Projeto de Pesquisa

An IT Recommender System Through Knowledge Graph

Patrick Anderson Matias de Araújo

patrick.araujo@students.ic.unicamp.br

Gian Franco Joel Condori Luna

g234826@dac.unicamp.br

Projeto de Pesquisa

Um Sistema Recomendador de TI Usando Grafos de Conhecimento

Patrick Anderson Matias de Araújo

Gian Franco Joel Condori Luna

patrick.araujo@students.ic.unicamp.br

g234826@dac.unicamp.br

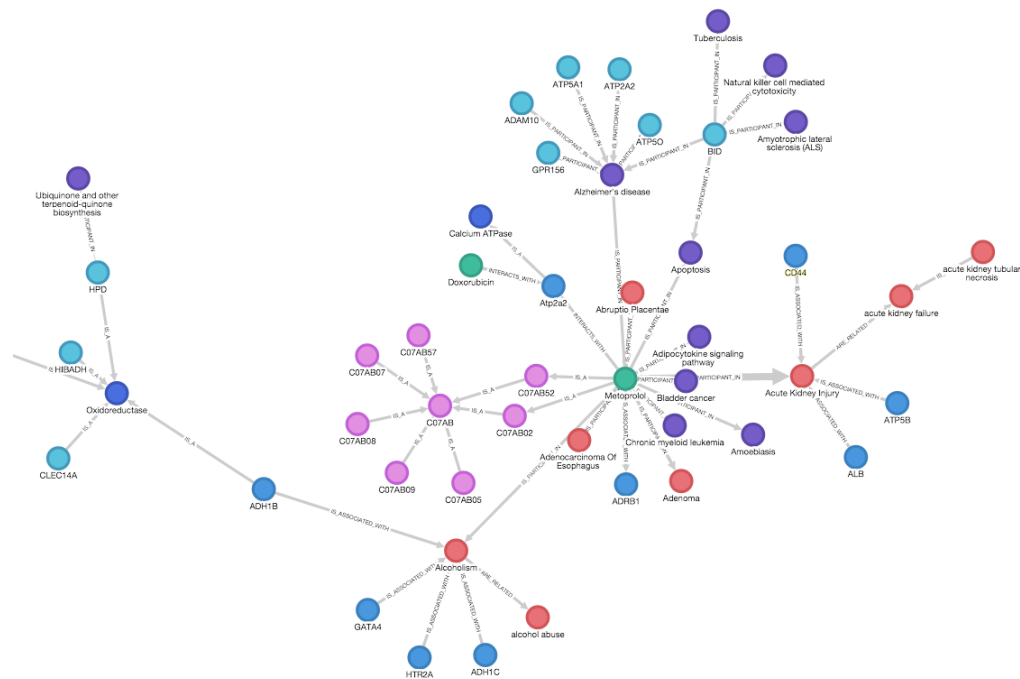# INTRODUÇÃO

- Comércio;
- Troca de mercadorias (escambo);
- Negócio.

- 1946

INTRODUÇÃO

US$ 4,5 TRILHÕES

- Muitas opções;
- Escolher a empresa com a melhor solução;
- Dificuldade de escolha.

INTRODUÇÃO: SISTEMA RECOMENDADOR

INTRODUÇÃO: GRAFO DE CONHECIMENTO

TRABALHOS CORRELATOS

Author	Topic / Focus	Recommender System	Technologies Used
Awangga et al. (2019)	Family Planning in Indonesia	No	XML, OWL, and RDF.
Khallouki, Abatal and Bahaj	Smart Tourism	Yes	OWL, RDFS, DAML+OIL, SPARQL.
Danyaro, Jaafar e Liew (2012)	Meteorological and Oceanographic data	No	XML, OWL, RDF, and SPARQL.
Jain, Mehla and Agarwal (2018)	Provide better response about seismic emergences	Yes	OWL, SQL, JSON, and SPARQL.
Chen et al. (2020)	Protein Entities	No	XML, RESTful, RDF, OWL, SPARQL.

Table 1. The main works that implements ontologies

Ontology design based on data family planning field officer using OWL and RDF.

Rolly Maulana Awangga[*1], Setiawan Assegaff[2], Syafrial Fachri Pane[3], and Muhammad Firman Kahfi[4]

[1,3,4]Applied Bachelor Program in Informatics Engineering, Politeknik Pos Indonesia, Bandung Indonesia
[2]Magister of Information System, STIKOM Dinamika Bangsa, Jambi, Indonesia
[*]Corresponding author, e-mail: awangga@poltekpos.ac.id

Abstract

Population density in Indonesia is ranked fourth in the world. The impact of a large population will affect the level of welfare of the community to decrease, and the number of unemployment is increasing so that the state makes Family Planning Program (PLKB) to control the rate of population growth. Problems in the PLKB program are on knowledge management and mapping from data contraception, counseling and planning so that this research using Ontology method will aim to do mapping with knowledge management and Ontology design shows represented data to relate and describes the resources contained in family planning data. This research approach the representation of ontology that is validated through model transformation from family planning data to ontology design using OWL and RDF which are useful for data processing and representing data to be utilized by field officers in educating the public and eradicating negative issues about family planning programs.

Keyword: Family Planning, Ontology, Resources Description Framework, Field Officer, Web Ontology Language, Representation.

An Ontology-based Context awareness for Smart Tourism Recommendation System

Hajar Khallouki
Hassan 1 University, FSTS
Settat
Morocco
hajar.khallouki@gmail.com

Ahmed Abatal
Hassan 1 University, FSTS
Settat
Morocco
abatalahmed@gmail.com

Mohamed Bahaj
Hassan 1 University, FSTS
Settat
Morocco
mohamedbahaj@gmail.com

ABSTRACT

Smart tourism concept appeared with the development of Smart Cities. Bringing Smartness into Tourism needs a dynamic and interconnected system on which information relating to tourism activities could be exchanged in real time. In this paper, we introduce a new approach for designing mobile tourism recommendation system using context awareness. The proposed approach combines Internet of Things (IoT) technologies with semantic web services to predict the tourist real-time context and provide the suitable services.

KEYWORDS

Smart tourism, tourism recommendation system, context awareness, IoT, semantic web, real-time context.

1 INTRODUCTION

With the evolution of mobile devices (laptops, tablets, smartphones, etc.), it becomes possible to offer online services to people anytime and anywhere. Such online services are peculiarly utile for people in places they have never been to before.

Mobile tourism systems are making it possible for tourists to plan and experience tourism. Travelers could simply use their

The context awareness is a key element for creating smart applications. Several aspects of context aware solutions are elaborated and applied in recommendation systems.

The collected contextual data is generally raw information from different distributed sources, which needs to be interpreted. Based on ontologies, it is possible to elaborate knowledge models which will be enriched by this raw data and thus not only to increase their level of semantic representation but especially of being able to use them to make automatic service recommendation for tourism.

The key contributions of this paper include:

- Building the smart tourism recommendation system.
- Context aware-based ontology modeling.
- Ontology evaluation and validation using a set of SPARQL queries.
- Describing the recommendation process by proposing an algorithm.

The rest of the paper is organized as follows. Section 2 introduces a description of related work on smart tourism and context aware solutions. Section 3 presents the system architecture and its components. Section 4 outlines the context modeling and reasoning through ontology. Finally, Section 5 concludes the paper.

An RDF Model for Meteorological and Oceanographic Information Systems

Kamaluddeen Usman Danyaro[1], Jafreezal Jaafar[1], Mohd Shahir Liew[2]

[1]Department of Computer & Information Sciences
[2]Department of Civil Engineering
Universiti Teknologi PETRONAS, Seri Iskandar, 31750 Tronoh,Perak, Malaysia
kamaluddeen_g01325@utp.edu.my jafreez@petronas.com.my, shahir_liew@petronas.com.my

Abstract—This paper suggests an RDF Model for effective management of distributed data on the Semantic Web. Resource Description Framework (RDF) is a data model that unifies aptly structured data. The Web is facing great challenge in data retrieval due to the increase and request of information from different sources. This brings the problem of information overload. In minimizing this problem, we provide a knowledge model for structuring data which is the first step of building Semantic Web. In particular, to elucidate the model, Meteorological and Oceanographic (MetOcean) data have been used. We present the result using SPARQL query method. Thus, the finding shows that the model is ideal for querying and data representation of distributed data.

Keywords-Resource Description Framework (RDF);
RDFModel; Semantic Web; MetOcean

I. INTRODUCTION

Since the beginning of Resource Description Framework (RDF) in 1990s, various processes have been followed in order to come up with the best model or structure for language specification. In 2004, there have been extended RDF specifications for representing Semantic data which was governed by World Wide Web Consortium (W3C). RDF is an essential root of Semantic Web that relates to different things over the networks. It manages and deals with the distributed data which provides ways for data representation [9]. Nevertheless, RDF is a language for specifying languages. It is also one of the three

and the *human*. This provides common syntax that makes up documents. Although, XML and database provide the consistency of each Web page which initiated the data model [9]. However, XML as a data model alone cannot handle huge amount of data as well as system interoperability. This is because that XML provides syntax not Semantics for data representation on the Web. Nevertheless, many Web database systems are not on link data model structure. In particular, the database of Meteorological and Oceanographic (MetOcean) information system is built on XML and XML metadata only [12].

MetOcean is a meteorological and oceanographic industry that deals with a lot of data and metadata. Many companies and research industries rely on its data which has been built on XML schema. However, these have less capabilities of describing data or metadata in meaningful way and later bring the problem of information overload.

The objective of this paper is to define and describe the distributed data of MetOcean in meaningful way. In view of this, we built a network graph store or a triplestore. A triplestore is a semantic repository that stores graphs [7]. In other words, it is an RDF database built for semantic data. More specifically, the paper attempts to present the semantic Web vocabularies' viewpoints of RDF that will be used in Meteorological data.

An Ontology Based Earthquake Recommendation System

Sarika Jain[1], Sonia Mehla[1], and Apoorv Gaurav Agarwal[2]([✉])

[1] National Institute of Technology, Kurukshetra, India
jasarika@nitkkr.ac.in, soniamehla9@gmail.com
[2] Birla Institute of Technology, Mesra, India
apoorv.agarwal.aga@gmail.com

Abstract. According to geographical statistics, 54% of the land in India is vulnerable to earthquakes, and several populous regions fall under Zone 4 or 5 in levels of seismicity. It is believed that a comprehensive earthquake recommendation system will significantly reduce the number of emergency services that are dispensed when disaster strikes and will also effectively minimize casualties. Various techniques have been employed to successively create such a recommendation system which provides efficiency in managing the dispensing of the emergency services. The paper focuses on determining the immediate course of action that should be taken during emergencies like earthquake to mitigate the damages to life and property using Ontology supported rule-based reasoning and case-based reasoning, i.e. the actions are determined using similar cases from the past, if any, and validated through existing rules and experience.

Keywords: Earthquake · Earthquake recommendation system
Ontology · Rule-based reasoning · Case-based reasoning

SCIENTIFIC DATA

Check for updates

OPEN

ARTICLE

Protein ontology on the semantic web for knowledge discovery

Chuming Chen [ID][1][✉], Hongzhan Huang[1], Karen E. Ross[2], Julie E. Cowart [ID][1], Cecilia N. Arighi [ID][1], Cathy H. Wu[1,2] & Darren A. Natale[2]

The Protein Ontology (PRO) provides an ontological representation of protein-related entities, ranging from protein families to proteoforms to complexes. Protein Ontology Linked Open Data (LOD) exposes, shares, and connects knowledge about protein-related entities on the Semantic Web using Resource Description Framework (RDF), thus enabling integration with other Linked Open Data for biological knowledge discovery. For example, proteins (or variants thereof) can be retrieved on the basis of specific disease associations. As a community resource, we strive to follow the Findability, Accessibility, Interoperability, and Reusability (FAIR) principles, disseminate regular updates of our data, support multiple methods for accessing, querying and downloading data in various formats, and provide documentation both for scientists and programmers. PRO Linked Open Data can be browsed via faceted browser interface and queried using SPARQL via YASGUI. RDF data dumps are also available for download. Additionally, we developed RESTful APIs to support programmatic data access. We also provide W3C HCLS specification compliant metadata description for our data. The PRO Linked Open Data is available at https://lod.proconsortium.org/.

Objetivo Geral

▶ - Criação de um sistema de recomendação para o setor de tecnologia da informação brasileiro por meio de texto não estruturado.

Objetivos específicos

- Realização de web scraping de páginas web do setor de TI brasileiro.

- Faça uma mineração de texto dos dados obtidos por web scraping.

- Criar uma ontologia sobre o contexto de compradores e vendedores de serviços de TI adicionando as informações obtidas na mineração de texto.
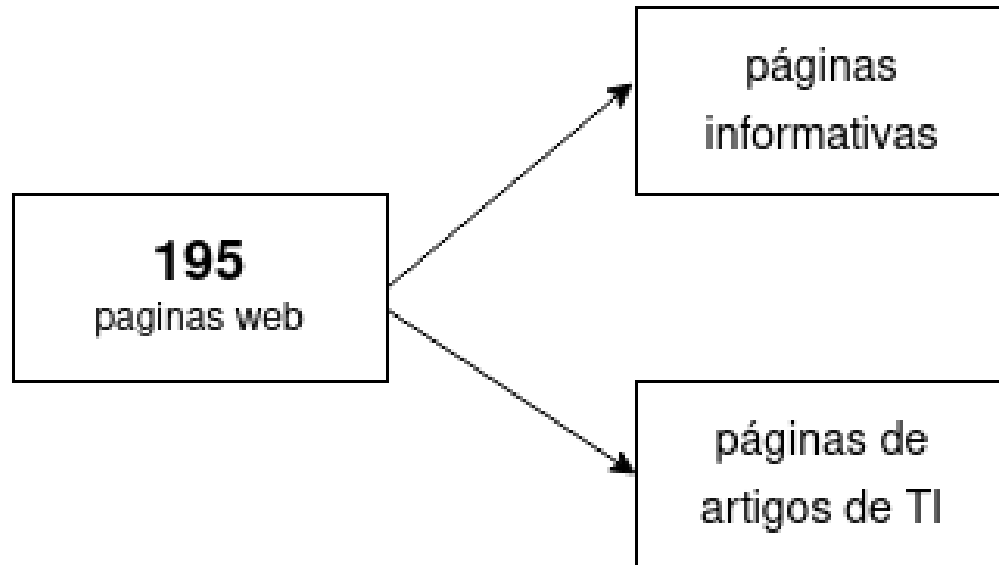
Web scraping
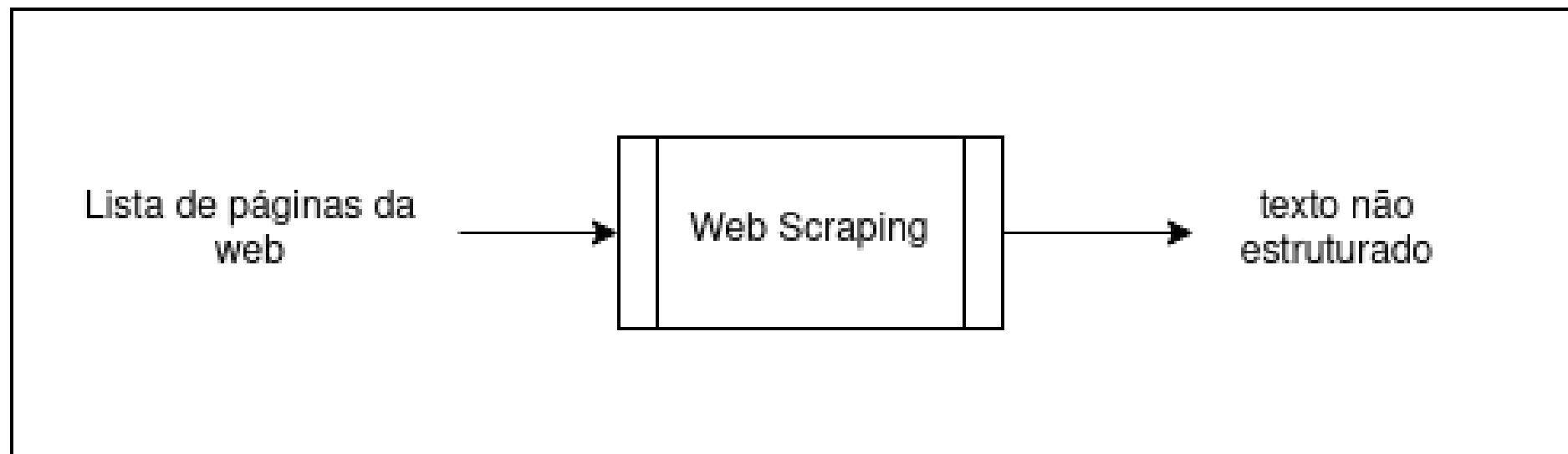
Listas de paginas da web

- ▶ - Buscas na web e também utilizamos nosso conhecimento de empresas dedicadas a este setor no Brasil.

Scraping

	Easy to use	Asynchronous	Big extraction data	Better Performance
Beautiful Soup	yes	no	no	no
Selenium	yes	no	no	yes
Scrapy	no	yes	yes	yes

Table 1: Comparison between the most used tools for web scraping.

Web scraping

Paginas informativas

Páginas de artigos de TI

Dataset TI

Estrutura do dataset coletado

title	paragraph	url	date
Gerenciamento total e inteligente de chamados ...	Registro Manual Autoatendimento Facebook Dire...	https://www.█ .com/plataforma-neo/geren...	05/07/2022, 16:20:40
Otimize seu atendimento de chamados via Facebook	Registro Manual Autoatendimento Facebook Aprov...	https://www. .com/plataforma-neo/faceb...	05/07/2022, 16:20:40
Atenda chamados também pelo Instagram!	Registro Manual Autoatendimento Facebook Direc...	https://www. com/plataforma-neo/insta...	05/07/2022, 16:20:41
O que seu cliente precisa... em alguns cliques!	Registro Manual Autoatendimento Ao adotar a F...	https://www. .com/plataforma-neo/autoa...	05/07/2022, 16:20:41
Potencialize a performance do seu time com o m...	Diminua a quantidade de carrinhos abandonados...	https://www.█ com/plataforma-neo/chat/	05/07/2022, 16:20:41
Interaja com seus clientes por vídeo em tempo ...	Registro Manual Autoatendimento Facebook Dire...	https://www. .com/plataforma-neo/video...	05/07/2022, 16:20:41
Gerencie seus chamados de qualquer lugar com o...	Registro Manual Autoatendimento Facebook Dire...	https://www.█ .com/plataforma-neo/telef...	05/07/2022, 16:20:41
Atenda seu cliente 24 horas por dia, 7 dias po...	Reduza custos e aumente resultados com o módul...	https://www. com/plataforma-neo/chatbot/	05/07/2022, 16:20:41
Seu atendimento ao cliente presente também no ...	Registro Manual Autoatendimento Facebook O Wha...	https://www.█ .com/plataforma-neo/whats...	05/07/2022, 16:20:41
NaN	Email: comercial@topwavetecnologia.com.br Ins...	https://www. .com.br	05/07/2022, 16:20:41
A solução que o seu atendimento precisa. A exp...	WhatsApp Chatbot Chat Workflow Registro Manua...	https://www.neoassist.com/plataforma-neo/	05/07/2022, 16:20:41

Mineração de texto

texto não estruturado → **Mineração de texto** →
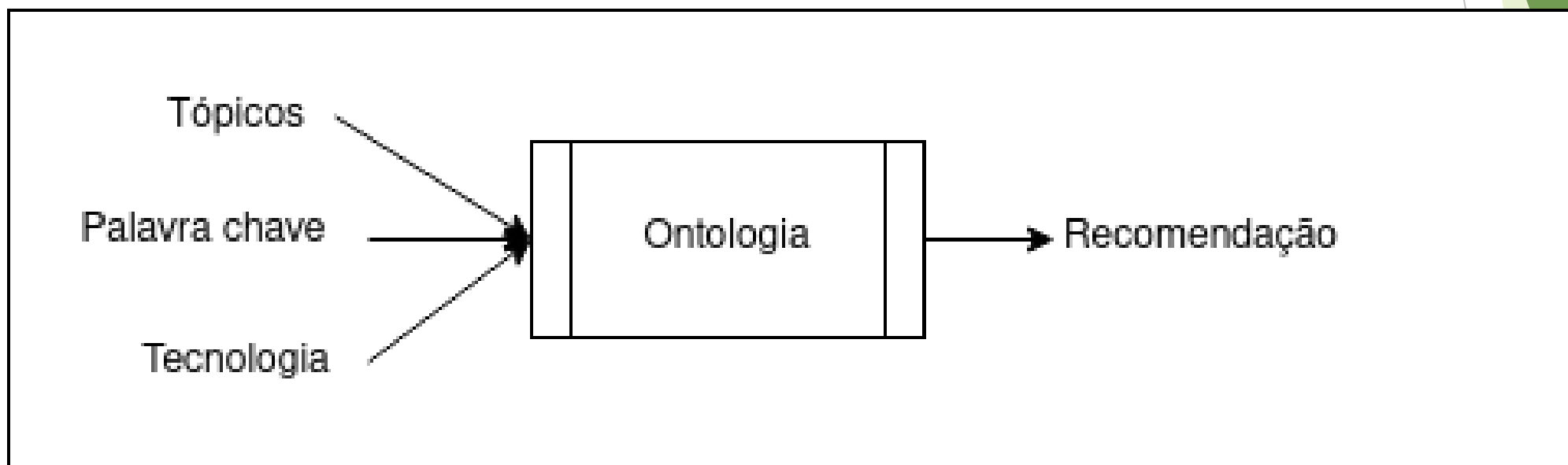- Tópicos
- Palavra chave
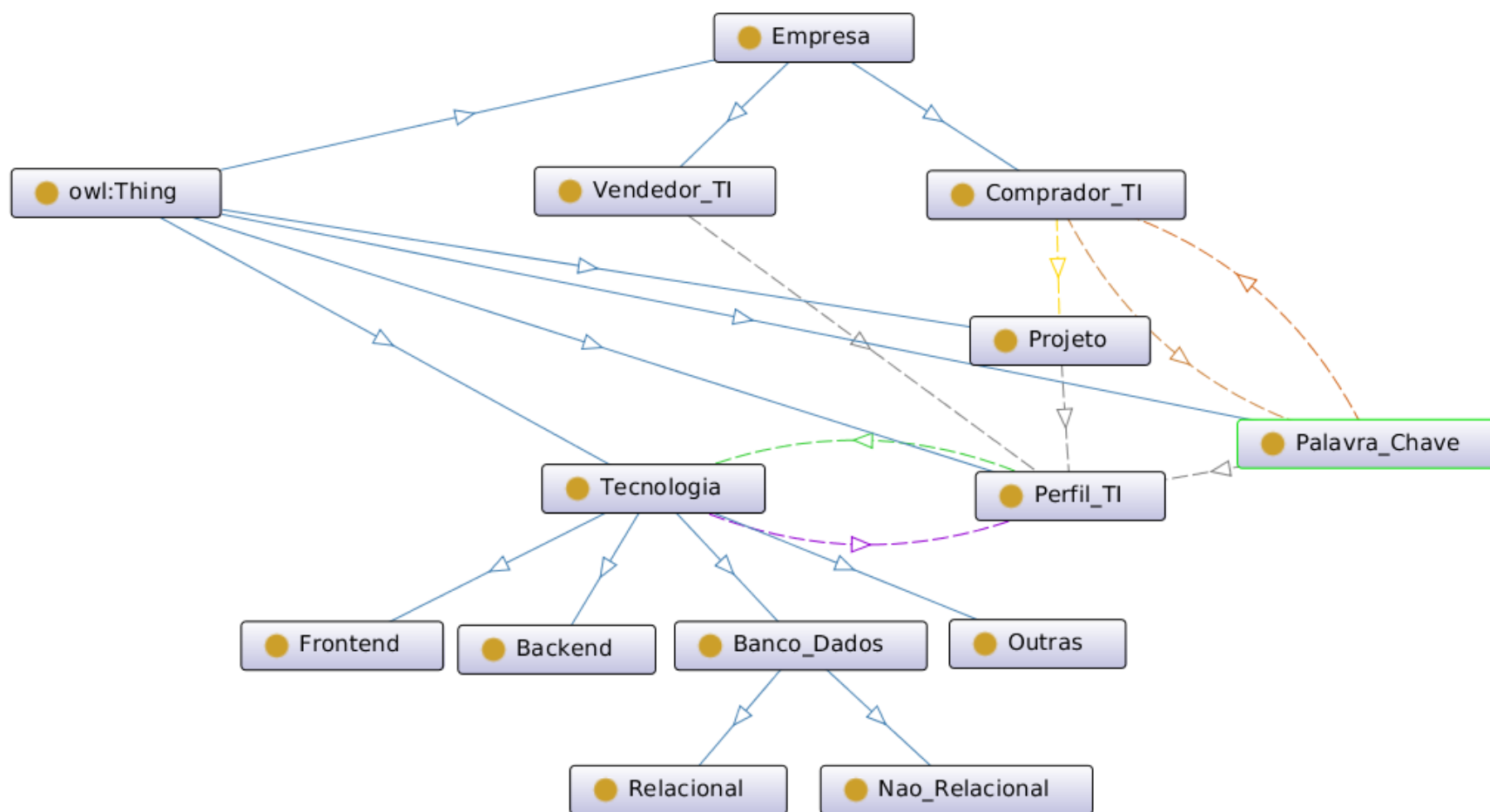- Tecnologia
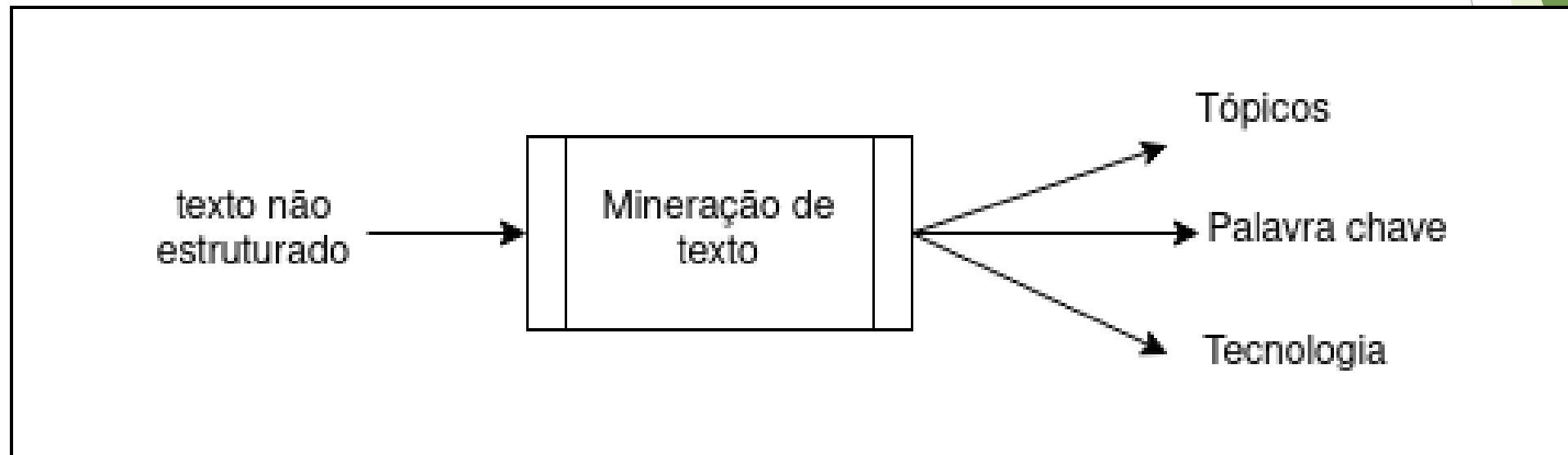
Topicos
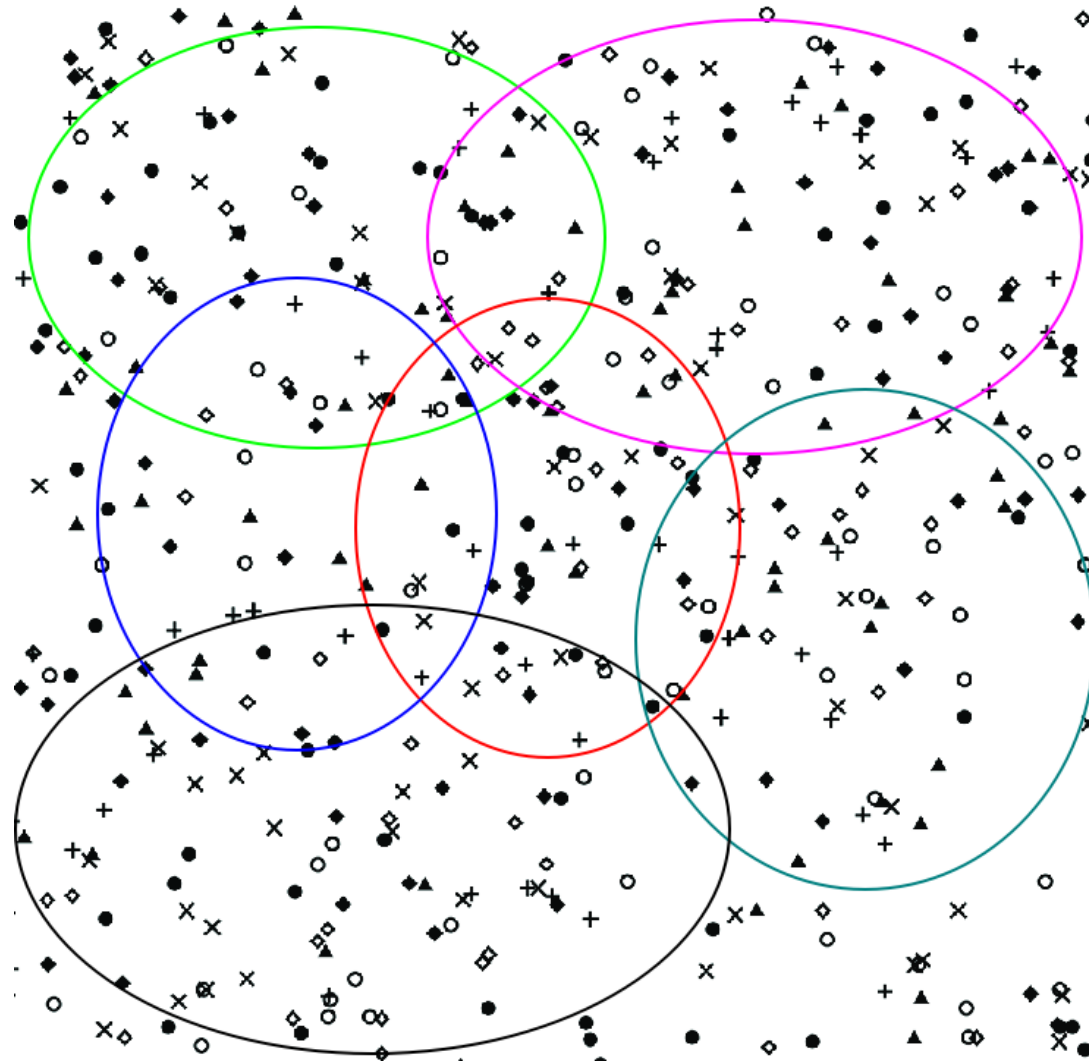
Palavras chave

Tecnologia

- Dados obtidos do dataset.

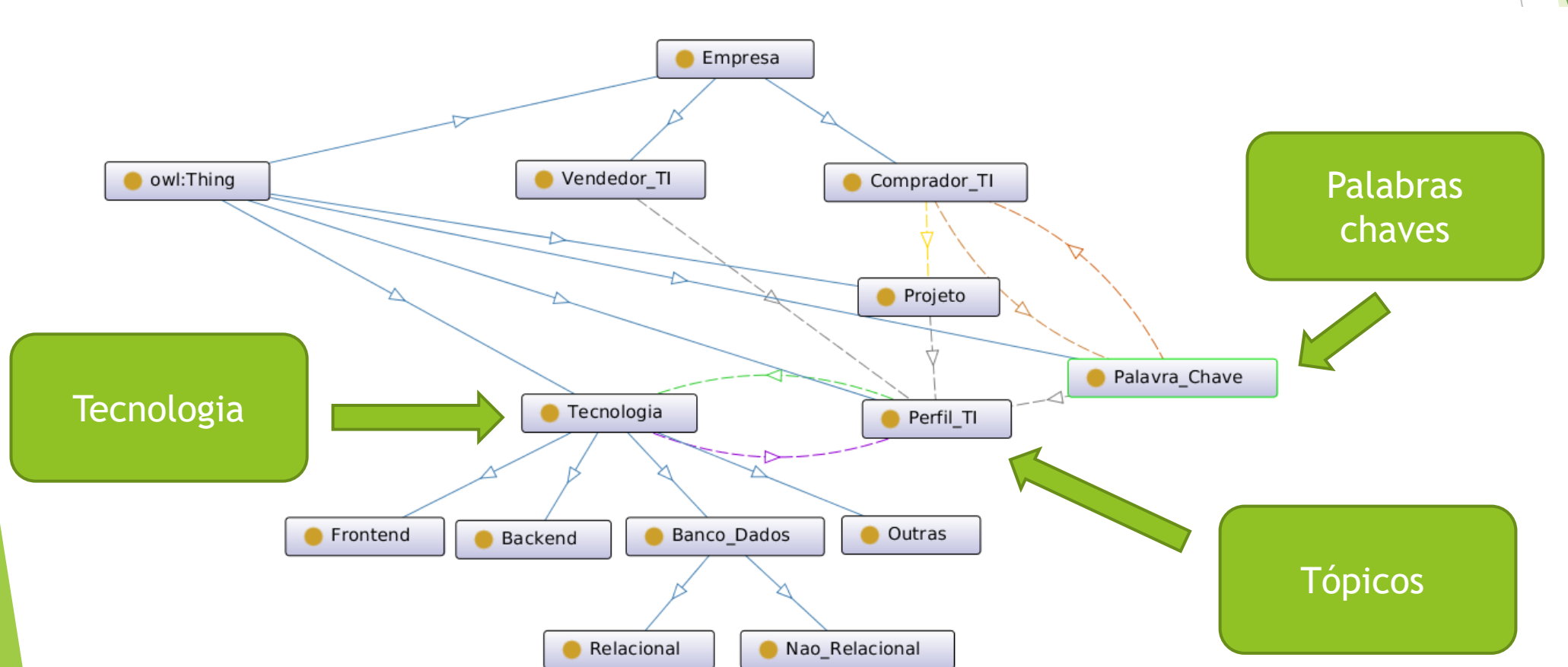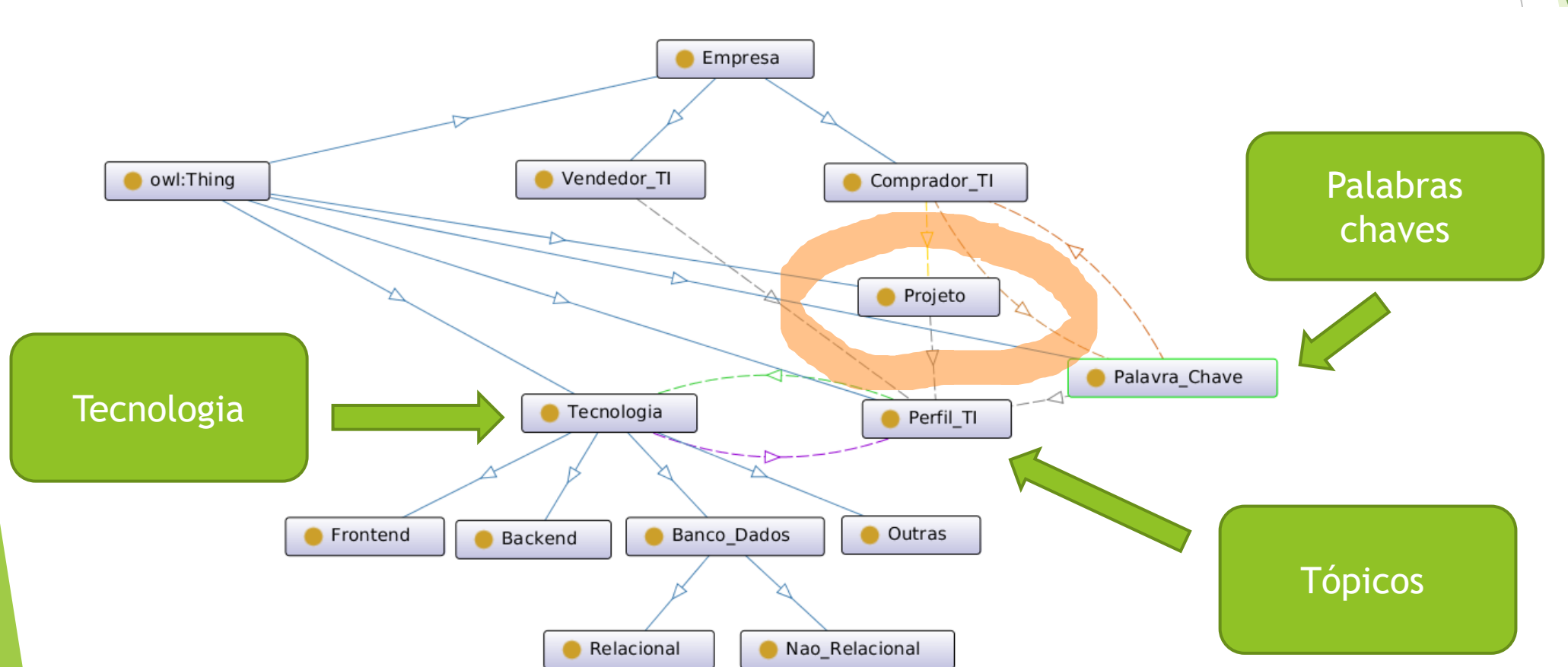- Dados de pessoas que conhecem este setor.

Ontologia

Grafo de conhecimento

Grafo de conhecimento

Grafo de conhecimento

Exemplo do resumo

Preciso fazer uma página web que possa ajudar minha empresa a ter uma melhor presença entre meus usuários. Para o qual precisarei de um servidor onde possa armazená-lo e trabalharei com PHP, MYSQL e HTML
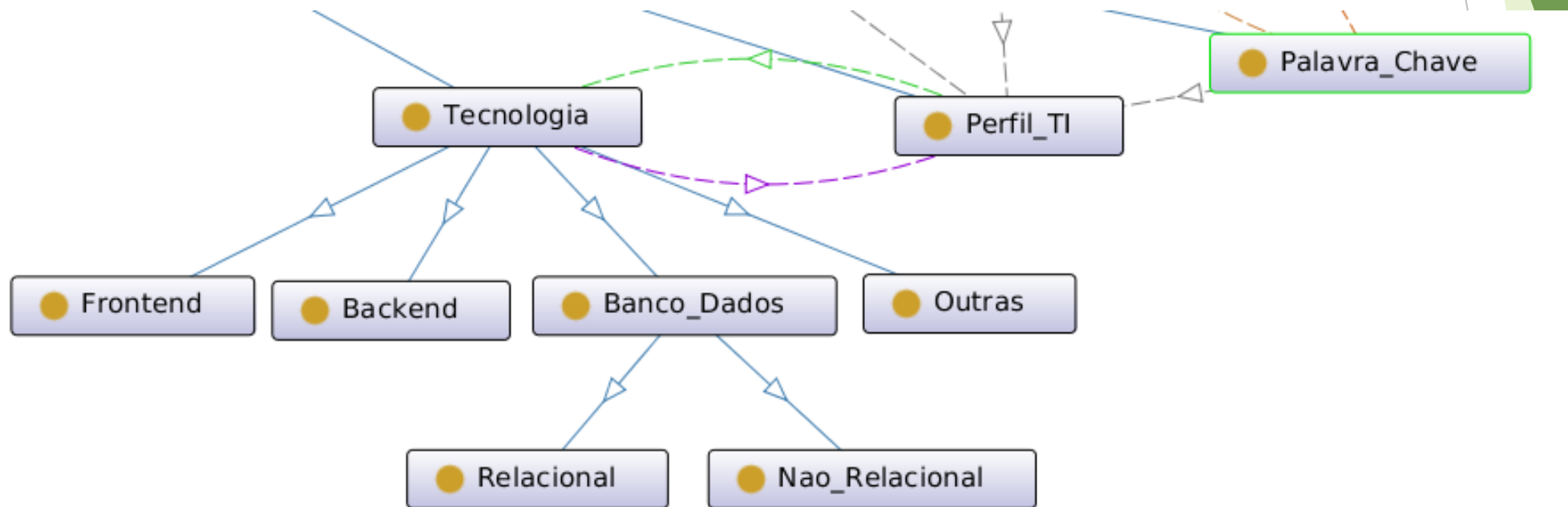
Exemplo do resumo

Preciso fazer uma página web que possa ajudar minha empresa a ter uma melhor presença entre meus usuários. Para o qual precisarei de um servidor onde possa armazená-lo e trabalharei com PHP, MYSQL e HTML
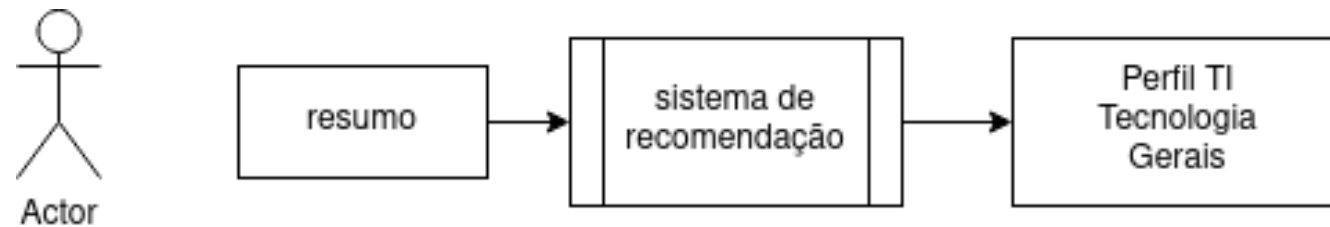
Palavras chave
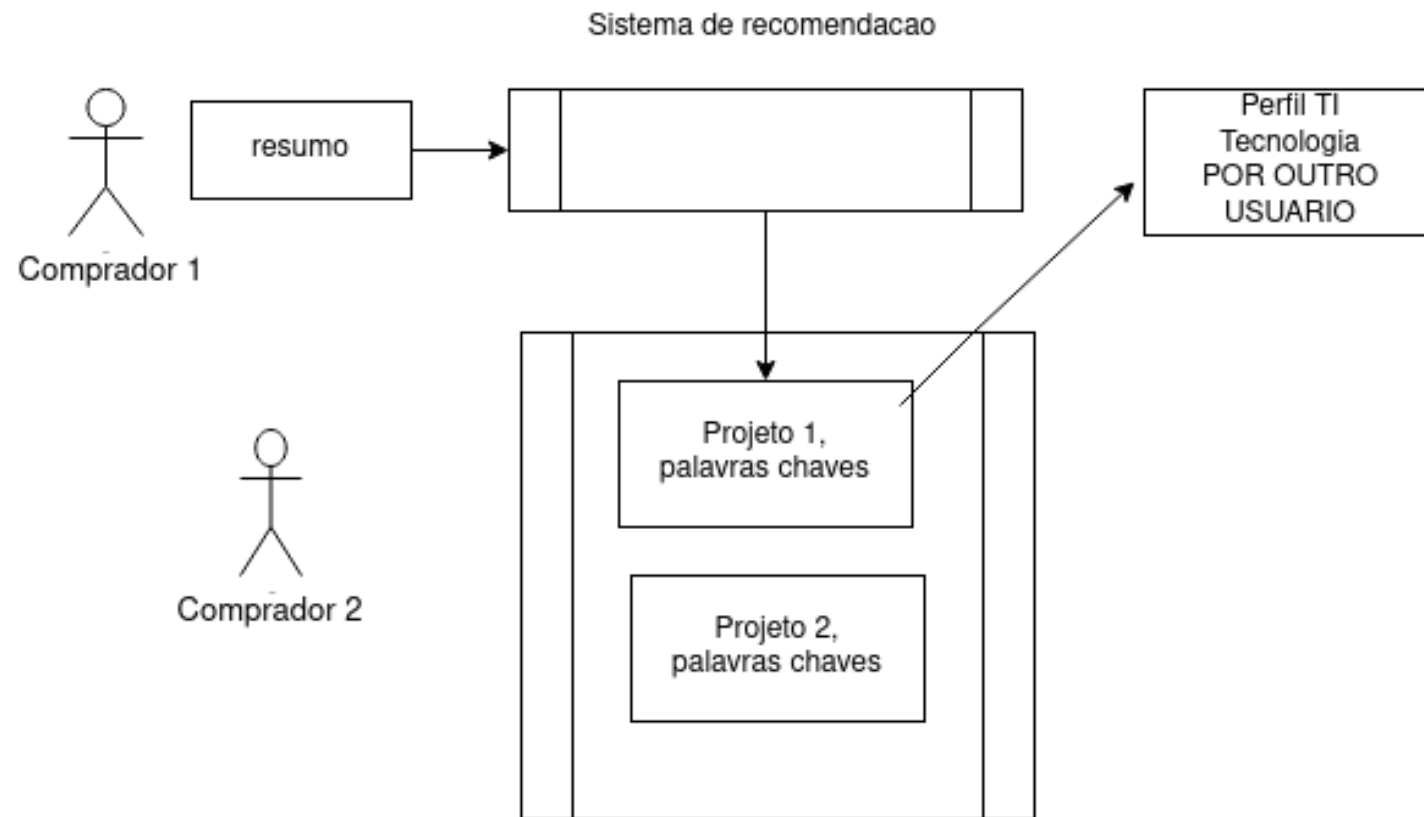
Fluxo de recomendação

Recomendações

- Geral

Recomendações

- Por outro usuario

Recomendações

- Personalizada
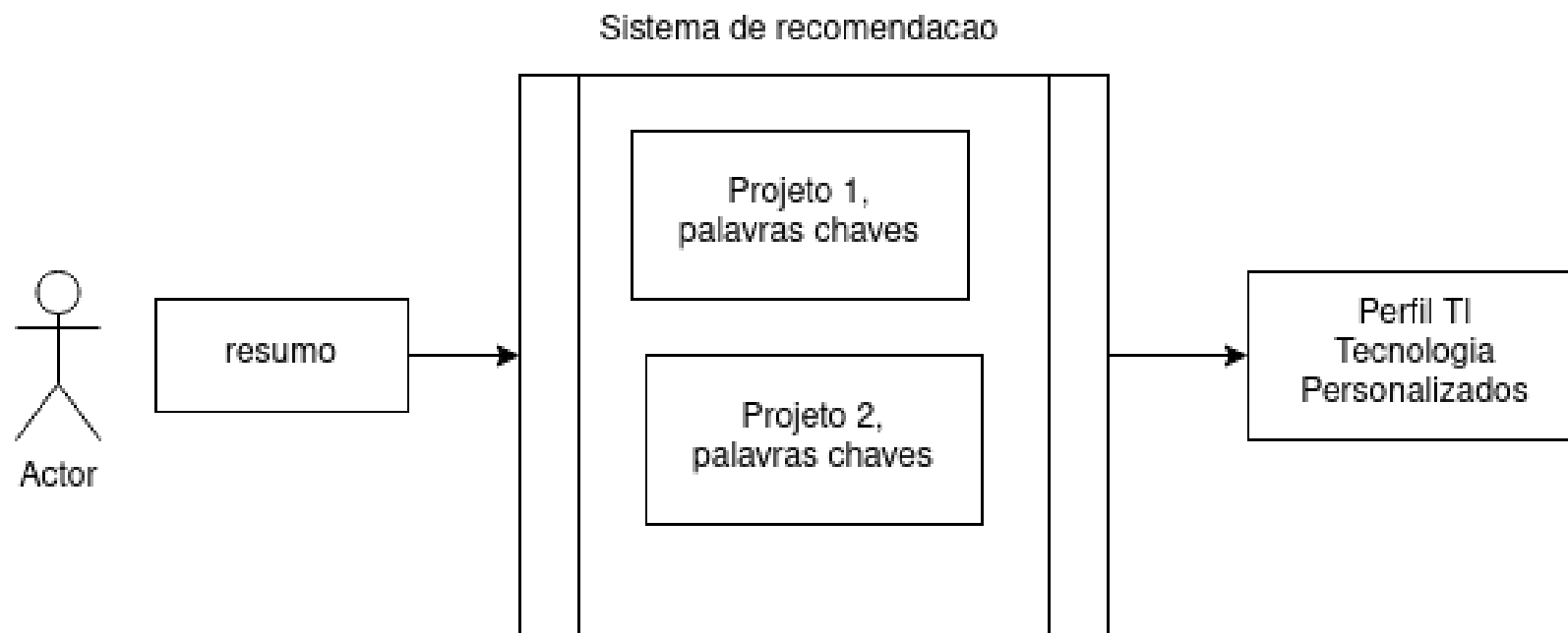
METODOLOGIA

▶ Para atingir o proposto por esse projeto e alcançar os objetivos, é necessário seguir alguns procedimentos metodológicos (atividades):

I. Realizar um levantamento bibliográfico;

II. Investigar o contexto sobre serviços de TI e procurar por páginas web sobre o contexto;

III. *Web Scrapping* nas páginas procuradas;

IV. Explorar algoritmos de clusterização;

V. Estudar, criar e implementar uma ontologia no domínio de serviços de TI, checando sua consistência;

VI. Análise dos resultados obtidos;

VII. Integrar e testar o grafo de conhecimento com o sistema recomendador usando as consultas criadas;

VIII. Escrever relatórios e artigos científicos sobre a implementação e os seus resultados;

CRONOGRAMA

▶ Baseado nos procedimentos metodológicos apresentados, nessa Tabela 2 é apresentado o cronograma para as atividades mencionadas

Tabela 2. Cronograma das atividades previstas

Year \ Month Activity	2022				2023											
	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
I	X	X	X	X												
II	X	X	X													
III		X	X													
IV		X	X	X	X	X										
V		X	X	X	X	X	X									
VI					X	X	X	X	X							
VII									X	X	X	X	X			
VIII											X	X	X	X	X	X

```
+function(a){
[b]())}}var c=function(b){this.element=a(b)};c.VERSION
pdown-menu)"),d=b.data("target");if(d||(d=b.attr("href"),d=d&&d.replace(/.*(?=#[^\s]*$)/,""))
a"),f=a.Event("hide.bs.tab",{relatedTarget:b[0]}),g=a.Event("show.bs.tab",{relatedTarget:e[
ultPrevented()){var h=a(d);this.activate(b.closest("li"),c),this.activate(h,h.parent(),functi
igger({type:"shown.bs.tab",relatedTarget:e[0]})})})}},c.prototype.activate=function(b,d,e){fun
>.active").removeClass("active").end().find('[data-toggle="tab"]').attr("aria-expanded",!1)
a-expanded",!0),h?(b[0].offsetWidth,b.addClass("in")):b.removeClass("fade"),b.parent(".dropd
.find('[data-toggle="tab"]').attr("aria-expanded",!0),e&&e()}var g=d.find("> .active"),h=e&&
")||!!d.find("> .fade").length);g.length&&h?g.one("bsTransitionEnd",f).emulateTransitionEnd
var d=a.fn.tab;a.fn.tab=b,a.fn.tab.Constructor=c,a.fn.tab.noConflict=function(){return a.fn.
how")};a(document).on("click.bs.tab.data-api",'[data-toggle="tab"]',e).on("click.bs.tab.dat
e strict";function b(b){return this.each(function(){var d=a(this),e=d.data("bs.affix"),f="o
ypeof b&&e[b]()})}var c=function(b,d){this.options=a.extend({},c.DEFAULTS,d),this.$target=
,a.proxy(this.checkPosition,this)).on("click.bs.affix.data-api",a.proxy(this.checkPosition
ll,this.pinnedOffset=null,this.checkPosition()};c.VERSION="3.3.7",c.RESET="affix affix-top
tate=function(a,b,c,d){var e=this.$target.scrollTop(),f=this.$element.offset(),g=this.$t
bottom"==this.affixed)return null!=c?!(e+this.unpin<=f.top)&&"bottom":!(e+
=c&&e<=c?"top":null!=d&&i+j>=a-d&&"bot
RESET).addCl
```

REFERÊNCIAS

▶ Ali, Farman, et al. (2019) "Transportation Sentiment Analysis Using Word Embedding and Ontology-Based Topic Modeling". In: *Knowledge-Based Systems*, vol. 174, page 27–42. https://doi.org/10.1016/j.knosys.2019.02.033.

▶ Awangga, Rolly Maulana, et al. (2019) "Ontology design based on data family planning field officer using OWL and RDF". In: *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 17, page 161, https://doi.org/10.12928/telkomnika.v17i1.9237.

▶ Berners-Lee, T., Hendler, J., & Lassila, O. (2001) "The Semantic Web. A New Form of Web Content That Is Meaningful to Computers Will Unleash a Revolution of New Possibilities". In: *Scientific American*, 284, pages 1-5.

▶ Chen, C., Huang, H., Ross, K.E. et al. (2020) Protein ontology on the semantic web for knowledge discovery. In: *Sci Data* 7, 337. https://doi.org/10.1038/s41597-020-00679-9.

▶ Dos Reis, Julio Cesar; Martins, Tulio Brandão Soares. (2022) "Handling Multi-chapter Inconsistencies in DBpedia Evolution". In: *SIMPÓSIO BRASILEIRO DE BANCO DE DADOS (SBBD)*, 37, 2022, Búzios. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, pages 128-137. ISSN 2763-8979. https://doi.org/10.5753/sbbd.2022.224307.

▶ Hitzler, Pascal. (2021) "A Review of the Semantic Web Field". Communications of the ACM, vol. 64, pages 76–83. https://doi.org/10.1145/3397512.

▶ Introduction — Owlready2 0.36 documentation. (2022). https://owlready2.readthedocs.io/en/v0.37/intro.html.

▶ Jain, S., Mehla, S., Agarwal, A.G. (2019). An Ontology Based Earthquake Recommendation System. In: Luhach, A., Singh, D., Hsiung, PA., Hawari, K., Lingras, P., Singh, P. (eds) *Advanced Informatics for Computing Research*. ICAICR 2018. Communications in Computer and Information Science, vol 955. Springer, Singapore. https://doi.org/10.1007/978-981-13-3140-4_30

REFERÊNCIAS

▶ Jupp, Simon, et al. (2016) "The Cellular Microscopy Phenotype Ontology". In: *Journal of Biomedical Semantics*, vol. 7, page 28. https://doi.org/10.1186/s13326-016-0074-0.

▶ K. U. Danyaro, J. Jaafar and M. S. Liew; (2012) "An RDF model for meteorological and oceanographic information systems". In: *International Conference on Computer & Information Science (ICCIS)*, 2012, page. 480-484, https://doi.org /10.1109/ICCISci.2012.6297293.

▶ Kontopoulos, Efstratios, et al. (2013) "Ontology-Based Sentiment Analysis of Twitter Posts". In: *Expert Systems with Applications*, vol. 40, pages 4065–74. https://doi.org/10.1016/j.eswa.2013.01.001.

▶ Kume, S., Masuya, H., Kataoka, Y., & Kobayashi, N. (2016). Development of an Ontology for an Integrated Image Analysis Platform to enable Global Sharing of Microscopy Imaging Data. In: *ISWC2016 The 15th International Semantic Web Conference*. https://doi.org/10.48550/arXiv.2110.10407.

▶ Michael Uschold (2001). Where are the semantics in the semantic web. In: *AI Magazine*, 24, 2003.

▶ Noy, Natasha, et al. (2019) "Industry-Scale Knowledge Graphs: Lessons and Challenges". In: *Communications of the ACM*, vol. 62, pages 36–43. https://doi.org/10.1145/3331166.

▶ Protégé. (2022). https://protege.stanford.edu/.

▶ Semantic Web - W3C. (2022). https://www.w3.org/standards/semanticweb/.