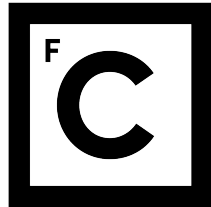


UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE INFORMÁTICA



**Ciências**  
**ULisboa**

**ISIAAW - Interpretação Semântica de Imagens na  
Avaliação da Acessibilidade Web**

**Carlos Miguel Ribeiro Duarte**

**Mestrado em Informática**

Dissertação orientada por:  
Prof. Doutor Carlos Alberto Pacheco dos Anjos Duarte  
Prof. Doutor Luis Manuel Pinto da Rocha Afonso Carriço

2018



## **Agradecimentos**

Em primeiro lugar, quero agradecer aos meus orientadores, Prof. Carlos Duarte e Prof. Luís Carriço, por me orientarem neste trabalho e por sua vez ajudado a evoluir ao longo do trabalho.

Em segundo quero agradecer a minha família pela educação que me transmitiram e pela orientação que me deram ao longo da vida. Agradeço também todos os meus amigos que me acompanharam nesta etapa, dando um especial agradecimento ao André Monteiro, João Vicente, Tiago Moucho, Inês de Matos e Ana Salvado. Estes foram a minha fonte de inspiração e a minha motivação para concluir esta etapa tão importante. Agradeço também à Teresa Monteiro, por ter me ajudado na escrita do documento final e ajudando na minha dificuldade de escrita, sem ela este documento não seria escrito.

Por ultimo quero agradecer aos grupos "7cumes" e aos "Escaladores da Nossa Senhora dos Escaladores Aflitos", pelos bons fins de semanas passados com estes grupos a escalar, dando assim forças para realizar este trabalho. Obrigado a todos pelo vosso apoio.



*“A dor é a fraqueza a sair do corpo”*

*Nuno Curto - 7cumes*



## Resumo

O uso da Web tem vindo aumentar devido ao facto dos utilizadores recorrerem à Web para vários fins do seu quotidiano: saúde, finanças ou lazer. Este recurso é usado por utilizadores com características muito diferentes. Desta forma, os recursos existentes na Web devem estar preparados para estes e, no entanto, não estão. Vários problemas de acessibilidade são encontrados em vários sítios Web, causando assim que utilizadores com limitações visuais, auditivas ou outras, tenham grande dificuldade a usar este recurso. Contudo, têm vindo a surgir formas de combater a falta de acessibilidade em conteúdos Web. Por exemplo, WCAG 2.1 da iniciativa WAI, que estabelece um conjunto de recomendações para quem cria conteúdos para a Web. Um dos problemas mais comuns de acessibilidade é a falta de coerência ou a falta da propriedade texto alternativo (alt) nas imagens. De forma a solucionar este problema, ISIAAW foi implementando como um serviço capaz de analisar imagens informativas e o seu texto alternativo, classificando a sua relação. O ISIAAW recorre a 3 importantes módulos: **Filtragem**, avalia o tamanho dos textos alternativos, pois estes têm que ser breves; **Extração**, extrai informação chave das imagens de forma a converter conteúdo não textual em textual; **Avaliação**, calcula a relação entre a informação chave extraída e a descrição da imagem. Este serviço foi desenhado de forma a ser modular, dando assim a capacidade de substituir ou remover qualquer módulo/ferramenta externa sem afetar outros módulos. Por último, ISIAAW foi implementado como um serviço Web de forma a possibilitar a sua integração em várias aplicações, independentemente da sua linguagem e arquitetura. Neste trabalho o serviço foi integrado num avaliador de acessibilidade Web, QualWeb, com a implementação da técnica H37 das diretrizes WCAG 2.0, aumentando assim a capacidade de avaliação deste avaliador.

**Palavras-chave:** Acessibilidade Web, Avaliação Semântica, QualWeb, Relação Imagem-Texto.



## Abstract

The use of the Web has been increasing, due to the fact that users use the Web for various purposes: health, finances or leisure. The Web is used by users with very different characteristics, so the existing Web resources should be prepared for them. But, often, these resources are not prepared. Several accessibility problems are found in various Web pages, thus affecting users with certain disabilities (visual or auditory), limiting their ability to use the features. However, certain resources have been emerging to combat the lack of accessibility in Web content. An example are the WCAG 2.1 of the WAI initiative. One of the most common accessibility issues is the lack of coherence or lack of alternative text (alt) property in the images. In order to solve this problem, ISIAAW was implemented as a service capable of analyzing the alternative text of informative images, classifying their relation. ISIAAW resorts to three important modules: **Filtering**, evaluates the size of alternative texts, as these have to be brief; **Extraction**, extracts key information from images in order to convert non-textual content into textual; **Evaluation**, calculates the relationship between the extracted key information and the image description. This service has been designed to be modular, thus giving the ability to replace or remove any module/third party without affecting other modules. Finally, ISIAAW was implemented as a Web service in order to allow its integration in several applications, regardless of its language and architecture. In this work the service was integrated in a Web accessibility evaluator, QualWeb, with the implementation of the H37 technique of the WCAG guidelines, thus increasing the evaluation capacity of this evaluator.

**Keywords:** Web Accessibility, Semantic Analysis, QualWeb, Relation Image and Text



# Conteúdo

<b>Lista de Figuras</b>	<b>xiv</b>
<b>Lista de Tabelas</b>	<b>xvii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	1
1.2 Objetivos . . . . .	3
1.3 Metodologia de Investigação . . . . .	3
1.4 Estrutura do Documento . . . . .	4
<b>2 Trabalho Relacionado</b>	<b>7</b>
2.1 Acessibilidade Web . . . . .	7
2.1.1 Diretrizes para a acessibilidade Web . . . . .	8
2.1.2 WCAG 2.1 . . . . .	9
2.1.3 Técnicas com carácter semântico . . . . .	12
2.1.4 Avaliadores de Acessibilidade Web . . . . .	12
2.2 Texto Alternativo . . . . .	15
2.2.1 Características de uma boa descrição . . . . .	18
2.3 Análise Conteúdo Semântica . . . . .	18
2.3.1 Ontologias . . . . .	19
2.3.2 Aplicações de relação semântica . . . . .	21
2.4 Extração de informação . . . . .	22
2.4.1 Aplicações de Extração de Texto . . . . .	22
2.4.2 Aplicações de Extração de Imagens . . . . .	24
2.5 Sumário . . . . .	25
<b>3 ISIAAW</b>	<b>27</b>
3.1 Arquitetura e Desenho . . . . .	27
3.2 Serviço Web . . . . .	28
3.3 Aplicações externas . . . . .	30
3.4 Algoritmo . . . . .	37
3.5 Módulos . . . . .	40

3.5.1	Filtro do Tamanho dos alts . . . . .	40
3.5.2	Módulo Concept Extraction . . . . .	42
3.5.3	Módulo Domain Extraction . . . . .	44
3.5.4	Módulo MetaInfo Extraction . . . . .	46
3.5.5	Módulo Relation Evaluation . . . . .	47
3.6	Sumário . . . . .	53
<b>4</b>	<b>Avaliação</b>	<b>57</b>
4.1	Dados de teste . . . . .	57
4.2	Resultados . . . . .	59
4.2.1	Análise por domínio . . . . .	61
4.3	Sumário . . . . .	62
<b>5</b>	<b>Integração no QualWeb</b>	<b>63</b>
5.1	QualWeb . . . . .	63
5.1.1	Implementação da técnica H37 . . . . .	64
5.2	Alterações nos Ambientes do QualWeb . . . . .	67
5.3	Sumário . . . . .	69
<b>6</b>	<b>Conclusão</b>	<b>71</b>
6.1	Limitações . . . . .	72
6.2	Trabalho Futuro . . . . .	73
<b>A</b>	<b>Distribuição dos resultados baseados no domínio</b>	<b>75</b>
<b>B</b>	<b>Resultados do serviço ISIAAW</b>	<b>81</b>
<b>C</b>	<b>Exemplos Dados de Treino</b>	<b>85</b>
<b>D</b>	<b>Exemplos Dados de Teste</b>	<b>87</b>
	<b>Bibliografia</b>	<b>93</b>





# Lista de Figuras

2.1	Técnicas suficientes para o critério Conteúdo Não Textual . . . . .	11
2.2	Processo de avaliação da técnica h37 . . . . .	11
2.3	Exemplo de uma imagem e do seu alt . . . . .	16
2.4	Exemplo de uma imagem funcionais . . . . .	17
2.5	Exemplo de Ontologia usada pelo DBpedia . . . . .	20
2.6	Exemplo Clarifai . . . . .	25
2.7	Exemplo Clarifai com o modelo Comida . . . . .	26
3.1	Arquitetura ISIAAW . . . . .	28
3.2	Imagem exemplo . . . . .	29
3.3	Exemplo de pedido do ISIAAW . . . . .	29
3.4	Imagem exemplo . . . . .	32
3.5	Resultado . . . . .	33
3.6	Exemplo de pedido ao Swoogle . . . . .	35
3.7	Fluxograma da aplicação ISIAAW . . . . .	38
3.8	Exemplo do conjunto de treino . . . . .	39
3.9	Distribuição da qualidade das descrições no conjunto de treino . . . . .	40
3.10	Gráfico da qualidade das descrições em função do número de palavras . . . . .	42
3.11	Imagem exemplo : One Story Wedding . . . . .	45
3.12	Resultado da relação entre Bride e Wedding . . . . .	45
3.13	Imagem Bordeaux . . . . .	47
3.14	Evolução da correlação baseada no valor limite do Swoogle . . . . .	51
3.15	Evolução da correlação baseada no valor limite do Swoogle [0.2,0.3] . . . . .	52
4.1	Exemplo do conjunto de teste . . . . .	58
5.1	Arquitetura do QualWeb . . . . .	64
5.2	Estado atual da interface do utilizador . . . . .	69
A.1	Valores obtidos da avaliação da categoria comida . . . . .	75
A.2	Valores obtidos da avaliação da categoria celebridades . . . . .	76
A.3	Valores obtidos da avaliação da categoria ciência . . . . .	76
A.4	Valores obtidos da avaliação da categoria desporto . . . . .	77

A.5	Valores obtidos da avaliação da categoria tecnologia . . . . .	77
A.6	Valores obtidos da avaliação da categoria viagem . . . . .	78
A.7	Valores obtidos da avaliação da categoria casamento . . . . .	78
A.8	Valores obtidos da avaliação da categoria outros . . . . .	79





# Lista de Tabelas

2.1	Diretrizes e Critérios de sucesso do princípio Percetível . . . . .	10
2.2	Lista de técnicas WCAG com componente semântica . . . . .	12
2.3	Tabela de avaliadores . . . . .	15
2.4	Métricas de relação semântica . . . . .	20
2.5	Ferramentas de extração de Informação em texto . . . . .	23
3.1	API da aplicação ISIAAW . . . . .	28
3.2	Modelos do Clarifai e as suas descrições . . . . .	31
3.3	Resultado do SpaCy . . . . .	35
3.4	Parâmetros do Swoogle . . . . .	36
3.5	Relação do número de palavras e a qualidade das descrições . . . . .	41
3.6	Tabela de Correlação dos valores de limite do Swoogle . . . . .	54
3.7	Resultados na alteração dos valores de acréscimo . . . . .	55
4.1	Tabela resumo da primeira fase de classificação . . . . .	59
4.2	Tabela representativa da concordância entre avaliações . . . . .	59
4.3	Tabela sumário da classificação final . . . . .	59
4.4	Tabela de previsão/classificação . . . . .	60
4.5	Tabulação cruzada entre as avaliações dos peritos e o serviço . . . . .	61
4.6	Resultados das medidas simétricas . . . . .	61
4.7	Tabela resumo das categorias . . . . .	62
5.1	Critérios de sucesso e técnicas implementadas no QualWeb (técnicas com componente semântica apresentadas a negrito) . . . . .	65
5.2	Tabela de parâmetros do avaliador QualWeb . . . . .	68
B.1	Resultados do serviço ISIAAW (60 de 149) . . . . .	81
B.2	Resultados do serviço ISIAAW (120 de 149) . . . . .	82
B.3	Resultados do serviço ISIAAW (149 de 149) . . . . .	83



# Capítulo 1

## Introdução

De acordo com o sítio [1], estima-se que 40% da população mundial tem acesso à Internet. No entanto este crescimento no número de utilizadores cria novos problemas, pois estes utilizadores podem ter algum tipo de limitação que dificulta o uso deste recurso. Em 2006 já se verificava um número elevado de utilizadores com limitações, como por exemplo cognitivas, visuais e auditivas. Em 2011 a Organização Mundial da Saúde identificou uma percentagem elevada de população que apresenta algum tipo de limitação [2]. Estes utilizadores, devido às suas limitações, podem não interagir da mesma forma que os restantes, sendo assim necessário focar as aplicações e recursos Web para a inclusão de todos.

### 1.1 Motivação

Este trabalho aborda duas áreas: acessibilidade Web e semântica. Relativamente à primeira, como mencionado acima, aproximadamente 40% da população mundial tem acesso à Internet [1]. Na última década assistiu-se a um aumento de utilizadores que recorrem a este recurso para vários fins, por exemplo, educação, comércio, saúde. Estes tiram partido da Internet através de conjuntos de hardware (computadores, smartphones) e software (navegadores, aplicações). Quando se fala de facilidade de uso da Internet, há que ter em mente que cada utilizador é único. Desta unicidade resulta uma gama de características a considerar, em que algumas delas se traduzem em limitações. Exemplificando: neurológicas, cognitivas, físicas, visuais, auditivas e/ou fala.

A Organização Mundial da Saúde afirma que pelo menos 15% da população sofre de algum tipo de limitação [2]. Consciente destas limitações a Internet devia promover um ambiente de inclusão, independentemente das características dos utilizadores, do seu hardware ou mesmo do seu software. Contudo, há ainda um grande caminho a percorrer [3].

Muitos dos recursos existentes na Internet não estão preparados para satisfazer as necessidades de alguns utilizadores, obrigando-os assim a uma navegação complexa [4]. Utilizadores com certas características requerem o uso de tecnologias de assistência. Estas

permitem que utilizadores com algumas limitações naveguem na Web. por exemplo os *ScreenReaders* permitem aos utilizadores invisuais ouvirem a página. Porém estas tecnologias dependem das características das páginas: contexto da página, tamanho do texto, cor, contraste e conteúdos. Esta última característica implica conhecer as limitações dos utilizadores de forma a implementar conteúdo acessível.

Com o intuito de combater as deficiências dos conteúdos Web, a WAI (Web Accessibility Initiative) [5], uma iniciativa da W3C (World Wide Web Consortium), desenvolveu um conjunto de recursos e diretrizes. Estes permitem o desenvolvimento de conteúdo Web acessível. Por outro lado, a existência deste tipo de recurso permite a avaliação da acessibilidade dos conteúdos de forma objetiva. Um dos recursos mais importantes criados pela WAI foi WCAG 1.0 (Web Content Accessibility Guidelines) em 1999 [5], que padronizou as regras para a construção de conteúdos Web em termos de acessibilidade. A sua versão mais recente, WCAG 2.1, fornece conjuntos de técnicas, normas e exemplos de forma a implementar conteúdo Web acessível. Devido à sua relevância como diretrizes de acessibilidade este recurso é muitas vezes usado em ferramentas de avaliação. Estas, como o QualWeb [6], permitem que o conteúdo Web criado seja examinado, de forma a ser possível a sua validação em termos de acessibilidade num pequeno espaço de tempo.

A avaliação processa-se através da verificação da estrutura do documento a nível sintático [6, 7]. Esta organização sintática verifica, por exemplo, se as *tags* de cabeçalho (*h1, h2, h3, ...*) estão organizados hierarquicamente<sup>1</sup>. No entanto, estas ferramentas de avaliação não avaliam a relação semântica entre elementos. No caso de imagens, por exemplo, é necessário verificar se a informação transmitida na descrição da imagem está relacionada com a imagem e vice-versa. De acordo com o estudo realizado pelo grupo Simple Accessibility [8], verifica-se que em muitos casos as imagens apresentadas contêm descrições inapropriadas ou mesmo erradas que leva a uma incoerência entre conteúdos. É ainda indicado neste estudo que grande parte dos problemas poderiam ser minimizados caso existisse um padrão para uma boa descrição. No entanto, este padrão não é de fácil definição pois as características de uma boa descrição são muitas vezes dependentes das características dos indivíduos.

Normalmente o processo de avaliação semântico é efetuado manualmente, uma vez que ainda existem poucos algoritmos capazes de o fazer [9]. Devido ao facto de algumas páginas Web sofrerem de incoerência semântica e sendo necessário tornar a Web mais acessível, a criação de um serviço que consiga avaliar automaticamente a relação semântica entre elementos tornou-se fundamental e indispensável para o futuro da acessibilidade.

---

<sup>1</sup><https://www.w3.org/TR/WCAG20-TECHS/G141.html>

## 1.2 Objetivos

A integração de técnicas de avaliação com cariz semântico em ferramentas de avaliação de acessibilidade resulta normalmente num pedido ao utilizador para avaliar este tipo de técnicas manualmente. Esta situação representa uma desvantagem que limita a aplicabilidade destas ferramentas de avaliação. Com o intuito de fornecer uma forma de avaliar a relação entre vários tipos de elementos, este trabalho tem como objetivo implementar um mecanismo capaz de avaliar a relação entre imagens e o seu texto descritivo. Este mecanismo deve ser capaz de avaliar e afirmar com algum nível de certeza se os dois formatos do mesmo conteúdo estão ou não relacionados. A construção deste mecanismo proporciona às ferramentas de avaliação a capacidade de avaliar técnicas com cariz semântico. Em seguida são apresentados os objetivos específicos a alcançar:

- Avaliação e compreensão dos serviços online de extração de informação em conteúdos não-textuais e textuais;
- Estudo e compreensão de regras para a construção de uma boa descrição de imagens na *Web*;
- Desenvolvimento de um algoritmo de avaliação semântica com base nas regras de uma boa descrição;
- Implementação de um mecanismo de avaliação da relação com o uso de aplicações de terceiros.
- Avaliação do desempenho do mecanismo implementado.
- Integração do mecanismo desenvolvido no avaliador de acessibilidade *Web*, QualWeb.

## 1.3 Metodologia de Investigação

Para atingir os objetivos enunciados, seguiu-se uma metodologia composta pelas seguintes etapas:

### **Etapa 1 - Estudo do trabalho relacionado e de ferramentas existentes**

**Estudo de trabalho relacionado:** Antes da implementação do serviço foi conduzido um estudo detalhado nas áreas da análise semântica e da acessibilidade Web. Nesta última, o foco foi nas técnicas de acessibilidade WCAG.

**Identificação de ferramentas de análise semântica:** Posteriormente ao estudo do estado da arte foi realizado um levantamento de ferramentas que podem ser usadas para o serviço à implementar.

## Etapa 2 - Desenho e implementação do serviço

**Criação de elementos de treino:** Construção de um formulário para recolher avaliações de peritos, com o objetivo de produzir um conjunto de treino para otimizar o desempenho do serviço.

**Mecanismo de avaliação:** Desenho de um mecanismo capaz de avaliar a relação entre vários conteúdos, textuais e não textuais.

**Desenho e implementação do serviço ISIAAW:** Implementação do mecanismo desenhado recorrendo às ferramentas e serviços disponíveis já estudados.

## Etapa 3 - Validação do serviço

**Criação de elementos de teste:** Recolha de novas avaliações de peritos, com o objetivo de produzir um conjunto de teste para validar o desempenho do serviço.

**Avaliação do serviço:** Avaliação do desempenho do serviço recorrendo ao conjunto de teste.

## Etapa 4 - Integração do serviço num avaliador de acessibilidade

**Integração do serviço no QualWeb:** Integrar o serviço ISIAAW no QualWeb, complementando-o com a possibilidade de avaliação de uma técnica com cariz semântico.

## Etapa 5 - Escrita da dissertação

**Dissertação Final:** Escrita do relatório final, descrevendo todas as tarefas, objetivos e resultados obtidos com a implementação do serviço ISIAAW.

## 1.4 Estrutura do Documento

O documento está organizado pelos seguintes capítulos:

- **Capítulo 2 - Trabalho Relacionado:** Neste capítulo é apresentado uma análise do estado da arte da Acessibilidade Web, da Análise Semântica, Extração de Informação de Conteúdos digitais, heurísticas de qualidade entre descrição e imagem.
- **Capítulo 3 - ISIAAW:** Implementação do serviço de relação semântica com o uso de várias ferramentas. Análise de várias métricas existentes para relações semânticas, estrutura do sistema, análise do algoritmo e explicação das ferramentas usadas.
- **Capítulo 4 - Avaliação:** Serão descritos os dados de treino obtidos pela avaliação de três peritos e os resultados finais da aplicação.

- **Capítulo 5 - Integração no QualWeb:** Implementação das alterações realizadas no QualWeb Evaluator de forma a integrar o sistema ISIAAW. Serão detalhadas as alterações realizadas na linha de comando e na interface do utilizador. Por último apresenta-se a implementação da técnica H37.
- **Capítulo 6 - Conclusão:** Serão descritas as conclusões obtidas deste trabalho, limitações encontradas e trabalho futuro.



# Capítulo 2

## Trabalho Relacionado

Este capítulo aborda a evolução tecnológica e documental na área da acessibilidade *Web* e da análise de conteúdo semântico. Será tratado em primeiro lugar a evolução dos padrões de acessibilidade, fornecendo também uma visão geral das ferramentas de avaliação. Em segundo lugar, serão discutidas as características fundamentais de uma descrição de imagem coerente, que seja capaz de substituir a imagem. Posteriormente, apresenta-se a discussão sobre a área da semântica e acessibilidade e como se podem interligar, assim como, aspetos relacionados com Ontologias, Prospecção de dados e Redes Neurais. Por fim será feito um levantamento dos serviços que vão possibilitar a construção do serviço ISIAAW.

### 2.1 Acessibilidade Web

Uma Web acessível significa que sites, ferramentas e tecnologias são desenhados e desenvolvidos de forma a mitigar as diferenças entre utilizadores. Este recurso acompanha o nosso quotidiano através da Educação, Trabalho, Governo, Comércio, Saúde, Lazer entre outros. A Web acessível tem como fundamento principal a inclusão dos utilizadores como um todo, independentemente do hardware, software, língua, cultura, localização ou características físicas [10].

O tema Internet para todos (Web for All) tem vindo a crescer na última década, devido a várias iniciativas culturais e legislativas. Em 2016 as Nações Unidas proclamam que o acesso à Informação e Tecnologias é um direito básico à vida humana [11]. Por outro lado alguns países como a Itália e os Estados Unidos declararam leis que obrigam sites governamentais e de saúde a ter um nível de acessibilidade que possibilite um acesso idêntico [12] a pessoas com diferentes capacidades. No entanto, sites de comércio e de lazer não são obrigados a implementar essas regras o que pode gerar um acesso desigual.

A implementação de regras de acessibilidade beneficiam os utilizadores com as seguintes limitações graves/não graves [13]:

- Utilizadores com limitações auditivas, cognitivas, neurológicas, motoras, fala ou

visuais;

- Utilizadores com limitações temporárias, por exemplo perda de visão momentânea, lesões físicas, entre outras;
- Utilizadores provenientes de países sub-desenvolvidos;
- Utilizadores de áreas rurais;
- Utilizadores com ligação de Internet lenta ou com *bandwidth* limitada (apresentação da descrição em vez da imagem, faz com que seja necessário menor *bandwidth*).

Paralelamente, organizações de Direitos Humanos, governo e empresas tentam impulsionar a criação de recursos acessíveis a todos. No entanto, os designers muitas vezes desconhecem técnicas de criação de conteúdo acessível, sendo necessário investir na formação para o desenvolvimento de conteúdos acessíveis. Para isso o grupo W3C (World Wide Web Consortium) criou a iniciativa WAI (Web Accessibility Initiative) [14]. O objetivo será desenvolver especificações, diretrizes, técnicas e recursos de suporte que descrevem como implementar conteúdo acessível.

### 2.1.1 Diretrizes para a acessibilidade Web

Uma boa acessibilidade depende de três grandes componentes [5]:

- **Web Content:** refere-se à parte do Website, incluindo texto, imagens, forms, scripts, etc.
- **User Agents:** Software usado pelos utilizadores para acederem a conteúdo *Web*: Browsers, multimedia players e plug-ins.
- **Authoring tools** - Software ou serviço usado pelos utilizadores para a produção de conteúdo Web, redes sociais, blogs, wikipedia.

Com o levantamento destas componentes mais importantes para uma boa acessibilidade torna-se necessário o desenvolvimento de recursos. A WAI-W3C desenvolveu um conjunto de especificações e técnicas para cada um dos componentes acima referidos:

- **WCAG (Web Content Accessibility Guidelines):** Incluem diretrizes e técnicas para construção de *websites* e aplicações para utilizadores com limitações, assim como para utilizadores idosos.
- **UAAG (User Agent Accessibility Guidelines):** A UAAG explica como devem ser feitos os *browsers* e *media players* acessíveis. As características dos browsers são um ponto importante para utilizadores idosos e com limitações, pois devem possuir características que possibilitem uma boa navegação, idealmente sem o uso de tecnologias de assistência.

- **ATAG (Authoring Tool Accessibility Guidelines):** Ferramentas de autoria são usadas para a criação de conteúdo. Exemplos destas ferramentas são Editores de HTML e *Content Management Systems* (CMS). Estas ferramentas permitem ao utilizador adicionar conteúdo à Web - tal como *blogs*, partilha de fotografias, *online forum* e redes sociais. Estas têm que ser acessíveis a todos os utilizadores.

O objetivo destas recomendações é aperfeiçoar as práticas de codificação dos *designers* para construírem *sites* de elevada qualidade que melhoram a experiência de todos os utilizadores. As diretrizes WCAG são as mais usadas para a construção de *websites*. A sua primeira versão, WCAG 1.0, foi publicada em 1999 [15]. Esta apresenta alguns exemplos e normas que devem ser seguidas para obter conteúdos acessíveis. No entanto, com a evolução das tecnologias tornou-se incompleta e pouco significativa, devido à falta de princípios e técnicas, como consequência do aparecimento de outras linguagens e ferramentas de construção de sites e aplicações [5]. Com o passar dos anos foi redigida uma versão mais moderna, denominada de WCAG 2.0. Nesta, as informações são mais detalhadas, combatendo assim a referida falta de princípios, resultado do uso de tecnologias de conteúdo dinâmico.

Na secção seguinte analisam-se as diretrizes WCAG na medida em que estas se direcionam para o conteúdo existente na página, que será o foco deste trabalho.

### 2.1.2 WCAG 2.1

WCAG, Web Content Accessibility Guidelines, é um documento que explica como devem ser construídos conteúdos *Web* mais acessíveis para utilizadores com limitações. Por conteúdo Web entende-se a informação existente na página, que pode ser informação natural (texto, imagens e sons), ou código, *scripts* e conteúdo dinâmico.

WCAG 2.0 foi publicado em 11 Dezembro de 2008 [5], mas surgiu uma nova versão (WCAG 2.1) em Junho de 2018. Este documento está organizado em quatro princípios:

- **Perceçível:** A informação e os componentes da *interface* do utilizador têm de ser apresentados de forma a que o utilizador os possa percecionar;
- **Operável:** Os componentes da *interface* de utilizador e a navegação têm de ser operáveis;
- **Compreensível:** A informação e a utilização da *interface* de utilizador têm de ser compreensíveis;
- **Robusto:** O conteúdo deve ser suficientemente robusto para ser interpretado de forma fiável por uma ampla variedade de *User Agents*, incluindo as tecnologias de apoio.

No total este documento apresenta treze Diretrizes divididas em vários princípios. Um dos principais objetivos das diretrizes é garantir que os conteúdos estejam diretamente acessíveis ao maior número possível de utilizadores. Cada Diretriz contempla um conjunto de critérios de Sucesso. Estes critérios descrevem concretamente os requisitos para obter conformidade com essa Diretriz. Cada Critério de Sucesso é apresentado como uma declaração, que pode ser verdadeira ou falsa quando testada com um determinado conteúdo Web. Todos os Critérios de Sucesso do WCAG 2.1 são concebidos como critérios testáveis, para determinar, objetivamente, se os conteúdos cumprem os Critérios de Sucesso. Cada critério corresponde a um nível de conformidade que é representado da seguinte forma: A, AA, e AAA (na Tabela 2.1 apresentam-se os critérios de sucesso para o princípio Perceptível). Nível AAA são critérios que beneficiam alguns utilizadores, enquanto os de nível A beneficiam um maior número.

Tabela 2.1: Diretrizes e Critérios de sucesso do princípio Perceptível

Princípio	Diretrizes	Critério de Sucesso	Nível de Conformidade
1. Perceptível	1.1 Alternativas em Texto	1.1.1 Conteúdo Não Textual	A
	1.2 Média Dinâmica ou Contínua	1.2.1 Conteúdo só de áudio e só de vídeo (pré-gravado)	A
		1.2.2 Legendas (pré-gravadas)	A
		1.2.3 Audiodescrição ou Alternativa em Multimédia (pré-gravada)	A
		1.2.4 Legendas (em direto)	AA
		1.2.5 Audiodescrição (em direto)	AA
		1.2.6 Língua Gestual	AAA
		1.2.7 Audiodescrição Alargada (pré-gravada)	AAA
		1.2.8 Alternativa em Multimédia (pré-gravada)	AAA
		1.2.9 Só áudio (em direto)	AAA
	1.3 Adaptável	1.3.1 Informações e Relações	A
		1.3.2 Sequência com Significado	A
		1.3.3 Características Sensoriais	A
	1.4 Distinguível	1.4.1 Utilização da Cor	A
		1.4.2 Controlo de Áudio	A
		1.4.3 Contraste (Mínimo)	AA
1.4.4 Redimensionar texto		AA	
1.4.5 Imagens de Texto		AA	
1.4.6 Contraste (Melhorado)		AAA	
1.4.7 Som Baixo ou Ausência de Som de Fundo		AAA	
1.4.8 Apresentação Visual		AAA	
1.4.9 Imagens de Texto (sem exceção)		AAA	

**Técnicas de alternativa em texto abreviado para utilizar nas técnicas de tipo suficiente acima**

1. [H36: Utilizar atributos alt em imagens utilizadas como botões Submeter](#) (HTML)
  2. [H2: Combinar a imagem adjacente e os links de texto para o mesmo recurso](#) (HTML)
  3. [H37: Utilizar atributos alt em elementos img](#) (HTML)
  4. [H35: Fornecer alternativas em texto em elementos applet](#) (HTML)
  5. [H53: Utilizar o corpo do elemento object](#) (HTML)
  6. [H24: Fornecer alternativas em texto para os elementos area dos mapas de imagens](#) (HTML)
  7. [H86: Fornecer alternativas em texto para arte ASCII, emoticons e leetspeak](#) (HTML)
  8. [H30: Fornecer texto do link que descreva a finalidade de um link para elementos anchor](#) (HTML)
- Nota: Consulte as [Noções sobre o Critério de Sucesso 2.4.4 Finalidade do Link \(Em Contexto\)](#).
9. [G196: Utilizar uma alternativa em texto num item de um grupo de imagens que descreva todos os itens do grupo](#)

Figura 2.1: Técnicas suficientes para o critério Conteúdo Não Textual

De forma a fornecer abordagens e exemplos de como cumprir as diretrizes, o grupo de trabalho WCAG identificou técnicas de tipo suficiente para cada Critério de Sucesso. Estas técnicas apresentam conjuntos de exemplos de sucesso, definições e testes capazes de avaliar a técnica. Na Figura 2.1 é apresentado um conjunto de Técnicas Suficientes para o Critério de Sucesso - Conteúdo Não Textual.

Focando-se na técnica H37: Utilizar atributos alt em elementos img, na Figura 2.2 é apresentado o teste necessário para esta técnica ter sucesso. O sucesso desta valida que o critério é satisfeito. Este procedimento pode ser dividido em dois grupos: procedimentos sintáticos, normalmente são passos que requerem a avaliação da organização do código (PASSO 1), e procedimentos semânticos, passos que requerem avaliar relações entre um ou mais elementos (PASSO 2 e 3).

#### Testes

---

##### *Procedimento*

1. Examine cada elemento `img` existente no conteúdo.
2. Verifique se cada elemento `img` que transmite um significado inclui um atributo `alt`.
3. Se a imagem incluir palavras importantes para compreender o conteúdo, as palavras são incluídas na alternativa em texto.

##### *Resultados Esperados*

Os passos 2 e 3 são verdadeiros.

Figura 2.2: Processo de avaliação da técnica h37

### 2.1.3 Técnicas com carácter semântico

Como pode ser observado na Figura 2.2 as técnicas descritas no WCAG podem ter dois tipos: sintáticos ou semânticos. De forma a estudar as técnicas WCAG com cariz semântico foi feito um levantamento destas, apresentadas na tabela 2.2, resultando num total de 33 técnicas: 5 técnicas de relação entre imagem e texto, 9 técnicas de relação texto e texto, 8 técnicas de relação entre texto e *link* e 11 técnicas de relação entre vários elementos, listas e tabelas.

Tabela 2.2: Lista de técnicas WCAG com componente semântica

Técnicas	Descrição
h2	Combinar a imagem adjacente e os links de texto para o mesmo recurso
h25	Fornecer um título utilizando o elemento title
h30	Fornecer texto de link que descreva a finalidade de um link para os elementos anchor
h33	Supplementing link text with the title attribute
h36	Utilizar atributos alt em imagens utilizadas como botões Submeter
h37	Utilizar atributos alt em elementos img
h39	Using caption elements to associate data table captions with data tables
h42	Utilizar h1-h6 para identificar cabeçalhos
h43	Utilizar os atributos id e headers para associar células de dados a células de cabeçalho em tabelas de dados
h44	Utilizar elementos label para associar etiquetas de texto a controlos de formulário
h45	Utilizar longdesc
h48	Utilizar ol, ul e dl para listas
h49	Utilizar a marcação semântica para assinalar texto especial ou realçado
h53	Utilizar o corpo do elemento object
h54	Utilizar o elemento dfn para identificar a ocorrência com a definição de uma palavra
h57	Utilizar atributos language no elemento html
h58	Utilizar atributos language para identificar alterações no idioma humano
h64	Utilizar o atributo title da frame e os elementos iframe
h65	Utilizar o atributo title para identificar controlos de formulário quando o elemento label não puder ser utilizado
h69	Fornecer elementos heading no início de cada secção de conteúdo
h70	Utilizar elementos frame para agrupar blocos de material repetido
h73	Utilizar o atributo summary do elemento table para fornecer uma descrição geral das tabelas de dados
h75	Garantir que as páginas Web sejam corretamente formadas
h77	Identificar a finalidade de um link utilizando o texto do link em conjunto com os respetivos itens da lista
h78	Identificar a finalidade de um link utilizando o texto do link em conjunto com o respetivo parágrafo
h79	Identificar a finalidade de um link utilizando o texto do link em conjunto com a respetiva célula de tabela e cabeçalhos de tabela associados
h80	Identificar a finalidade de um link utilizando o texto do link em conjunto com o elemento heading precedente
h81	Identificar a finalidade de um link numa lista encaixada utilizando o texto do link em conjunto com o item da lista principal sob o qual a lista está encaixada
h83	Utilizar o atributo target para abrir uma nova janela a pedido do utilizador e indicá-lo no texto do link
h85	Utilizar OPTGROUP para agrupar elementos OPTION dentro de um SELECT
h89	Utilizar o atributo title para fornecer ajuda contextualizada
h90	Indicar os controlos de formulário obrigatórios
h97	Agrupar links relacionados usando o elemento nav

### 2.1.4 Avaliadores de Acessibilidade Web

O recurso WCAG é importante para a construção de aplicações e avaliação de conteúdo *Web* em termos de acessibilidade. A avaliação de conteúdo *Web* pode ser realizada de três formas: Avaliação com Utilizadores, com Especialistas e com Avaliadores de Acessibilidade Web [16]

#### Utilizadores

Uma avaliação com utilizadores permite obter dificuldades de acessibilidade em contextuais reais. No entanto, afigura-se difícil a reunião de representantes de vários grupos por forma a obter uma avaliação significativa e representativa.

## Especialistas

Avaliações feitas por especialistas recorrem tipicamente às WCAG para avaliarem os *sites*. Estes normalmente apresentam relatórios detalhados das falhas e erros encontrados nas suas avaliações. Embora tenham mais facilidade em encontrar problemas ou falhas, consomem muito mais tempo, comparativamente aos Utilizadores, para efetuar uma avaliação completa. Este facto deve-se à complexidade dos parâmetros utilizados por estes.

## Avaliadores de Acessibilidade Web

Avaliadores de acessibilidade *Web* são programas de *software* ou serviços *online* que ajudam a determinar se o conteúdo **Web** cumpre diretrizes de acessibilidade. Estas ferramentas identificam rapidamente problemas de acessibilidade. Podem ser usadas em todas as fases do processo de desenvolvimento. No entanto, não têm a capacidade de avaliar todas as técnicas propostas no WCAG. Estas ferramentas não têm a capacidade de avaliar as técnicas de cariz semântico, gerando erros ou avisos, e por isso requerem que seja feita uma avaliação manual.

Em [17] apresenta-se uma lista completa de ferramentas de avaliação. As características principais deste tipo de ferramentas são as seguintes:

- Diretrizes: Diferentes tipos de organizações podem pedir diferentes tipos de padrão de acessibilidade (WCAG, SECTION 508, etc.).
- Língua: Linguagem que suportam;
- Tipo de Ferramenta: *Plugin*, *Software*, Serviço ou Linha de Comando;
- Formatos suportados: Maior parte das ferramentas avaliam a acessibilidade de conteúdo HTML. No entanto, outras avaliam tecnologias como, WAI-ARIA, CSS, SVG e PDF;
- Resultados: Relatórios, Avaliação *Step-by-step*, *In-Page feedback* e *Page Transformation*.

Aproximadamente vinte seis destas ferramentas usam o WCAG como padrão de acessibilidade (Guideline), que evidencia a relevância deste recurso. De forma a tornar mais eficaz o estudo destas ferramentas foi aplicada uma filtragem usando os seguintes campos:

- **Tecnologia** - HTML e CSS: Estas tecnologias são muitas vezes encontradas em conteúdo digital e responsáveis pela visualização de conteúdo *Web*.
- **Tipo de Ferramenta** - Aplicações *Online*: Uma Aplicação *Online* é de fácil acesso, não necessitando de instalar qualquer tipo de programa ou aplicação. Este tipo de ferramentas é normalmente acedida usando um *browser*

- **Estado da Aplicação:** Algumas das ferramentas existentes na lista acima encontram-se *offline* ou os seus avaliadores não estão em funcionamento.
- **Credenciais:** Foram retiradas as aplicações que necessitam de algum tipo de registo para ser possível utilizar essa ferramenta.

Com o uso destes filtros foram escolhidas 10 aplicações. Destas foram estudadas as seguintes características:

- **Pós-Processamento:** A existência de milhões de *sites* construídos em Javascript [18] permite a utilização de conteúdo dinâmico na página. O uso do Javascript permite a alteração do estado dos *sites* sem carregar de novo a página, por meio de um *script*. O uso deste tipo de conteúdos influencia a acessibilidade da página. O conteúdo da página apresentado ao utilizador pode ser diferente do que está presente no código fonte.
- **Versão WCAG:** Dá alguma perspetiva sobre a versão utilizada pelas ferramentas para proceder a avaliação, já que a versão 1.0 se apresenta desatualizada;
- **Análise Semântica:** A capacidade de avaliar técnicas com relação semântica é um ponto importante para o trabalho a ser realizado. Esta característica permite ter uma ideia sobre se algum dos avaliadores de acessibilidade tem a capacidade de avaliar ou não.

A tabela 2.3 apresenta os resultados do estudo. Como é possível verificar, quase nenhuma das ferramentas *online* existentes tem a capacidade de analisar semanticamente conteúdos *Web*. Somente o QualWeb apresenta capacidade de verificar/avaliar semanticamente conteúdos *Web*, focando-se somente em relações entre imagens e as suas descrições [9]. As ferramentas sem capacidade de realizar análise semântica, pedem aos utilizadores para efetuarem uma avaliação manual. Desta forma, desenvolver um mecanismo de análise semântica é fundamental para o futuro destas ferramentas.

Tabela 2.3: Tabela de avaliadores

Ferramenta	Pos-Processamento	Versão WCAG	Análise semântica
AccessMonitor	Não	2.0	Manual
AChecker	Não	Ambas	Erro
CynthiaSays	Não	2.0	Manual
A-Tester	Não	2.0	Erro
Hera	Não	1.0	Manual
ASqatasun	Sim	Ambas	Não
Tanaguru	Sm	Ambas	Manual
QualWeb	Sim	2.0	Parcial
Wave	Sim	2.0	Manual
Tawdis	Sim	2.0	Erro

## 2.2 Texto Alternativo

Com o crescimento do uso da Internet e das suas funcionalidades, o uso de imagens na *Web* tem vindo a aumentar exponencialmente, em particular graças ao uso de redes sociais. Em média são feitos por dia 300 milhões de *uploads* de imagens no *facebook* [19] e 52 milhões no *instagram* [20]. O uso deste recurso tem dois objetivos: estético ou informativo. No entanto, este não é o único recurso não-textual existente. Outros conteúdos textuais podem ser gráficos, vídeo ou mesmo som. A grande variedade de conteúdos não-textuais faz com que alguns utilizadores sejam excluídos devido a limitações, por exemplo, visuais, no caso de imagens, ou auditivas, no caso de vídeo e som. Cada tipo de recurso não-textual é suportado por um conjunto de técnicas, WCAG, para solucionar a barreira existente. Às imagens, de forma a possibilitar aos utilizadores a perceção destes conteúdos, deve associar-se uma propriedade denominada de texto alternativo, *alt* [21]. Esta propriedade deve ter a capacidade de substituir totalmente a imagem caso esta não possa ser carregada, devido a fatores externos ou internos, já descritos na Secção 2.1.

A Simply Accessible em 2015 [8] procurou estudar de que modo a propriedade *alt* é usada. Num conjunto de 100 problemas com imagens foram obtidos os seguintes resultados:

- 66.33% das imagens não tinham a propriedade *alt*;
- 20.33% das imagens tinham *alt* redundante ou inadequado;
- 10.25% das imagens com *alt* eram decorativas, logo não precisavam de *alt*;
- 2.96% das imagens tinham *alt*, mas repetiam conteúdo já existente;

Neste estudo é possível verificar que frequentemente esta propriedade é mal implementada ou mesmo ignorada, acrescentando ainda que muitos dos casos encontrados têm como principal origem a falta de comunicação entre a equipa técnica e equipa de informação. Outro fator de extrema importância presente no estudo prende-se com a dificuldade

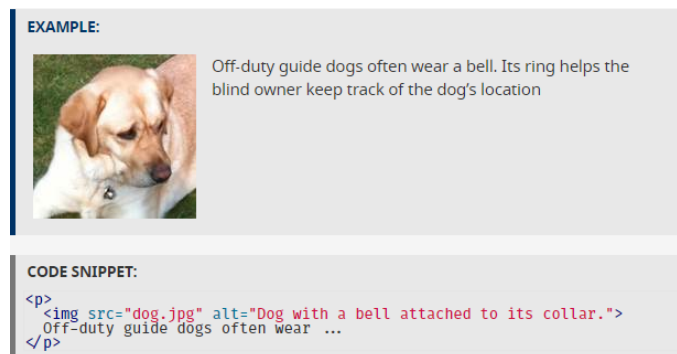


Figura 2.3: Exemplo de uma imagem e do seu alt

na elaboração de uma boa descrição. As imagens normalmente estão divididas em grupos de acordo com a função que desempenham. A W3C apresenta uma divisão dos tipos de imagens [22], acrescentando ainda um conjunto de exemplos de texto alternativo para cada. Estes tipos são:

- Imagens Informativas;
- Imagens Decorativas;
- Imagens Funcionais.

## Imagens Informativas

Imagens informativas [23] têm como objetivo transmitir informação extra ao utilizador. Desta forma, o texto alternativo deve transmitir o significado ou o conteúdo exibido. Contudo, este não pode ser uma descrição literal da imagem, ou seja, os textos alternativos devem ser breves. Em algumas situações, pode ser necessário uma descrição literal detalhada, mas somente quando o conteúdo/ação da imagem é importante, como no caso de obras de arte ou pinturas ilustrativas.

Estas imagens requerem um texto alternativo pequeno com o máximo de **140 caracteres** com os **conceitos mais relevantes**, que deve apresentar o **mesmo contexto da página**. Caso o contexto seja diferente, esta deve ser considerada decorativa, pois apresenta informação redundante, ou mesmo fora de contexto. A Figura 2.3 apresenta um cão a usar um sino. Esta é complementada pelo texto adjacente onde é apresentado o objetivo do sino. Um texto curto e breve no alt é suficiente para descrever a informação que é apresentada na imagem e que não é descrita no texto. Neste caso, o texto alternativo ideal seria “Cão com um sino preso na coleira”. A localização do sino não é descrita no texto adjacente. A imagem é somente usada para fornecer uma ideia ao leitor da localização do sino.

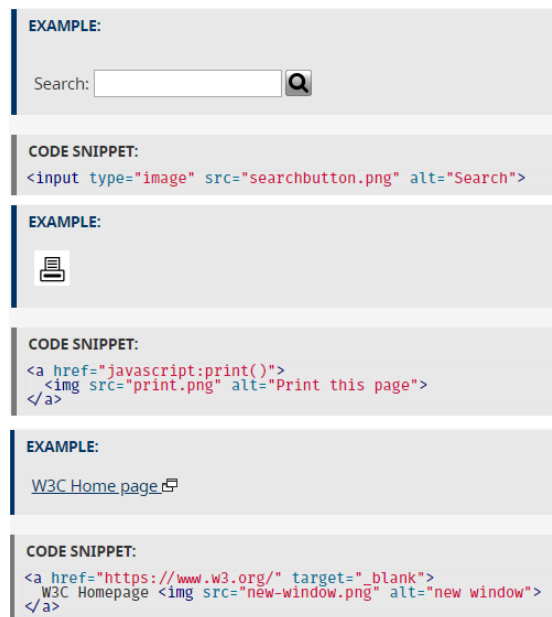


Figura 2.4: Exemplo de uma imagem funcionais

## Imagens Decorativas

Imagens decorativas [24] **não adicionam informações ao conteúdo de uma página**. Por exemplo, as informações fornecidas pelo texto adjacente são suficientes ou a imagem pode ser incluída para tornar os sítios visualmente mais apelativos. Nestes casos, um texto alternativo nulo (vazio) deve ser fornecido (**alt =""**) para que a propriedade possa ser ignorada pelas tecnologias de assistência, como *Screen Readers*. A existência de texto neste tipo de imagens adicionaria confusão auditiva ao leitor ou poderia distrair os utilizadores caso o contexto fosse diferente. A escolha entre imagem decorativa e informativa é da responsabilidade do autor. Este deve conseguir avaliar se a imagem inserida necessita ou não de ter um alt que a complemente.

## Imagens Funcionais

Imagens funcionais [25] são usadas para iniciar ações em vez de transmitir informações. Estas são usadas para substituir botões, *links* e outros elementos interativos. O texto alternativo destas imagem deve **transmitir a ação que será iniciada** (o propósito da imagem), em vez de uma descrição desta. Por exemplo, como mostrado nos exemplos apresentados na Figura.2.4, a alternativa de texto deve ser “pesquisa” em vez de “lupa”, “imprimir esta página” em vez de “impressora” ou “Abrir nova janela” em vez de “Home page”. Valores alt ausentes ou vazios criam problemas significativos para os utilizadores que usam Screen Reader porque estas imagens são essenciais para o funcionamento do conteúdo.

### 2.2.1 Características de uma boa descrição

O recurso apresentado anteriormente pela WAI [22] apresenta um conjunto de exemplos para a criação de boas descrições dependendo do tipo de imagem. As imagens informativas são um bom exemplo para definir os parâmetros de uma boa descrição e por isso se constituíram como uma das preocupações principais deste trabalho.

A Universidade de Leicester [26] apresenta um conjunto de características e alguns pontos que se devem evitar na descrição, como o uso de “Uma imagem de” e o uso de descrições redundantes. WebAIM [27] apresenta um documento bastante detalhado acerca do modo como deve ser feita a descrição de imagens informativas. Este documento tem como base as características do WAI [22]. Contudo um ponto importante para a criação de uma boa descrição é o contexto da página. Diferentes contextos, levam a conceitos importantes diferentes. A lista que se segue resume as características de uma boa descrição somente para imagens Informativas:

- Descrição deve ser curta, com o máximo de 140 caracteres;
- Os conceitos mais relevantes da imagem devem ser descritos no texto alternativo;
- Quanto maior for o número de conceitos da imagem no texto alternativo, maior é a probabilidade de esta ser uma boa descrição;
- O texto alternativo deve ter o mesmo contexto da página;
- Evitar o uso de *keyword-stuffing*, por exemplo “pan, pancake, cake, sugar, ...”;
- O texto existente não deve ser redundante;
- Evitar o uso de frases como “images of” ou “graphic of”.

## 2.3 Análise Conteúdo Semântica

Na secção 2.1 foi apresentado um conjunto de recursos e ferramentas indispensáveis para a avaliação/implementação de conteúdo acessível. Posteriormente na secção 2.2 discute-se como as imagens existentes na *web* são complementadas por textos alternativos inadequados ou não existentes, criando barreiras de acessibilidade a vários utilizadores. Outro ponto importante é a existência de um padrão que defina uma boa descrição. Este permite avaliar a qualidade das descrições. Um dos pontos importantes deste padrão é saber se conceitos chave da imagem estão relacionados com parte da descrição. A relação entre texto e imagens é uma tarefa complexa e tem sido discutida em várias áreas científicas, como na saúde [28] e na biologia [29]. Averiguar que dois objetos estão relacionados, é tarefa fácil para um humano, pois este é capaz de detetar padrões e reconhecer ligações

entre elementos. Contudo, no caso das máquinas esta apresenta-se como uma tarefa complexa [30], pois estas não têm a capacidade de aprender e detetar padrões sem o uso de algoritmos complexos, como por exemplo árvores de decisão, redes neuronais ou outro tipo de algoritmo de prospeção. Outra forma mais simples de avaliar/descobrir relações entre conceitos é o uso de Ontologias.

### 2.3.1 Ontologias

A Ontologia [31] é um modelo de dados que guarda informação sobre Entidades e as suas relações, dentro de um dado domínio. Uma entidade é constituída por um conjunto de propriedades, por exemplo nomes comuns, nomes científicos ou características como a cor do pelo. Normalmente este tipo de modelo é usado em áreas como a inteligência artificial, saúde, biologia e Web (WEB 3.0) como forma de representar conhecimento no mundo digital [31].

O Wornet [32] é uma Ontologia que apresenta palavras da língua inglesa com relações lexicais com outras palavras, como por exemplo antónimos, sinónimos e outros. Foram surgindo outros projetos baseados no Wornet. Como por exemplo IndoWordnet, para a língua Indo [33], Germanet, para a língua germânica [34] e PULO (OpenWordNET-PT), para a língua portuguesa [35]. As Ontologias descritas previamente estão focadas em relações entre palavras. Outros exemplos mais complexos são o DBpedia [31], o Yago [36] e o FreeBase [37]. Estas, ao contrário das Ontologias tipo Wordnet, representam informações mais complexas como por exemplo relações entre várias pessoas ou mesmo objetos.

As ontologias permitem obter relações entre Entidades ou conceitos. No entanto, é necessário obter métricas capazes de qualificar objetivamente a relação [38]. Estes métodos são chamados de Métodos Baseados em Conhecimento. Usam Ontologias para calcular a distância entre conceitos num grafo. Mas existem métodos que não usam Ontologias, utilizando em alternativa grandes volume de texto. Ao contrário dos métodos baseados em conhecimento, estes calculam a probabilidade de dois conceitos estarem relacionados. Uma maior probabilidade significa que estão relacionados. De seguida serão descritos os dois tipos de métodos.

## Métodos Baseados em Conhecimento

Métodos baseados em Conhecimento normalmente usam Ontologias de forma a capacitar a medição da relação. O uso destes possibilita encontrar relações entre conceitos. No exemplo da Figura 2.5, *movie* e *novel* estão relacionados pois ambos são do tipo produto, por outro lado são bastantes diferentes pois ambos tem diferentes entidades. Estas relações são possíveis devido à existência de ontologias como o Wordnet [32], que usa relação entre palavras, ou a DBpedia [31], que usa relações entre entidades. Ao longo

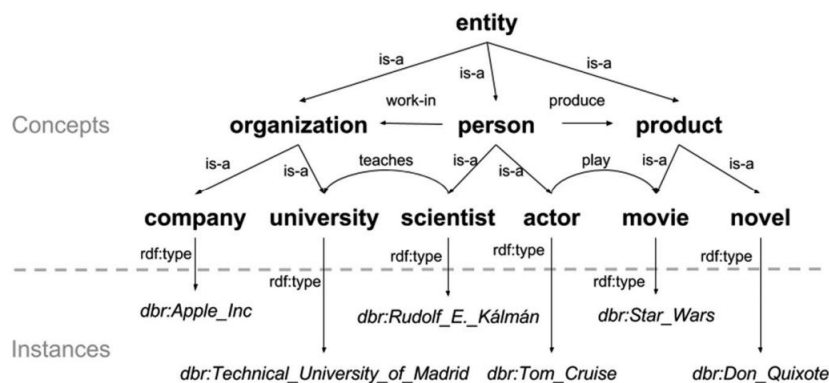


Figura 2.5: Exemplo de Ontologia usada pelo DBpedia

das ultimas décadas têm vindo a surgir conjuntos de métricas capazes de calcular relações entre elementos em Ontologias. Estas métricas estão divididas em dois subgrupos, *Path-based* e *IC-Based*. *Path-based* medem relação usando a distância entre os dois conceitos; *IC-Based* acrescentam valores probabilísticos de dois elementos que se encontram no mesmo contexto. A Tabela.2.4 apresenta um conjunto de métricas bastante usadas para calcular objetivamente a relação entre conceitos. Em [39] é apresentado um estudo que avalia o desempenho e a precisão das métricas apresentadas na tabela. Neste estudo, as métricas com melhor desempenho são as de **lin** e **wpath**.

Tabela 2.4: Métricas de relação semântica

Métrica	Medida (Valores)	Tipo
Path	0 - 1	Path-Based
Wu and Palmer (wup)	0 - 1	Path Based
Leacock and Chodorow (lch)	0 - 10	Path Based
Resnik (res)	0 - 1	IC - Based
Jiang and Conrath (jcn)	0 - 8	IC - Based
Lin (lin)	0 - 1	IC - Based
Wpath	0 - 1	Path-Based

### Métodos Baseados em *Corpus*

Métodos baseados em corpus normalmente estão associados ao uso de tecnologias de *data mining*. Estes medem a relação semântica entre conceitos com o uso de grandes coleções de texto, que são atribuídos a estes mecanismos como dados de treino. Vários trabalhos têm sido desenvolvidos de forma a possibilitar a relação de conceitos como por exemplo, Pointwise Mutual Information [40], Normalised Google Distance [41], LS (Latent Semantic) [42] e Explicit semantic Analysis [43]. No entanto, estes algoritmos requerem uma grande quantidade de dados de treino de forma a estabelecer a relação entre vários elementos [44].

### 2.3.2 Aplicações de relação semântica

Atualmente existe um conjunto de aplicações capazes de usar as tecnologias e métricas já descritas no Capítulo 2.3. O uso destas fornece capacidades de analisar semanticamente vários conceitos sem o desenvolvimento de aplicações complexas. Estas ferramentas são:

#### TensorFlow (Word2Vec)

É uma *framework* de construção de redes neuronais, gratuita e *openSource*. Tem uma ampla gama de algoritmos de relações entre palavras e frases. Por exemplo, os modelos Word2Vec [45] e LS [42]. No entanto, o uso destes modelos requer grande quantidade de dados de treino de forma a gerar algoritmos de confiança. É possível a obtenção de modelos já treinados. No entanto, estes modelos são bastante pesados e normalmente são implementados para um dado domínio, por exemplo conceitos relacionados com biológicas, animais e outros.

#### Swoogle

Swoogle<sup>1</sup> é uma Aplicação REST, completamente gratuita e não tem limite de chamadas. No entanto, o tempo de resposta é lento. Esta é capaz de relacionar frases simples [46]. Para calcular a relação existente entre conteúdos, usa métricas baseadas em conhecimento e baseadas em Corpus [42].

#### WordnetSimilarity

Wordnet<sup>2</sup> tem um conjunto de APIs não oficiais que podem ser usadas para medir relações entre palavras, como é apresentado em [2]. Esta *framework* usa exclusivamente métodos baseados em conhecimento. Os algoritmos usados por esta API são os seguintes: Path [47], Wup [48], Lch [49] e Lin [50].

#### Sematch

Sematch<sup>3</sup> é uma aplicação capaz de relacionar conceitos, entidades e organizações com níveis de confiança altos [51]. Usa Ontologias para realizar relações entre elementos, mais especificamente o DBpedia e Yago.

Das ferramentas estudadas serão usadas as seguintes aplicações: Swoogle, que permite relacionar frases, pois será necessário em alguns casos fazer uma relação rápida entre duas frases sem ser necessário aprofundar cada palavra da descrição. Em caso de ser necessário aprofundar a relação será usado o Wordnet e Sematch. Ambas as aplicações

<sup>1</sup><http://swoogle.umbc.edu/SimService/api.html>

<sup>2</sup><http://www.nltk.org/howto/wordnet.html>

<sup>3</sup><http://gsi-upm.github.io/sematch/>

permitem calcular a relação entre palavras, acrescentando ainda a capacidade do Sematch de avaliar relações entre entidades, como por exemplo Bill Gates com Microsoft. Tensor-Flow não será utilizado devido ao facto de não ter sido possível encontrar um modelo que fornecesse o mesmo nível de desempenho que as outras aplicações.

## 2.4 Extração de informação

No Capítulo 2.3 foram apresentados vários mecanismos capazes de relacionar conceitos. Porém, estes estão limitados a relacionar somente palavras/frases. Estes mecanismos não são capazes de relacionar conteúdos não-textuais, como as imagens, frequentemente usados na *Web* [51, 32]. Esta limitação gera a necessidade de extrair, de outra forma, a informação mais relevante dos textos e imagens. Desta forma, é imperativo que haja uma extração dos conceitos chave, de forma a ser possível usar as ferramentas referidas no capítulo anterior.

Várias tecnologias têm vindo a abordar a extração de conteúdo chave de texto e imagens, com o uso de algoritmos de prospeção. Mas o desenvolvimento destas tecnologias requer grandes quantidades de dados de treino e de tempo, de forma a desenvolver modelos de extração bons [52, 53]. No entanto, como o objetivo do trabalho não é a criação de mecanismos de extração, serão usadas aplicações já existentes para o efeito. Estas aplicações já foram treinadas para a extração de conceitos em texto, resumir texto, extração de palavras-chave em imagens ou identificação de pessoas. Estas aplicações são descritas nos tópicos seguintes: **Aplicações de Extração de Texto** e **Aplicações de Extração de Imagem**:

### 2.4.1 Aplicações de Extração de Texto

Várias tecnologias têm vindo a surgir de forma a extrair informação em texto graças aos avanços na área NLP (Natural Language Processing) [54]. Esta tem tido sucesso em gerar métodos capazes de extrair informação de conteúdo textual. Os métodos desenvolvidos para extração estão divididos em cinco grupos:

1. Extração de Conceitos ou entidade : Extração de palavras-chave, ou entidades como pessoas, locais ou produtos, de textos.
2. Sumarizar: Resumir grandes textos em textos mais breves. Estes métodos usam diretamente a Extração de conceitos de forma a extrair as palavras-chave, e posteriormente fazem os textos com estas palavras.
3. Categorização e Tópicos: Métodos capazes de extrair o tema de corpus de texto de grandes dimensões.

4. Extração de Sentimentos: Estes são capazes de extrair sentimentos de texto, como raiva, ódio, alegria, outros.
5. Extração da Língua: Extração do dialeto existente no texto.

Tabela 2.5: Ferramentas de extração de Informação em texto

Ferramenta	Conceitos	Entidade	Sumarizar	Sentimentos	Tópico	Língua	Chamadas
Text Analytics	Sim	Não	Não	Sim	Não	Sim	5000 Calls/Mês
Aylien	Sim	Não	Sim	Sim	Sim	Não	30000 Calls/Mês
Indico	Sim	Não	Não	Sim	Não	Não	10000 Calls/Mês
Dandelion	Sim	Sim	Não	Sim	Sim	Sim	30000 Calls/Mês
Ambiverse	Não	Sim	Não	Não	Não	Não	1000 Calls/Mês
Monkey Learn	Sim	Sim	Não	Não	Sim	Não	1000 Calls/Mês
Rosette Text	Sim	Não	Não	Não	Não	Não	10000 Calls/Mês
SpaCy	Sim	Sim	Sim	Não	Não	Não	Ilimitado

A tabela 2.5 apresenta um conjunto de aplicações que usam os métodos apresentados anteriormente. O objetivo principal do uso destas aplicações para este trabalho é a extração de palavras-chave – colunas **Conceitos** e **Entidades**. Também relevante para este trabalho é o número de chamadas gratuitas – coluna **Chamadas** – que estes serviços oferecem. Como é possível deduzir pela tabela. As únicas ferramentas que fornecem modos de extração de **Conceitos** e **Entidades** simultaneamente são a API *Dandelion*, a *Monkey Learn* e *SpaCy*. No entanto, com a leitura da coluna **Chamadas** é possível concluir que a aplicação *Spacy* oferece um número ilimitado de chamadas. *Dandelion* oferece, com uma conta gratuita 30000 chamadas, enquanto a aplicação *Monkey* oferece 1000 chamadas. Outra vantagem do uso do *SpaCy* é que esta aplicação é local, ao contrário das anteriores que são aplicações externas, fornecendo assim um sistema independente para a prospeção de informação chave em descrições.

## 2.4.2 Aplicações de Extração de Imagens

A extração de conteúdo em imagens e vídeos é apresentado como um grande desafio para o computador. Este tem de ser capaz de detetar padrões e de reconhecer objetos. Tecnologias de *Image Mining* têm surgido de forma a capacitar a máquina na deteção de objetos [52]. Devido ao facto de não ser do âmbito deste trabalho a criação de algoritmos de prospeção de imagens, não serão construídos tais algoritmos. No entanto, de forma a usar o poder destes algoritmos de extração serão usadas aplicações existentes na *web*:

- Cloud Vision<sup>4</sup>
- Computer Vision<sup>5</sup>
- Clarifai<sup>6</sup>
- Aylien text Analysis API<sup>7</sup>

**Cloud Vision:** Provavelmente uma das melhores aplicações para extração de informação de imagens. Esta API é capaz de extrair conceitos, locais, tópicos das imagens, logos e deteção de entidade. No entanto, esta aplicação tem limite de 10 chamadas por mês com uma conta gratuita.

**Computer Vision:** Serviço de extração de informação desenvolvido pela Microsoft Azure. Este é capaz de extrair conceitos, locais, texto em imagens, entidades e conteúdo dos vídeos. Esta aplicação fornece 1000 chamadas por mês. No entanto, está limitada à deteção de cor e de conceito. De forma a ser possível a usar as capacidades totais desta aplicação é necessário pagar uma mensalidade.

**Clarifai:** Aplicação capaz de detetar conceitos das imagens e analisar vídeos. Está limitada a 5000 chamadas por mês. Esta aplicação extrai um par conceitos valor. Esse valor numérico representa a probabilidade deste conceito existir na imagem. A figura 2.6 representa um exemplo de imagem e o seu resultado.

Esta aplicação fornece um conjunto de domínios diferentes, como por exemplo, casamento, logótipos, comida, celebridades, etc. O uso destes domínios permite limitar os conceitos que são retirados da imagem. Por exemplo, o modelo Comida, quando usado em imagens extrai conceitos relacionados com comida, Figura 2.7. Por último, esta aplicação fornece funcionalidades de criação de modelos pelo utilizador, permitindo assim a criação de novos domínios.

**Aylien Text Analysis API:** Igual ao Clarifai, contudo a extração de conceitos de imagens só pode ser utilizada com contas não gratuitas. Porém, o modelo da Aylien ainda

---

<sup>4</sup><https://cloud.google.com/vision/>

<sup>5</sup><https://azure.microsoft.com/en-us/services/cognitive-services/computer-vision/>

<sup>6</sup><https://www.clarifai.com>

<sup>7</sup><http://docs.aylien.com>

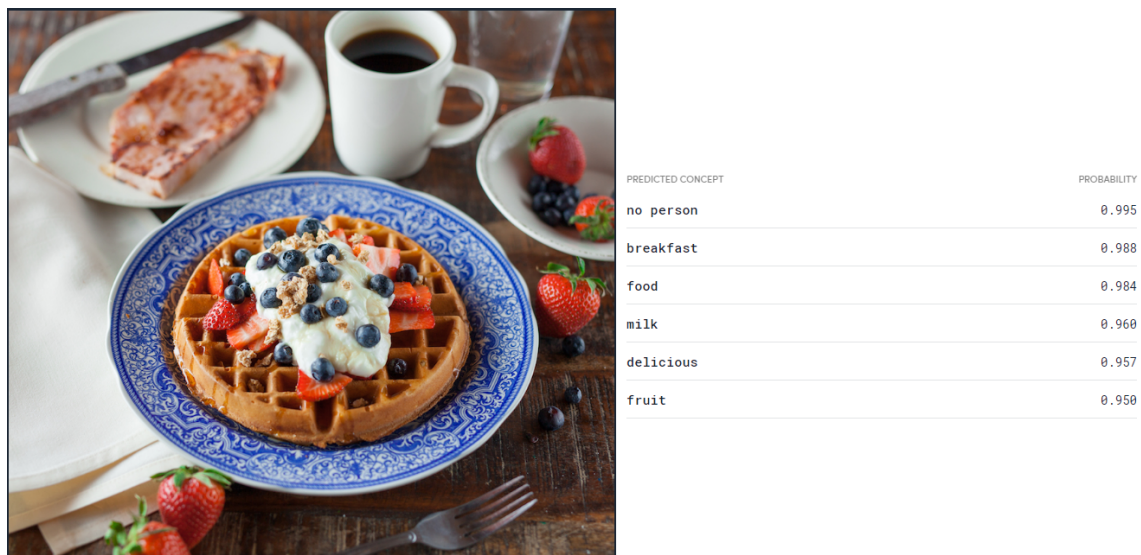


Figura 2.6: Exemplo Clarifai

é muito recente e tem uma aprendizagem fraca, levando assim à extração de conceitos com confianças muito baixas ou mesmo errados.

Das ferramentas estudadas será utilizada o Clarifai. Esta aplicação fornece uma grande quantidade de domínios e de funcionalidades gratuitas comparativamente com a Cloud Vision e a Computer Vision. A aplicação Clarifai fornece todas as suas funcionalidades, limitando o uso da aplicação a 5000 chamadas por mês, acrescentando ainda a possibilidade de usar domínios para limitar os conceitos retirados das imagens. Permite também a criação de *custom models*, através dos quais o utilizador pode criar os seus próprios domínios. Em comparação com a aplicação Aylien-text, o Clarifai produz resultados mais confiantes, dado que a primeira ainda usa as suas funcionalidade em Beta.

## 2.5 Sumário

Neste capítulo apresenta-se uma visão geral das várias áreas importantes para este projeto. A primeira secção, 2.1, estuda a Acessibilidade *Web*. Composta por uma introdução ao tema da acessibilidade *Web*, apresenta um estudo acerca das várias diretrizes de acessibilidade, mais direcionada para as diretrizes de conteúdo *Web*, WCAG. Estas fornecem exemplos e técnicas para produção de conteúdo *Web* acessível e de qualidade. Também é feito um levantamento das técnicas WCAG com cariz semântico, passo importante para o trabalho a realizar. Por último são estudados os avaliadores de acessibilidade *Web* e as suas funcionalidades focando-se na capacidade destes avaliarem conteúdo semântico. Dos avaliadores estudados somente o QualWeb tem capacidade de avaliar técnicas de cariz semântico (h37).

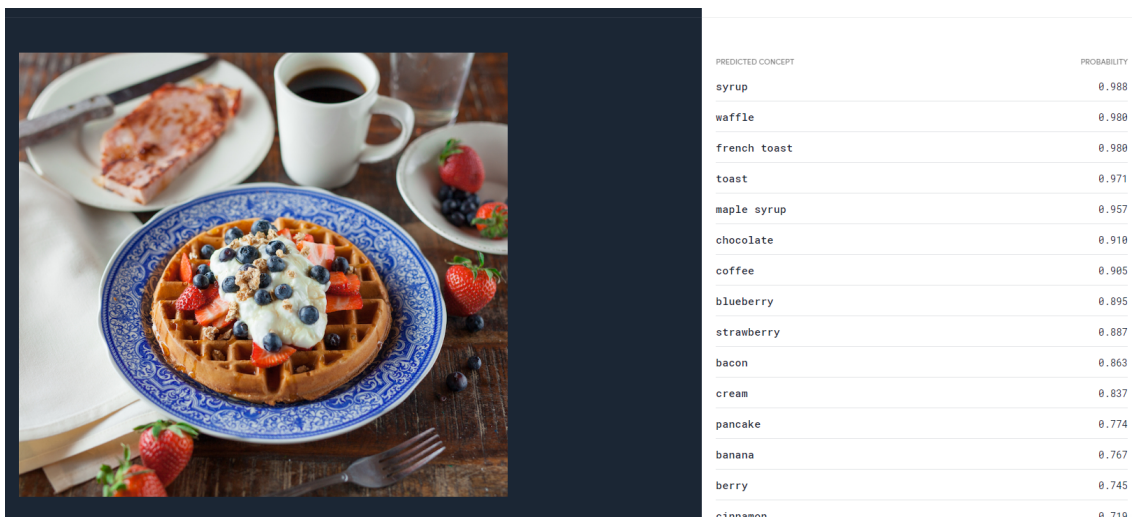


Figura 2.7: Exemplo Clarifai com o modelo Comida

Tendo em conta que o projeto somente vai tratar de relações semânticas entre imagem e a propriedade alt, na secção 2.2 é discutido o problema da existência de más descrições. Faz-se referência ao estudo do Simple Accessibility [26] no qual são apresentadas as maiores falhas do uso desta propriedade. Posteriormente são discutidas as diferenças entre os vários tipos de imagens (Informativas, Decorativas ou Funcionais). Cada tipo de imagem tem um conjunto de características que são necessárias serem transmitidas com o uso da alternativa em texto. Ainda dentro desta secção é feito um levantamento das características ideais para uma boa descrição de imagens Informativas, dando assim a possibilidade de avaliar objetivamente cada descrição usando um conjunto de heurísticas.

Na secção 2.3 a noção de relação entre palavras é explorada. São estudadas várias tecnologias para relacionar vários elementos, como o uso de Ontologias. Ainda dentro destas tecnologias são estudadas as várias métricas para medir relações entre elementos/palavras. No fim desta secção foram estudadas as ferramentas existentes com a capacidade de relacionar palavras.

Por último, na secção 2.4 é feito um levantamento de um conjunto de ferramentas existentes para a extração de informação, quer de texto quer de imagens. Estas ferramentas serão a fonte de informação que possibilita o uso das aplicações estudadas na secção anterior.

# Capítulo 3

## ISIAAW

Neste capítulo, são apresentadas as características do avaliador semântico. Inicialmente será descrita a sua arquitetura, e o serviço de Web implementado. Serão também introduzidas as aplicações externas usadas como o Clarifai e o Swoogle, detalhando as funcionalidades e o objetivo destas aplicações. Por último, são descritos os vários módulos, dando ênfase ao objetivo e aos parâmetros mais importantes destes.

### 3.1 Arquitetura e Desenho

O ISIAAW foi criado como um Serviço *Web*, tornando-se assim uma ferramenta capaz de ser usada em vários sistemas, independentemente da sua arquitetura e linguagem de programação. Além da sua universalidade, o aspeto de modularidade também foi planeado. A implementação de um sistema modular permite que sejam introduzido novos módulos ou a substituição de módulos existentes, com o objetivo de melhorar ou expandir as funcionalidades deste serviço.

A Figura 3.1 ilustra a arquitetura do serviço, dividido em 4 áreas: Serviço Web, Controlador, Módulos e Aplicações Externas. O Serviço *Web* é a área com o objetivo de receber e responder aos pedidos recebidos. O Controlador controla o fluxo de trabalho e funciona como agente mediador entre os vários módulos. Estes estão divididos em 3 conjuntos:

- **Filtros:** Módulo responsável por filtrar a descrição do texto que vem com a imagem;
- **Extração:** Módulos responsáveis pela extração de conceitos chave das imagens;
- **Avaliação:** Módulo responsável pela avaliação e obtenção de um valor de relação.

A última área representa as aplicações externas. Estas são serviços *online* que cada módulo utiliza.

Numa etapa inicial, cada solicitação ao Serviço *Web* é enviada para o Controlador. Este é responsável por evocar cada módulo numa dada ordem. Esta operacionaliza-se

pela sequência que se apresenta: Filtros, Extração e Avaliação. No final deste processo, o Controlador é responsável por enviar a resposta para o componente serviço *Web*.

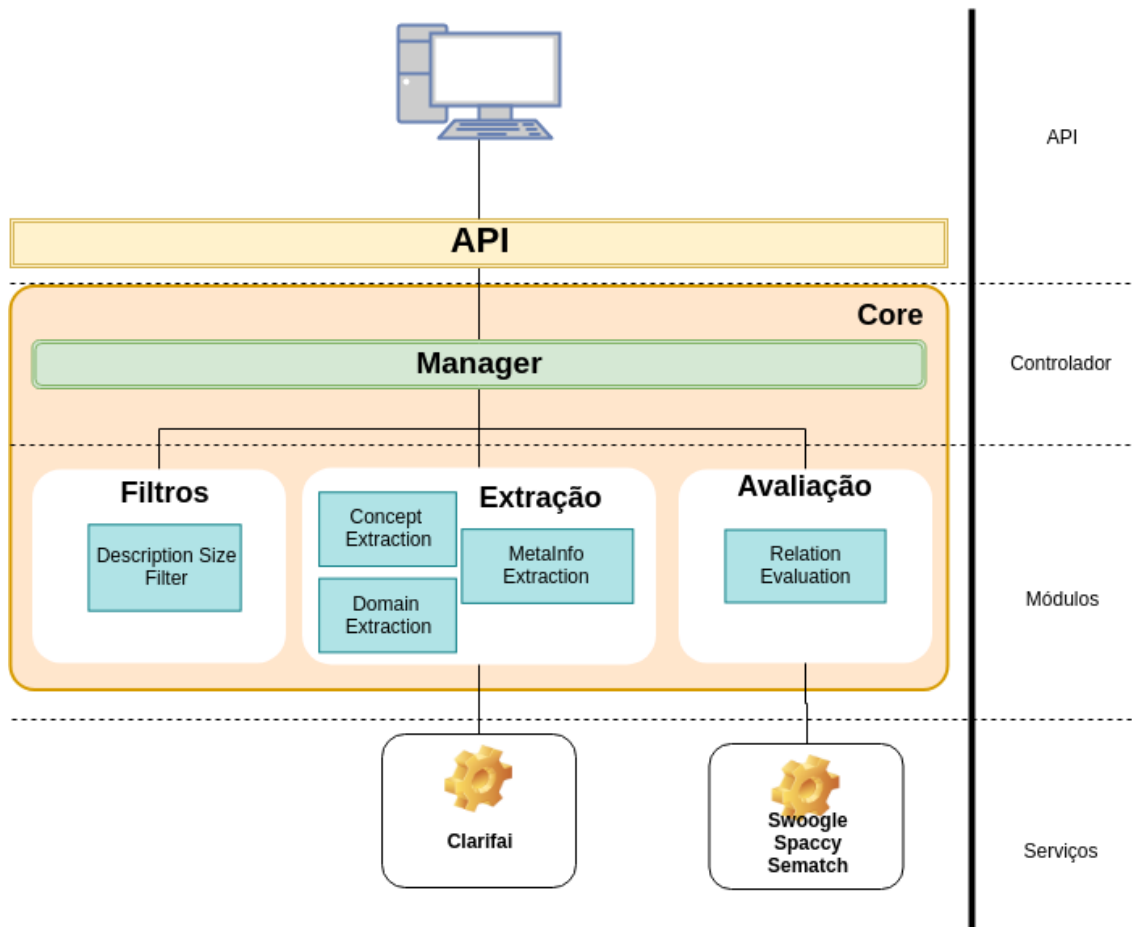


Figura 3.1: Arquitetura ISIAAW

### 3.2 Serviço Web

Como foi definido, ISIAAW funciona como um serviço *Web*, recebendo pedidos HTTP quando é necessário uma avaliação semântica. A Tabela 3.1 descreve este serviço e como executar o pedido para uma avaliação.

Recurso	Tipo	Parâmetros	Codificação dos parâmetros
/isiaaw/	GET	URL: Link da imagem ALT: Descrição da imagem	BASE64

Tabela 3.1: API da aplicação ISIAAW

Ambos os parâmetros são obrigatórios. O parâmetro URL corresponde ao recurso da imagem que se deseja avaliar e o parâmetro alt corresponde à descrição da imagem.



Figura 3.2: Imagem exemplo

```
http://127.0.0.1:5000/warsi?url=aHR0cHM6Ly9pY2hl
Zi0xLmJiY2kuY28udWsvbmV3cy82OTUvY3BzcHJvZHBiLz
E2OTcyL3Byb2RlY3Rpb24vXzEwMTEwMzUyOV9wYXN0YV9nZ
XR0eWltYWdlcy05NDUwOTgxMTguanBn&alt=V29tYW4gZWFO
aW5nIHBhc3RhCg== -
```

Figura 3.3: Exemplo de pedido do ISIAAW

Ambos os parâmetros têm de ser codificados em base64 para serem avaliados. Caso não o sejam, será apresentada uma mensagem de erro. O pedido enviado é do tipo GET, pois não se verificou a necessidade de ocultar qualquer dos parâmetros enviados. No entanto, se o pedido ultrapassar os 2048 caracteres não é possível proceder à avaliação, pois este é o limite máximo do GET. Por fim, o serviço responde ao cliente com o resultado da avaliação em formato JSON. Esta resposta apresenta o resultado da relação entre os dois conteúdos. Esse valor varia entre 0 e 1, sendo 0 nada relacionado e 1 muito relacionado.

A figura 3.2 é usada como exemplo para um pedido de avaliação. A imagem apresentada tem como url=<https://ichef-1.bbci.co.uk/news/womanEatingPasta.jpg> e o seu alt="Woman eating pasta". Convertendo estes dois parâmetros em base64 obtemos o pedido representado na Figura 3.3.

```
1 {"Score":0.76254,"IMG": "https://ichef-1.bbci.co.uk/news/695/cpsprodpb
/16972/production/_101103529_pasta_gettyimages-945098118.jpg", "ALT"
:"Woman eating pasta"}
```

Código 3.1: Exemplo de resposta do ISIAAW

No código 3.1 é apresentada a resposta ao pedido feito à aplicação. Esta resposta é apresentada em formato JSON com três campos: o resultado da relação entre imagem e descrição no campo *Score*, o campo *IMG* que representa o link da imagem e o campo *ALT* que representa o texto alternativo da imagem. Neste caso específico, o ISIAAW conseguiu concluir que esta imagem e o *alt* correspondente estão relacionadas, pois o valor obtido está próximo de 1.

### 3.3 Aplicações externas

Para ser possível alcançar algumas das funcionalidades propostas para este trabalho foram considerados serviços *Web* já existentes. Estas funcionalidades estão relacionadas com a extração de informação de conteúdo textual e não textual. O uso destes serviços permite reduzir o esforço de implementação, permitindo assim que o trabalho defina como prioridade a funcionalidade principal deste sistema: avaliação semântica entre imagens e as suas descrições. Consequentemente, um dos desafios a dar resposta era de encontrar ferramentas capazes de extrair informação chave dos vários tipos de conteúdo. Como foi indicado no ponto 2.4, existe um conjunto de aplicações para extração de informação. Para a escolha das aplicações foram seleccionadas as seguintes condições:

- **Sem limitações:** As aplicações devem permitir o uso de todas as funcionalidades;
- **Sem custo adicional:** Devido ao facto deste trabalho não ser financiado é imperativo usar aplicações gratuitas;
- **Capacidade de personalização:** Possibilidade de personalizar as ferramentas como a criação de modelos, que permitem os utilizadores implementar domínios de conceitos;
- **API em Python:** A aplicação será implementada em Python. Desta forma as aplicações têm de ser capazes de comunicar com esta linguagem;
- **Bom desempenho:** As aplicações usadas devem satisfazer um certo nível de eficácia na extração ou na relação entre elementos.

Com a identificação das condições necessárias para a escolha das aplicações, foram usadas as seguintes aplicações. Para a extração de informação em imagens e texto foi usado o Clarifai e o SpaCy respetivamente. Ambas aplicações são gratuitas e têm APIs para Python. O uso da aplicação SpaCy é gratuito e sem limitações. O Clarifai está limitado a 5000 chamadas. No entanto, comparativamente às aplicações enumeradas no ponto 2.4.2 esta é a única ferramenta que permite a criação de modelos personalizados pelo utilizador. Serão usadas as aplicações Swoogle, Wordnet e Sematch para a avaliação semântica. O Swoogle permite a avaliação entre expressões, o Sematch avalia a relação

entre entidades (Bill Gates e Microsoft), e por último o Wordnet é usado para a obtenção de relações entre duas palavras. Estas são totalmente gratuitas e sem limitações de uso. Em seguida, serão identificadas as funcionalidades e o objetivo destas.

## Clarifai

O serviço Clarifai tem a capacidade de extrair informação de imagens por meio de modelos de reconhecimento. O seu uso requer que seja feito a instalação do pacote "clarifai". Torna-se necessário a atribuição de uma chave de utilizador. Esta permite que o cliente use as funcionalidade deste serviço que pode ser obtido através da criação de uma conta gratuita no sítio do Clarifai.

Para iniciar o serviço basta chamar a função `ClarifaiAPP()`, Código 3.2 (Linha 1). Nesta linha é imperativo enviar a chave do utilizador como parâmetro pois sem esta não é possível utilizar a ferramenta. Posteriormente à iniciação do Clarifai, é escolhido o modelo, usando a função `app.models.get()`. Estes modelos representam domínios, que redirecionam a previsão do Clarifai para a extração de conceitos mais específicos. O serviço fornece no total oito modelos: Genérico, Comida, Viagem, e outros representados na Tabela 3.2. Como já descrito, estes modelos direcionam a previsão para um domínio, como por exemplo uma imagem enviada para o Clarifai com o modelo food. A sua previsão vai dirigir os conceitos extraídos para o domínio Comida.

Modelo	Chave	Descrição
Vestuário	apparel	Reconhece objetos de moda
Celebridades	celeb-v1.3	Identifica celebridades em imagens
Cor	color	Determina as cores dominantes
Comida	food-item-v1.0	Reconhecimento de ingredientes
Genérico	general-v1.3	Modelo de reconhecimento generalizado
Texturas e padrões	Textures & Patterns	Identificação de texturas e padrões
Viagem	travel-v1.0	Identificação de conceitos de viagem
Casamentos	wedding-v1.0	Identificação de conceitos de casamento

Tabela 3.2: Modelos do Clarifai e as suas descrições

Por fim, para submeter a imagem para reconhecimento, é usada a função `predict_by_filename()` ou `predict_by_url()`, caso seja ficheiro ou URL. Por exemplo a previsão da Figura 3.4, pela função `predict`, recebe uma resposta em formato JSON representada na Figura 3.5. A parte mais importante, desta resposta encontra-se no campo `outputs.data.concepts`. Neste encontra-se uma lista de elementos. Cada elemento tem o nome do conceito encontrado (`name`) e a probabilidade de se encontrar na imagem (`value`).

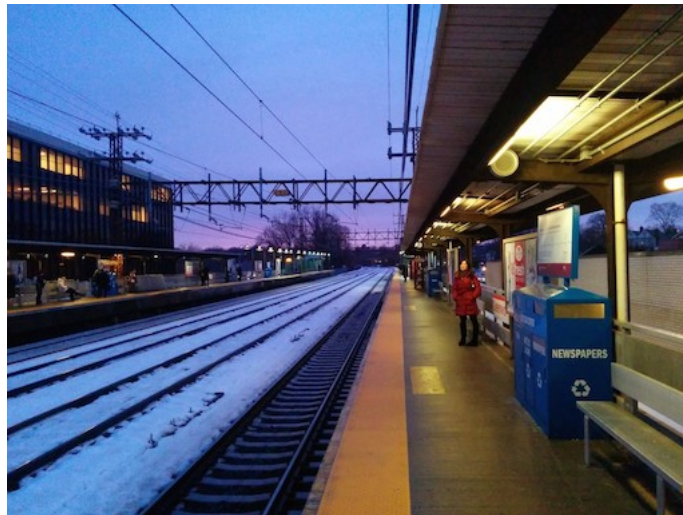


Figura 3.4: Imagem exemplo

### Uso no ISIAAW

Atualmente, o sistema usa o Clarifai na sua totalidade, com exceção dos modelos de textura e cor. O modelo *General* (general-v1.3) é sempre usado nas imagens submetidas para avaliação dando assim o conjunto base de conceitos da imagem. Posteriormente, poderão ser usados outros modelos, caso algum dos conceitos extraídos esteja relacionado com um dos modelos. Ou seja, se um dos conceitos extraídos pelo modelo *General* for o conceito *food*, será utilizado o modelo *Food*, de forma a ser possível extrair mais informação da imagem.

```

1 app = ClarifaiApp(api_key='USER_API_KEY')
2 app.models.get("general-v1.3")
3 tries = 0
4 def conceptExtraction_General ( img , minProb = 0.6, maxConcepts = 40,
   isFile = False):
5     list = []
6     try:
7         #Extracao de conceitos de Ficheiro ou URL
8         if isFile:
9             j = json.dumps(model.predict_by_filename(img, max_concepts=
              maxConcepts))
10        else:
11            j = json.dumps(model.predict_by_url(img, max_concepts =
              maxConcepts))
12        #Descompactar resultados da previsao
13        tt = json.loads(j)
14        if len(tt['outputs'][0]['data']) > 0:
15            for x in range(0, len(tt['outputs'][0]['data']['concepts'])):
16                name = tt['outputs'][0]['data']['concepts'][x]['name']
17                value = tt['outputs'][0]['data']['concepts'][x]['value']
18                if value >= minCof:
19                    \\Preencher estrutura
20                    list.append(dict(img=img, concept= name , probability = value
              , SK= None))

```

```
21     return list
22     #Em caso de erro faz se uma nova tentativa de previsao ate serem
    realizadas 3 tentativas
23 except Exception as e:
24     print e
25     if tries < 3:
26         tries += 1
27         list = getImageWords (img, maxConcepts = maxConcepts, tries =
    tries, isFile = isFile)
28     else:
29         tries = 0
30     return -1
31 return list
```

Código 3.2: Código usado no ISIAAW

```
{
  "status": {
    "code": 10000,
    "description": "Ok"
  },
  "outputs": [
    {
      "data": {
        "concepts": [
          {
            "id": "ai_HLmqFqBf",
            "name": "train",
            "app_id": null,
            "value": 0.9989112
          },
          {
            "id": "ai_fv1BqXZR",
            "name": "railway",
            "app_id": null,
            "value": 0.9975532
          },
          {
            "id": "ai_Xxjc3MhT",
            "name": "transportation_system",
            "app_id": null,
            "value": 0.9959158
          },
          {
            "id": "ai_6kTjGfF6",
            "name": "station",
            "app_id": null,
            "value": 0.992573
          }
        ], ...
      }
    }
  ]
}
```

Figura 3.5: Resultado

## SpaCy

SpaCy é uma biblioteca de *software* de código aberto para processamento avançado de linguagem natural (NLP), escrita nas linguagens de programação Python e Cython. A biblioteca é publicada sob a licença do MIT e atualmente oferece modelos estatísticos de redes neurais para a língua inglesa, alemã, espanhola, portuguesa, francesa, italiana, holandesa e multi-língua. Para usar este **software** é instalado o pacote “spacy” usando o pip. Antes de ser possível utilizar este software é necessário fazer download de um dos modelos que representam os dialetos<sup>1</sup>. Basta escrever a seguinte linha de comando, `python -m spacy download en`, sendo **en** o modelo da língua inglesa. As funcionalidades fornecidas pelo spaCy são as seguintes:

- Part-of-speech tagging: Identificação das classes de palavras;
- Named entity: Identificação de Entidades como Pessoas, Locais ou mesmo Organizações no texto;
- Tokenization: Separação de texto em segmentos de tamanho reduzido;
- Dependency Parse: Identificação de dependências no texto.

```
1 spaCy = spacy.load('xx')
2
3 ...
4
5 def __namedEntityOnText(txt):
6     doc = spaCy(txt)
7
8     entitys = list()
9     for ent in doc.ents:
10        m = dict()
11        m["Text"] = ent.text
12        m["Label"] = ent.label_
13        entitys.append(m)
14    return entitys
```

Código 3.3: Código do Spacy usado no ISIAAW

### Uso no ISIAAW

Das funcionalidades descritas, somente a funcionalidade de Named Entity será usada para a aplicação ISIAAW. Uma entidade é um “objeto do mundo real” ao qual é atribuído um nome. Por exemplo, uma pessoa, um país, um produto ou mesmo um título de um livro. Para ser extraída esta informação, é preciso pedir uma previsão ao spaCy. O Código 3.3 apresenta o funcionamento do Named Entity. Antes de ser possível a previsão

---

<sup>1</sup><https://spacy.io/models/>

das entidades, é necessário iniciar o spaCy, usando a função `spacy.load()`, dando como parâmetro o modelo linguístico desejado. Posteriormente para pedir uma previsão é enviado o texto usando a função `spaCy()`. No fim da previsão, o spaCy responde com uma estrutura de dados que apresenta o nome da entidade e o seu tipo. Por exemplo, se for enviado para o spaCy a seguinte frase: “Apple is looking at buying U.K. startup for \$1 billion”, o resultado da previsão será o indicado na Tabela 3.3 .

Nome	Tipo
Apple	ORG
U.K	GPE
\$1 billion	MONEY

Tabela 3.3: Resultado do SpaCy

## Swoogle

Como apresentado na secção 2.3, Swoogle é uma aplicação capaz de comparar semanticamente duas frases. Atualmente o serviço tem 3 funcionalidades:

- Top-N Similarity: Apresenta N palavras semelhantes à palavra recebida;
- Phrase Similarity: Avalia a semelhança semântica entre dois nomes abreviados ou frases verbais;
- STS Similarity: Avalia a semelhança semântica entre duas frases.

Para pedir uma avaliação basta fazer uma solicitação HTTP (Figura 3.6)

```
http://swoogle.umbc.edu/StsService/GetStsSim?operation=api&phrase1=a%20small%20violin%20is%20being%20played%20by%20a%20girl&phrase2=a%20child%20is%20performing%20on%20a%20tiny%20instrument
```

Figura 3.6: Exemplo de pedido ao Swoogle

Neste pedido podemos especificar o *corpus* e o tipo de avaliação desejada. O *corpus* pode ter dois valores, *webBase* (Refined Stanford WebBase) e *gigawords English Gigawords Corpus*. Relativamente ao tipo de avaliação pode indicar-se *relation* (relação) e *concept* (conceitos). *Relation* avalia a relação semântica existente entre elementos. *Concept* avalia a semelhança entre todas as palavras. Os parâmetros podem ser adicionados ao URL usando `corpus=` , para a escolha do corpus e `type=`, para a escolha do tipo de avaliação. A Tabela 3.4 resume os parâmetros possíveis para o pedido.

Parâmetro	Chave	Valores possíveis	Obrigatório
Função	Operação	top_sim/api/STS	Sim
Frase 1	phrase1	Texto	Sim
Frase 2	phrase2	Texto	Sim
Corpo	corpus	webBase / gigawords	Não
Tipo de avaliação	type	relation / concept	Não

Tabela 3.4: Parâmetros do Swoogle

### Uso no ISIAAW

No ISIAAW serão usadas 2 das 3 funcionalidade do Swoogle. Estas são a top-n Similarity e a Phrase Similarity. A primeira é usada como verificador. O objetivo deste é verificar se os conceitos retirados da imagem, via Clarifai, são conhecidos pelo Swoogle. A segunda funcionalidade será usada para avaliar a relação semântica entre os conceitos extraídos e a descrição da imagem. Os parâmetros usados para esta avaliação são o *web-base* como *corpus* e o tipo de avaliação *relation*. Foram selecionados estes parâmetros devido aos seguintes fatores:

- Relacionar conceitos e avaliar as relações, desta forma foi usado o *relation* do *swoogle*;
- Foi escolhido o *webbase* devido a este ter obtido melhor resultados comparativamente ao *gigawords*. [46]

```

1 req = "http://swoogle.umbc.edu/SimService/GetSimilarity"
2
3 # phrase1 = Frase 1 a comparar
4 # phrase2 = Frase 2 a comparar
5 # type = 0 (concept similarity) ou 1 (relation similarity)
6 # corpus = webbase (WebBaseCorpus) ou gigawords (English Gigawords
  Corpus)
7 def swoogle(s1, s2, type='relation', corpus='WebBase'):
8     response = ""
9     try:
10        params = {'operation': 'api', 'phrase1': s1.lower(),
11                 'phrase2': s2.lower(), 'type': type, 'corpus': corpus}
12
13        response = get(req, params=urlencode(params))
14        r = float(response.text.strip())
15        if r < 0:
16            r = 0.0
17        return r
18    except:
19        print 'Error in getting similarity for %s: %s' % ((s1, s2), response)
20        return -1

```

Código 3.4: Código do Swoogle usado no ISIAAW

## Sematch

Sematch é uma aplicação desenvolvida para a avaliação semântica em Ontologias, nomeadamente a Wikipédia e o Yago. Comparando o sematch com o swoogle, o primeiro permite relacionar Entidades, ao contrario do swoogle que não é capaz de as avaliar. Para instalar o pacote Sematch usa-se o pip. Esta aplicação permite também relacionar conceitos usando as métricas apresentadas na secção 2.3. No entanto, estas não serão usadas no trabalho, dado que só relacionam palavras e não entidades. Para realizar uma avaliação entre entidades, é usada a função `sim.word_similarity(word1, word2)`. `word1` e `word2` correspondem a dois conceitos, que obrigatoriamente têm de corresponder a entidades. Caso contrário o valor da relação é 0. O valor desta avaliação é retornado num valor numérico entre 0 e 1, que representa a relação.

### Uso no ISIAAW

A aplicação Sematch é usada de forma a complementar a aplicação Swoogle. Com a aplicação Sematch o serviço ISIAAW é capaz de avaliar semanticamente pessoas, organizações, locais e outros, que não sucede com o Swoogle. Desta forma é possível usar a informação extraída pelo spaCy para complementar a avaliação.

## 3.4 Algoritmo

ISIAAW é um serviço Web, tendo como objetivo avaliar a relação entre as imagens e as correspondentes descrições. Para ser possível a avaliação da relação entre estes conteúdos, estes tem que passar por 3 passos (Figura 3.7): Filtragem, Extração e Avaliação.

O primeiro passo do ISIAAW é a filtragem (Módulo Filtragem). Este é responsável por filtrar e avaliar a qualidade da descrição com base no tamanho. Esta avaliação vai produzir um peso, entre 0 e 1, que posteriormente vai ser usado na avaliação.

O segundo passo é a extração de conceitos chave nas imagens. Este passo está dividido em três extrações:

- **Extração de conceitos base** (Módulo Concept Extraction): conceitos gerais que não indicam nenhum domínios em específico.
- **Extração de Dominio** (Módulo Domain Extraction): extração do domínio
- **Extração de MetaDados** (Módulo MetaInfo Extraction): extração dos metadados da imagem, caso estes existam

Estas extrações tem como objetivo fornecer conceitos chave para ser possível a avaliação entre texto e imagem.

O último passo do ISIAAW é a avaliação (Módulo Relation Evaluation). Este junta os resultados dos passos anteriores para obter um valor de relação entre descrição e imagem. Esta avaliação está dividida em 3 ordens:

- Primeira Ordem: Avaliação dos conceitos chave
- Segunda Ordem: Avaliação dos domínios
- Terceira Ordem: Avaliação dos MetaDados

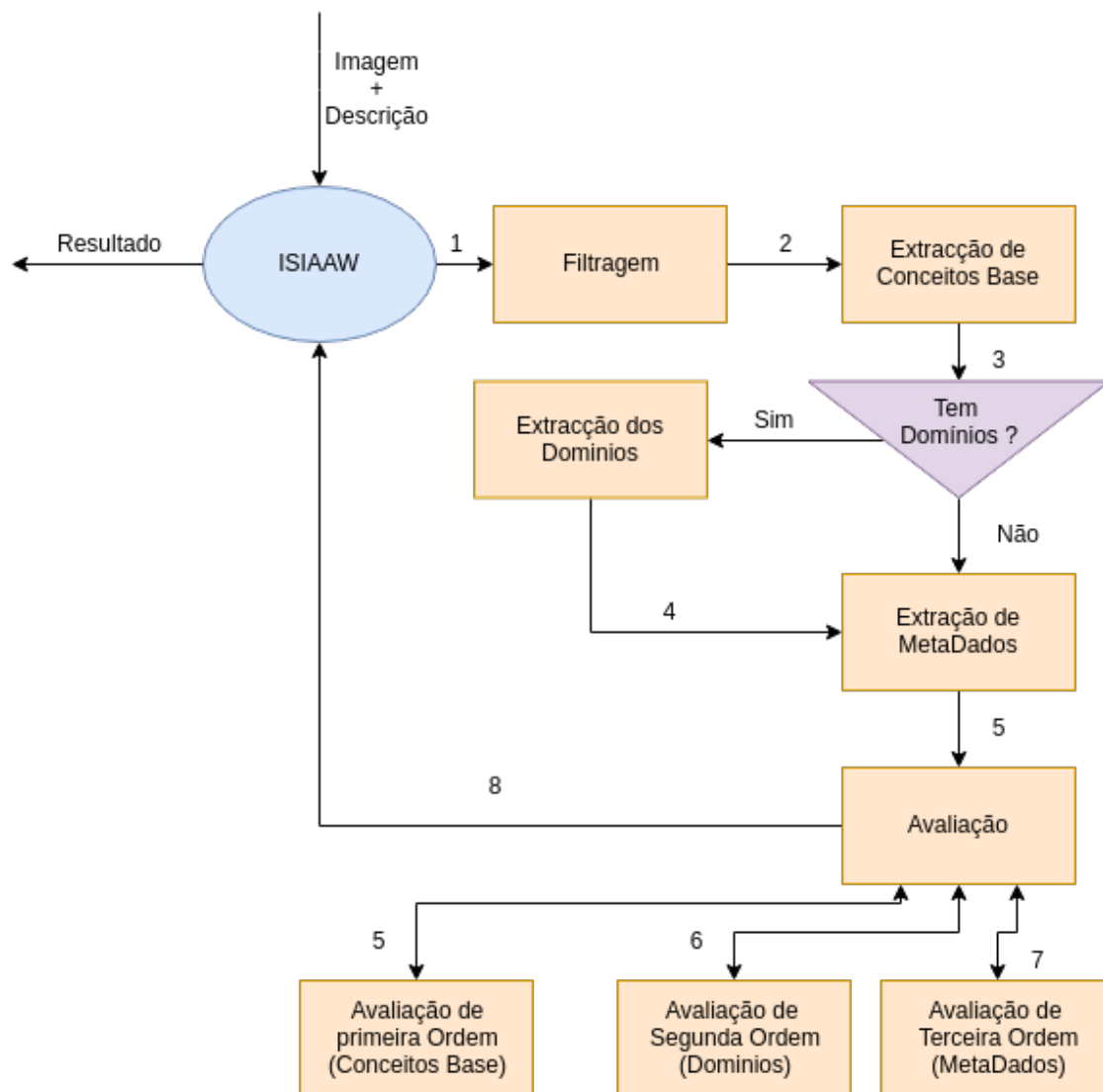


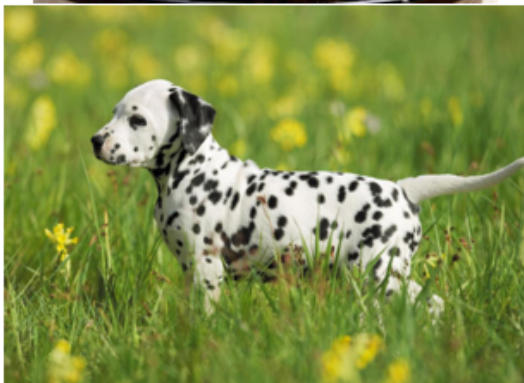
Figura 3.7: Fluxograma da aplicação ISIAAW

No fim dos 3 passos é obtido um valor de relação entre a descrição e imagem entre 0 e 1. Nos vários passos já enumerados existe um conjunto de parâmetros que podem ser alterados de forma a alterar o valor de retorno do serviço. Com o intuito de otimizar estes foi construído um conjunto de treino. A

partir de vários sítios de notícias e blogues foram extraídos 45 conjuntos de imagens e os correspondentes textos alternativos que podem ser visualizados no seguinte link: [https://docs.google.com/spreadsheets/d/1rN3m1fer\\_NvERpq4T6qEw6oqfWqHRrgErHNtcz590XE/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1rN3m1fer_NvERpq4T6qEw6oqfWqHRrgErHNtcz590XE/edit?usp=sharing)



Vegetable burger



Dalmatians suffer from Hyperuricemia due to in-breeding



José Mourinho

Figura 3.8: Exemplo do conjunto de treino

A imagem 3.8, apresenta um exemplo das imagens e descrições retiradas. Após este levantamento foi realizada uma avaliação de forma a avaliar a relação entre o texto alternativo e a imagem. A avaliação consistiu em classificar, numa escala entre 1 e 4, se o texto alternativo e a imagem estavam relacionadas, com base nas heurísticas levantadas na Secção 2.2.1. Os resultados desta avaliação podem ser observados na coluna Classificação do link acima referido. Desta avaliação foram obtidos os seguintes valores de média e desvio padrão:

$$Media = 2.15$$

$$DesvioPadrao = 1.106$$

Por último, na Figura 3.9 é apresentada a distribuição das classificações. Nesta é possível verificar que existe um número maior de ocorrências de más descrições que boas descrições, o que também já era indicado pela média inferior a 2.5.

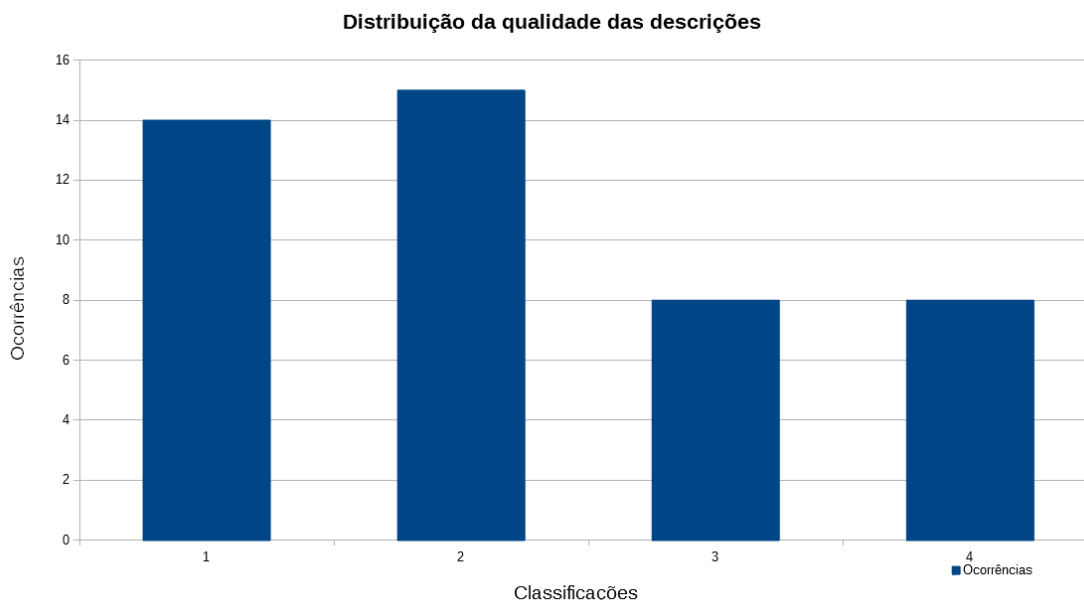


Figura 3.9: Distribuição da qualidade das descrições no conjunto de treino

## 3.5 Módulos

Nesta secção serão descritos os módulos implementados, detalhando os objetivos e o algoritmo. Em cada módulo também serão discutidos alguns parâmetros importantes e como se chegou a certos valores. Por último será também discutido o algoritmo de avaliação usado. Este será discutido após a introdução dos módulos.

### 3.5.1 Filtro do Tamanho dos alts

O primeiro módulo a ser usado é o Filtro do Tamanho dos alts. Este tem como objetivo avaliar se o texto alternativo/descrição vai contra duas das heurísticas descritas na Secção 2.2:

- Texto alternativo breve;
- Não ultrapassar os 140 caracteres.

Caso alguma destas heurísticas seja infringida, o módulo deve penalizar o resultado final da avaliação. Para ser obtido uma forma justa de penalização, foi necessário estudar

a relação entre o número de palavras e a qualidade das descrições medida de acordo com a avaliação efetuada ao conjunto de treino. Desta forma optou-se por juntar as várias descrições em intervalos de palavras. Estes intervalos correspondem ao número de palavras existentes na descrição. Com a criação desses intervalos foram analisados os conjuntos de dados classificados positivamente. Só as classificações positivas foram estudadas, dado que o objetivo desta análise é avaliar a relação entre boa descrição e o número de palavras, de forma a descobrir qual é o número ideal de palavras numa “boa” descrição. A partir desta análise foram obtidos os resultados apresentados na Tabela 3.5.1.

<b>Intervalo</b>	<b>Boas Descrições</b>	<b>Frequência</b>	<b>Filtro</b>
<b>0</b>	0	0	0,135
<b>1</b>	1	0,062	0,246
<b>2-8</b>	6	0,4	0,916
<b>9-13</b>	5	0,333	0,832
<b>14-18</b>	3	0,188	0,554
<b>16-20</b>	1	0,062	0,356
<b>Total</b>	16		

Tabela 3.5: Relação do número de palavras e a qualidade das descrições

Com o levantamento das boas descrições em cada intervalo, foi calculada a frequência de uma boa descrição, apresentada na coluna **Frequência**. Com a obtenção de frequências é possível detetar que as boas descrições possuem, mais frequentemente, entre 2 a 13 palavras, pois as frequências mais altas encontram-se nesse intervalo. Posteriormente ao cálculo da frequência é necessário obter um valor teórico da penalização. Para ser possível a obtenção desse valor, foram usados os valores da frequência obtidos e normalizados. Com a normalização é possível obter um valor entre 0 e 1. Esta normalização encontra-se na coluna **Filtro**, indicando a qualidade da descrição baseada no seu tamanho. Este valor será multiplicado pelo valor obtido pelo módulo Relation Evaluation, como pode ser observado na fórmula:

$$Relacao = Filtro * RE$$

$$RE = Valor\ Obtido\ via\ Relation\ Evaluation$$

Usando a multiplicação, é possível verificar que valores altos, próximos de 1, têm pouco impacto. Por outro lado, valores baixos, próximos de 0, fazem que o valor final da relação aproxime-se de 0, indicando assim uma má descrição. A figura 3.10 representa o gráfico da tabela com os intervalos de palavras sobre a qualidade. Este gráfico apresenta também a variação do valor da qualidade ao longo dos intervalos.

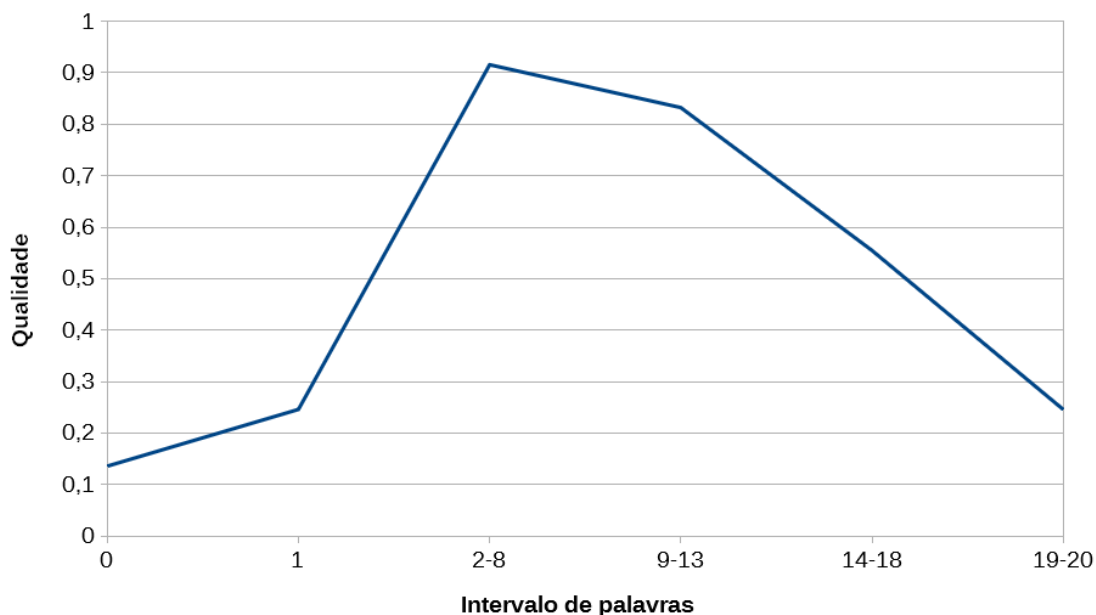


Figura 3.10: Gráfico da qualidade das descrições em função do número de palavras

### 3.5.2 Módulo Concept Extraction

O segundo passo do serviço é a extração de conteúdo das imagens. O módulo Concept Extraction é capaz de extrair conceitos de imagens usando o serviço Clarifai. Como já referido, o Clarifai permite a escolha de vários modelos. No entanto, para este módulo só é usado o modelo General, devido ao facto de ser necessário uma ideia generalizada da imagem antes de se detalhar, com o uso de outros domínios. A capacidade de extração deste módulo é altamente dependente do serviço Clarifai. No entanto é possível alterar o serviço de extração alterando o código. Este módulo tem como objetivo responder a duas das heurísticas descritas na Secção 2.2:

- Os conceitos mais relevantes da imagem devem ser descritos no texto alternativo;
- Quanto maior for o número de conceitos da imagem na descrição, maior é a probabilidade de esta ser uma boa descrição.

Como é óbvio, este módulo em separado não consegue responder a estas heurísticas. No entanto, devido à capacidade de extração de conceitos, fornece uma parte da solução para ambas.

A estrutura de resposta deste módulo pode ser observada no código 3.6. O resultado contém os seguintes campos: `Img`, `Concept` e `Probability`. O campo `Img` corresponde ao URL da imagem, `Concept` corresponde ao nome do conceito extraído e o campo `Probability` corresponde à probabilidade deste conceito fazer parte da imagem. O resultado final deste módulo produz uma lista com os conceitos e probabilidades.

Para ser pedido ao módulo para iniciar uma extração é usada a função **conceptExtraction\_General()** (Código 3.5). Esta tem vários parâmetros de entrada. A saber:

- **img** - Representa o link/URI da imagem a avaliar;
- **minProb** (Opcional) - Representa a probabilidade mínima de cada conceito;
- **maxConcepts** (Opcional) - Representa o número máximo de conceitos a extrair;
- **isFile** (Opcional) - Representa se o parâmetro **img** é um ficheiro ou um link.

Em termos de execução a função tem 3 partes importantes. A primeira é a previsão da imagem, que pode ser observada no intervalo de linhas 8-11, nas quais é usada a função **predict()** para ficheiros e para *URLs*. Na segunda é realizada a descompactação dos resultados e transformados os resultados na estrutura de dados já referenciada anteriormente. Na terceira, caso de algum dos passos tenha sido interrompido devido a uma falha na previsão ou na leitura do JSON é realizado um novo pedido a esta função. Atualmente o serviço realiza no máximo 3 tentativas. Caso sejam ultrapassada é enviado um código de erro (-1) como resposta. Terminado o uso desta função, em caso de sucesso, é enviada uma estrutura de dados igual à apresentada no Código 3.6.

```

1 app = ClarifaiApp(api_key='USER_API_KEY')
2 app.models.get("general-v1.3")
3 tries = 0
4 def conceptExtraction_General ( img , minProb = 0.6, maxConcepts = 40,
   isFile = False):
5     list = []
6     try:
7         #Extracao de conceitos de Ficheiro ou URL
8         if isFile:
9             j = json.dumps(model.predict_by_filename(img, max_concepts=
              maxConcepts))
10        else:
11            j = json.dumps(model.predict_by_url(img, max_concepts =
              maxConcepts))
12        #Descompactar resultados da previsao
13        tt = json.loads(j)
14        if len(tt['outputs'][0]['data']) > 0:
15            for x in range(0, len(tt['outputs'][0]['data']['concepts'])):
16                name = tt['outputs'][0]['data']['concepts'][x]['name']
17                value = tt['outputs'][0]['data']['concepts'][x]['value']
18                if value >= minCof:
19                    \\Preencher estrutura
20                    list.append(dict(img=img, concept= name , probability = value
              , SK= None))
21        return list
22        #Em caso de erro faz se uma nova tentativa de previsao ate serem
        realizadas 3 tentativas
23    except Exception as e:
24        print e
25        if tries < 3:
26            tries += 1

```

```
27     list = getImageWords (img, maxConcepts = maxConcepts, tries =
        tries, isFile = isFile)
28     else:
29         tries = 0
30         return -1
31     return list
```

Código 3.5: Código de extração de conceitos

### 3.5.3 Módulo Domain Extraction

O terceiro passo do serviço reporta-se à deteção de domínios e à extração de conceitos relacionados. Tal como o módulo anterior, este usa o Clarifai de forma a extrair conceitos. Ao contrário do Concept Extraction este usa os restantes modelos do Clarifai. Atualmente o serviço usa os seguintes modelos:

- Vestuário (apparel)
- Celebidades (celeb-v1.3)
- Comida (food-item-v1.0)
- Viagem (travel-v1.0)
- Casamentos (wedding-v1.0)

Dos modelos implementados pelo Clarifai somente os modelos de Cor e Textura & Padrões não foram utilizados. Com efeito estes modelos não detetam a probabilidade dos conceitos mas sim, a quantidade de cor ou padrão que existe na imagem.

Antes do uso do Clarifai, para o reconhecimento dos domínios, é verificado se alguns dos conceitos extraídos pelo Concept Extraction estão relacionados com os nomes dos modelos. Para ser possível verificar a relação entre conceitos e os modelos, é usado o Swoogle. Após o cálculo da relação, se o valor for superior a 0.5 é utilizado o modelo.

Por exemplo, com o uso da imagem<sup>2</sup> apresentada na Figura 3.11, são extraídos os seguintes conceitos (Código 3.6) pelo módulo Concept Extraction: bride, groom, engagement, ceremony, marriage e outros. Quando o controlo passa os conceitos detetados para o Domain Extraction, este tenta verificar se existe algum tipo de relações com os modelos, usando o Swoogle como ferramenta. Relacionando os conceitos pelo nome dos modelos, existem casos onde a relação é superior a 0.5. Por exemplo, o conceito Bride e o modelo Wedding estão relacionados. Usando o Swoogle para verificar a sua relação é obtido o valor 0.596, como é apresentado na Figura 3.12. Ao identificar esta relação é pedido ao

---

<sup>2</sup>Link da Imagem: <http://www.onestoryweddings.com/wp-content/uploads/Orange-County-Wedding-Photographer-23-600x460.jpg>



Figura 3.11: Imagem exemplo : One Story Wedding

Clarifai para realizar uma previsão usando o modelo Wedding, apresentando os seguintes conceitos que figuram no Código 3.6.

Os novos conceitos extraídos são acrescentados aos já encontrados pelo serviço. Caso estes já existam na lista de conceitos do Concept Extraction, são ignorados.

Phrase 1:

Phrase 2:

Type:  Concept Similarity (noun phrase only)  Relation Similarity

Corpus:  Refined Stanford WebBase corpus  LDC English Gigawords Corpus (American newswire services only)

The similarity score is 0.59620804

Figura 3.12: Resultado da relação entre Bride e Wedding

```

1 {img="http://www.onestoryweddings.com/wp-content/uploads/Orange-County-
  Wedding-Photographer-23-600x460.jpg", concept="bouquet",
  probability=0.988},
2 {img="http://www.onestoryweddings.com/wp-content/uploads/Orange-County-
  Wedding-Photographer-23-600x460.jpg", concept="bride", probability
  =0.986},
3 {img="http://www.onestoryweddings.com/wp-content/uploads/Orange-County-
  Wedding-Photographer-23-600x460.jpg", concept="groom", probability
  =0.970},
4 {img="http://www.onestoryweddings.com/wp-content/uploads/Orange-County-
  Wedding-Photographer-23-600x460.jpg", concept="flowers",
  probability=0.941},
5 {img="http://www.onestoryweddings.com/wp-content/uploads/Orange-County-
  Wedding-Photographer-23-600x460.jpg", concept="love", probability
  =0.932},
6 ...
7 }

```

Código 3.6: Conceitos extraídos do serviço Clarifai com o modelo Wedding

### 3.5.4 Módulo MetaInfo Extraction

O último módulo de extração de informação fornece métodos de extração informação dos metadados da imagem. Com o passar dos anos têm vindo a surgir formas de adicionar informação nos metadados das imagens. Estas informações podem ser conceitos, localização e outros. Este módulo tem como função principal a extração dessa informação. Caso esta exista, conceitos ou nomes de localizações, são adicionadas à lista de conceitos. No entanto, ao contrário do uso do Clarifai este não é capaz de dar uma percentagem que corresponda à probabilidade de estes conceitos estarem corretos. Desta forma a probabilidade destes conceitos será iniciada a 1, tendo em conta a seguinte premissa: estes metadados são introduzidos por pessoas que são capazes de detetar os conceitos chave e as localizações corretas das imagens.

Devido à escassez de imagens com metadados, não é possível usar só o módulo MetaInfo. Por exemplo, nos dados de treino usados somente 4 imagens em 45 têm metadados, o que poderia dificultar a avaliação semântica dos conteúdos. Desta forma, este módulo deve ser usado de forma a complementar a extração de informação dos outros módulos.

Para ser realizada a extração desta informação é usado o pacote Pillow do python. Tendo a Figura 3.13 como exemplo, usando a função `_getexif()` são obtidos os Metadados apresentados no Código 3.7. É de notar que não está apresentada a totalidade do resultado devido à quantidade de metadados presentes nesta imagem. Ao estudar os dados obtidos com o uso do Pillow é possível detetar vários campos importantes. Estes são: Country, Keywords, Image Description, City e Sub-Location. Todos estes, exceto o Image Description, são adicionados aos conceitos extraídos pelo MetaInfo Extraction.

```
1 Image Name[2,5]: 20101005-FRA-TG-395.jpg
2   Created Date[2,55]: 20101005
3   ImageDescription[2,55]: Ripe black grapes in vineyard
4   and the town of St Emilion in the Bordeaux
5   wine region of France
6   City[2,90]: St Emilion
7   Province State[2,95]: Bordeaux
8   Country[2,101]: France
9   Keyword[2,25]: Bordeaux
10  Keyword[2,25]: Bordelais
11  Keyword[2,25]: France
12  Keyword[2,25]: French Scenes
13  Keyword[2,25]: French Vineyard
14  Keyword[2,25]: French chateau
15  Keyword[2,25]: French countryside
16  Keyword[2,25]: French mansion
17  Keyword[2,25]: Rural scene
18  Keyword[2,25]: Saint Emilion
19  ...
20  Keyword[2,25]: wine production
21  Keyword[2,25]: wine region of France
22  Keyword[2,25]: winemaking
```



Figura 3.13: Imagem Bordeaux

```
23 Keyword[2,25]: winery
24 Sub-Location[2,92]: Bordeaux
25 Country Code[2,100]: FRA
26 unknown[2,221]: 0:0:0:-00001
```

Código 3.7: Metadados da Imagem Bordeaux

### 3.5.5 Módulo Relation Evaluation

Com a fase da extração de conceitos concluída, é necessário avaliar a relação semântica entre a imagem e os seus conceitos. No Código 3.8 é apresentado o pseudocódigo deste módulo. Este para realizar a avaliação dos conteúdos tem que realizar 5 passos, que podem ser observados no código assinalados com cardinal (#). Estes passos são:

1. Verificação;
2. Avaliação dos conceitos base;
3. Avaliação dos conceitos do domínio;
4. Avaliação dos conceitos de Metadados;
5. Penalização.

```
1 ConceitosBase = ConceptExtraction
2 ConceitosDominio = DomainExtraction
3 ConceitosMeta = MetaExtraction
```

```
4 Filtro = Peso do tamanho da descricao
5 Crd = 0
6 Crs = 0
7 Csw = 0
8 #Verificacao
9 for conceito in ConceitosBase:
10     if conceito nao eh reconhecido pelo swoogle:
11         conceito eh apagado da lista ConceitosBase
12
13 #Avaliacao dos conceitos Base
14 for conceito in ConceitosBase:
15     if (conceito esta presente na descricao):
16         Incrementa Crd
17         Salta as restantes relacoes
18
19     if (conceito tem um sinonimo na descricao):
20         Incrementa Crs
21         Salta as restantes relacoes
22
23     if (conceito tem relacao superior a 0.5 via Swoogle):
24         Incrementa Csw
25         Salta as restantes relacoes
26 eval = (Crd + Crs + Csw)/Tamanho do ConceitosBase
27
28 #Avaliacao dos conceitos dominio
29 for conceito in DomainConcepts:
30     if (conceito esta presente na descricao):
31         eval Incrementa 0.45
32     if (conceito tem um sinonimo na descricao):
33         eval Incrementa 0.25
34     if (conceito tem relacao superior a 0.21 via Swoogle):
35         eval Incrementa 0.1
36
37 #Avaliacao dos conceitos Meta
38 entidades = Extracao de entidades via SpaCy (Locais, Pessoas, etc...)
39 for conceito in MetaConcepts:
40     if (conceito esta presente na descricao):
41         eval Incrementa 0.45
42     for entidade in entidades:
43         if (conceito tem relacao com entidade via Sematch):
44             eval Incrementa 0.25
45
46 #Peso do tamanho da descricao
47 eval = Filtro * eval
48 if eval > 1:
49     eval = 1
```

Código 3.8: Pseudocódigo do Relation Evaluation

## Verificação

O primeiro passo, tem como objetivo verificar se todos os conceitos extraídos pelo Concept Extraction existem no vocabulário do Swoogle. Este passo permite ultrapassar

alguns problemas com o uso da ferramenta Swoogle, que em vários casos apresentava valores de relação 0 devido ao facto dos conceitos extraídos não fazerem parte do vocabulário. Esses conceitos podem estar relacionados, mas devido à inexistência no vocabulário do Swoogle, podiam influenciar negativamente o resultado final. Desta forma, caso o Swoogle não conheça algum conceito, este é retirado da lista. Ao retirar este conceito estamos, a garantir que todos os conceitos da lista dão o valor correto de relação.

Para ser realizada essa verificação, é enviado o conceito para o método `swoogleKnows()` (Código 3.9). Este método retorna **True** se for conhecido e **False** se não for. Esta avaliação é feita usando a primeira funcionalidade do Swoogle, Top-N Similarity. Caso seja detetado o Warning na resposta, significa que o conceito não é reconhecido e posteriormente retirado.

```

1 req = "http://swoogle.umbc.edu/SimService/GetSimilarity"
2
3 def swoogleKnows( word, corpus= "WebBase" ):
4     response = ""
5
6     params = {'operation': 'top_sim', 'phrase1': word, 'corpus': corpus}
7
8     response = get( req, params=urllib.urlencode( params ) )
9     if "Warning:" not in response.text:
10         return True
11     return False

```

Código 3.9: Verificação dos conceitos no Swoogle

### Avaliação dos conceitos base

O segundo passo do Relation Evaluation avalia a relação entre a descrição e os conceitos extraídos pelo Concept Extraction. Esta avaliação é feita usando 3 formas de relação distintas:

1. **Relação direta:** Detetar se o conceito está na descrição.
2. **Relação via Sinónimo:** Detetar se algum sinónimo do conceito está na descrição
3. **Relação via Swoogle:** Avaliar se a relação do conceito + descrição pelo Swoogle é superior a 0.25.

Para o conceito estar relacionado basta ser validado por uma das avaliações. Caso o conceito tenha sucesso num dos tipos de relação, as posteriores não são avaliadas. Por exemplo, se um conceito for validado na Relação direta, este não realiza a avaliação da relação via Sinónimo e a relação via Swoogle.

**Relação direta (RD):** Avalia se o conceito extraído se encontra na descrição. Tome-mos como exemplo a seguinte descrição: “Liberian and United States national flags with

similarities being flown on a roadside in downtown Monrovia, Liberia” e o conceito **flag**. Quando o conceito é avaliado, este retorna **True**, porque o conceito **flag** encontra-se na descrição. Devido aos conceitos extraídos pelo Clarifai estarem na forma singular, alguns métodos foram implementados de forma a ser possível a verificação no plural. Estes acréscimos são os seguintes [55]:

- Juntar um S ao fim da palavra;
- Caso o conceito termine em -S, -SS, -CH, -SH, -X, -Z e -O é acrescentado -ES;
- Caso o conceito termine em -Y é substituído por -IES;
- Conceitos que terminam em -F ou -FE são trocados por -VES;

Em caso de um conceito se encontrar na descrição é incrementado um contador (RD). Este é usado para contar o número de relações diretas encontradas. Este será posteriormente usado na fórmula final deste passo.

**Relação via Sinónimo (RSY):** Esta segunda avaliação tenta averiguar se existe algum sinónimo na descrição, utilizando para isso o pacote WordNet do Python. Este é capaz de devolver uma lista de sinónimos da palavra. Posteriormente esses sinónimos passam por uma avaliação igual à relação direta. No entanto, o sucesso desta relação será incrementado no contador de relação via sinónimos (RSY).

**Relação via Swoogle (RSW):** Este passo tem como objetivo identificar alguma relação que não tenha sido detetada pelas relações anteriores. Este usa o serviço de Phrase Similarity do Swoogle, que devolve um valor entre 0 e 1, representando a relação entre os *inputs*, neste caso a descrição da imagem e os conceitos extraídos pelo Concept Extraction. De forma a obter um valor limite que corresponda a uma boa relação foi executado o algoritmo de primeira ordem, alterando o limite do Swoogle. Os resultados podem ser observados na Tabela 3.6, que se encontra na página 54. A coluna correlação indica a correlação entre a avaliação do sistema e a avaliação manual para o mesmo conjunto. Observando a Figura 3.14 é possível verificar a evolução da correlação com a alteração do limite, dando uma especial atenção ao intervalo [0.2,0.3], no qual se encontra o ponto máximo. A Figura 3.15 apresenta esse intervalo, de forma a ser possível verificar melhor a evolução da correlação. Como é possível verificar pelas Figuras e Tabela o ponto máximo de correlação é de 0.351 no limite 0.21, sendo este o valor que foi utilizado para ser o limite de uma má relação. Caso a relação entre a descrição e os conceitos seja superior a 0.21, significa que existe relação e desta forma será incrementado o contador (RSW).

### Cálculo da Baseline

No fim das várias avaliações serão usados os contadores para a seguinte equação:

$$(RD + RSY + RSW)/nConcepts$$

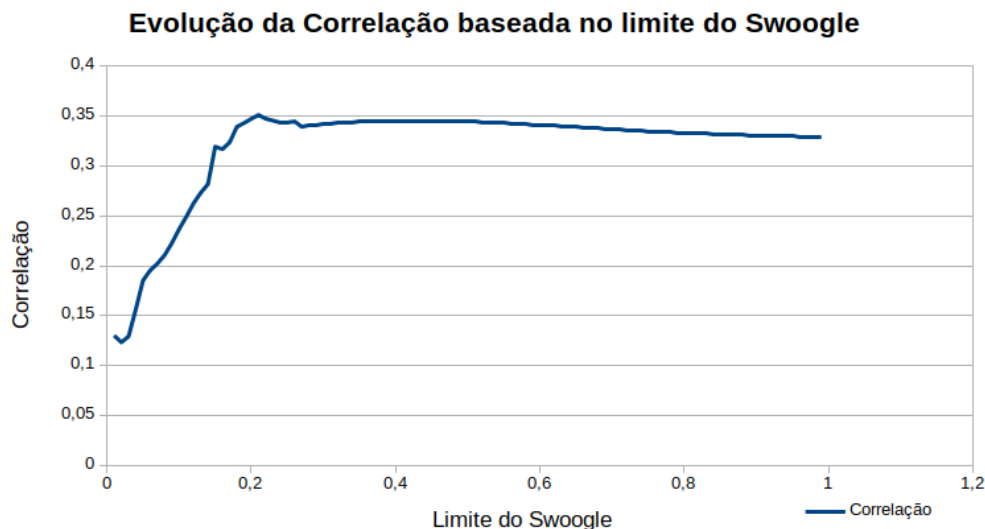


Figura 3.14: Evolução da correlação baseada no valor limite do Swoogle

O uso desta fórmula permite avaliar se os conceitos, na sua totalidade, estão descritos no texto alternativo. Segundo a fórmula quanto mais conceitos forem detetados em cada tipo de relação, maior será o valor da relação. Deste modo, obtém-se a resposta a uma das heurísticas levantada na Secção 2.2 (Conceitos existente na imagem devem ser descritos no texto alternativo). Terminado este passo, o valor da relação deverá situar-se entre 0 e 1, sendo que o valor 0 significa que não existe relação, enquanto que 1 significa que todos os conceitos extraídos estão na descrição da imagem.

### Avaliação dos conceitos dos domínios

A avaliação de segunda ordem avalia os conceitos extraídos pelo módulo Domain Extraction. O objetivo deste é beneficiar a relação final, caso exista alguma relação entre os conceitos do domínio e a descrição. É de notar que qualquer conceito já existente na lista de conceitos base será excluído desta avaliação. A fórmula de beneficiar as descrições é feita da seguinte forma:

- Relação direta: Aumento de 0.35;
- Relação via Sinónimo: Aumento de 0.15;
- Relação via Swoogle: Aumento de 0.1;

Os valores de acréscimo foram obtidos executando o serviço, alterando o valor de acréscimo entre 0 e 1 com um incremento de 0.05 nos vários parâmetros. Disponibiliza-se uma parcela dos resultados na Tabela 3.7. Não é apresentada a totalidade dos resultados devido a dimensão desta (8000 linhas). Contudo, estão representados os melhores conjuntos de acréscimo ordenados pela sua correlação entre a avaliação do sistema e a avaliação

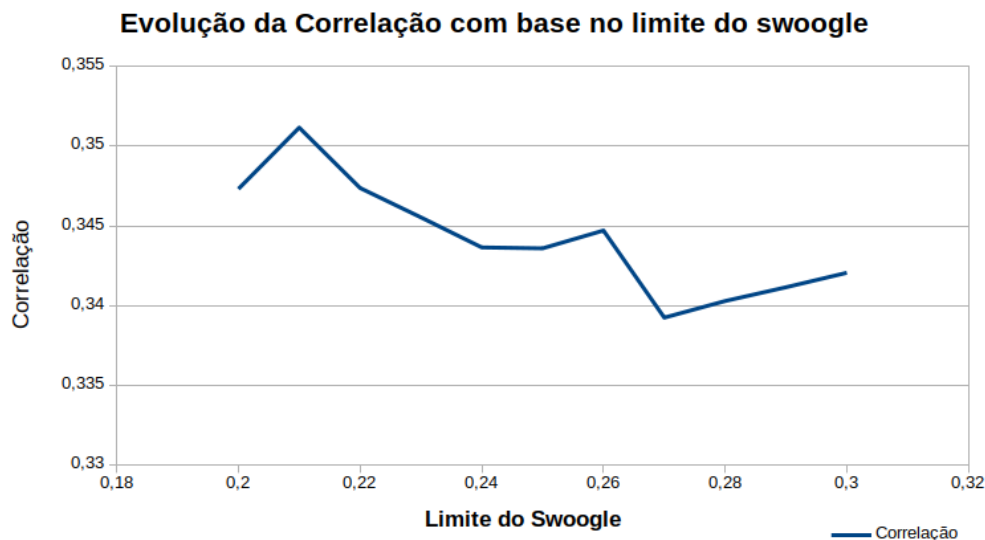


Figura 3.15: Evolução da correlação baseada no valor limite do Swoogle [0.2,0.3]

manual. Como é possível observar, a melhor correlação é do conjunto 0,35 para a relação direta, 0,15 para a relação via sinónimos e 0,1 para a relação via Swoogle. Este conjunto teve uma correlação de 0,55. Devido a este facto, foram escolhidos os valores tabulados, pois a correlação deste conjunto encontra-se mais próxima de 1 comparativamente aos outros.

### Avaliação dos conceitos dos Metadados

A última avaliação foca-se na deteção de conceitos extraídos pelo módulo MetaInfo Extraction. Este tenta solucionar um problema encontrado ao longo do desenvolvimento do serviço. Em diversas situações em que a imagem representava um local, as ferramentas de extração não têm a capacidade de extrair tal informação. No entanto, as imagens, em alguns casos, têm informações extra guardadas nos metadados que normalmente são introduzidos pelo utilizador ou por máquinas fotográficas. Usando esta informação, é possível determinar a localização onde a imagem foi captada, obtendo assim informação extra para a avaliação. Nesta são utilizadas as seguintes verificações:

- **Relação direta:** Detetar se o conceito está na descrição;
- **Relação de entidades:** Detetar se existe alguma relação com as entidades extraídas da descrição.

Na **Relação de entidades**, são extraídas, usando o SpaCy, as entidades existentes no texto. De seguida, é usado o Sematch de forma a obter uma relação entre entidades. Caso esta relação seja superior a 0,5, é incrementada a avaliação final. Os valores de acréscimo destas duas relações são 0,35 para relação direta e 0,15 para relação via Sematch. Estes

valores foram obtidos a par do estudo referido no ponto anterior, agora representado na Tabela 3.7.

## 3.6 Sumário

Neste capítulo é apresentada uma visão geral do serviço ISIAAW. Este foi implementado como um Serviço *Web* de forma a ser possível a sua integração noutros sistemas. Estes sistemas para usarem o serviço devem realizar um pedido HTTP, com o link da imagem e a sua descrição. ISIAAW encontra-se dividido em 3 componentes: Controlador, Módulos e Serviços *Web*. O Controlador monitoriza todos os módulos. Os Módulos correspondem ao núcleo da aplicação, responsáveis pela filtragem, extração de informação e avaliação semântica. Por último, as ferramentas externas (Serviços *Web*) são os mecanismos usados pelos módulos para extração de informação e relação semântica. Cada ferramenta usada nesta aplicação é detalhada neste capítulo, assim como cada módulo usado, os parâmetros mais importantes e o valor aplicado para os parâmetros dos módulos, de forma a obter um serviço fiável. Ainda neste capítulo é apresentada a arquitetura e desenho, evidenciando as características de modularidade, com o intuito de produzir um sistema que pode ser alterado facilmente. Em termos de arquitetura são apresentados os módulos existentes: Extração, Filtragem e Avaliação. Extração tem como objetivo extrair a informação das imagens, Filtragem valida se o tamanho da descrição é ideal e por último a Avaliação, valida a relação entre imagem e descrição.

Swoogle relation	Correlação	Swoogle relation	Correlação
0,01	0,130078890417023	0,51	0,34436560376335
0,02	0,123359353865202	0,52	0,344180546537927
0,03	0,129145627031773	0,53	0,343844703666124
0,04	0,156384467385371	0,54	0,343536961634819
0,05	0,185351989124744	0,55	0,343096539513557
0,06	0,19548457657491	0,56	0,342698545094457
0,07	0,202277293012554	0,57	0,342278918649419
0,08	0,210785385186135	0,58	0,341896938241914
0,09	0,222778499447726	0,59	0,341539602102253
0,1	0,236795372741426	0,6	0,341110625665687
0,11	0,249162358036678	0,61	0,340705932780383
0,12	0,262645001236913	0,62	0,340252241886778
0,13	0,273269759844427	0,63	0,339927255467089
0,14	0,281780117690658	0,64	0,33962195778197
0,15	0,319393156256893	0,65	0,3391030154076
0,16	0,316852482380092	0,66	0,338607064651559
0,17	0,323602842392307	0,67	0,338132839674447
0,18	0,339305813368333	0,68	0,337679164538811
0,19	0,343053010756775	0,69	0,337244945315624
0,2	0,347315057662699	0,7	0,336829163015129
0,21	0,351166450979816	0,71	0,336394153251537
0,22	0,347376211517202	0,72	0,335976588659728
0,23	0,345532712401518	0,73	0,335575603189283
0,24	0,343645897240586	0,74	0,335190386166462
0,25	0,343590528718459	0,75	0,334820177914655
0,26	0,344708310661932	0,76	0,334464265787348
0,27	0,339239803412628	0,77	0,334121980568632
0,28	0,340281787597775	0,78	0,333792693201858
0,29	0,341147128578543	0,79	0,333475811811813
0,3	0,342058113403883	0,8	0,333170778989915
0,31	0,342692300985529	0,81	0,332877069315555
0,32	0,343295920876375	0,82	0,332594187089781
0,33	0,343666694337879	0,83	0,33232166426026
0,34	0,343980007987875	0,84	0,332059058518833
0,35	0,344662312221763	0,85	0,331805951555012
0,36	0,344703883581958	0,86	0,331561947450636
0,37	0,344700663023711	0,87	0,331326671202444
0,38	0,34488070266899	0,88	0,331099767360773
0,39	0,344790644040511	0,89	0,330880898773792
0,4	0,344776288056625	0,9	0,330669745427795
0,41	0,344814179144762	0,91	0,330466003375043
0,42	0,344886725398609	0,92	0,330269383741482
0,43	0,344990019906659	0,93	0,330079611807447
0,44	0,344995774051621	0,94	0,329896426155129
0,45	0,345031286342333	0,95	0,329719577877185
0,46	0,34509350918082	0,96	0,329548829841399
0,47	0,345096825257867	0,97	0,329383956006806
0,48	0,34493288404355	0,98	0,32922474078709
0,49	0,344714667075923	0,99	0,329070978457473

Tabela 3.6: Tabela de Correlação dos valores de limite do Swoogle

<b>DIRECT</b>	<b>SYN/ENTI</b>	<b>SWOOGLE</b>	<b>Correlação</b>
0,35	0,15	0,1	0,553
0,35	0,1	0,05	0,536
0,35	0,15	0	0,519
0,3	0,1	0	0,502
0,3	0,1	0,1	0,492
0,4	0,1	0	0,492
0,4	0,1	0,05	0,482
0,4	0,1	0,1	0,492
0,45	0,1	0	0,482
0,45	0,1	0,05	0,485
0,45	0,1	0,1	0,486
0,5	0,1	0	0,484
0,5	0,1	0,05	0,482
0,5	0,1	0,1	0,475
0,05	0	0	0,472
0,05	0	0,05	0,472
0,05	0	0,1	0,471
0,2	0,15	0	0,461
0,2	0,15	0,05	0,461
0,2	0,15	0,1	0,461
0,25	0,15	0	0,4575
0,25	0,15	0,05	0,455
0,25	0,15	0,1	0,454
0,3	0,15	0	0,449
0,3	0,15	0,05	0,445
0,3	0,15	0,1	0,445
0,35	0,15	0	0,437

Tabela 3.7: Resultados na alteração dos valores de acréscimo



# Capítulo 4

## Avaliação

Neste capítulo serão descritos os dados de teste e os resultados finais da avaliação do serviço. Os pontos importantes focados serão: precisão e a exatidão do sistema. Por último foram realizadas avaliações mais detalhadas nas categorias propostas pelo Clarifai.

### 4.1 Dados de teste

Para avaliar o serviço foi feito um novo levantamento de imagens e correspondentes alternativas de texto. As fontes destas imagens são diversas, provenientes de blogs, revistas *online*, notícias e outros, tendo sido selecionadas 149 imagens. Depois deste levantamento foi realizada uma avaliação pericial com três utilizadores. Estes avaliaram se o texto alternativo podia ou não substituir a imagem correspondente com base nas heurísticas levantadas na Secção 2.2.1, isto é, se a descrição estava relacionada com a imagem. Para ser capaz de obter um valor objetivo da relação foi escolhida uma escala entre 1 a 4. O valor 1 significa que a descrição não substitui a imagem, enquanto o oposto, 4, significa que substitui totalmente a imagem. As imagens usadas e as classificações obtidas podem ser observadas no seguinte link ou na figura 4.1 onde é apresentado um exemplo das imagens extraídas: [https://docs.google.com/spreadsheets/d/1qmrVAvUdp8LzxQM\\_lainb-30edghUamAsZoP2XaD9W0/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1qmrVAvUdp8LzxQM_lainb-30edghUamAsZoP2XaD9W0/edit?usp=sharing)

Na Tabela 4.1 são apresentados os valores da média, desvio padrão, o número de classificações positivas e o número de classificações negativas. A partir da análise da tabela é possível verificar que as avaliações dos peritos são variadas. A leitura da linha **Classificações positivas** e **negativas** apresenta que o Perito 1 classificou mais conjuntos negativos do que positivos, ao contrário dos Peritos 2 e 3.

Na Tabela 4.2, é apresentada a taxa de concordância entre os peritos e a taxa de concordância total. Entre as avaliações, o grupo com maior concordância é entre o Perito 1 e o Perito 2 com 64%; o grupo com menor concordância encontra-se no grupo Perito 1 e 3 com aproximadamente 45%. Contudo a concordância entre todos os participantes foi



Baby eels, or angulas, are one of Spain's most expensive foods



National network forest fire lookout posts manned from today



Helmet vanga's killer blue beak



Phil Gaimon

Figura 4.1: Exemplo do conjunto de teste

aproximadamente de 34%.

	<b>Perito 1</b>	<b>Perito 2</b>	<b>Perito 3</b>
<b>Média</b>	2.395973154	2.570469799	2.758389262
<b>Desvio Padrão</b>	1.038572252	1.13456167	0.8826480942
<b>Classificações Positivas (3-4)</b>	67	84	101
<b>Classificações Negativas (1-2)</b>	82	65	48

Tabela 4.1: Tabela resumo da primeira fase de classificação

<b>Combinações</b>	<b>Concordância</b>	<b>Discordância</b>	<b>Total</b>	<b>Taxa de Concordância</b>
Perito 1 - Perito 2	96	53	149	64.43%
Perito 2 - Perito 3	75	74	149	50.35%
Perito 1 - Perito 3	67	82	149	44.967%
<i>Todos os Peritos</i>	50	99	149	33.56%

Tabela 4.2: Tabela representativa da concordância entre avaliações

Devido ao facto da concordância ter valores muito baixos, foi feita uma segunda avaliação. Ao contrário da avaliação anterior que foi feita individualmente, esta foi realizada em conjunto. Com os 3 Peritos reunidos, foi pedido que discutissem as classificações das 99 imagens que tiveram valor diferente, chegando a um consenso para estes casos. Os resultados finais podem ser observados no link referenciado anteriormente, na coluna **Classificação Final**. Na tabela 4.3, é apresentado o resumo das classificações finais: 95 conjuntos foram avaliados positivamente e 54 negativamente.

<b>Média</b>	2.711409396
<b>Desvio Padrão</b>	1.054560892
<b>Classificações Positivas</b>	95
<b>Classificações Negativas</b>	54

Tabela 4.3: Tabela sumário da classificação final

## 4.2 Resultados

Para avaliar o ISIAAW foram usados os resultados obtidos no conjunto de imagens e descrições apresentados acima. Foi executado o serviço com as imagens e as correspondentes descrições. Os resultados desta execução podem ser observados no anexo B. Os resultados da avaliação correspondem a uma correlação de 0.698, o que representa uma correlação forte entre a avaliação dos Peritos e a avaliação do serviço.

Como muitas vezes é útil ter uma classificação binária (i.e., boa ou má descrição) as classificações foram divididas em dois grupos: "más descrições" e "boas descrições". No primeiro grupo encontram-se as avaliações do serviço inferiores a 0.5 ou as avaliações de

1 e 2 dos Peritos. No segundo grupo encontram-se as avaliações do serviço superiores ou iguais a 0.5 e as avaliações de 3 e 4 dos Peritos. Sendo a verdade dada pela classificação dos peritos construiu-se a tabela de erro apresentada na Tabela 4.4.

	<b>Verdadeiros</b>	<b>Falsos</b>	<b>Total</b>
<b>Positivos</b>	38	4	42
<b>Negativos</b>	72	35	107
<b>Total</b>	110	39	149

Tabela 4.4: Tabela de previsão/classificação

Na tabela o eixo horizontal representa a previsão do serviço (ISIAAW); o eixo vertical representa a classificação dos peritos. De seguida foram calculadas as seguintes métricas:

- Precisão (Precision): 0.905;

$$Precision = VP/(VP + FP)$$

$$38/(38 + 4) = 0.905$$

- Revogação (Recall): 0.52;

$$Recall = VP/(VP + FN)$$

$$38/(38 + 35) = 0.52$$

- Exatidão (Accuracy): 0.738;

$$Accuracy = VP + VN/(VP + VN + FP + FN)$$

$$38 + 72/(38 + 72 + 4 + 35) = 0.738$$

- F-Measure: 0.66

$$F - Measure = 2 * ((Precision * Recall)/(Precision + Recall))$$

$$2 * ((0.905 * 0.52)/(0.905 + 0.52)) = 0.66$$

- Especificidade (Specificity): 0.897

$$Specificity = VN/VN + FN$$

$$72/(72 + 4) = 0.897$$

Em termos de precisão o serviço tem uma alta probabilidade de avaliar boas descrições corretamente (90% de precisão). Por forma a obter um valor de relação correto em boas e más descrições foi registada uma exatidão de 0.738, o que é um resultado positivo. Em termos das descrições negativas, o serviço obteve uma especificidade de 0.897, o que se afigura como outro ponto importante, pois a ferramenta é capaz de identificar más descrições corretamente.

De forma a concluir se existe alguma concordância entre o sistema e a avaliação dos peritos foi utilizado o teste Cohen's Kappa. A tabela 4.5 apresenta a tabela cruzada entre as classificações do sistema e os peritos. Com a realização do teste de Kohen, usando a ferramenta SPSS, foi obtido o resultado apresentado em 4.6.

		ISIAAW		Total
		Não	Sim	
Peritos	Não	72	4	76
	Sim	35	38	73
Total		107	42	149

Tabela 4.5: Tabulação cruzada entre as avaliações dos peritos e o serviço

	Valor	Erro Padrão	Significância Aproximada
<b>Medida de concordância Kappa</b>	0.472	0.066	0.000

Tabela 4.6: Resultados das medidas simétricas

O teste de Cohen's Kappa foi usado para determinar a existência de concordância entre o sistema e os peritos sobre se 149 textos alternativos exibiam boas ou más descrições das correspondentes imagens. Foi obtida uma concordância moderada entre as avaliações,  $k = 0.472$ ,  $p < 0.0005$ .

Juntamente com os resultados apresentados anteriormente, é possível afirmar que o serviço ISIAAW demonstra potencial para avaliar corretamente relações entre imagens e os seus textos alternativos.

### 4.2.1 Análise por domínio

De forma a ter uma visão mais aprofundada do sistema foi analisada a capacidade de avaliação de cada categoria a que as imagens usadas pertenciam. A tabela 4.7 representa a média das avaliações do sistema e a correlação entre o serviço e os peritos no conjunto de teste para cada categoria. Como é possível concluir com a análise desta tabela, as categorias Casamento e Tecnologia são as que apresentam melhor correlação. Outro ponto importante é a correlação do modelo comida, no qual é possível verificar um valor fraco de correlação. Com a deteção deste valor é imperativo ter um especial cuidado quando este modelo é usado.

Categoria	Peritos		Serviço		Correlação
	Média	Desvio Padrão	Média	Desvio Padrão	
<b>Celebridades</b>	3,154	0,801	2,462	1,506	0,626
<b>Comida</b>	2,65	0,988	2,6	1,188	0,363
<b>Ciência</b>	2,875	0,641	2,5	1,069	0,509
<b>Desporto</b>	2,174	0,984	1,522	0,846	0,509
<b>Tecnologia</b>	1,917	1,084	1,583	1,165	0,827
<b>Viagem</b>	2,444	0,698	1,815	0,786	0,56
<b>Casamento</b>	1,8	1,095	2,2	1,643	0,993
<b>Outros</b>	2,439	1,097	1,878	0,927	0,624

Tabela 4.7: Tabela resumo das categorias

No anexo A são apresentados os gráficos com os valores da avaliação do serviço e dos peritos para cada imagem de cada categoria, de forma a obter uma visão geral da diferença entre o valor das avaliações.

### 4.3 Sumário

Neste capítulo foi apresentado um estudo para averiguar o desempenho do serviço implementado, após a otimização de cada módulo. Para estudar o desempenho do sistema foi pedido a 3 peritos em acessibilidade para responderem a um formulário em que avaliavam os textos alternativos de 149 imagens. As avaliações dos peritos foram então utilizadas para avaliar o desempenho do sistema. No fim do estudo foi possível perceber que o sistema tem capacidade de avaliar semanticamente a relação entre uma imagem informativa e a sua descrição.

# Capítulo 5

## Integração no QualWeb

Neste capítulo será descrito como foi realizada a integração do sistema ISIAAW no Avaliador de acessibilidade Web, QualWeb. Far-se-á uma pequena introdução ao QualWeb e ao seu funcionamento. Posteriormente a esta introdução são referenciadas as alterações realizadas na API do avaliador e na *interface* da aplicação. Será também descrita a integração de outros sistemas no QualWeb como o SISSY (Avaliação semântica de menus) [56] e o SCREW (Avaliação semântica de imagens) [9].

### 5.1 QualWeb

Como referenciado na Secção 2.1, o QualWeb é um avaliador de acessibilidade Web, capaz de avaliar a acessibilidade de páginas *Web* e criar relatórios de erros/falhas de acessibilidade. É um avaliador implementado em node.js. Devido a este facto, o QualWeb fornece um sistema totalmente modular. A arquitetura do QualWeb pode ser observada na Figura 5.1, na qual são apresentadas as cinco componentes:

- Environments (Ambiente)
- Formatters (Formatar)
- QualWeb Evaluator
- Web Server (Servidor *Web*)
- Techniques (Técnicas)

O componente QualWeb Evaluator é responsável pela avaliação da acessibilidade das páginas *Web*, através das funcionalidades fornecidas pelo componente Techniques (técnicas WCAG 2.0), e utiliza o componente Formatters para adaptar os resultados em formatos específicos, como relatórios EARL ou CSVs. Finalmente, o QualWeb Evaluator é aplicado nos diferentes ambientes (Environments) que instância os tipos de ambientes que podem usar o avaliador.

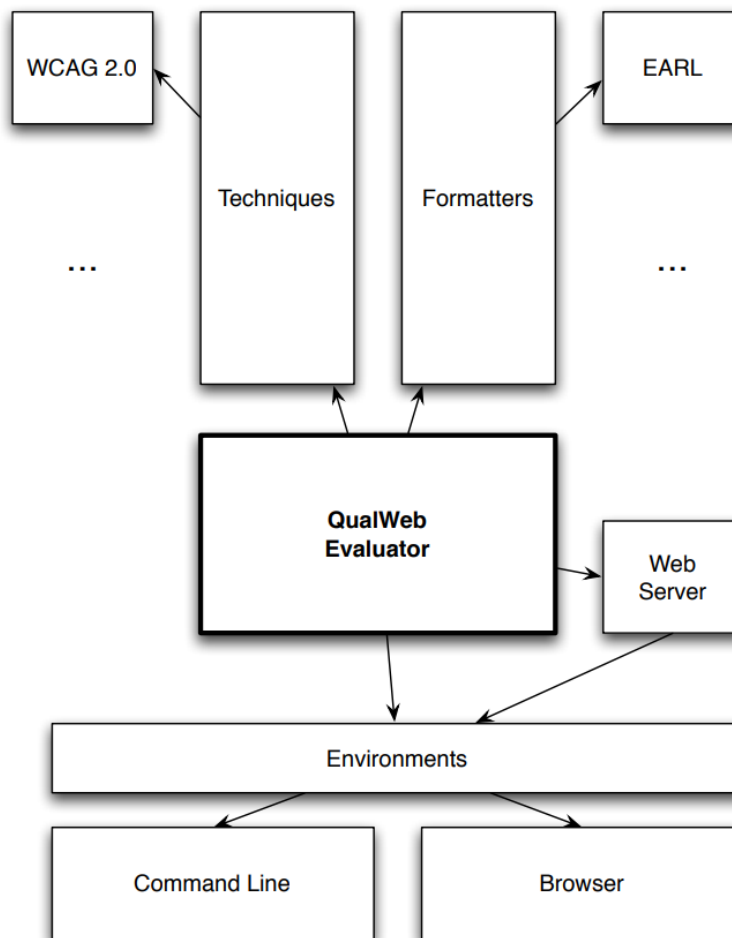


Figura 5.1: Arquitetura do QualWeb

Como indicado, o QualWeb usa as funcionalidades do componente Techniques para executar a avaliação. Este componente tem implementado um conjunto de técnicas WCAG, que podem ser observadas na Tabela 5.1. Nesta tabela também são enfatizadas as técnicas com cariz semântico, que se encontram a negrito. Como já referenciado o QualWeb foi implementado em node.js, o que fornece modularidade em vários componentes, como por exemplo a inserção de novas técnicas de acessibilidade é feita de forma simples. Para implementar novas técnicas, basta produzir a técnica e criar uma referência para esta. Para este trabalho foi implementado a técnica h37 e esta será a forma de integrar o serviço ISIAAW no avaliador, descrita na secção seguinte.

### 5.1.1 Implementação da técnica H37

Devido ao ISIAAW ser um sistema implementado em Python, a integração deste no QualWeb não pode ser direta, pois este último está implementado em node.js. Contudo, a modularidade do sistema QualWeb permite que seja possível integrar várias ferramentas

<b>Critérios de Sucesso</b>	<b>Técnicas</b>
1.1.1	<b>H2</b> H24 H30 H35 <b>H36 H37 H44 H45</b> H46 <b>H53 65</b> H67
1.2.3	<b>H53</b>
1.2.8	H46 <b>H53</b>
1.3.1	<b>H39 H44</b> H51 H63 H71 <b>H73</b> C22
1.3.2	C6 C8
1.4.1	C15
1.4.4	C12 C13 C14 C17 C20 C22 C28
1.4.5	C6 C8 C6 C12 C13 C14 C22
1.4.8	C6 C8 C12 C13 C14 C19 C20 C24 C21 C23 C25
1.4.9	C6 C8 C12 C13 C14 C22
2.1.3	H91
2.2.4	<b>H2</b>
2.4.1.	<b>H64</b> C6
2.4.2	<b>H25</b>
2.4.4	H24 <b>H30 H33</b> C7
2.4.5	H59
2.4.7	C15
2.4.9	<b>H2</b> H24 H30 H33 C7
3.1.1	<b>H57</b>
3.1.6	H62
3.2.2	H32 H84
3.2.5	H76
3.3.2	<b>H44</b>
3.3.5	<b>H89</b>
4.1.1	H93 H94
4.1.2	<b>H44 H64</b> H91

Tabela 5.1: Critérios de sucesso e técnicas implementadas no QualWeb (técnicas com componente semântica apresentadas a negrito)

neste. A integração do sistema ISIAAW foi realizada implementando a técnica H37, que posteriormente foi introduzida no módulo Techniques. De forma a ser usado o ISIAAW, o QualWeb tem que realizar um pedido HTTP para o serviço. Pelo facto do ISIAAW ter sido desenhado como um serviço *Web* independente, é possível que possa ser utilizado facilmente por qualquer avaliador, independentemente da linguagem usada. Esta integração com o QualWeb foi realizada com a alteração da técnica H37 (Utilizar atributos alt em elementos img) que já se encontra implementada no avaliador. Esta técnica tem o seguinte procedimento:

## Procedimento

1. Examine cada elemento **img** existente no conteúdo.
2. Verifique se cada elemento **img** que transmite um significado inclui um atributo **alt**.

3. Se a imagem incluir palavras importantes para compreender o conteúdo, as palavras são incluídas na alternativa em texto.

## Resultado Esperado

Passo 2 e 3 verdadeiros

Devido à inexistência de mecanismos de avaliação semântica no QualWeb (Original), esta técnica estava implementada da seguinte forma: Validava se o **alt** tinha algum texto; caso tivesse a técnica dava um *Warning*; caso não tivesse dava erro de acessibilidade. Este aviso aconselhava o utilizador que validasse a relação entre o alt e a imagem manualmente.

Antes da chamada do ISIAAW, é necessário averiguar 2 casos:

1. Verificar se o *link* da imagem tem (<http://>) ou (<https://>)
2. Verificar se a imagem não está no formato GIF

Ao longo do estudo foi verificado que muitas imagens não tinham <http://> nos links, impossibilitando o uso do Clarifai. Desta forma, foi adicionado, em caso de falta, <http://> ao link da imagem. A segunda verificação deve-se ao Clarifai não ter capacidade de extrair informação de imagens no formato GIF. Depois destas verificações é possível chamar o ISIAAW. A execução desta chamada ao serviço faz que o QualWeb receba um valor de relação entre o alt e a imagem. Caso o valor seja superior ou igual a 0.5 o avaliador considera sucesso; caso contrário será considerado um erro de acessibilidade. Esta facilidade de integração de novas funcionalidades no QualWeb permite que seja uma ferramenta importante para a análise de conteúdos web.

A técnica implementada para o ISIAAW pode ser observada no Código 5.1. Quando o valor da relação for superior ou igual a 0.5 significa que existe relação entre a imagem e a descrição, e desta forma a técnica é satisfeita, caso contrário, quando o valor é inferior a 0.5 a técnica falha. Esta avaliação é possível verificar na linha 20-25 do código apresentado. No código 5.2 é apresentado a função implementada para chamar o serviço ISIAAW.

Por permitir que sejam introduzidas novas técnicas rapidamente e de forma simples, ao mesmo tempo em que foi realizada a integração do ISIAAW no QualWeb foram introduzidos outros dois serviços. O primeiro é o SCREW [9], que tem o mesmo objetivo que o ISIAAW: avaliar a relação semântica entre imagem e descrição. O segundo, o SISSY [56], tem como objetivo estudar a relação semântica entre menus. Ambos os projetos foram integrados no QualWeb.

```

1 function inspectH37(c,u)
2 {
3   if( c.attrs != undefined
4     && c.attrs.src != undefined
5     && c.attrs.src != ""
6     && c.attrs.src.indexOf("data:") == -1
7     && c.attrs.src.indexOf(".gif") == -1
8   )
9   {
10    var domain
11    var imageurl = c.attrs.src
12    if(imageurl.indexOf("/") < 0 && imageurl.indexOf("www") != -1){
13      imageurl = "http:"+imageurl;
14    }
15    if(imageurl.indexOf("http") < 0 && imageurl.indexOf("www") < 0 ){
16      var urlsplit = u.split("/");
17      domain = urlsplit[0]+"//"+urlsplit[2];
18      imageurl = domain + imageurl;
19    }
20    if (checkDescription(u, imageurl, c.attrs.alt) >= 0.5){
21      return addElementH37(u, "img", 'passed', c, 2);
22    }
23    else{
24      return addElementH37(u, "img", 'failed', c, 1, imageurl+";");
25    }
26  }
27 }

```

Código 5.1: Técnica H37 implementada para o QualWeb

```

1 function checkDescription(url, imageurl, des)
2 {
3   var txt = imageurl;
4   var value = 0
5   txt = b64EncodeUnicode(imageurl);
6   des = b64EncodeUnicode(des);
7   var runRequest = true;
8   pedido = host+"?alt="+encodeURIComponent(description)+"&img="+encodeURIComponent(txt);
9   while(runRequest){
10    var res = srequest('GET', pedido)
11    value = res.getBody("utf8");
12  }
13  return value;
14 }

```

Código 5.2: Função checkDescription

## 5.2 Alterações nos Ambientes do QualWeb

Com a integração das novas funcionalidades do QualWeb é necessário fornecer meios para ativar ou desativar estas nos vários ambientes. Para isto foi necessário realizar alterações nos ambientes do QualWeb, Linha de Comando e Browser.

### Ambiente - Linha de Comando

Foi alterada a linha de comando pertencente ao componente Ambiente, usada pelo QualWeb, com o intuito de ser possível ao utilizador declarar as funcionalidades a executar. Os parâmetros que podem ser usados na linha de comando são apresentados na Tabela 5.2:

Parâmetros	Código	Descrição
URL	-url	URL a testar
Ficheiro	-file	Ficheiro com um ou mais URLs
Javascript	-js	Ativa a função AfterProcessing
Versão Mobile	-m	Ativa a função avaliação Mobile
CSS	-css	Ativa a avaliação de Css
Reparação	-r	Ativa a reparação do erros de acessibilidade
Gravar	-s	Grava resultado para um local
Semântica	-sm	Avaliar relações semânticas
Menus	-ms	Avaliar menus

Tabela 5.2: Tabela de parâmetros do avaliador QualWeb

Os parâmetros adicionados ao sistema encontram-se nas últimas duas linhas da tabela. Estes fornecem ao utilizador formas de especificar o projeto/projetos que deseja ativar. O primeiro parâmetro (**Semântica**), representa os trabalhos ISIAAW e SCREW. A este pode ser atribuído os valores 0, 1 e 2: 0 (Nenhuma aplicação), 1 (SCREW), 2 (ISIAAW). O segundo, **Menus**, representa o trabalho SISSY. A este pode ser atribuído os valores: 0 (Desligado) e 1 (Ligado). Estas adições permitem que o utilizador tenha a possibilidade de combinar vários projetos numa única avaliação. Contudo, é de notar que as aplicações ISIAAW e SCREW não podem ser executadas na mesma avaliação devido ao facto de ambas estarem baseadas na mesma técnica.

### Ambiente - Browser

Por último, as alterações realizadas no ambiente Browser correspondem a alterações na interface do utilizador. De forma a fornecer meios para o utilizador desativar e ativar as novas funcionalidades foram adicionadas duas *checkbox* (Figura 5.2). A primeira *checkbox*, **Semantic Relation**, apresenta as funcionalidades de avaliação semântica integradas. A segunda corresponde às funcionalidades de avaliação de menus, **Menus**. É necessário enfatizar que só pode ser ativada uma funcionalidade de cada grupo. Os serviços ISIAAW e SCREW não podem ser executados na mesma avaliação, devido ao facto de estes corresponderem ao mesmo tipo de avaliação, ou seja, relações semânticas entre imagens e texto alternativo. Estas foram as únicas alterações necessárias no *front-end* do avaliador, consequência do ambiente Browser chamar o ambiente Linha de comando quando quer executar uma avaliação.

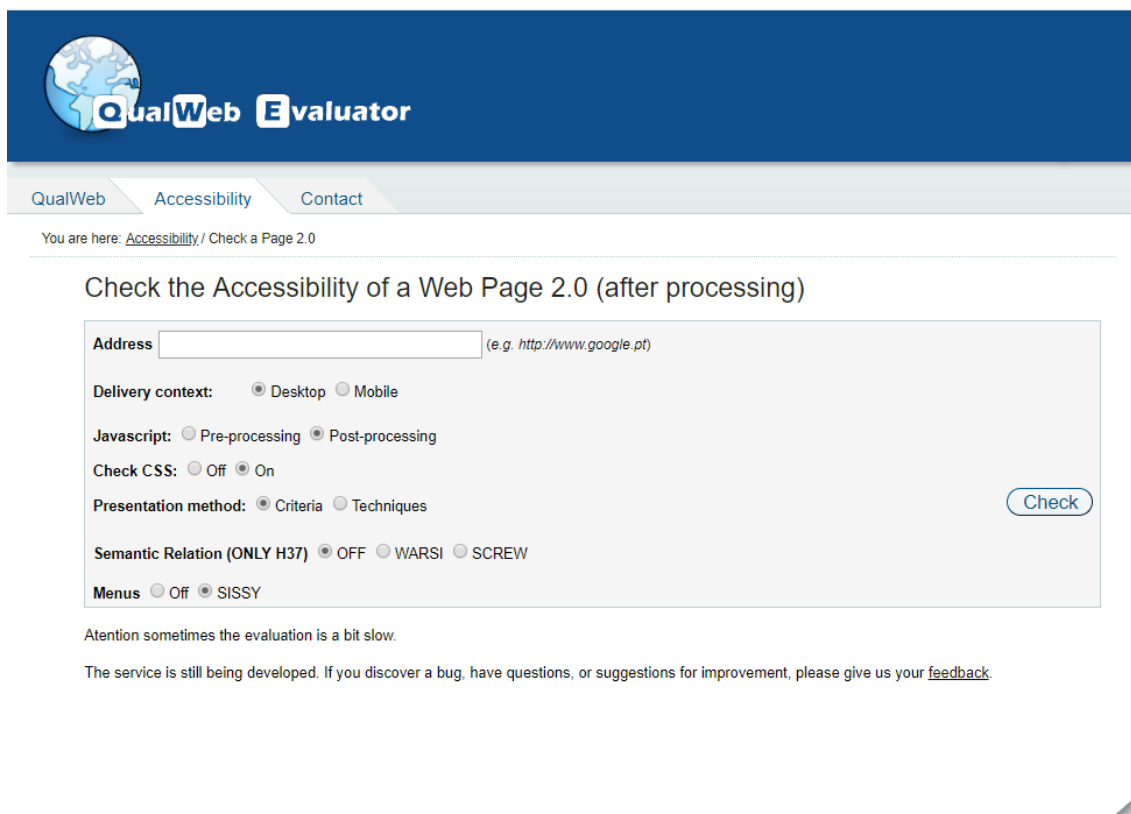


Figura 5.2: Estado atual da interface do utilizador

### 5.3 Sumário

Neste capítulo é descrito o desenho e o funcionamento do QualWeb como ferramenta de avaliação de acessibilidade Web e as alterações necessárias para a integração do serviço ISIAAW e outros serviços como o SCREW e o SISSY. Para ser possível a integração dos sistemas foi preciso alterar 2 componentes. O primeiro componente foi o módulo “Técnicas” onde foi necessário implementar uma técnica WCAG para ser capaz de chamar o sistema ISIAAW. O segundo componente alterado foi o “Ambiente”. Este foi alterado de forma a fornecer ao utilizador formas de acionar os vários sistemas usando um conjunto de *flags*, linha de comando, ou botões, no caso da aplicação no browser.



# Capítulo 6

## Conclusão

A *Web* é uma ferramenta imprescindível para a sociedade. Esta é usada por um grande conjunto de utilizadores com várias características. No entanto, existe um conjunto de utilizadores com certas limitações que dificultam o uso desta (limitações visuais, auditivas, motoras ou cognitivas). Como averiguado neste trabalho, existe um número considerável de páginas com problemas de acessibilidade. Um dos problemas mais comuns é a falta de relação entre as imagens e a propriedade do texto alternativo (alt). Esta propriedade tem como objetivo descrever o conteúdo da imagem. Num estudo da Simple Accessibility [52], vários problemas foram encontrados em imagens em vários sítios, sendo possível constatar que grande parte das imagens tem *alts* errados ou redundantes o que pode causar a certos utilizadores confusão, quando as ferramentas de assistência as leem. Assim, este trabalho visa criar um serviço capaz de avaliar a relação existente entre estes dois elementos, imagem e a sua descrição.

O ISIAAW é apresentado como um serviço *Web* capaz de avaliar a relação existente entre esses dois conteúdos. Este serviço foi implementado de forma modular, dando assim a capacidade de remoção ou integração de novos módulos. No trabalho atual foram implementados 3 módulos importantes:

- **Filtro:** Avalia se o tamanho da descrição é válido e as descrições devem ser breves;
- **Extração:** Extrai conteúdos chave de imagens;
- **Avaliação:** Avalia a relação entre descrição e imagem.

Cada módulo faz uso de serviços externos como o Clarifai, SpaCy, Sematch e o Wordnet. Estes serviços fornecem formas de extrair informação de imagens e relacionar conteúdo textual. Foram usados dois conjuntos de imagens e descrições. O primeiro conjunto foi utilizado para otimizar parâmetros em cada módulo, de forma a produzir um sistema fiável. O segundo foi utilizado para avaliar o desempenho do serviço. Ambos os conjuntos dependem da avaliação realizada por peritos de acessibilidade, e são responsáveis por avaliar cada conjunto com base na relação entre a descrição e a imagem. Os resultados

finais deste trabalho mostram que o ISIAAW é um serviço capaz de avaliar semanticamente imagens e as suas descrições. Este apresenta resultados bastante satisfatórios em alguns domínios, como por exemplo: Celebidades, Viagem, Tecnologia e Casamentos. Contudo, no domínio Comida, é encontrada alguma dificuldade em avaliar corretamente. Desta forma é necessário ter um especial cuidado quando este domínio é utilizado. No final, o serviço é capaz de avaliar qualquer imagem e descrição, tendo no conjunto de teste obtido 0.73 de exatidão. Outro ponto a favor deste sistema é a correlação entre a sua classificação e a classificação dada por peritos de acessibilidade ser de 0.698, valor que representa uma correlação forte.

Por fim o ISIAAW foi integrado no QualWeb, fornecendo assim a capacidade de avaliação semântica na acessibilidade Web. Em simultâneo, foi realizada a integração de dois outros componentes. O primeiro componente pertencia ao projeto SCREW (com o mesmo intuito que o ISIAAW - avaliação semântica entre imagem e as suas descrições). O segundo foi o projeto SISSY. Este tem o intuito de avaliar menus semanticamente. Por último foi alterada a página de utilizador do QualWeb fornecendo assim forma dos utilizadores ativarem e desativarem estes componentes.

## 6.1 Limitações

O serviço apresentado neste documento apresenta algumas limitações. A saber:

- **Limitações linguísticas:** Serviço foi direcionado para a língua inglesa, pois grande parte do conteúdo *Web* encontra-se em inglês. Alguns serviços externos usados permitem a alteração da língua. Contudo, ferramentas como o Clarifai, extração de conceitos chave, só realizam previsões em inglês;
- **Alta Dependência:** Os módulos implementados no serviço ISIAAW estão dependentes dos serviços externos para realizarem as suas tarefas;
- **Domínios:** Os parâmetros usados na aplicação podem não ser ideais para todo os contextos de imagem como pode ser verificado na avaliação das categorias no capítulo de avaliação (Domínio Comida);
- **Capacidade de resposta:** Atualmente este serviço permite no total 5 pedidos simultâneos, desta forma caso o sistema seja disponibilizado publicamente é preciso configurá-lo de forma diferente.
- **Contexto da página:** O contexto da página não é uma variável desta avaliação. É necessário um cuidado especial pois contextos diferentes requerem descrições diferentes.

- **Tempo de Avaliação:** Tempo de avaliação de cada imagem ronda os 2 segundos. Logo a existência de uma página com um elevado número de imagens faz com que seja necessário muito tempo para a sua avaliação. Este tempo é devido ao facto da aplicação Swoogle ter um desempenho lento. De forma a diminuir o tempo de resposta quando o sistema usa o swoogle, as relações avaliadas são guardadas numa base de dados de forma a agilizar futuras avaliações com o mesmo par (Descrição, Conceito).

## 6.2 Trabalho Futuro

Na sequência da secção anterior, Secção 6.1, perspectiva-se melhorar o serviço através dos seguintes pontos:

- Integrar vários modelos de língua no serviço ISIAAW. Atualmente o serviço ISIAAW carece de outros modelos linguísticos limitando assim a avaliação semântica para páginas Web em língua inglesa;
- Integrar novas formas de relação. O serviço ISIAAW apresenta apenas avaliação de relações entre texto e imagem. Desta forma é imperativo fornecer algoritmos para a relação de texto texto e de texto vídeo;
- Integrar serviços internos de relação semântica. É necessário diminuir a dependência de serviços externos para realizar a avaliação semântica, devido ao facto do Swoogle, serviço usado para relação, ter um desempenho lento;
- Integrar serviços internos de extração de informação em conteúdo textual e não textual. A implementação de um serviço interno de extração de conteúdo chave permite que seja feita a personalização da previsão obtida por este serviço;
- Desenvolver domínios de forma a promover uma melhor categorização nos contextos das imagens. Nos resultados finais do serviço foram detetados alguns domínios com correlação baixa. Desta forma é importante desenvolver melhores domínios;
- Melhorar o tempo de resposta do ISIAAW. As melhorias de tempos de resposta estão dependentes dos pontos anteriores;
- Contribuir para a melhoria dos tempos de resposta do QualWeb. Desenvolver processos de melhoria no avaliador QualWeb com o objetivo de diminuir o tempo de avaliação dos conteúdos *Web*;
- Expandir as funcionalidades do serviço ISIAAW com a criação de um módulo de reparação. A existência de um avaliador de relação semântica permite que este

detete se a descrição é válida ou não. Porém caso a descrição não seja válida estes devem recomendar formas de reparação.

# Apêndice A

## Distribuição dos resultados baseados no domínio

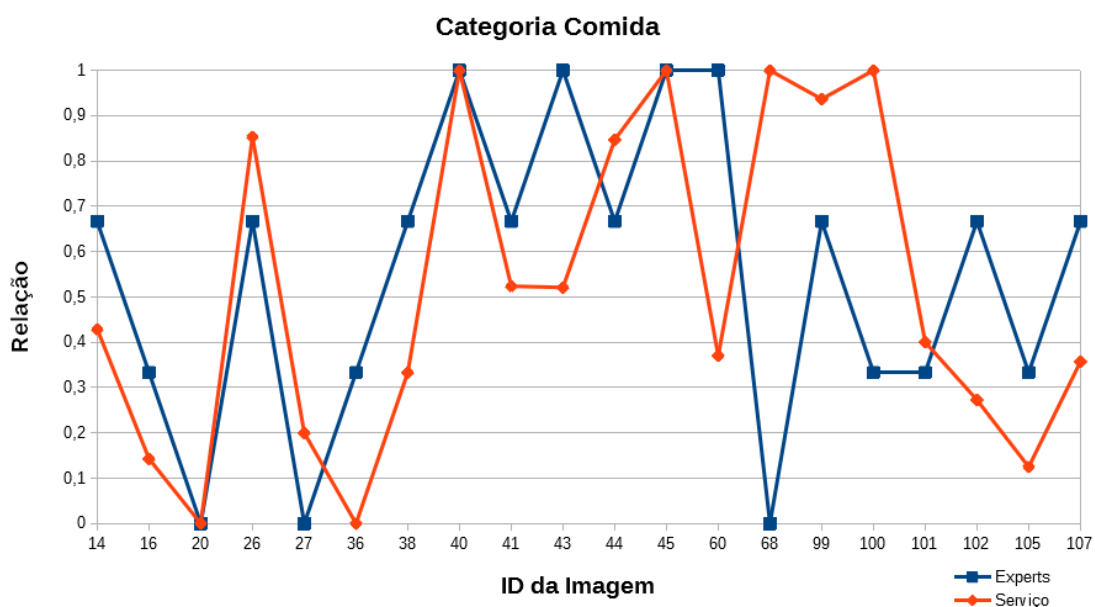


Figura A.1: Valores obtidos da avaliação da categoria comida

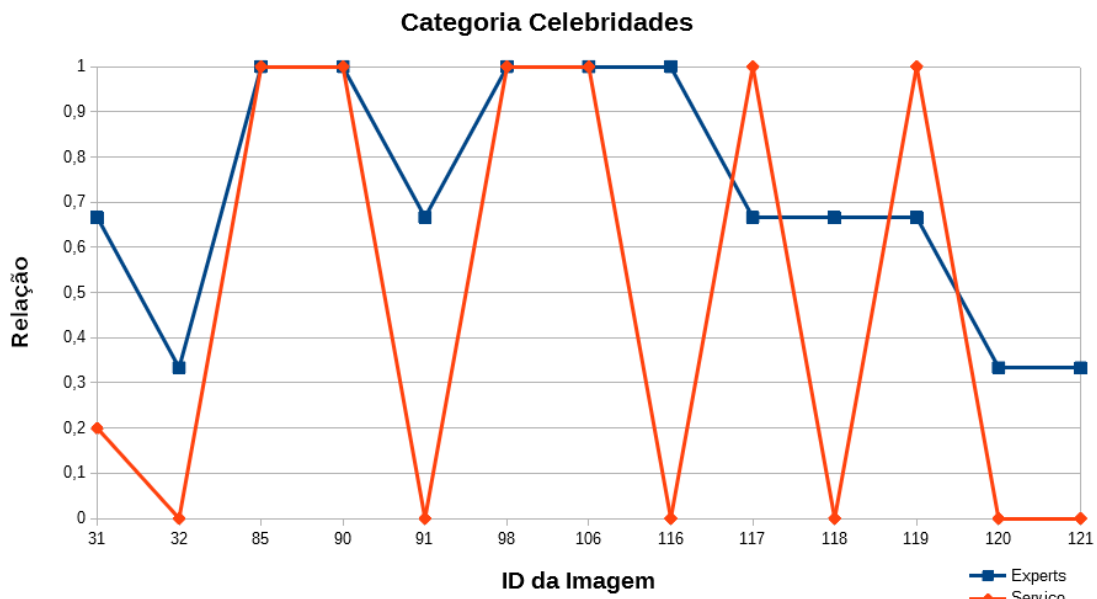


Figura A.2: Valores obtidos da avaliação da categoria celebridades

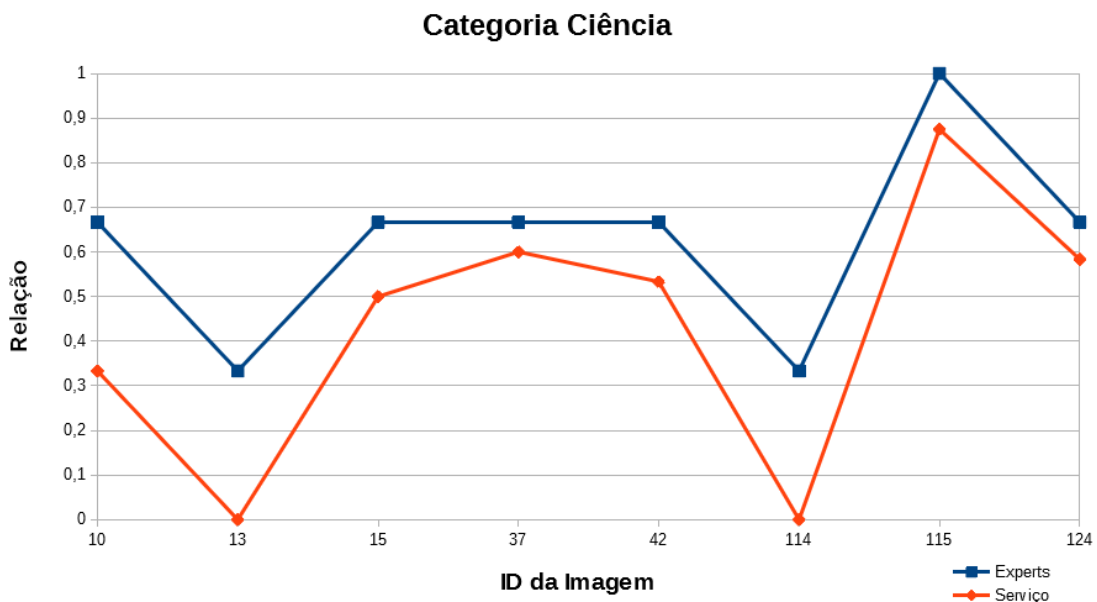


Figura A.3: Valores obtidos da avaliação da categoria ciência

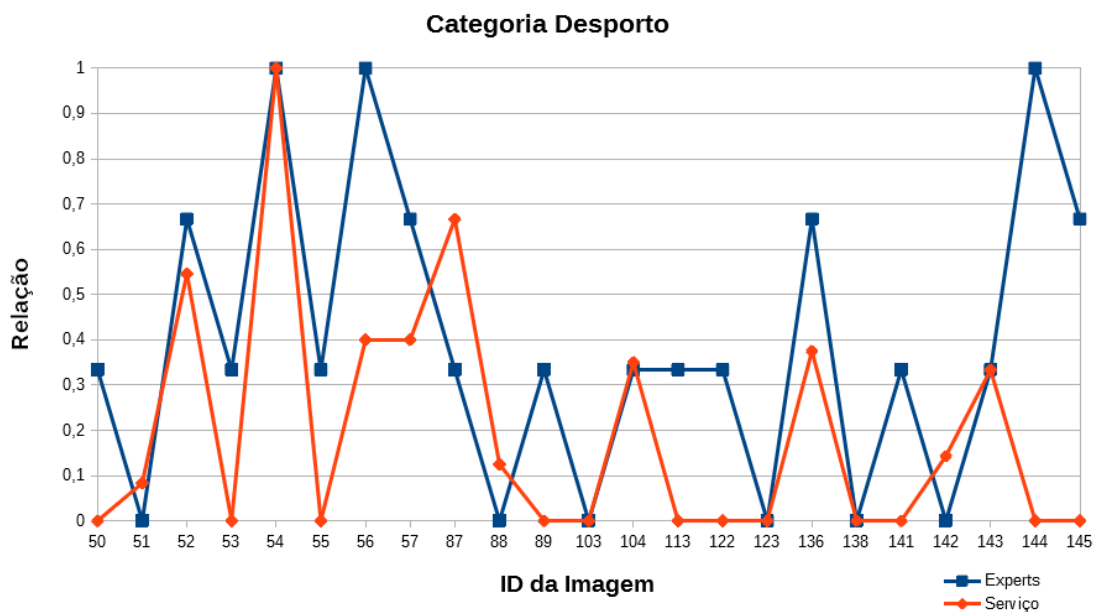


Figura A.4: Valores obtidos da avaliação da categoria desporto

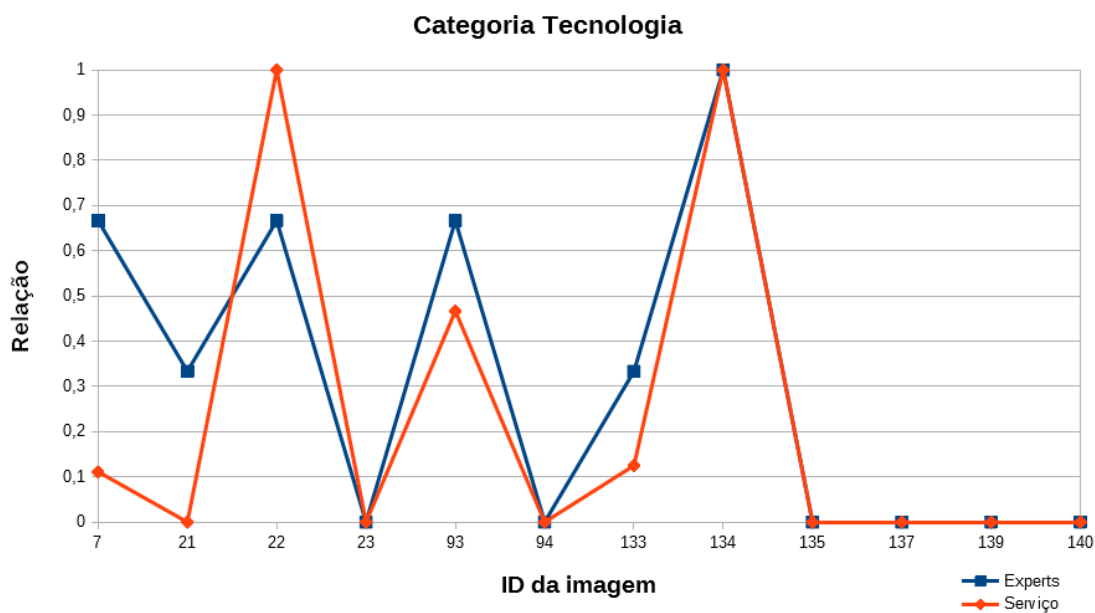


Figura A.5: Valores obtidos da avaliação da categoria tecnologia

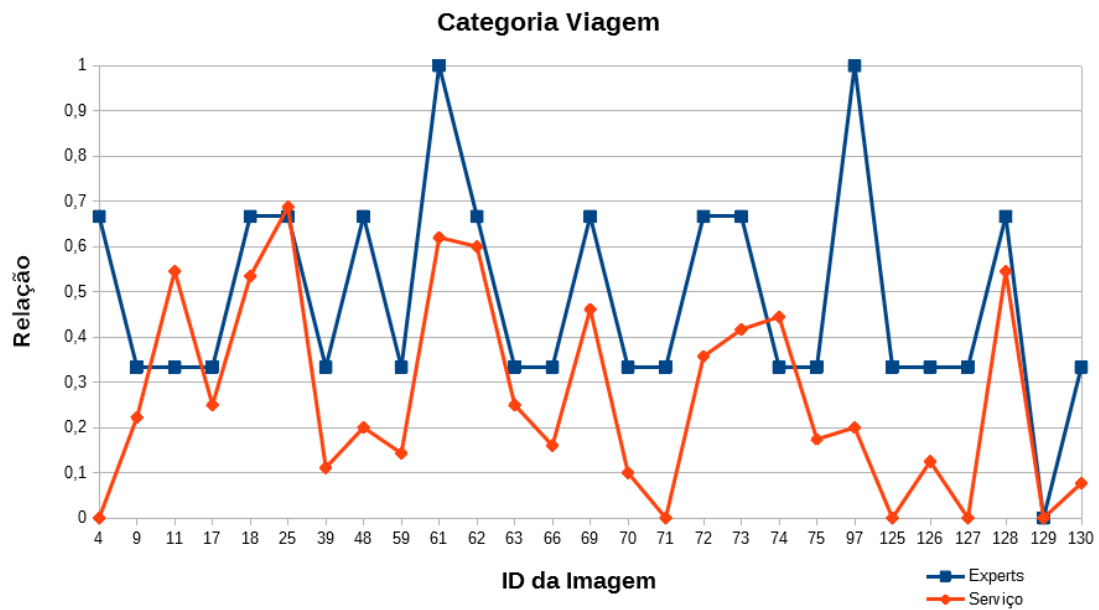


Figura A.6: Valores obtidos da avaliação da categoria viagem

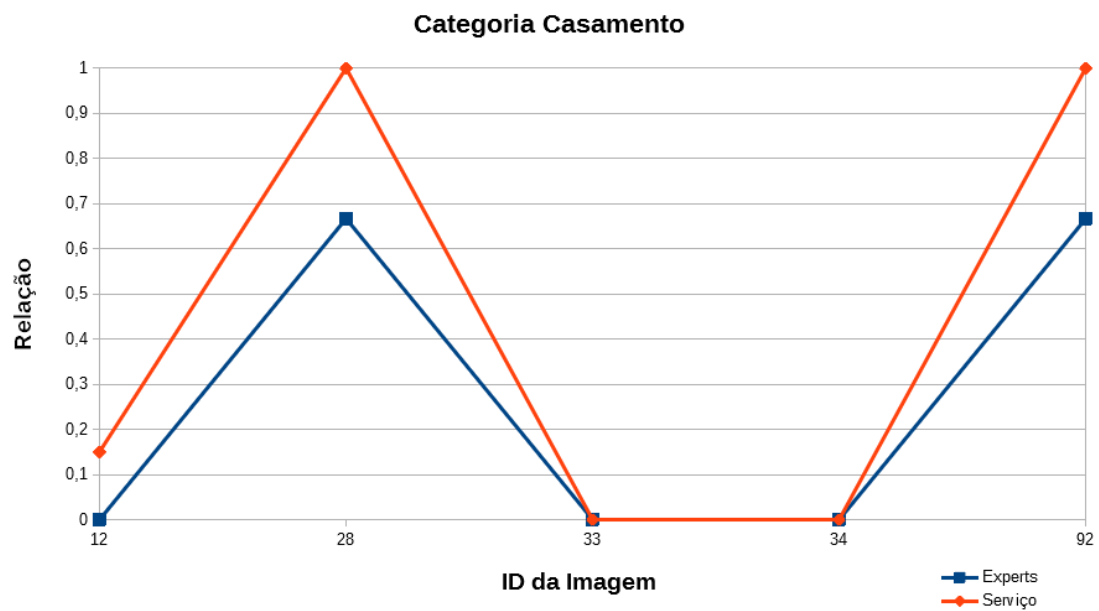


Figura A.7: Valores obtidos da avaliação da categoria casamento

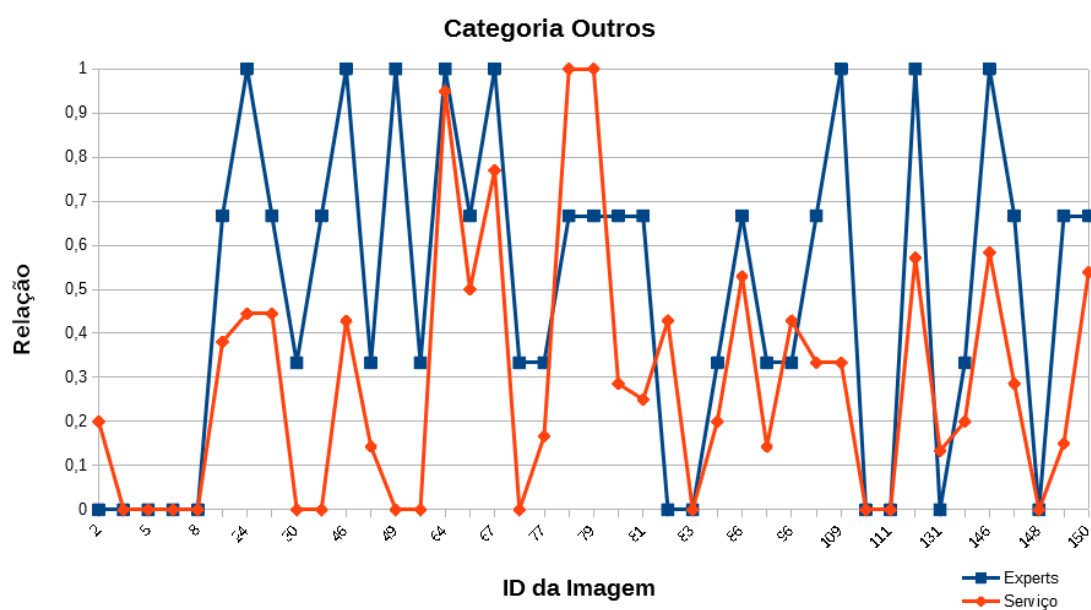


Figura A.8: Valores obtidos da avaliação da categoria outros



# Apêndice B

## Resultados do serviço ISIAAW

<b>Imagem</b>	<b>Tipo</b>	<b>Experts</b>	<b>Serviço</b>	<b>Imagem</b>	<b>Tipo</b>	<b>Experts</b>	<b>Serviço</b>
1	Travel	0	0,2	31	celebr	0,333	0
2	Travel	0	0	32	Wedding	0	0
3	Travel	0,667	0	33	Wedding	0	0
4	Travel	0	0	34	Other	0,667	0
5	Travel	0	0	35	Food	0,333	0
6	tech	0,667	0,111	36	Science	0,667	0,6
7	Travel	0	0	37	Food	0,667	0,333
8	Travel	0,333	0,222	38	Travel	0,333	0,111
9	Science	0,667	0,333	39	Food	1	1
10	Travel	0,333	0,545	40	Food	0,667	0,524
11	Wedding	0	0,15	41	Science	0,667	0,533
12	Science	0,333	0	42	Food	1	0,521
13	Food	0,667	0,429	43	Food	0,667	0,847
14	Science	0,667	0,5	44	Food	1	1
15	Food	0,333	0,143	45	Other	1	0,429
16	Travel	0,333	0,25	46	Other	0,333	0,143
17	Travel	0,667	0,535	47	Travel	0,667	0,2
18	Other	0,667	0,381	48	Other	1	0
19	Food	0	0	49	Sport	0,333	0
20	tech	0,333	0	50	Sport	0	0,083
21	tech	0,667	1	51	Sport	0,667	0,545
22	tech	0	0	52	Sport	0,333	0
23	Other	1	0,444	53	Sport	1	1
24	Travel	0,667	0,688	54	Sport	0,333	0
25	Food	0,667	0,853	55	Sport	1	0,4
26	Food	0	0,2	56	Sport	0,667	0,4
27	Wedding	0,667	1	57	Other	0,333	0
28	Other	0,667	0,444	58	Travel	0,333	0,143
29	Other	0,333	0	59	Food	1	0,37
30	celebr	0,667	0,2	60	Travel	1	0,62

Tabela B.1: Resultados do serviço ISIAAW (60 de 149)

<b>Imagem</b>	<b>Tipo</b>	<b>Experts</b>	<b>Serviço</b>	<b>Imagem</b>	<b>Tipo</b>	<b>Experts</b>	<b>Serviço</b>
61	Travel	0,667	0,6	91	Wedding	0,667	1
62	Travel	0,333	0,25	92	tech	0,667	0,467
63	Other	1	0,949	93	tech	0	0
64	Other	0,667	0,5	94	Other	0,333	0,143
65	Travel	0,333	0,16	95	Other	0,333	0,429
66	Other	1	0,77	96	Travel	1	0,2
67	Food	0	1	97	celebr	1	1
68	Travel	0,667	0,462	98	Food	0,667	0,937
69	Travel	0,333	0,1	99	Food	0,333	1
70	Travel	0,333	0	100	Food	0,333	0,4
71	Travel	0,667	0,357	101	Food	0,667	0,273
72	Travel	0,667	0,417	102	Sport	0	0
73	Travel	0,333	0,444	103	Sport	0,333	0,35
74	Travel	0,333	0,174	104	Food	0,333	0,125
75	Other	0,333	0	105	celebr	1	1
76	Other	0,333	0,167	106	Food	0,667	0,357
77	Other	0,667	1	107	Other	0,667	0,333
78	Other	0,667	1	108	Other	1	0,333
79	Other	0,667	0,286	109	Other	0	0
80	Other	0,667	0,25	110	Other	0	0
81	Other	0	0,429	111	Other	1	0,571
82	Other	0	0	112	Sport	0,333	0
83	Other	0,333	0,2	113	Science	0,333	0
84	celebr	1	1	114	Science	1	0,875
85	Other	0,667	0,529	115	celebr	1	0
86	Sport	0,333	0,667	116	celebr	0,667	1
87	Sport	0	0,125	117	celebr	0,667	0
88	Sport	0,333	0	118	celebr	0,667	1
89	celebr	1	1	119	celebr	0,333	0
90	celebr	0,667	0	120	celebr	0,333	0

Tabela B.2: Resultados do serviço ISIAAW (120 de 149)

<b>Imagem</b>	<b>Tipo</b>	<b>Experts</b>	<b>Serviço</b>
121	Sport	0,333	0
122	Sport	0	0
123	Science	0,667	0,583
124	Travel	0,333	0
125	Travel	0,333	0,125
126	Travel	0,333	0
127	Travel	0,667	0,545
128	Travel	0	0
129	Travel	0,333	0,077
130	Other	0	0,133
131	Other	0,333	0,2
132	tech	0,333	0,125
133	tech	1	1
134	tech	0	0
135	Sport	0,667	0,375
136	tech	0	0
137	Sport	0	0
138	tech	0	0
139	tech	0	0
140	Sport	0,333	0
141	Sport	0	0,143
142	Sport	0,333	0,333
143	Sport	1	0
144	Sport	0,667	0
145	Other	1	0,583
146	Other	0,667	0,286
147	Other	0	0
148	Other	0,667	0,15
149	Other	0,667	0,538

Tabela B.3: Resultados do serviço ISIAAW (149 de 149)



# **Apêndice C**

## **Exemplos Dados de Treino**



# **Apêndice D**

## **Exemplos Dados de Teste**



# Bibliografia

- [1] “Number of Internet Users (2016) - Internet Live Stats.” <http://www.internetlivestats.com/internet-users/>. [Online; accessed 18-Set-2018].
- [2] A. Officer and A. Posarac, “World report on disability,” *World Health Organisation*, 2011.
- [3] “Limitations of Web Accessibility.” <http://evengrounds.com/blog/limitations-of-web-accessibility>, 2009. [Online; accessed 28-Nov-2017].
- [4] C. Brown, “Assistive technology computers and persons with disabilities,” *Communications of the ACM*, vol. 35, no. 5, pp. 36–45, 1992.
- [5] L. G. Reid and A. Snow-Weaver, “Wcag 2.0: a web accessibility standard for the evolving web,” in *Proceedings of the 2008 international cross-disciplinary conference on Web accessibility (W4A)*, pp. 109–115, ACM, 2008.
- [6] N. Fernandes, R. Lopes, and L. Carriço, “On web accessibility evaluation environments,” in *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*, p. 4, ACM, 2011.
- [7] G. Gay and C. Q. Li, “Achecker: open, interactive, customizable, web accessibility checking,” in *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*, p. 23, ACM, 2010.
- [8] “The internet as a human right.” <https://simplyaccessible.com/article/three-pitfalls-text-alternatives/>, 2016. [Online; accessed 5-Jul-2018].
- [9] C. Duarte, I. Matos, and L. Carriço, “Semantic content analysis supporting web accessibility evaluation,” in *Proceedings of the Internet of Accessible Things*, p. 22, ACM, 2018.
- [10] W. W. W. Consortium *et al.*, “Accessible rich internet applications (wai-aria) 1.0,” *W3C*, 2014.

- [11] “Three common accessibility pitfalls for developers: text alternatives.” <https://www.brookings.edu/blog/techtank/2016/11/07/the-internet-as-a-human-right/>, 2016. [Online; accessed 28-Nov-2017].
- [12] A. Olalere and J. Lazar, “Accessibility of us federal government home pages: Section 508 compliance and site accessibility statements,” *Government Information Quarterly*, vol. 28, no. 3, pp. 303–309, 2011.
- [13] “Introduction to Web Accessibility.” <https://www.w3.org/WAI/fundamentals/accessibility-intro/>, 2017. [Online; accessed 28-Nov-2017].
- [14] “Web Accessibility Initiative (WAI).” <https://www.w3.org/WAI/>, 2009. [Online; accessed 29-Nov-2017].
- [15] S. Lawrence and C. L. Giles, “Accessibility of information on the web,” *Nature*, vol. 400, no. 6740, pp. 107–107, 1999.
- [16] C. Bailey, E. Pearson, and V. Gkatzidou, “Measuring and comparing the reliability of the structured walkthrough evaluation method with novices and experts,” in *Proceedings of the 11th Web for All Conference*, p. 11, ACM, 2014.
- [17] “Web Accessibility Evaluation Tools List.” <https://www.w3.org/WAI/ER/tools/>, 2018. [Online; accessed 28-Nov-2017].
- [18] “Javascript Usage Statistics.” <https://trends.builtwith.com/docinfo/Javascript>, 2017. [Online; accessed 28-Nov-2017].
- [19] “top-15-valuable-facebook-statistics.” <https://zephoria.com/top-15-valuable-facebook-statistics/>, 2017. [Online; accessed 28-Nov-2017].
- [20] “How-many-photos-are-being-uploaded-on-Instagram-daily.” <https://www.quora.com/How-many-photos-are-being-uploaded-on-Instagram-daily>, 2017. [Online; accessed 28-Nov-2017].
- [21] “Understanding Success Criterion 1.1.1: Non-text Content.” <https://www.w3.org/WAI/WCAG21/Understanding/non-text-content.html>, 2018. [Online; accessed 28-Apr-2018].
- [22] “Images Concepts.” <https://www.w3.org/WAI/tutorials/images/>, 2017. [Online; accessed 28-Nov-2017].

- [23] “Informative Images.” <https://www.w3.org/WAI/tutorials/images/informative/>, 2016. [Online; accessed 28-Nov-2017].
- [24] “Decorative Images.” <https://www.w3.org/WAI/tutorials/images/decorative/>, 2016. [Online; accessed 28-Nov-2017].
- [25] “Functional Images.” <https://www.w3.org/WAI/tutorials/images/functional/>, 2016. [Online; accessed 28-Nov-2017].
- [26] “Alt Text.” <https://moz.com/learn/seo/alt-text>, 2017. [Online; accessed 28-Nov-2017].
- [27] “Alternative Text.” <https://webaim.org/techniques/alttext/>, 2017. [Online; accessed 28-Nov-2017].
- [28] B. T. McInnes and T. Pedersen, “Evaluating measures of semantic similarity and relatedness to disambiguate terms in biomedical text,” *Journal of biomedical informatics*, vol. 46, no. 6, pp. 1116–1124, 2013.
- [29] M. Barbieri, *The organic codes: an introduction to semantic biology*. Cambridge University Press, 2003.
- [30] X. Wu, X. Zhu, G.-Q. Wu, and W. Ding, “Data mining with big data,” *IEEE transactions on knowledge and data engineering*, vol. 26, no. 1, pp. 97–107, 2014.
- [31] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann, “Dbpedia—a crystallization point for the web of data,” *Web Semantics: science, services and agents on the world wide web*, vol. 7, no. 3, pp. 154–165, 2009.
- [32] G. A. Miller, “Wordnet: a lexical database for english,” *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.
- [33] P. Bhattacharyya, “Indowordnet,” in *The WordNet in Indian Languages*, pp. 1–18, Springer, 2017.
- [34] B. Hamp, H. Feldweg, *et al.*, “Germanet—a lexical-semantic net for german,” in *Proceedings of ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*, pp. 9–15, 1997.
- [35] V. de Paiva, A. Rademaker, and G. de Melo, “Openwordnet-pt: An open Brazilian Wordnet for reasoning,” in *Proceedings of COLING 2012: Demonstration Papers*, (Mumbai, India), pp. 353–360, The COLING 2012 Organizing Committee, Dec. 2012. Published also as Techreport <http://hdl.handle.net/10438/10274>.

- [36] J. Hoffart, F. M. Suchanek, K. Berberich, and G. Weikum, "Yago2: A spatially and temporally enhanced knowledge base from wikipedia," *Artificial Intelligence*, vol. 194, pp. 28–61, 2013.
- [37] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor, "Freebase: a collaboratively created graph database for structuring human knowledge," in *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pp. 1247–1250, AcM, 2008.
- [38] P. Resnik, "Using information content to evaluate semantic similarity in a taxonomy," *arXiv preprint cmp-lg/9511007*, 1995.
- [39] A. Budanitsky and G. Hirst, "Evaluating wordnet-based measures of lexical semantic relatedness," *Computational Linguistics*, vol. 32, no. 1, pp. 13–47, 2006.
- [40] K. W. Church and P. Hanks, "Word association norms, mutual information, and lexicography," *Computational linguistics*, vol. 16, no. 1, pp. 22–29, 1990.
- [41] R. Gligorov, W. ten Kate, Z. Aleksovski, and F. Van Harmelen, "Using google distance to weight approximate ontology matches," in *Proceedings of the 16th international conference on World Wide Web*, pp. 767–776, ACM, 2007.
- [42] T. K. Landauer and S. T. Dumais, "A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge.," *Psychological review*, vol. 104, no. 2, p. 211, 1997.
- [43] E. Gabrilovich and S. Markovitch, "Computing semantic relatedness using wikipedia-based explicit semantic analysis.," in *IJcAI*, vol. 7, pp. 1606–1611, 2007.
- [44] P. D. Turney and P. Pantel, "From frequency to meaning: Vector space models of semantics," *Journal of artificial intelligence research*, vol. 37, pp. 141–188, 2010.
- [45] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, pp. 3111–3119, 2013.
- [46] L. Han, A. L. Kashyap, T. Finin, J. Mayfield, and J. Weese, "Umbc\_ebiquity-core: Semantic textual similarity systems.," in *\*SEM@ NAACL-HLT*, pp. 44–52, 2013.
- [47] R. Rada, H. Mili, E. Bicknell, and M. Blettner, "Development and application of a metric on semantic nets," *IEEE Transactions on systems, man, and cybernetics*, vol. 19, no. 1, pp. 17–30, 1989.

- [48] Z. Wu and M. Palmer, “Verbs semantics and lexical selection,” in *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, pp. 133–138, Association for Computational Linguistics, 1994.
- [49] C. Leacock and M. Chodorow, “Combining local context and wordnet similarity for word sense identification,” *WordNet: An electronic lexical database*, vol. 49, no. 2, pp. 265–283, 1998.
- [50] D. Lin *et al.*, “An information-theoretic definition of similarity.,” in *Icml*, vol. 98, pp. 296–304, 1998.
- [51] G. Zhu and C. A. Iglesias, “Computing semantic similarity of concepts in knowledge graphs,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 1, pp. 72–85, 2017.
- [52] J. Zhang, W. Hsu, and M. L. Lee, “Image mining: Issues, frameworks and techniques,” in *Proceedings of the Second International Conference on Multimedia Data Mining*, pp. 13–20, Springer-Verlag, 2001.
- [53] W. Fan, L. Wallace, S. Rich, and Z. Zhang, “Tapping the power of text mining,” *Communications of the ACM*, vol. 49, no. 9, pp. 76–82, 2006.
- [54] A.-H. Tan *et al.*, “Text mining: The state of the art and the challenges,” in *Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases*, vol. 8, pp. 65–70, sn, 1999.
- [55] “PLURAL DOS SUBSTANTIVOS.” <https://www.ef.edu.pt/guia-de-ingles/gramatica-inglesa/plural-dos-substantivos/>. [Online; accessed 29-Nov-2017].
- [56] C. Duarte, A. Salvado, M. E. Akpınar, Y. Yeşilada, and L. Carriço, “Automatic role detection of visual elements of web pages for automatic accessibility evaluation,” in *Proceedings of the Internet of Accessible Things, W4A '18*, (New York, NY, USA), pp. 21:1–21:4, ACM, 2018.