

Recomendação de Classes de Ativos baseada no Perfil do Investidor com uso de Ontologia de Investimentos

MO656 - Introdução à Web Semântica

André Gomes Regino¹ (230252), Eduardo Yuji Sakabe¹ (166810),
Fabio Fogliarini Brolesi¹ (023718), Paula Jeniffer dos Santos Viriato¹ (234831)

¹Instituto de Computação – Universidade Estadual de Campinas (Unicamp)
Campinas – SP – Brasil

{a230252, e166810, f023718, p234831}@dac.unicamp.br

Resumo. O crescimento expressivo no número de investidores na Bolsa de Valores brasileira é incontestável nos últimos anos. A diminuição da taxa das corretoras, a democratização de investimentos em títulos públicos, bem como a popularização do pregão online para ativos de renda variável contribuíram para o atual estado pujante de negociação de ativos. Novas aplicações automatizadas foram criadas aproveitando esta demanda a fim de acompanhar e gerenciar os ativos, uma vez que a complexidade do gerenciamento da compra e venda de ativos financeiros cresce em razão da quantidade de operações feitas com estes ativos. Entretanto, não há um padrão ou base de dados que centraliza informações sobre estes ativos de forma confiável, aberta e escalável, forçando estas aplicações a criarem cada uma sua própria base de dados, muitas vezes difícil de manter e demandando constante atualização. O objetivo deste relatório é o de descrever uma base de conhecimento do tema de investimentos brasileiros, bem como exibir uma aplicação prática em funcionamento usando esta base. A aplicação é capaz de traçar o perfil de investimento do usuário a partir de um formulário de questões a ele fornecido. Adicionalmente, a aplicação usará as respostas fornecidas pelo usuário para recomendar classes de ativos que mais se encaixam ao perfil. Resultados preliminares demonstram que a ontologia desenvolvida pode ser usada como base de conhecimento para um conjunto de aplicações da área de investimentos.

1. Introdução

O mercado de investimentos pessoa física no Brasil cresce a passos largos, com crescimento de mais de 43% no número de investidores somente no primeiro semestre de 2021 como apontado pela B3¹. Além disso, possui ainda um grande espaço para crescer, já que somente 3% dos brasileiros investem na bolsa, enquanto países como Estados Unidos e Japão esse índice chega a 55% e 45%, respectivamente².

Este crescimento é acompanhado pelo volume de investimento aplicado no mercado de *fintechs*³, gerando fomento a criação de novas aplicações para o mercado de investimentos pessoais.

¹https://www.b3.com.br/pt_br/noticias/porcentagem-de-investidores-pessoa-fisica-cresce-na-b3.htm

²<https://valorinveste.globo.com/mercados/renda-variavel/bolsas-e-indices/noticia/2021/01/15/apenas-3percent-dos-brasileiros-investiram-em-acoes-em-2020-e-media-aplicada-caiu-31percent.ghtml>

³<https://www.pwc.com.br/pt/estudos/setores-atividade/financeiro/2021/pesquisa-fintech-deep-dive-2020.html>

Naturalmente, o investidor brasileiro busca mais informação sobre ativos e passa a adotar uma postura protagonista em relação as suas finanças pessoais⁴. Porém, não existe padronização para a caracterização de diferentes tipos de investimentos, cada corretora e banco cria e realiza a manutenção de seu próprio banco de dados para atender seus usuários, indicando classes de investimento de acordo com perfis extraídos de questionários de *Suitability*⁵, um processo altamente ineficiente comparado a uma base de dados unificada.

Este trabalho objetiva a construção de uma ontologia e um *Personal Financial Recommender System* (PFRS), um agente que recomendará classes de investimentos baseando-se no perfil do usuário e nas entidades da ontologia. Esta aplicação faz parte de um projeto de pesquisa que visa construir uma ontologia genérica para investimentos financeiros brasileiros.

Além da problemática mercadológica aportada anteriormente, esta aplicação foi escolhida pois requer que a ontologia capture propriedades genéricas e básicas para diferentes tipos de investimento. Sendo capaz de cumprir tarefas entre diferentes classes de investimentos e podendo estender seu teste para múltiplas aplicações como recomendado por [Zhang et al.] para a construção de uma ontologia financeira genérica.

Para a utilização do sistema, primeiro o usuário responde a perguntas que avaliam seu perfil de investidor, baseado no questionário de *Suitability* do *Credit Suisse*. Depois, o sistema busca por entidades com propriedades que satisfazem as respostas dadas pelo usuário. Por fim, é realizado uma ponderação baseada por pesos, que avaliam a importância de cada pergunta, que resulta na indicação de uma carteira financeira por classes de investimentos para o usuário.

Na seção 1 explicamos as decisões tomadas para cada etapa do sistema, como o ambiente de desenvolvimento, a construção e população da ontologia, elaboração do questionário e o processo de escolha de classes de ativos baseando-se nas respostas do usuário.

Na seção 4 discutimos os resultados do nosso projeto, desde uma análise crítica do processo de desenvolvimento a efetividade da ontologia e do PFRS.

Por fim, na seção 5 concluímos os aprendizados e descobertas do projeto e discutimos possíveis próximos passos visando a construção de uma ontologia genérica.

2. Metodologia

Nossa metodologia é exemplificada na Figura 1.

2.1. Decisões de projeto

2.1.1. Ambiente de Desenvolvimento

Primariamente, optamos por realizar o desenvolvimento no *Google Colab*, um notebook de ciência de dados, pois a ferramenta é versátil para o uso colaborativo. Além disso, seria

⁴<https://exame.com/exame-in/raio-x-anbima-pandemia-abriu-fosso-entre-investidor-das-classes-a-b-e-c/>

⁵https://www.investidor.gov.br/menu/Menu_Investidor/funcionamento_mercado/Suitability.html

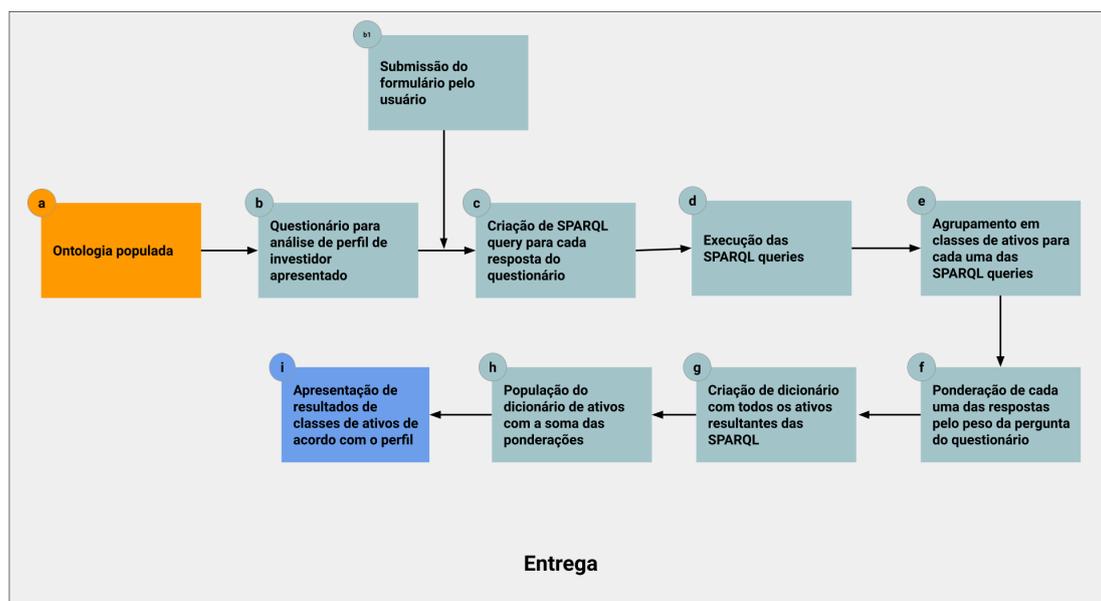


Figura 1. Etapas específicas da metodologia. Em laranja, o início da etapa; em azul, destacamos o fim da etapa.

possível construir o formulário junto ao código, sendo uma interface funcional. Priorizou-se construir um protótipo funcional que utilizasse da ontologia ao desenvolvimento de uma interface arrojada para a aplicação.

2.1.2. Construção e População da ontologia

Para o presente trabalho, a construção da ontologia foi realizada buscando as propriedades gerais presentes em ativos financeiros, principalmente aquelas que são fatores de risco. As classes foram definidas baseando-se no site da B3, dividindo os investimentos em duas principais classes, renda fixa⁶ e renda variável⁷. Depois foram analisadas possíveis propriedades para cada tipo de ativo, havendo um estudo prévio sobre as propriedades válidas para cada um.

Populamos a ontologia manualmente (etapa (a)), criando estruturas e indivíduos para conseguirmos evoluir na solução do problema de encontrar o perfil do investidor baseando-se nas respostas do formulário dadas pelo usuário.

2.1.3. Questionário para análise de perfil

Para a etapa (b) temos uma estrutura de perguntas baseadas na ontologia, que terá perguntas múltipla escolha afim de identificar o perfil do usuário.

⁶https://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/cotacoes/renda-fixa/

⁷https://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/

Foram estudados os questionários de *Suitability*, que analisam o perfil do investidor, dos bancos Itaú e *Credit Suisse* e das corretoras Rico, XP Investimentos, Warren. O questionário elaborado para o projeto é baseado no questionário do *Credit Suisse*⁸, uma vez que este era o único dentre os analisados, que era transparente em relação a metodologia do questionário, ou seja, como cada pergunta impacta no computo final da classificação do perfil do investidor. Os pesos de cada pergunta foram utilizados para a ponderação sobre cada pergunta em nosso sistema 2.1.7 e as perguntas que não eram de múltipla escolha foram removidas.

Na etapa (b1) temos a submissão do questionário pelo usuário.

2.1.4. Criação das queries

Após a submissão do formulário, os cortes para a SPARQL query são montadas nesta etapa (c) a partir das respostas selecionadas pelo usuário, na etapa (b1). Durante esta etapa, cada resposta do usuário corresponde a uma SPARQL query que procura entre as entidades da ontologia aquelas que possuam propriedades compatíveis com a query.

2.1.5. Execução das queries

Uma vez montadas as consultas na etapa (c) então elas serão executadas na etapa (d) retornando as classes das entidades encontradas. Por exemplo, caso o usuário respondesse na primeira pergunta que pretende utilizar seus investimentos no curto prazo (até um ano) é inferido que este necessita de um investimento com alta liquidez e uma data próxima de vencimento. Desta maneira, a SPARQL query é montada visando procurar entidades com estas propriedades. Supondo que a query retorne a entidade Tesouro Selic 2021, o programa retorna a classe da entidade, Tesouro Direto, para a pergunta. Vale ressaltar que é possível que uma pergunta mais de uma classe de investimento como resposta. Dessa maneira, cada pergunta retorna uma lista de classes de investimentos.

2.1.6. Agrupamento

Com as consultas executadas na etapa (d), as classes são exibidas com as porcentagens de cada tipo de ativo (renda fixa, renda variável e seus subconjuntos) na etapa (e) a partir das respostas do questionário.

2.1.7. Ponderação

Uma vez com a lista de classes de ativos retornados, nesta etapa (f), para o cálculo das porcentagens um peso w_i é atribuído para cada uma das n perguntas de maneira que $\sum_{i=0}^n w_i = 1$. Logo, cada pergunta possui um diferente peso, baseado no questionário do *Credit Suisse*, que impacta diretamente na proporção que os investimentos retornados pela pergunta terão na carteira final recomendada.

⁸https://www.cshg.com.br/site/publico/cs/DocumentacaoCadastral/PessoaFisica/Formulario_Suitability.pdf

Supondo um questionário de apenas duas perguntas, que as respostas foram $p_1 = [tesouro_direto, criptomoeda]$ e $p_2 = [acao_brasileira, criptomoeda]$ e que $w_1 = 0.6$ e $w_2 = 0.4$, a recomendação dada pelo nosso sistema será investir 30% em Tesouro Direto, 20% em ações brasileiras e 50% em criptomoedas.

2.2. Dicionário de dados

Com as respostas definidas, e com as proporções de cada ativo para cada pergunta discriminados, criamos um dicionário de chave-valor na etapa (g), considerando todos os ativos de todas as perguntas retornadas.

2.3. População do dicionário

Com o dicionário definido, somamos as proporções dos ativos e colocamos elas dentro das chaves referentes a cada classe da nossa ontologia nesta etapa (h). Uma vez com o resultado final, realizamos uma normalização dos dados afim de todos os ativos serem proporcionais às suas porcentagens iniciais e ao mesmo tempo a soma ser igual a 1.

2.4. Apresentação

O resultado dos cálculos no dicionário avaliando as classes de ativo é exibido na última etapa (i), onde um gráfico de pizza é mostrado com cada uma das classes de ativos retornadas e suas proporções de acordo com a resposta que o usuário deu a cada uma das perguntas levando em consideração os pesos dessas perguntas, como demonstrado em 2.

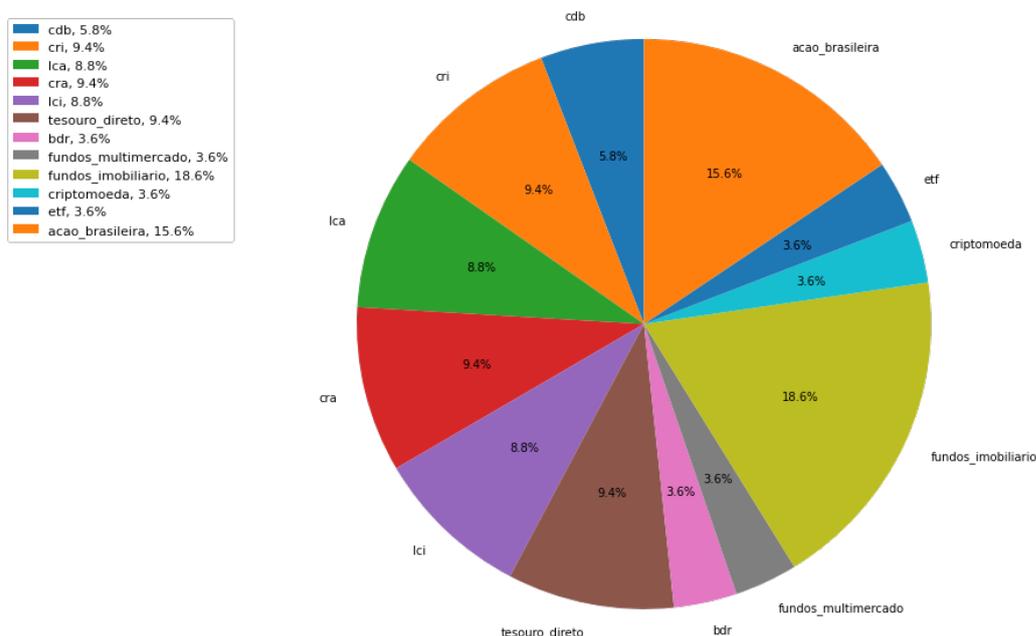


Figura 2. Gráfico da carteira recomendada pelo sistema.

3. Aplicação em Uso

O objetivo desta seção é relatar os passos de uso e aquisição das recomendações de carteiras de investimentos. Será apresentado o questionário identificador de perfil de investimento, a conversão das respostas neste questionário para filtros de consultas SPARQL,

os critérios para identificação dos perfis de investimento, e exemplos de carteiras geradas para diferentes perfis de investidores.

3.1. Conversão das Respostas em Consultas

Abaixo é apresentada uma tabela mapeando os filtros selecionados por cada resposta do questionário e que são integrados dentro das consultas SPARQL à ontologia desenvolvida neste trabalho. As consultas retornam recomendações de categorias de investimento de acordo com os filtros selecionados com a resposta a cada pergunta. Cada questão contribui com uma porcentagem específica da carteira final, baseando-se no questionário do *Credit Suisse*.

Tabela 1 - Filtros de Acordo com as Respostas

Q.	A.	Liquidez	Vencimento	Tributação	Volatilidade	Divi.	Renta.
1.a.	A	>8 dias	>= 04/2023	-	-	-	-
	B	>5 dias	>= 10/2023	-	-	-	-
	C	>5 dias	>= 10/2027	-	-	-	-
1.b.	A	-	-	Longo Prazo	Baixa	-	-
	B	-	-	Médio Prazo	Moderada	-	-
	C	-	-	Curto Prazo	Alta	-	-
1.c.	A	-	-	-	Baixa	-	-
	B	-	-	-	Moderada	-	-
	C	-	-	-	Alta	-	-
2.b.	A	-	-	-	Baixa	-	-
	B	-	-	-	Moderada	-	-
	C	-	-	-	Alta	-	-
2.c.	A	>0 dias	>= 04/2023	-	Baixa	-	-
	B	>5 dias	>= 10/2023	-	Moderada	-	-
	C	>8 dias	>= 04/2027	-	Alta	-	-
2.d.	A	-	-	-	-	<1.0	>0%
	B	-	-	-	-	>1.0	>50%
	C	-	-	-	-	>1.1	>100%
2.e.	A	-	-	-	Baixa	-	-
	B	-	-	-	Moderada	-	-
	C	-	-	-	Alta	-	-

Q.: Questão

A.: Alternativa

Divi.: Dividendo

Renta.: Rentabilidade

Como dito anteriormente, os resultados adquiridos por consultas SPARQL para cada uma das respostas ao questionário identificador de perfil de investimento, contribui em um certo peso na carteira recomendada. A seguir são apresentados os pesos de cada questão na concepção da carteira de investimentos final:

Tabela 2 - Ponderação de Cada Questão

	1.a.	1.b.	1.c.	2.b.	2.c.	2.d.	2.e.
Peso	30	20	40	20	40	60	10
Porcentagem	13.6%	9.09%	18.2%	9.09%	18.2%	27.3%	4.54%

3.2. Identificação dos Perfis de Investimento

Os perfis de investimentos são gerados baseados em critérios estabelecidos pelo *Credit Suisse*. Na tabela abaixo podemos observar as diferentes pontuações para as respostas de cada questão. A pontuação de cada resposta é somada às outras respostas da mesma seção, sendo que neste questionário existem duas seções (1 e 2).

Tabela 3 - Pontuações para cada alternativa nas diferentes questões

	A	B	C
1.a	0	10	30
1.b	0	10	20
1.c	0	20	40
2.b	0	10	20
2.c	0	20	40
2.d	20	40	60
2.e	0	10	20

Por meio da soma das pontuações em cada seção, é caracterizado um perfil para cada seção em específico, sendo que ao final é considerado o perfil **mais conservador** identificado. Por exemplo, se um usuário foi caracterizado como agressivo na seção 1 do questionário e como moderado na seção 2, ao final do questionário seu perfil será moderado, por ser o mais conservador. Abaixo é apresentado um mapeamento das possíveis pontuações na seção 1 e na seção 2 com o perfil caracterizante do usuário respondente. As pontuações da seção 1 estão na vertical à esquerda e as pontuações da seção 2 estão na parte superior.

Tabela 4 - Mapeamento entre pontuações e Perfis de Investimento

S1 / S2	n <= 13 (1)	n <= 43 (2)	n <= 87 (3)	n <= 117 (4)	n > 117 (5)
n <= 13 (1)	Conservador	Conservador	Conservador	Conservador	Conservador
n <= 30 (2)	Conservador	Moderado	Moderado	Moderado	Moderado
n <= 54 (3)	Conservador	Moderado	Arrojado	Arrojado	Arrojado
n <= 72 (4)	Conservador	Moderado	Arrojado	Agressivo	Agressivo
n > 72 (5)	Conservador	Moderado	Arrojado	Agressivo	Super agressivo

3.3. Exemplos de Carteiras

Nesta subseção serão exemplificados alguns resultados para diferentes respostas ao questionário proposto, ou seja, as diferentes carteiras propostas pelo sistema projetado para dois dos perfis de usuários possíveis.

3.3.1. Perfil Conservador

Respostas do usuário:

- Pretendo utilizar um percentual relevante dos meus investimentos no curto prazo (até 1 ano).
- Vou precisar resgatar valores acima dos rendimentos financeiros ao longo dos próximos anos para manter o meu padrão de vida.
- Eu tenho uma carteira conservadora.
- Eu não acredito que, em períodos de mais de cinco anos, produtos de maior risco sejam mais atrativos do que produtos de menor risco.
- Venderia imediatamente caso os investimentos caíssem mais de 10%.
- Não estou disposto a sofrer flutuações negativas nos meus investimentos, mesmo em períodos curtos (1 mês).
- Eu não considero que minha formação acadêmica e/ou experiência profissional me ofereça conhecimento sobre o mercado financeiro.

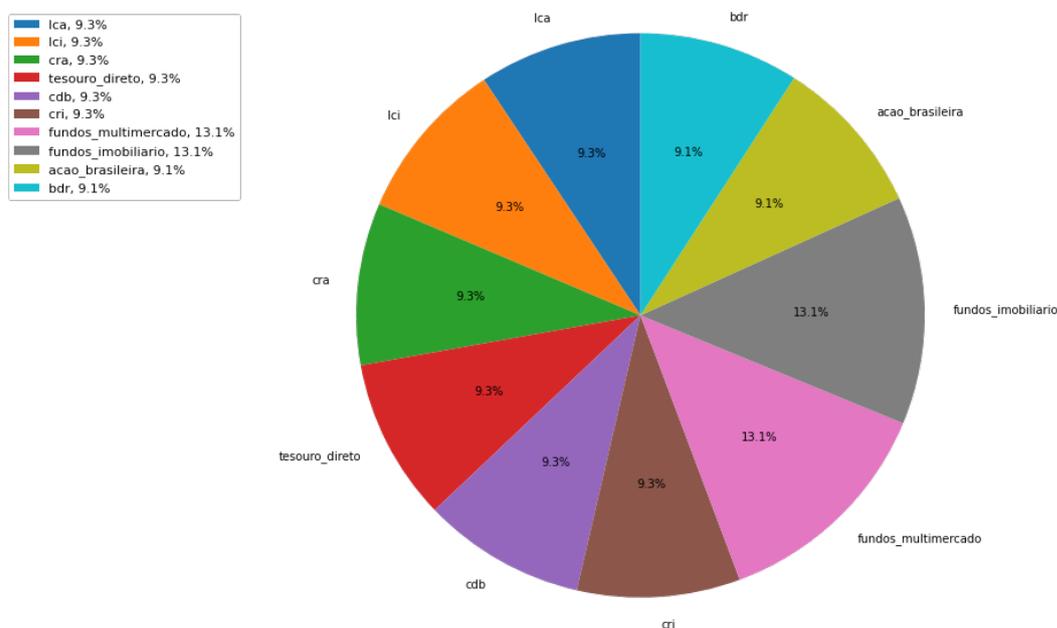


Figura 3. Exemplo de Carteira no Perfil Conservador

3.3.2. Perfil Super Agressivo

Respostas do usuário:

- Não tenho previsão de utilizar meu investimento no curto e médio prazo.
- Estou fazendo constantes aportes na carteira de investimento, aumentando o patrimônio financeiro.
- Eu tenho uma carteira mais agressiva em certas instituições financeiras.
- Eu acredito que, em períodos de mais de cinco anos, produtos de maior risco são mais atrativos do que produtos de menor risco.
- Entendo que o meu portfólio está sujeito a sofrer flutuações dessa magnitude (mais de 10%).

- Busco retornos absolutos e admito flutuações negativas dos meus investimentos em períodos superiores a 12 meses.
- Eu considero que minha formação acadêmica e/ou experiência profissional me ofereça conhecimento sobre o mercado financeiro.

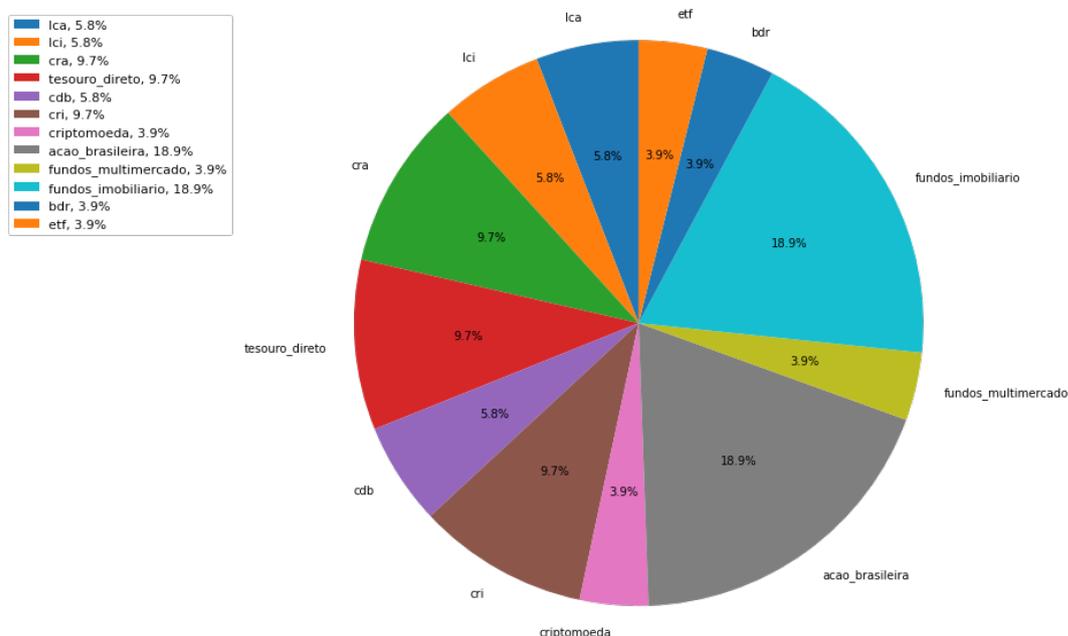


Figura 4. Exemplo de Carteira no Perfil Super Agressivo

4. Discussão

Fizemos uma análise de como respostas predominantes em cada uma das seções afeta no resultado do perfil de investimento. Apenas o perfil *Agressivo* não é exibido na tabela abaixo, por não ser possível com respostas predominantes e ser muito próximo do perfil *Super Agressivo*.

Tabela 5 - Perfil Identificado com Respostas Predominantes

Seção 1	Seção 2	Resultado
A	A	Conservador
A	B	Conservador
A	C	Conservador
B	A	Moderado
B	B	Arrojado
B	C	Arrojado
C	A	Moderado
C	B	Arrojado
C	C	Super Agressivo

Observando as carteiras geradas para diferentes perfis de investidores, é possível verificar as diferenças de pesos das carteiras são coerentes com o que é esperado. São recomendadas porcentagens maiores de investimentos em renda variável para perfis agres-

sivos em detrimento à perfis conservadores, o que é um bom indício da efetividade da nossa solução.

Porém a nossa solução ainda recomenda porcentagens altas de investimentos em renda variável para perfis conservadores, o que não esperado nestes casos. Investidores conservadores possuem uma tendência clara em optar por investimentos de renda fixa, e por tanto com menor risco. Outro fato que não corresponde à realidade é que todas as categorias de investimentos são recomendadas, mesmo que em pequenas porcentagens. Estes problemas podem ter x motivações: a inserção de alguns dados sintéticos que não correspondem à realidade, ponderações das questões não adequadas para a solução proposta, e uma não limitação dos resultados retornados nas consultas.

Em trabalhos futuros é esperado que tais problemas sejam superados, por meio da inserção exclusiva de dados reais, identificação de uma ponderação adequada à solução e a limitação dos resultados das consultas. Também é esperado que o trabalho seja estendido para o problema de balanceamento de carteiras, auxiliando o usuário em recomendações de compra e venda de ativos. Os dados que compõem a ontologia criada por este trabalho poderiam ser utilizados para diversos fins, como o preenchimento de DARF (Documento de Arrecadação de Receitas Federais) em situações de venda de ativos, ou na inserção de dados necessários em declarações de imposto de renda.

A ontologia de investimentos desenvolvida é capaz de agregar informações estratégicas para diferentes níveis de investidores, e sua disponibilização pública auxiliaria muito tais usuários. São informações que dificilmente são encontradas centralizadas e relacionadas, o que evidencia ainda mais a necessidade de uma ontologia como a projetada neste trabalho.

5. Conclusão

É inegável o papel das ontologias para a comunicação e interoperabilidade entre sistemas. Estas podem auxiliar na construção de sistemas robustos da área de investimentos em um contexto brasileiro. Este trabalho teve como objetivo criar uma ontologia de investimentos brasileiras, bem como uma aplicação modelo que consultasse esta ontologia afim de apresentar o perfil e algumas classes de ativos recomendadas ao usuário da aplicação.

Com o preenchimento de um questionário disponibilizado ao usuário, foi possível indicar, fazendo uso de SPARQL query *templates*, classes de ativos condizentes às respostas dadas pelo usuário. Para a construção da versão inicial da ontologia foram catalogadas 15 classes de ativos, 25 propriedades e 118 instâncias, que são os ativos financeiros brasileiros. Com o desenvolvimento do trabalho foi possível constatar que é possível utilizar uma ontologia única e robusta para responder queries sobre o mercado financeiro brasileiro.

Referências

- [Zhang et al.] Zhang, Z., Zhang, C., and Ong, S. S. Building an ontology for financial investment. In Leung, K. S., Chan, L.-W., and Meng, H., editors, *Intelligent Data Engineering and Automated Learning — IDEAL 2000. Data Mining, Financial Engineering, and Intelligent Agents*, Lecture Notes in Computer Science, pages 308–313. Springer.