #13 - Estrutura de repetição for	cL4YDtFnCt4	27 - Education

Fonte: Elaborado pelo autor, com base na pesquisa realizada

A Tabela 4 apresenta os resultados do experimento 2. Percebe-se que a busca realizada por meio do aplicativo retorna resultados diferentes da pesquisa realizada diretamente por meio da plataforma. A exemplo, apenas os experimentos Exp.1-8 e Exp.2-8 apresentaram os mesmos resultados na mesma ordem de aparição.

A Tabela 5 apresenta os resultados do experimento 3. Nota-se que para este experimento, encontram-se vídeos presentes nas consultas de ambos os experimentos anteriores,

Tabela 4 – Resultados da busca pelo aplicativo.

#Exp	Título	idVideo	id - Categoria
Exp.2-1	Anitta - Medicina (Official Music Video)	UhurAzTl5gs	10 - Music
	QUERO MEDICINA, E AGORA? - Débora Aladim	7NKlihwokyk	27 - Education
	COMO É O CURSO DE MEDICINA? Trilha de profissões	DSBHxRcMrzI	27 - Education
Exp.2-2	Discovery Brasil Inteligência Artificial - IBM	W95YIM5-iPk	2 - Autos Vehicles
	A história da Inteligência Artificial - TecMundo	Lhu8bdmkMCM	28 - Science & Technology
Exp.2-3	Qual a diferença entre Inteligência Artificial, Machine Learning, Data Science, Deep Learning, etc?	ccZ2pyr3YDw	28 - Science & Technology
	Dribles Mais Humilhantes Da História Do Futebol	CWWOW7T82Hw	17 - Sports
	GAROTADA VIRA JOGO NA RAÇA E CONQUISTA O TETRA MUNDIAL PARA NOSSA SELEÇÃO HD 17/11/2019	piJycdsGDEI	10 - Music
Exp.2-4	DESAFIO DOS GOLS DE FALTA IMPOSSÍVEIS NO FUTEBOL	aaIZk0BhI50	17 - Sports
	GRINGOS PROVANDO CERVEJAS BRASILEIRAS	t0tvIA1hUGU	19 - Travel Events
	CERVEJA EM CASA	hXwi3oucImo	24 - Entertainment
Exp.2-5	Atitude 67 - Cerveja De Garrafa (Fumaça Que Eu Faço)	cPjhBow3dLw	10 - Music
	Jozyanne- Herança (legendado)	OQGMGWv3_Z4	10 - Music
	HERANÇA	gSkMDV_LOIE	23 - Comedy
Exp.2-6	Tudo o que você deveria saber sobre herança... mas certamente não sabe	ITfdWxHc2tA	25 - News & Politics
	Qual é a melhor linguagem de programação? #HipersPontoTube	Uh-GNH-t89w	28 - Science & Technology
	Curso online de programação para iniciantes - AULA GRÁTIS: O que é uma linguagem de programação?	PmefpISZ7Ew	27 - Education
Exp.2-7	Qual Linguagem de Programação você deve aprender primeiro?	Yqq7IsxDGZc	27 - Education
	Java Tutorial for Beginners [2019]	eIrMbaQSU34	27 - Education
	Learn Java in 14 Minutes (seriously)	RRubcjpTkks	27 - Education
Exp.2-8	JAVA ESTÁ MORTO? Pontos positivos e negativos da linguagem!	2lgPfXrtVyo	28 - Science & Technology
	Estruturas de Repetição 1 - Curso de Algoritmos #9 - Gustavo Guanabara	U5PnCt58Q68	27 - Education
	13 - Lógica de Programação - Estruturas de Repetição (Loop) - ENQUANTO	6BLB0fBqzlg	28 - Science & Technology
	Curso Python #13 - Estrutura de repetição for	cL4YDtFnCt4	27 - Education

Fonte: Elaborado pelo autor, com base na pesquisa realizada

contudo, também encontram-se vídeos que não são da área educacional, tais como os resultados: Exp.3-1 códigos: 7NKlihwokyk e UkhSLsDgj4M; Exp.3-3 códigos: 7ZUx6OvMrwo e PHigOIqh5SY. É possível observar que para algumas expressões pesquisadas, a categoria do vídeo pode auxiliar durante a pesquisa, tal como nos resultados: Exp.3-3 código: urr6pnKfLI0; Exp.3-4 códigos: q2A2Q3Vd_C0 e 9C2Q6gi8KJo; e Exp.3-5. Essa melhora deve-se ao fato que anteriormente nenhuma dessas pesquisas havia obtido um resultado de vídeo pertencente a área educacional. Diante destes resultados, aponta-se que a categoria “Education” da própria

Tabela 5 – Resultados da busca pelo aplicativo e *categoryID = Education*.

#Exp	Título	idVideo	id - Categoria
Exp.3-1	QUERO MEDICINA, E AGORA? - Débora Aladim	7NKlihwokyk	27 - Education
	COMO É O CURSO DE MEDICINA? Trilha de profissões	DSBHxRcMrzI	27 - Education
Exp.3-2	Você REALMENTE quer MEDICINA?	UkhSLsDgj4M	27 - Education
	Inteligência artificial	UhA_ZgI-otM	27 - Education
	Inteligência Artificial - Introdução à Inteligência Artificial - Visão Geral da área	Bcw5YZA-Avw	27 - Education
Exp.3-3	Tendencias de la Inteligencia Artificial para 2020	CFvy6zSsOEc	27 - Education
	Turma da Mônica FUTEBOL NOVO !! Cebolinha x Cascão #TiaCris #TURMADAMÔNICA #Turmada-Monica	7ZUx6OvMrwo	27 - Education
Exp.3-4	Tutorial de Banner para Futebol no Photoshop o futebol que nao foi	urr6pnKfLI0	27 - Education
	7 BENEFÍCIOS DA CERVEJA	PHigOlqh5SY	27 - Education
	Receita Cerveja Heineken - Clone?	IMWu3nbejsY	27 - Education
	Como Fazer Cerveja por R\$ 2 o LITRO	q2A2Q3Vd_C0	27 - Education
Exp.3-5	Advogado esclarece dúvidas sobre o direito de herança	9C2Q6gI8KJo	27 - Education
	HERANÇA LIGADA AO SEXO Parte 01 - HERANÇA SEXUAL - Genética - Prof. Kennedy Ramos	7WZAGI8FmlA	27 - Education
Exp.3-6	ELA TEM DIREITO A MINHA HERANÇA	1o7weCgWFCE	27 - Education
	Curso online de programação para iniciantes - AULA GRÁTIS: O que é uma linguagem de programação?	OwjJpzvFIQQ	27 - Education
	Qual Linguagem de Programação você deve aprender primeiro?	PmefpISZ7Ew	27 - Education
Exp.3-7	Qual Linguagem de Programação estudar para ser um Hacker?	Yqq7IsxDGZc	27 - Education
	Java Tutorial for Beginners [2019]	GIB8oz2lizs	27 - Education
	Learn Java in 14 Minutes (seriously)	eIrMbAQSU34	27 - Education
	Learn Java 8 - Full Tutorial for Beginners	RRubcjpTkks	27 - Education
Exp.3-8	grEKMHGyYns	grEKMHGyYns	27 - Education
	Estruturas de Repetição 1 - Curso de Algoritmos #9 - Gustavo Guanabara	U5Pnct58Q68	27 - Education
	Curso Python #13 - Estrutura de repetição for	cL4YDfFnCt4	27 - Education
	Programação C - Aula 07 - while, do-while, for - Estruturas de repetição	rCFh-tvoXIc	27 - Education

Fonte: Elaborado pelo autor, com base na pesquisa realizada

plataforma não deve ser considerada como critério determinante a ser utilizado para identificação dos vídeos educacionais, uma vez que em alguns casos pode melhorar a qualidade da busca e em outros piora-lá.

Durante a realização dos experimentos, deparou-se com dois importantes questionamentos, a saber: 1) a categoria de um vídeo pertence ao canal ao qual o vídeo se encontra? 2) caso inclua-se a expressão “*Education*” na pesquisa realizada por meio do site do *YouTube*, os primeiros resultados da pesquisa serão vídeos educacionais?

Quanto ao primeiro questionamento, verificaram-se os metadados dos seguintes vídeos, a saber: Resumo de História: ERA VARGAS (Débora Aladim) (código: ZcTDWBqUju8; Categoria: 22 - *People & Blogs*), e QUERO MEDICINA, E AGORA? - Débora Aladim (código:

7NKlihwokyk; Categoria: 27 - *Education*), ambos do canal denominado Débora Aladim. Neste sentido, conclui-se que a categoria é associada a um vídeo, e não a um canal, uma vez que um canal pode conter vídeos em categorias distintas.

Quanto ao segundo questionamento, pesquisaram-se na plataforma as seguintes expressões, a saber: “herança educação”; “futebol educação”; “java educação”. Os três primeiros resultados para as pesquisas são apresentados a seguir.

- **“herança educação”**: MINISTRO DA EDUCAÇÃO EXPÕE HERANÇA MALDITA DO PT E COMO CORRIGIR, DETONA LULA E MÉTODO PAULO FREIRE, 22 - People & Blogs; Maior herança maldita do PT é a educação - Marco Antonio Villa, 25 - News & Politics; Quebrando Paradigmas sobre Heranças Espirituais I - Pastora Tânia Tereza, 29 - no Category;
- **“futebol educação”**: Futebol Lúdico - Elementos do futebol na educação física, 27 - Education; Educação Física - Aula 10.2 - Fundamentos do futebol, 27 - Education; NOVAS REGRAS DO FUTEBOL 2019 COMPLETO, 17 - Sports;
- **“java educação”**: Inclusão Social a partir da Educação Java - Jean Vargas, 27 - Education; Vídeo - Curso Online de Java nas Tecnologias JSP/Servlet JSP - Portal Educação 12/07/2010, 27 - Education; Vídeo - Curso Onlinde de Java e o paradigma da Orientação a Objetos - Portal Educação, 27 - Education;

Percebe-se que os vídeos retornados possuem em seu título e/ou descrição ambos os termos da expressão de busca utilizada. Por exemplo, o vídeo NOVAS REGRAS DO FUTEBOL 2019 COMPLETO, foi o terceiro resultado da pesquisa por “futebol educação”. Isto se deve ao fato do canal ser “Dicas Educação Física”, e sua descrição iniciar-se com a expressão “INSCREVA-SE NO CANAL DICAS EDUCAÇÃO FÍSICA Curso Futebol Moderno”.

No caso do vídeo Inclusão Social a partir da Educação Java - Jean Vargas, percebe-se que o mesmo possui ambos os termos pesquisados. Contudo, apesar de estar categorizado em “*Education*”, o mesmo não apresenta um viés educacional, pois se trata de vídeo-resposta de agradecimento do proprietário a um e-mail recebido.

Diante dos resultados, nota-se que a utilização de um termo adicional na pesquisa não auxilia a obtenção de resultados mais precisos/refinados. Neste sentido, pensa-se ser necessário abordar outras formas de identificação de vídeos educacionais na plataforma.

3.5 Considerações Finais e Trabalhos Futuros

O *YouTube* é uma plataforma muito popular, com um crescente acervo de vídeos sobre os mais diversos temas. Porém, apesar do crescimento constante de conteúdos disponíveis, nem sempre os conteúdos sugeridos atendem a necessidade do usuário, demandando diversas visualizações e/ou extensivas buscas. Devido a esses problemas, um estudo específico sobre o mecanismo de busca desta plataforma torna-se necessário.

Por meio do estudo realizado, observaram-se três pontos críticos relacionados ao mecanismo de busca da plataforma *YouTube*, saber: 1) falta de transparência quanto à categorização

de um vídeo; 2) informações incorretas quanto à categorização dos vídeos; 3) impossibilidade de se buscar vídeos por suas categorias.

Observa-se que apesar de possuir uma categoria por vídeo, algumas delas, não são relevantes ou são erroneamente definidas. Desta maneira, não se deve considerar as categorias dos vídeos como determinantes para a identificação dos mesmos. Neste sentido, pensa-se ser necessário propor novas maneiras de se classificar os vídeos da plataforma.

Este trabalho é uma contribuição para área de identificação de vídeos educacionais na plataforma *YouTube*. Para a continuidade desta pesquisa, está sendo desenvolvido um sistema *Web*, que utiliza os metadados e os comentários dos vídeos para categorizá-los, bem como para mensurar o nível de qualidade dos mesmos.

REFERÊNCIAS

- ABU-EL-HAIJA, S.; KOTHARI, N.; LEE, J.; NATSEV, P.; TODERICI, G.; VARADARAJAN, B.; VIJAYANARASIMHAN, S. Youtube-8m: A large-scale video classification benchmark. **arXiv preprint arXiv:1609.08675**, 2016.
- ADERMANN, N.; BOGGIANO, J. Composer: Dependency manager for php. **URL: <https://getcomposer.org/>**, 2020. Último acesso em 16 de junho de 2020.
- BRAGA, J.; MENEZES, L. **Objetos de aprendizagem, volume 1: introdução e fundamentos**. UFABC, 2014. v. 1. 145 p. Disponível em: <<http://proec.ufabc.edu.br/uab/metdesOA2/2014-BRAGA-livro-oa-v1.pdf>>.
- BURGESS, J.; GREEN, J. **YouTube e a revolução digital**. [S.l.: s.n.], 2009. 24 p.
- CARVALHO, H. C. F. B.; PITANGUI, C. G.; ASSIS, L.; VIVAS, A. Educavídeos: Um sistema de recomendação de objetos de aprendizagem de vídeos educacionais do youtube. In: CONGRESSO BRASILEIRO DE ENSINO SUPERIOR A DISTÂNCIA, XVII., 2020, Goiânia. **2020 XVII Congresso Brasileiro de Ensino Superior a Distância**. Goiânia, 2020.
- CARVALHO, H. C. F. B.; PITANGUI, C. G.; TRINDADE, E. A. C.; ASSIS, L. P. de; ANDRADE, A. V.; SOUZA, D. P. B. de. Categorização de vídeos educacionais do youtube por meio de comentários. **RENOTE**, v. 18, n. 2, p. 621–629, 2020.
- CARVALHO, H. F. B.; PITANGUI, C. G.; TRINDADE, E. C.; ASSIS, L. P. de; ANDRADE, A. V. Learning objects and youtube: an analysis of videos and their categories. In: IEEE. **2020 XV Conferencia Latinoamericana de Tecnologias de Aprendizaje (LACLO)**. [S.l.], 2020. p. 1–6.
- CAZELLA, S. C.; BEHAR, P.; SCHNEIDER, D.; SILVA, K. K. da; FREITAS, R. Desenvolvendo um sistema de recomendação de objetos de aprendizagem baseado em competências para a educação: relato de experiências. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. [S.l.: s.n.], 2012. v. 23, n. 1.
- COX, C. **Listen to this**. [S.l.]: C. Cox Publishing Company, 1922.
- GOMES, L. Vídeos didáticos: uma proposta de critérios para análise. **Revista Brasileira de Estudos Pedagógicos**, v. 89, n. 223, 2008.
- HEILBRON, F. D. C. **Understanding Human Activities at Large Scale**. Tese (Doutorado), 2019.
- IEEE. Ieee learning technology standards committee (ltsc): Ieee p1484.12.1 ieee standard for learning object metadata. 2002. Acessado em: 27 de março de 2019. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1032843>>.
- JÚNIOR, C. B.; DORÇA, F. **Uma Abordagem para a Criação e Recomendação de Objetos de Aprendizagem usando um Algoritmo Genético, Tecnologias da Web Semântica e uma Ontologia**. 2018. 1533 p.
- LERDORF, R. Php: O que é o php? **URL: https://www.php.net/manual/pt_BR/intro-what-is.php**, 2020. Último acesso em 16 de junho de 2020.

MENOLLI, A.; MALUCELLI, A.; REINEHR, S. Criação semi-automática de objetos de aprendizagem a partir de conteúdos da wiki. In: BRAZILIAN SYMPOSIUM ON COMPUTERS IN EDUCATION (SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO-SBIE), XXII., 2011, Aracaju. **Anais**. Aracaju, 2011.

MIRANDA, R. M. d. **GROA: um gerenciador de repositórios de objetos de aprendizagem**. Dissertação (mathesis), 2004.

MORAIS, E. A. M.; AMBRÓSIO, A. P. L. Mineração de textos. **Relatório Técnico–Instituto de Informática (UFG)**, 2007.

NASCIMENTO, P. do; BARRETO, R.; PRIMO, T.; GUSMÃO, T.; OLIVEIRA, E. **Recomendação de objetos de aprendizagem baseada em modelos de estilos de aprendizagem: Uma revisão sistemática da literatura**. 2017. 213 p.

OTTO, M.; THORNTON, J. Bootstrap. URL: <https://getbootstrap.com/>, 2020. Último acesso em 16 de junho de 2020.

PINHEIRO, R. R. A. *et al.* Sistema de recomendação de vídeos educacionais: um estudo de caso no you tube. Universidade Federal de Alagoas, 2018.

PONTES, W. L.; FRANÇA, R. M.; COSTA, A. P. M.; BEHAR, P. **Filtragens de Recomendação de Objetos de Aprendizagem: uma revisão sistemática do CBIE**. 2014. 549 p.

RESIG, J. Jquery. URL: <https://jquery.com/>, 2020. Último acesso em 16 de junho de 2020.

SORGALLA, J. Lity: lightweight, accessible and responsive lightbox. URL: <https://sorgalla.com/lity/>, 2020. Último acesso em 16 de junho de 2020.

THELWALL, M. Social media analytics for youtube comments: Potential and limitations. **International Journal of Social Research Methodology**, Taylor & Francis, v. 21, n. 3, p. 303–316, 2018.

VIEIRA, F. J. R.; NUNES, M. A. S. N. Dica: Sistema de recomendação de objetos de aprendizagem baseado em conteúdo. **Scientia Plena**, v. 8, n. 5, 2012.

WILEY, D. A. **Learning object design and sequencing theory**. Tese (Doutorado) — Brigham Young University, 2000.

YOUTUBE. **API Reference**. 2015. Acesso em: 08/05/2019. Disponível em: <<https://developers.google.com/youtube/v3/docs/?hl=pt-br>>.

YOUTUBE. **Youtube Insights 2017**. 2017. Acesso em: 16 de Abril de 2019. Disponível em: <<https://www.thinkwithgoogle.com/intl/pt-br/youtubeinsights/2017/introducao/>>.

YOUTUBE. **YouTube Insights**. 2019. Acesso em: 17 de Abril de 2019. Disponível em: <<https://youtubeinsights.withgoogle.com>>.

YOUTUBE. **Youtube para imprensa**. 2019. Acesso em: 16 de Abril de 2019. Disponível em: <<https://www.youtube.com/intl/pt-BR/yt/about/press/>>.

4 CATEGORIZAÇÃO DE VÍDEOS EDUCACIONAIS DO *YOUTUBE* POR MEIO DE COMENTÁRIOS

RESUMO

O Youtube é uma plataforma online que oferta Objetos de Aprendizagem sobre as mais diversas áreas do conhecimento. De maneira a facilitar a busca por conteúdos, o *Youtube* categoriza os seus vídeos em diversas classes. No entanto, já apontou-se que, de forma geral, a categorização dos vídeos não é relevante ou que as categorias estão erroneamente atribuídas aos vídeos. Este trabalho utiliza técnicas de Mineração de Texto com o objetivo de analisar os comentários dos vídeos e identificar se eles possuem potencial para serem utilizados para a categorização de vídeos educacionais. Resultados experimentais apontam que existem diferenças importantes nos vocabulários dos comentários dos vídeos educacionais e não educacionais.

Palavras-chave: categorização, mineração de texto, youtube.

ABSTRACT

Youtube is an online sharing video platform that offers Learning Objects on several knowledge areas. In order to facilitate the content search, Youtube categorizes its videos in different classes. However, it has already been pointed out that, in general, this categorization is not relevant or that the categories are erroneously assigned to videos. This work uses Text Mining techniques in order to analyze the comments of the videos and identify if they have potential to be used in order to categorize educational videos. Experimental results indicate that there are important differences in the vocabularies of the comments of educational and non-educational videos.

Keywords: categorization, text mining, youtube.

4.1 Introdução

O *Youtube*, uma plataforma criada por Chad Hurley, Steve Chen e Jawed Karim, foi lançado no final de junho de 2005 com o intuito de compartilhar vídeos na Internet. Segundo Burgess e Green (2009), possui uma interface simples e intuitiva, onde os usuários podem, sem elevado conhecimento técnico, assistir e publicar vídeos.

A plataforma está voltada para vídeos e possui conteúdos sobre diversas áreas do conhecimento. Segundo Heilbron (2019), a cada minuto são enviadas 400 horas de vídeos, e mais de 2 bilhões de usuários acessam a plataforma mensalmente (YOUTUBE, 2019b). Segundo Youtube (2017), 1,5 bilhões de pessoas conectadas pelo mundo acessam o *Youtube* mensalmente e 95% da população brasileira conectada acessa a plataforma ao menos uma vez por mês. No Brasil, 9 em cada 10 usuários do *Youtube* acessam a plataforma com o intuito de aprender algo novo e mais da metade acredita que o *Youtube* é o lugar onde encontra-se tudo o que se deseja ver e aprender (YOUTUBE, 2019a).

O *Youtube/Google* em parceria com a fundação Lemann desenvolveram um canal denominado *Youtube Edu* com a proposta de fornecer conteúdos educacionais gratuitos e de qualidade, em português. O conteúdo do canal é voltado para o Ensino Fundamental e Ensino Médio, e aborda temas das disciplinas de Biologia, Física, Química, Geografia, História, Língua Espanhola, Língua Inglesa, Língua Portuguesa, e Matemática. O canal possui vídeos e *playlists* segmentadas por curso e série, além de “aulões” para o Enem.

Apesar de apresentar diversos materiais sobre as áreas mencionadas, o conteúdo do *Youtube Edu* é limitado no sentido de que não contempla temáticas de cursos técnicos e de nível superior. Observa-se, adicionalmente, que, por ser um canal, o *Youtube Edu* é englobado pelo *Youtube*, ou seja, uma pesquisa no *Youtube* pode retornar vídeos do canal *Youtube Edu*, contudo, o inverso não é verdadeiro. Este é o motivo desta pesquisa se basear na utilização da plataforma *Youtube* e não no canal *Youtube Edu*.

O *Youtube*, sob o ponto de vista educacional, pode ser entendido como um repositório de Objetos de Aprendizagem (OA). De forma geral, os OA podem ser definidos como “qualquer recurso digital que possa ser reutilizado com o intuito de auxiliar na educação, desde que possa ser entregue pela rede, como imagens, vídeos, animações, textos e outros” (WILEY, 2000).

Neste sentido, a oferta de OA é realizada pela plataforma por meio do seu mecanismo de busca, i.e., busca-se por um conteúdo e a plataforma apresenta os OA condizentes à expressão de pesquisa utilizada. Essa “assistência” faz parte da área de estudos que aborda a Recomendação de Objetos de Aprendizagem. Segundo Pontes *et al.* (2014), a Recomendação de Objetos de Aprendizagem é utilizada para identificar usuários, armazenar suas preferências e recomendar itens que podem ser produtos, serviços e/ou conteúdos, de acordo com suas necessidades e interesses.

Apesar de disponibilizar conteúdos sobre as mais diversas áreas e possuir um extenso acervo de vídeos, é possível identificar um problema que se relaciona aos resultados retor-

nados pela plataforma ao se realizar uma pesquisa, pois, em muitos casos, o número de resultados retornado é muito elevado, sendo muitos destes de baixa qualidade e/ou não muito relacionados à pesquisa realizada.

Uma das formas de se reduzir o número de resultados da pesquisa e de, portanto, particularizar o retorno da mesma, se faz pela especificação da categoria dos vídeos pelos quais se busca. De fato, o *Youtube* categoriza seus vídeos, porém, não é possível realizar pesquisas por elas, e a maioria dos usuários não tem conhecimento sobre essa categorização.

No intuito de se utilizar as categorias dos vídeos para a pesquisa na plataforma e consequentemente refinar o conteúdo retornado, Carvalho *et al.* (2020a) desenvolveram um sistema *Web* que permitiu a realização de pesquisa de vídeos por categorias. Entretanto, os resultados apontam que muitas vezes a categorização dos vídeos não é relevante ou que as categorias estão erroneamente atribuídas aos vídeos, o que prejudica ainda mais a a qualidade da pesquisa. Neste sentido, apontou-se que não se deve considerar as categorias dos vídeos do *Youtube* como determinantes para a identificação dos mesmos (CARVALHO *et al.*, 2020c).

Este trabalho propõe uma nova forma de categorização de vídeos do *Youtube* por meio dos comentários postados pelos usuários. Neste sentido, analisam-se os comentários de diversos vídeos educacionais e não educacionais da plataforma. Os comentários são processados para a obtenção dos vocábulos e termos mais frequentes, com o intuito de identificar se os comentários podem ser utilizados para a categorização de vídeos educacionais.

Resultados preliminares demonstram diferenças entre os termos e vocábulos mais frequentes empregados nos vídeos educacionais e não educacionais. Observa-se, por exemplo, que termos como “melhor professor”, e “ótima aula” figuram frequentemente em comentários de vídeos educacionais, mas não ocorrem em comentários de vídeos não educacionais. De forma semelhante, verificou-se que os radicais “profes”, “aul”, e “obrig” figuram frequentemente em comentários de vídeos vídeos educacionais. Neste sentido, aponta-se que os comentários possuem potencial para serem utilizados com o intuito de categorizar os vídeos educacionais do *Youtube*.

O presente trabalho se organiza como segue. A seção 2 apresenta os principais trabalhos relacionados a presente pesquisa. A seção 3 descreve a metodologia experimental utilizada. A seção 4 discute os resultados obtidos. Por fim, a seção 5 apresenta as considerações finais e apontamentos a trabalhos futuros.

4.2 Trabalhos Relacionados

Esta seção aborda os trabalhos relacionados a esta pesquisa. Destaca-se que poucos trabalhos foram identificados referentes a categorização de vídeos educacionais do *Youtube*. Neste sentido, esta seção relaciona os trabalhos que abordam o tema de Recomendação de OA utilizando a *Wikipédia* e o *Youtube*, por se tratarem de pesquisas que mais se aproximam do problema abordado.

Menolli, Malucelli e Reinehr (2011) objetivam gerar OA com a Web 2.0, através da *Wikipedia*, utilizando tecnologias semânticas e o padrão *Learning Object Metadata* (LOM). Em

sua proposta, acessam os conteúdos inseridos na plataforma, e realizam a mineração de textos a fim de extrair e classificar os conteúdos de acordo com o padrão LOM. Com a utilização desse padrão, encontram os atributos e metadados da página, gerando um *XML-schema* com os metadados da página em questão. Concluem que essa abordagem é necessária, pois facilita a utilização dos conteúdos devido as ferramentas wikis não considerarem como o conteúdo será utilizado.

Abu-El-Haija *et al.* (2016) abordam a classificação de vídeos do *Youtube* com o objetivo de desenvolver um sistema de multi-classificação. A base de dados utilizada possui aproximadamente 8 milhões de vídeos englobando um total de 1,9 bilhão de quadros, e 500 mil horas de conteúdos. A pesquisa é realizada em duas etapas, a saber: 1) os rótulos dos vídeos são obtidos por meio do *Knowledge Graph entities*; 2) os vídeos são processados *frame a frame*, e categorizados por uma Rede Neural Convolutiva pré-treinada no ImageNet. O ImageNet é um banco de dados visual com diversos objetos/entidades já classificados. Por meio do processamento de mais de 50 anos de vídeos, provendo 2 bilhões de *frames*, e mais de 8 milhões de vídeos que podem ser modelados rapidamente em uma única máquina, o trabalho aponta sua contribuição no sentido de auxiliar o desenvolvimento de pesquisas sobre compreensão de vídeos.

Apesar da categorização em diversas classes, não foi encontrada uma categoria específica para vídeos educacionais. O trabalho cita a categoria “*Jobs & Education*” e nela estão universidades, salas de aulas, palestras, etc. Desta forma, um vídeo que possua imagens de um *campus* universitário, por exemplo, será enquadrado nessa categoria, e não necessariamente é um vídeo educacional.

Júnior e Dorça (2018) apresentam uma abordagem para criação e recomendação de OA por meio da plataforma *Wikipédia*. A abordagem é definida por três etapas, a saber: 1) enriquecimento da ontologia por meio dos metadados das seções wiki; 2) recomendação dos OA tratada como um Problema de Cobertura de Conjuntos, com a utilização de Algoritmo Genético; 3) uso de operações CRUD (*Create, Read, Update, Delete*). O trabalho conclui que a abordagem adotada resolve o problema da recomendação de OAs, retornando soluções de elevada qualidade.

Pinheiro *et al.* (2018) apresentam o *Easy Youtube*, um Sistema de Recomendação de OA baseado no *Youtube*. O funcionamento do sistema se faz em seis etapas, a saber: 1) enriquecimento de consultas - definição de temas pré-definidos, cadastrados por especialistas; 2) extração de vídeos - busca de vídeos, que pode ser realizada por meio de pesquisa ou de temas pré-definidos; 3) pré-processamento - tratamento dos textos (em português), com remoção de pontuação, espaços, etc.; 4) classificação - utilização de um algoritmo para classificar os vídeos considerados como educacionais e de qualidade; 5) engenho de recomendação - o sistema recebe os vídeos considerados “bons” e classifica-os; 6) coletor de *feedback* - o usuário avalia a recomendação fornecida pelo sistema por meio de notas, que variam de 1 a 7 estrelas. O trabalho aponta suas principais contribuições nos pontos: 1) o Sistema de Recomendação desenvolvido

pode ser utilizado como solução para vários domínios de aplicação; 2) o sistema serviu como prova de conceito para melhorar as recomendações, por meio de características do *Youtube*, como a avaliações dos usuários, e linguagem nativa do vídeo.

O trabalho apresentado não detalha partes importantes de pesquisa. Por exemplo, para a classificação de vídeos considerados de qualidade, afirma-se que utilizou-se um conjunto de treinamento de 100 vídeos, contendo avaliações realizadas por especialistas e alunos que já tiveram em sua grade curricular o assunto “Orientação a Objetos/Herança”. Porém, não se explica de que maneira esta análise foi realizada, e quais características dos vídeos foram consideradas. Outro ponto de questionamento, é o apontamento de que, devido ao prazo para a realização da pesquisa, o trabalho focou em “algumas características para o experimento”. Tais características não foram descritas.

Thelwall (2018) analisa os comentários de vídeos do *Youtube* relacionados a estilos de dança. A base de dados utilizada contém 36.702 vídeos. O trabalho objetiva identificar, por meio dos comentários postados nos vídeos da plataforma, os tipos de dança, relações quanto aos gêneros (masculino e feminino), sentimentos expressados, e discussões referente aos estilos. Utiliza-se, para tanto, o método denominado *Comment Term Frequency Comparison* (CTFC) na tentativa de identificação de subtópicos/subtemas das discussões sobre um tópico nos comentários do *Youtube*, questões de gênero, sentimentos, e relacionamento entre tópicos. O método utilizado define com sucesso diversas atitudes predominantes em homens e mulheres. Os 10 termos homem-associados foram: *shit, fuck, shuffle, man, fucking, cripp, dude, bro, shuffling, hardstyle*. Por sua vez, os 10 termos mulher-associados foram: *she, amazing, her, beautiful, cute, omg, belly, ballet, really, workout*. A análise de sentimentos forneceu ideias plausíveis dos motivos pelos quais as danças eram apreciadas. Os 10 termos positivos mais utilizados foram: *please, nice, wow, beautiful, loved, job (e.g. nice/great/good job), pretty, hope, perfect, keep (going/up the good work/it up)*. Por sua vez, os 10 termos negativos mais utilizados foram: *shit, fuck, killed, stupid, wtf, hate, idiot, dislike, die, dead*.

Os autores afirmam que os resultados podem servir como ponto de partida para análises mais aprofundadas sobre o tema, e que a pesquisa destacou as diferenças de gêneros, sentimentos, e subtópicos entre as danças. Também apontam que método utilizado pode ser útil para discutir, em larga escala, fenômenos específicos do *Youtube*, como também pode ser útil em outros contextos para fornecer uma análise exploratória inicial de um problema que não havia sido pesquisado anteriormente.

Carvalho *et al.* (2020c) realizam uma análise dos vídeos e das categorias utilizadas pela plataforma *Youtube*. Aponta-se que a plataforma apresenta 3 pontos críticos quanto ao seu mecanismo de busca e sua categorização, a saber: 1) falta de transparência quanto à categorização de um vídeo; 2) informações errôneas quanto à categorização dos vídeos, e 3) impossibilidade de se pesquisar vídeos por suas categorias. Devido a esses pontos críticos, o trabalho conclui que a categorização realizada pela plataforma não pode ser determinante para

a identificação dos vídeos. Ademais, sugerem-se que novas abordagens são necessárias para categorização de vídeos educacionais.

4.3 Metodologia

A metodologia adotada para o desenvolvimento do presente trabalho é apresentada na figura 7 e detalhada a seguir.



Fonte: Elaborado pelo autor

1. **Análise dos vídeos.** Esta etapa consistiu na identificação manual (por meio de visualização) de vídeos educacionais e não educacionais da plataforma. Para a análise dos vídeos, adotou-se a definição de vídeos educacionais proposta por (GOMES, 2008). Segundo o autor, um vídeo educativo é um “produto específico, produzido com intenção didático-pedagógica e que considera seu contexto de recepção como especialmente a escola e a sala de aula, sendo, portanto, intrinsecamente diferente dos vídeos de documentários, entrevistas, reportagens, etc.”. Adotou-se tal definição, pois, acredita-se, que a mesma se mostra abrangente o suficiente para a análise proposta neste trabalho, ao mesmo tempo em que pode ser utilizada para a categorização de um vídeo como educacional ou não.

Nesta etapa foram analisados 200 vídeos, sendo 100 educacionais, e 100 não educacionais. Esse número foi selecionado como ponto de partida para a possível identificação de um vocabulário específico nos comentários de vídeos educacionais do *Youtube*.

2. **Coleta dos comentários.** Esta etapa consistiu em armazenar os identificadores (*IDs*) dos vídeos selecionados na etapa anterior, e coletar os comentários utilizando o aplicativo *Mozdeh Big Data Text Analysis*, também utilizado em (THELWALL, 2018). O aplicativo foi utilizado apenas para a coleta dos comentários, pois apesar de possuir outras funcionalidades, estas são específicas para vocábulos em inglês.
3. **Pré-processamento dos comentários.** Esta etapa consistiu nas seguintes sub-etapas, a saber.
 - **Padronização dos dados:** transformação de todas as letras em maiúsculo para minúsculo.
 - **Remoção de caracteres especiais e *stopwords*:** no primeiro momento foram removidos os caracteres especiais como: !, #, @, dentre outros. Isso é necessário, por exemplo, para que comentários tais como, “muito bom!” e “muito bom”; “melhor professor!” e “melhor professor”, sejam considerados equivalentes. As *stopwords* podem ser consideradas palavras não relevantes para o texto. Estas podem ser artigos, preposições, advérbios, pronomes, e outras palavras auxiliares (MORAIS; AMBRÓSIO, 2007). Alguns exemplos de *stopwords* são “tem”, “isto”, “aos”, dentre outras.
 - **Normalização morfológica (*stemming*):** consiste em remover os prefixos e os sufixos dos vocábulos, mantendo apenas o radical da palavra, ou seja, fazendo com que palavras como “professor” e “professora”, sejam “reduzidas” a “profes”, por exemplo.
4. **Análise dos termos e vocábulos mais frequentes.** Esta etapa consistiu na identificação dos termos e vocábulos mais frequentes nos comentários. Tal passo possibilitou a análise dos comentários para a identificação das diferenças entre os comentários presentes nos vídeos educacionais e não educacionais.

4.4 Resultados

A base de dados é composta por 100 vídeos educacionais e 100 vídeos não educacionais. Os resultados dos termos mais frequentes nos comentários dos vídeos são apresentados na tabela 6. Por sua vez, a tabela 7 apresenta os vocábulos mais frequentes nos comentários dos vídeos.

Considerando os vídeos educacionais, têm-se um total de 77.095 comentários, e 1.002.070 vocábulos. Já os vídeos não educacionais contêm 81.464 comentários e 1.260.833 vocábulos.

Observa-se que os vídeos não educacionais apresentam termos como risadas, que não foram incluídos na tabela 6 por serem irrelevantes ao estudo. Neste sentido, apontam-se

Tabela 6 – Termos mais frequentes nos comentários dos vídeos.

Termos mais frequentes			
Educacional		Não Educacional	
Termo	Quantidade	Termo	Quantidade
muito bom	572	top	152
melhor professor	291	oi	117
ótima aula	138	primeiro	97
obrigado	127	ah é	95
amei	126	muito bom	91
eu	113	eu	61
show	103	salve	58

Fonte: Elaborado pelo autor, com base na pesquisa realizada

Tabela 7 – Vocábulo mais frequentes nos comentários dos vídeos.

Vocábulo mais frequentes			
Educacional		Não Educacional	
Vocábulo	Quantidade	Vocábulo	Quantidade
profes	16421	pra	6363
aul	11828	faz	5666
obrig	7544	músic	4020
melhor	6183	vai	3723
ajud	6087	tod	3658
vc	5073	melhor	3411
aprend	4974	vc	3365

Fonte: Elaborado pelo autor, com base na pesquisa realizada

que os termos “kkkkk” (cinco k’s) e “kkkkkk” (seis k’s), figuram, respectivamente, 65 e 59 vezes, e ocupariam, respectivamente, a sexta e oitava posição.

Ao analisar a tabela 6, percebem-se as diferenças entre os termos mais frequentes. O único termo utilizado nos comentários de ambas as classes de vídeos é “muito bom”, sendo utilizado 572 vezes nos comentários dos vídeos educacionais, e 91 vezes nos comentários dos vídeos não educacionais. Apesar de estar presente nos comentários de ambas as classes de vídeos, o número de vezes em que este termo foi utilizado é bastante diferente: mais de 6 vezes nos comentários dos vídeos educacionais quando comparada à utilização nos comentários dos vídeos não educacionais.

Outras palavras como “amei”, “eu” e “show”, apesar de figurarem mais frequentemente nos comentários dos vídeos educacionais, podem não ser muito indicadas para a utilização para a classificação, por serem muito genéricas. Já os termos como “melhor professor” e “ótima aula”, por serem mais específicos ao contexto educacional, podem ser indicados para a categorização de um vídeo.

Na tabela 7, percebe-se que as únicas palavras que figuram tanto nos comentários dos vídeos educacionais quanto nos comentários dos vídeos não educacionais são “melhor” e “vc”. Contudo, mesmo presente em ambas as classes de vídeos, o número de vezes em que

figuram é muito diferente: “melhor” foi utilizado 6183 vezes, e “vc” 5073 vezes nos comentários dos vídeos educacionais, em comparação a, respectivamente, 3411 vezes e 3365 vezes nos comentários dos vídeos não educacionais.

Os radicais “profes”, “aul”, “obrig”, “ajud”, e “aprend” figuram apenas nos comentários dos vídeos educacionais, sendo, portanto, potencialmente adequados para serem utilizados para a classificação de um vídeo.

Devido as marcantes diferenças entre os vocábulos e os termos mais frequentes nos comentários de vídeos educacionais e não educacionais, aponta-se que o uso dos comentários é uma possível alternativa para a categorização dos vídeos educacionais da plataforma.

4.5 Considerações Finais e Trabalhos Futuros

Este trabalho demonstrou que existem diferenças fundamentais entre os vocabulários dos comentários de vídeos educacionais e não educacionais do *Youtube*. Pôde-se identificar grandes variações nos termos mais frequentes e nas palavras mais utilizadas, demonstrando que os comentários apresentam grande potencial para serem utilizados para categorizar os vídeos educacionais da plataforma.

A continuidade dessa pesquisa se desenvolve em duas frentes, saber: 1) utilização de técnicas de Aprendizagem de Máquina para processar os comentários dos vídeos e classificá-los como educacionais ou não educacionais; 2) desenvolvimento de um sistema *Web* que utiliza os metadados e os comentários dos vídeos para categorizá-los, bem como para mensurar o nível de qualidade dos mesmos.

REFERÊNCIAS

- ABU-EL-HAIJA, S.; KOTHARI, N.; LEE, J.; NATSEV, P.; TODERICI, G.; VARADARAJAN, B.; VIJAYANARASIMHAN, S. Youtube-8m: A large-scale video classification benchmark. **arXiv preprint arXiv:1609.08675**, 2016.
- ADERMANN, N.; BOGGIANO, J. Composer: Dependency manager for php. **URL: <https://getcomposer.org/>**, 2020. Último acesso em 16 de junho de 2020.
- BRAGA, J.; MENEZES, L. **Objetos de aprendizagem, volume 1: introdução e fundamentos**. UFABC, 2014. v. 1. 145 p. Disponível em: <<http://proec.ufabc.edu.br/uab/metdesOA2/2014-BRAGA-livro-oa-v1.pdf>>.
- BURGESS, J.; GREEN, J. **YouTube e a revolução digital**. [S.l.: s.n.], 2009. 24 p.
- CARVALHO, H. C. F. B.; PITANGUI, C. G.; ASSIS, L.; VIVAS, A. Educavídeos: Um sistema de recomendação de objetos de aprendizagem de vídeos educacionais do youtube. In: CONGRESSO BRASILEIRO DE ENSINO SUPERIOR A DISTÂNCIA, XVII., 2020, Goiânia. **2020 XVII Congresso Brasileiro de Ensino Superior a Distância**. Goiânia, 2020.
- CARVALHO, H. C. F. B.; PITANGUI, C. G.; TRINDADE, E. A. C.; ASSIS, L. P. de; ANDRADE, A. V.; SOUZA, D. P. B. de. Categorização de vídeos educacionais do youtube por meio de comentários. **RENOTE**, v. 18, n. 2, p. 621–629, 2020.
- CARVALHO, H. F. B.; PITANGUI, C. G.; TRINDADE, E. C.; ASSIS, L. P. de; ANDRADE, A. V. Learning objects and youtube: an analysis of videos and their categories. In: IEEE. **2020 XV Conferencia Latinoamericana de Tecnologias de Aprendizaje (LACLO)**. [S.l.], 2020. p. 1–6.
- CAZELLA, S. C.; BEHAR, P.; SCHNEIDER, D.; SILVA, K. K. da; FREITAS, R. Desenvolvendo um sistema de recomendação de objetos de aprendizagem baseado em competências para a educação: relato de experiências. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. [S.l.: s.n.], 2012. v. 23, n. 1.
- COX, C. **Listen to this**. [S.l.]: C. Cox Publishing Company, 1922.
- GOMES, L. Vídeos didáticos: uma proposta de critérios para análise. **Revista Brasileira de Estudos Pedagógicos**, v. 89, n. 223, 2008.
- HEILBRON, F. D. C. **Understanding Human Activities at Large Scale**. Tese (Doutorado), 2019.
- IEEE. Ieee learning technology standards committee (ltsc): Ieee p1484.12.1 ieee standard for learning object metadata. 2002. Acessado em: 27 de março de 2019. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1032843>>.
- JÚNIOR, C. B.; DORÇA, F. **Uma Abordagem para a Criação e Recomendação de Objetos de Aprendizagem usando um Algoritmo Genético, Tecnologias da Web Semântica e uma Ontologia**. 2018. 1533 p.
- LERDORF, R. Php: O que é o php? **URL: https://www.php.net/manual/pt_BR/intro-what-is.php**, 2020. Último acesso em 16 de junho de 2020.

MENOLLI, A.; MALUCELLI, A.; REINEHR, S. Criação semi-automática de objetos de aprendizagem a partir de conteúdos da wiki. In: BRAZILIAN SYMPOSIUM ON COMPUTERS IN EDUCATION (SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO-SBIE), XXII., 2011, Aracaju. **Anais**. Aracaju, 2011.

MIRANDA, R. M. d. **GROA: um gerenciador de repositórios de objetos de aprendizagem**. Dissertação (mathesis), 2004.

MORAIS, E. A. M.; AMBRÓSIO, A. P. L. Mineração de textos. **Relatório Técnico–Instituto de Informática (UFG)**, 2007.

NASCIMENTO, P. do; BARRETO, R.; PRIMO, T.; GUSMÃO, T.; OLIVEIRA, E. **Recomendação de objetos de aprendizagem baseada em modelos de estilos de aprendizagem: Uma revisão sistemática da literatura**. 2017. 213 p.

OTTO, M.; THORNTON, J. Bootstrap. URL: <https://getbootstrap.com/>, 2020. Último acesso em 16 de junho de 2020.

PINHEIRO, R. R. A. *et al.* Sistema de recomendação de vídeos educacionais: um estudo de caso no you tube. Universidade Federal de Alagoas, 2018.

PONTES, W. L.; FRANÇA, R. M.; COSTA, A. P. M.; BEHAR, P. **Filtragens de Recomendação de Objetos de Aprendizagem: uma revisão sistemática do CBIE**. 2014. 549 p.

RESIG, J. Jquery. URL: <https://jquery.com/>, 2020. Último acesso em 16 de junho de 2020.

SORGALLA, J. Lity: lightweight, accessible and responsive lightbox. URL: <https://sorgalla.com/lity/>, 2020. Último acesso em 16 de junho de 2020.

THELWALL, M. Social media analytics for youtube comments: Potential and limitations. **International Journal of Social Research Methodology**, Taylor & Francis, v. 21, n. 3, p. 303–316, 2018.

VIEIRA, F. J. R.; NUNES, M. A. S. N. Dica: Sistema de recomendação de objetos de aprendizagem baseado em conteúdo. **Scientia Plena**, v. 8, n. 5, 2012.

WILEY, D. A. **Learning object design and sequencing theory**. Tese (Doutorado) — Brigham Young University, 2000.

YOUTUBE. **API Reference**. 2015. Acesso em: 08/05/2019. Disponível em: <<https://developers.google.com/youtube/v3/docs/?hl=pt-br>>.

YOUTUBE. **Youtube Insights 2017**. 2017. Acesso em: 16 de Abril de 2019. Disponível em: <<https://www.thinkwithgoogle.com/intl/pt-br/youtubeinsights/2017/introducao/>>.

YOUTUBE. **YouTube Insights**. 2019. Acesso em: 17 de Abril de 2019. Disponível em: <<https://youtubeinsights.withgoogle.com>>.

YOUTUBE. **Youtube para imprensa**. 2019. Acesso em: 16 de Abril de 2019. Disponível em: <<https://www.youtube.com/intl/pt-BR/yt/about/press/>>.

5 CONSIDERAÇÕES FINAIS

O Youtube é uma plataforma amplamente utilizada e com um acervo em expansão sobre os mais diversos temas. Desta maneira, estudos específicos sobre a plataforma deveriam ser realizados. O presente estudo demonstra alguns pontos críticos relacionados ao seu mecanismo de busca, a saber: 1) crescimento constante dos conteúdos disponíveis, nem sempre atendendo a necessidade do usuário ou demandando extensiva busca; 2) falta de transparência quando a categorização realizada; 3) categorizações incorretas; 4) impossibilidade de se buscar vídeos por suas categorias e, adicionalmente, é demonstrado maneiras de ultrapassar os problemas identificados.

O trabalho introduz o sistema EducaVídeos, um sistema de recomendação de objetos de aprendizagem que utiliza a API do Youtube e permite a utilização da categoria Education para refinar as buscas. O sistema demonstra que apesar da plataforma não possibilitar o uso das categorias é possível utilizar sua API e refinar as buscas. Observa-se que o sistema, para resultados preliminares, apresenta melhor qualidade no retorno das buscas quando é utilizado com o intuito de assistir vídeos educacionais, ou seja, para as pesquisas analisadas a busca com a categoria Education foi melhor para obter vídeos educacionais que a plataforma Youtube.

Posteriormente, é realizada análise de vídeos do Youtube e a comparação das buscas do Youtube através do navegador Google Chrome em modo anônimo, através do sistema desenvolvido de maneira padrão e através do sistema desenvolvido utilizando a categoria Education. Identifica-se, após a análise, que a plataforma apresenta categorizações errôneas em alguns vídeos, desta maneira, não deve-se considerar a categorização realizada pelo *Youtube* como determinante para a identificação de um vídeo. Portanto, pensa-se ser necessário propor novas maneiras de categorizar os vídeos da plataforma.

Por fim, com o intuito de propor maneiras de sobrepujar os problemas apresentados pensou-se no uso dos comentários como alternativa para a categorização dos vídeos. Para verificação da validade da proposta foi realizada a análise de vídeos educacionais e não educacionais da plataforma e coletados os comentários. Após análise e coleta dos comentários foram utilizadas técnicas de mineração de texto identificando diferenças significativas entre os vocabulários empregados pelos usuários em vídeos educacionais e não educacionais do *Youtube*. Devido as diferenças nos vocábulos mais utilizados e nos termos mais frequentes, constatou-se que os comentários apresentam grande potencial para serem utilizados como alternativa para a categorização dos vídeos do Youtube.

A continuidade dessa pesquisa se desenvolve em três frentes, saber: 1) utilização de técnicas de Aprendizado de Máquina para processar os comentários e classificar os vídeos como educacionais ou não educacionais; 2) desenvolvimento de um sistema de recomendação de objetos de aprendizagem que utiliza os metadados e os comentários dos vídeos para categorizá-los, bem como para mensurar o nível de qualidade dos mesmos; 3) uso de técnicas de aprendizado de máquina para identificação de vídeos quanto ao grau de escolaridade.

