



**Miguel Rangel Fernandes de Oliveira Pais**

Licenciatura em Engenharia Informática

## **Grupos Implícitos em Aplicações Interactivas**

Dissertação para obtenção do Grau de Mestre em  
Engenharia Informática

Orientadora : Carmen Pires Morgado, Prof<sup>ª</sup>. Auxiliar,  
Universidade Nova de Lisboa

Co-orientador : José Alberto Cardoso e Cunha, Prof. Catedrático,  
Universidade Nova de Lisboa

Júri:

Presidente: Prof. Doutor João Carlos Gomes Moura Pires

Arguente: Prof. Doutor José António Moinhos Cordeiro

Vogal: Prof<sup>ª</sup>. Doutora Carmen Pires Morgado



FACULDADE DE  
CIÊNCIAS E TECNOLOGIA  
UNIVERSIDADE NOVA DE LISBOA

**Outubro, 2011**



## **Grupos Implícitos em Aplicações Interactivas**

Copyright © Miguel Rangel Fernandes de Oliveira Pais, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.







# Agradecimentos

Em primeiro lugar, gostaria de agradecer à Professora Doutora Carmen Pires Morgado e ao Professor Doutor José Alberto Cardoso e Cunha, pelo seu apoio, disponibilidade, boa disposição, motivação, críticas e sugestões dadas durante a orientação desta dissertação, que enriqueceram, sem dúvida, este documento.

Gostaria ainda de agradecer a todas as pessoas que durante este último ano de trabalho me acompanharam e tornaram possível a realização desta dissertação, em especial à minha família, à minha namorada e aos meus amigos.





# Resumo

---

A sobrecarga de informação, resultante do crescimento da Internet, aumenta a necessidade de mecanismos eficientes capazes de identificar e gerir informação relevante para os utilizadores e grupos de utilizadores. Dado que os utilizadores podem ser agrupados com base nos interesses, características e interações que partilham, os grupos constituem estruturas adequadas ao suporte da personalização necessária. Trabalhos anteriores relacionados com o conceito de grupos implícitos, realizados em áreas distintas como as redes de computadores e *social media*, sugerem a utilidade do conceito para o desenvolvimento de aplicações baseadas na Internet, especialmente na forma como os conteúdos são partilhados e a personalização é obtida. Ao longo desta dissertação, investigou-se a identificação dinâmica de grupos implícitos, com base na geração de extensões aos perfis dos utilizadores suportada pela aplicação de ferramentas existentes de análise de conteúdos e disponíveis externamente ao sistema desenvolvido neste trabalho. A abordagem discutida estende os perfis básicos, definidos explicitamente pelos próprios utilizadores, com dados implícitos automaticamente inferidos os quais reflectem os interesses, as características e as interações que partilham com outros utilizadores. A fim de se poder avaliar a utilidade do conceito de grupos implícitos, como componente no desenvolvimento de aplicações, é apresentada uma ferramenta que tem como objectivo principal facilitar a integração do conceito nas referidas aplicações. Adicionalmente, é especificado um cenário de aplicação, no contexto de um Campus universitário, no âmbito do qual se integrou a ferramenta desenvolvida na plataforma Moodle. No cenário apresentado, o módulo de extensão da plataforma Moodle implementado utiliza a ferramenta desenvolvida para identificar e gerir grupos implícitos de utilizadores que partilham os mesmos interesses, características e interações. Os grupos identificados são posteriormente utilizados no suporte ao desenvolvimento de novas funcionalidades da plataforma Moodle.

**Palavras-chave:** Grupos Implícitos, Aplicações Interactivas, Internet, Personalização.

---



# Abstract

---

The information overload resulting from the growth of the Internet increases the need for efficient mechanisms to identify and manage relevant information for users and groups of users within web-based interactive applications. As users can be grouped according to shared interests, characteristics and interactions, groups form convenient structures to support the required personalization.

Previous work on the subject of implicit groups, carried out in areas such as computer networks and social media, suggests its usefulness in the development of web-based applications, namely in the way content is shared and user personalization is achieved. In this dissertation we investigate the dynamic identification of such type of groups based on the generation of extended user profiles and the application of content-analysis features, which can be provided by external tools to the system that was developed in this work. The discussed approach extends common explicit user profiles with implicit extracted data reflecting each user's interests, characteristics and the interactions established among users.

In order to draw conclusions about the use of implicit groups, we present a tool that aims to ease the integration of the concept in web applications. Finally an application scenario is discussed, in the context of a college campus, and the integration of the developed tool into the Moodle platform is described.

In the presented scenario, a Moodle extension module relies on the developed tool in order to identify and manage implicit groups of users sharing common interests, characteristics and interactions. The identified groups are then used to support new Moodle features, which teachers and students can use to share research and news articles with each other.

**Keywords:** Implicit Groups, Interactive Applications, Internet, Personalization.

---



# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Motivação . . . . .	2
1.2	Descrição e Contexto do Trabalho . . . . .	4
1.3	Solução Apresentada . . . . .	5
1.4	Principais Contribuições . . . . .	6
1.5	Organização da Dissertação . . . . .	7
<b>2</b>	<b>Trabalho relacionado</b>	<b>9</b>
2.1	Grupos . . . . .	9
2.1.1	Grupos Sociais . . . . .	10
2.1.2	Grupos e Perfis de Utilizadores . . . . .	13
2.1.3	Identificação de Grupos . . . . .	17
2.1.4	Grupos Implícitos . . . . .	18
2.1.5	Grupos Implícitos e Redes de Computadores . . . . .	22
2.1.6	Análise de Conteúdo . . . . .	23
2.2	Grupos e Aplicações Interactivas . . . . .	26
2.2.1	Software Colaborativo . . . . .	26
2.2.2	<i>Sites</i> de Redes Sociais . . . . .	28
2.3	Resumo dos Aspectos Essenciais . . . . .	30
<b>3</b>	<b>Conceptualização da Ferramenta</b>	<b>33</b>
3.1	Enquadramento . . . . .	33
3.2	Objectivos . . . . .	35
3.3	Conceitos . . . . .	35
3.3.1	Grupos Implícitos . . . . .	36
3.3.2	Perfis Estendidos . . . . .	36
3.3.3	Análise de Conteúdos . . . . .	36
3.3.4	Processo de Identificação . . . . .	37
3.3.5	Gestão de grupos . . . . .	39

3.4	Resumo dos Aspectos Essenciais . . . . .	40
<b>4</b>	<b>TIG : Uma Ferramenta de Grupos Implícitos</b>	<b>41</b>
4.1	Arquitectura . . . . .	41
4.1.1	Padrão Arquitectural MVC . . . . .	41
4.2	Implementação de um Protótipo . . . . .	48
4.2.1	Perfis Estendidos . . . . .	48
4.2.2	Identificação de Grupos baseados em Interesses . . . . .	48
4.2.3	Identificação de Grupos baseados em Características . . . . .	50
4.2.4	Identificação de Grupos baseados em Interacções . . . . .	51
4.2.5	Gestão Dinâmica de Grupos . . . . .	53
4.3	Interfaces Disponibilizadas . . . . .	54
4.3.1	Core API . . . . .	54
4.3.2	Interface do Utilizador . . . . .	56
4.3.3	API REST . . . . .	59
4.4	Conclusão . . . . .	68
<b>5</b>	<b>Utilização e Integração da Ferramenta</b>	<b>71</b>
5.1	Cenários de Aplicação Gerais . . . . .	71
5.2	Um Cenário Aplicacional . . . . .	73
5.3	Integração TIG-Moodle . . . . .	74
5.3.1	Registo de Utilizadores . . . . .	74
5.3.2	Pedidos e Respostas . . . . .	76
5.4	Aplicação Desenvolvida . . . . .	80
5.4.1	Módulo de Extensão . . . . .	81
5.4.2	Bibliotecas Moodle . . . . .	82
5.4.3	Funcionalidades Implementadas . . . . .	82
5.5	Utilização de Métricas . . . . .	88
5.5.1	Grupos baseados em características . . . . .	88
5.5.2	Grupos baseados em interacções . . . . .	88
5.5.3	Grupos baseados em interesses . . . . .	89
5.6	Conclusão . . . . .	93
<b>6</b>	<b>Avaliação de Resultados da Utilização</b>	<b>95</b>
6.1	Testes de Utilizadores e Inquéritos de Satisfação . . . . .	95
6.1.1	Testes de Utilizadores e Metodologia . . . . .	96
6.1.2	Inquérito de Satisfação . . . . .	99
6.2	Conclusões . . . . .	102
<b>7</b>	<b>Conclusões e Trabalho Futuro</b>	<b>105</b>

# Lista de Figuras

1.1	O acesso à Internet nos últimos cinco anos [retirado de ITU World Telecommunication/ICT Indicators database] . . . . .	2
2.1	Mapa mundial dos sites de redes sociais em Dezembro de 2010, [fonte : GoogleTrends] . . . . .	12
2.2	Vista geral do processo de identificação de grupos, [retirado de [Pal10]] . .	21
2.3	Vista do processo de extracção de meta-dados [retirado de [Reu11]] . . . .	25
3.1	Vista generalizada da geração de perfis estendidos. . . . .	37
4.1	Ilustração do padrão arquitectural de software MVC . . . . .	42
4.2	Pedido básico MVC, <a href="http://book.cakephp.org/view/10/Understanding-Model-View-Controller">http://book.cakephp.org/view/10/Understanding-Model-View-Controller</a> . . . . .	43
4.3	Visão das diferentes camadas arquitecturais da ferramenta TIG. . . . .	44
4.4	Relações entre modelos na ferramenta TIG (notação CakePHP) . . . . .	45
4.5	Processo de identificação de grupos implícitos de interesses e entidades envolvidas . . . . .	49
4.6	Grupos de características formados através dos perfis dos utilizadores A, B e C. . . . .	52
4.7	Diferentes interfaces da ferramenta TIG . . . . .	55
4.8	Vista geral da GUI: Espaço de Utilizador (a), Espaço de Grupo (b) . . . . .	57
4.9	Espaço de Administração : exemplo de configurações possíveis . . . . .	58
5.1	Ilustração da integração de TIG na plataforma Moodle. . . . .	75
5.2	Definição das credenciais de um utilizador (extensão Moodle) . . . . .	76
5.3	Diversos blocos existentes na plataforma Moodle . . . . .	81
5.4	Visualização do bloco apresentado inicialmente a cada utilizador. . . . .	83
5.5	Visualização do bloco após a introdução de credenciais de um utilizador. .	83
5.6	Vista da secção de grupos e do espaço de um grupo. . . . .	85
5.7	Vista das secções de interesses e amigos. . . . .	86

5.8	Vista da secção de <i>bookmarks</i> . . . . .	87
6.1	Utilizadores/inquiridos e respectivas faixas etárias . . . . .	96
6.2	Resultados obtidos para a questão 1 . . . . .	100
6.3	Resultados obtidos para a questão 2 . . . . .	100
6.4	Resultados obtidos para a questão 3 . . . . .	101
6.5	Resultados obtidos para a questão 4 . . . . .	101
6.6	Resultados obtidos para a questão 5 . . . . .	102
6.7	Comentários e opiniões dos utilizadores inquiridos. . . . .	102



# Lista de Tabelas

4.1	Funções da Core API implementadas no <i>UsersController</i> . . . . .	55
4.2	Funções da Core API implementadas no <i>GroupsController</i> . . . . .	55
4.3	Funções da Core API implementadas no <i>InterestsController</i> . . . . .	56
4.4	Funções da Core API implementadas no <i>MessagesController</i> . . . . .	56
4.5	Funções da Core API implementadas no <i>PostsController</i> e no <i>StatsController</i> . . . . .	56
6.1	Os 5 grupos implícitos de interesses com maior número de membros. . . . .	98



# Listagem

4.1	Pseudocódigo: Atribuição de interesses automática . . . . .	49
4.2	Pseudocódigo: Selecção de um interesse . . . . .	50
4.3	Função createAgeGroups: Criação de grupos de características . . . . .	51
4.4	Pseudocódigo: Actualização de propriedades de interacção . . . . .	53
4.5	Cabeçalho HTTP de autorização . . . . .	59
4.6	Utilizar conexões SSL . . . . .	59
4.7	Exemplo de pedido à API . . . . .	60
4.8	addUserToGroup . . . . .	60
4.9	delUserFromGroup . . . . .	61
4.10	edit . . . . .	61
4.11	add . . . . .	61
4.12	delete . . . . .	61
4.13	import_bookmark . . . . .	61
4.14	home . . . . .	62
4.15	index . . . . .	62
4.16	clearSuggestions . . . . .	62
4.17	settings . . . . .	62
4.18	view . . . . .	63
4.19	add . . . . .	63
4.20	delete . . . . .	63
4.21	select . . . . .	63
4.22	index . . . . .	64
4.23	add . . . . .	64
4.24	delete . . . . .	64
4.25	index . . . . .	65
4.26	view . . . . .	65
4.27	sendPM . . . . .	65
4.28	inbox . . . . .	65
4.29	outbox . . . . .	66

4.30	view . . . . .	66
4.31	delete . . . . .	66
4.32	edit . . . . .	66
4.33	addToGroups . . . . .	67
4.34	view . . . . .	67
4.35	edit . . . . .	67
4.36	delete . . . . .	67
4.37	updateStats . . . . .	68
5.1	Código: cliente REST na linguagem PHP . . . . .	76
5.2	Código: resposta a um pedido de <i>../groups/view.xml</i> . . . . .	78
5.3	Código: dados XML contidos na resposta à consulta de um determinado grupo implícito. . . . .	79
5.4	Código: extracto do ficheiro <i>block_implicit_groups.php</i> contendo a chamada da primitiva <i>../users/home.xml</i> . . . . .	84



# Introdução

Nos últimos cinco anos, o número de utilizadores da Internet duplicou (ver figura 1.1). Novos sistemas, aplicações e serviços são desenvolvidos a um ritmo elevado, aproveitando o potencial de um enorme conjunto de redes que interliga e abrange uma parte substancial da população mundial. A Internet adquiriu uma importância inquestionável na vida de pessoas e instituições, tornando-se num dos principais meios de comunicação, transferência de dados e consulta de informação. Alguns países, como a Espanha, Estónia e a Finlândia, declararam o acesso à Internet como um direito legal dos seus cidadãos [Uni11].

Os utilizadores, por sua vez, passam cada vez mais tempo *online*, usufruindo dos serviços oferecidos para realizar variadas tarefas e actividades, relacionadas com aspectos profissionais e pessoais. Na vertente profissional, por exemplo, a Internet oferece cada vez mais flexibilidade, nos horários de trabalho e localização dos funcionários, flexibilidade essa que é sustentada pela propagação de aplicações web, com acessos cada vez mais rápidos.

O baixo custo e a quase instantânea partilha de ideias, conhecimentos e capacidades, tornou mais fácil o trabalho em grupo, suportado pela criação de software e ambientes colaborativos. Além dos projectos colaborativos, a proliferação de blogues, comunidades de conteúdos, *sites* de redes sociais e outras comunidades virtuais, evidencia a importância das aplicações interactivas, no actual contexto de utilização da Internet.

À medida que os utilizadores aderem a comunidades virtuais e interagem com outros utilizadores, sistemas e aplicações, também divulgam informação sobre eles próprios. Quer seja através do preenchimento de perfis que descrevem os seus interesses e características, ou através das acções que realizam, revelam, explícita ou implicitamente, dados que permitem construir uma representação digital da sua identidade.

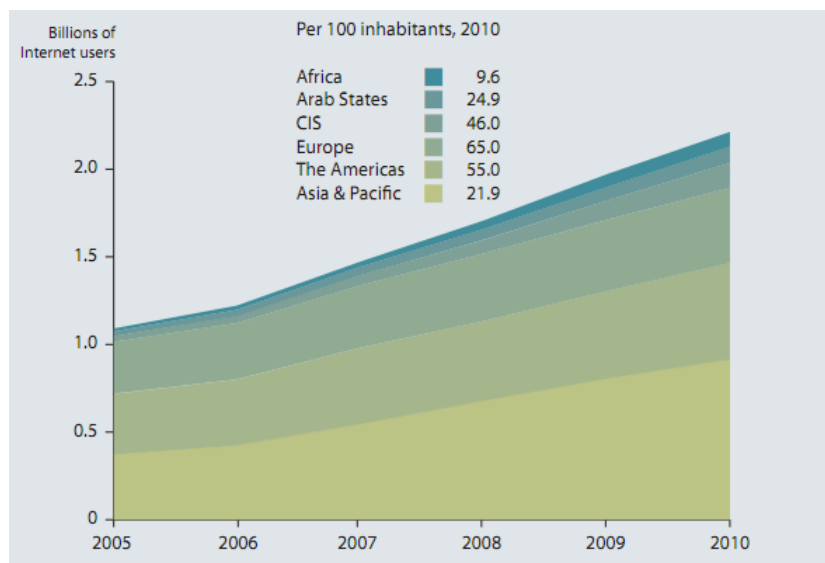


Figura 1.1: O acesso à Internet nos últimos cinco anos [retirado de ITU World Telecommunication/ICT Indicators database]

Na maioria das aplicações, a adesão dos indivíduos a grupos de utilizadores, é realizada explicitamente pelos próprios, originando grupos designados como explícitos, como acontece por exemplo, em Facebook [Fac11], GooglePlus [Goo11b], BigTent [BD11] e Meetup [Mee11]. Contudo, desde que a representação das suas identidades contenha dados suficientes, é possível identificar e formar grupos implícitos de utilizadores de forma automática, com base em características e interesses comuns, ou através do estudo das interações e relações existentes entre esses utilizadores, num determinado contexto.

## 1.1 Motivação

A actual proliferação e sucesso de comunidades virtuais, na forma de aplicações, serviços e plataformas, torna os conceitos fundamentais em que se baseiam, em interessantes temas de investigação. Como resultado do seu próprio sucesso, inúmeros utilizadores continuam a aderir a estas aplicações, por razões profissionais e/ou pessoais, de modo a procurar, processar e partilhar diversos tipos de informação.

À medida que o volume e a diversidade dos dados disponíveis aumentam, os utilizadores são confrontados com maiores dificuldades na descoberta e gestão da informação relevante para os respectivos contextos de utilização. De modo a solucionar estes problemas, as aplicações necessitam de mecanismos eficientes que permitam a personalização das suas funcionalidades centrada na identificação da informação relevante, para cada caso/cenário de utilização.

As aplicações associadas à designação *Social Media* revolucionaram a forma como as pessoas comunicam e interagem entre si, promovendo a construção e utilização de grupos ditos sociais. O conceito de grupo social está tradicionalmente associado a conjuntos

de indivíduos e suas interações, originando grafos sociais implícitos, formados com base nas relações estabelecidas entre cada utilizador e os respectivos contactos. Da análise destes conceitos, emerge a definição de grupos implícitos [CQL08, MCCC09, TMB09, Pal10], onde cada um destes grupos pode ser visto como um conjunto de indivíduos interrelacionados através de características, interesses e/ou interações comuns.

Ao contrário do que acontece com a formação de grupos explícitos, onde os utilizadores explicitamente definem as adesões aos grupos pretendidos, os grupos implícitos podem ser identificados e formados automática e dinamicamente, sem a intervenção dos utilizadores. A identificação de grupos implícitos de utilizadores requer, no entanto, informação pertinente sobre os utilizadores a serem agrupados, informação essa que é essencial para que os atributos comuns relevantes sejam identificados e traduzidos em novos grupos.

A título ilustrativo de propostas recentes, mencionam-se nos parágrafos seguintes, alguns exemplos de trabalhos relacionados (no capítulo 2 apresenta-se uma panorâmica mais extensa das diversas dimensões envolvidas).

O trabalho desenvolvido por Cutting et al. [CQL08], aborda a utilização do conceito de grupos implícitos como parte de um paradigma escalável e dissociado para a comunicação por mensagens IGM (Implicit Group Messaging), utilizado para conectar produtores de conteúdos a consumidores.

Em [RBDD<sup>+</sup>10], Roth et al. estuda o grafo social implícito formado pelas interações existentes entre utilizadores e seus grupos. Os autores apresentam ainda uma métrica baseada na análise de interações, cujo objectivo é estimar a afinidade de um utilizador aos seus contactos e grupos, concretizando a aplicação da mesma através de um algoritmo de sugestão de "amigos".

Gilbert e Karahalios propõem também, em [GK09], um modelo preditivo que mapeia dados contidos em *Social Media* em afinidades ou forças de ligação (*tie-strength*), analisando as relações e interações entre utilizadores. Os autores complementam a sua investigação, ilustrando como a modelação daquelas forças de ligação pode melhorar a concepção dos ambientes de suporte às aplicações, incluindo controlos de privacidade, encaminhamento de mensagens, introdução de novos amigos e priorização de informações.

Palsma [Pal10] explora a identificação de grupos, baseada na análise de curtas mensagens de texto e na agregação de conceitos, relacionados com os perfis dos utilizadores em *sites* de redes sociais. Teevan et al., em [TMB09, TDH10], abordam diversos métodos utilizados na identificação de grupos, incluindo o agrupamento de utilizadores baseado em índices de similaridade, juízos de valores, na identificação de utilizadores que produzem consultas semelhantes e em grupos de interesses baseados na afiliação de utilizadores a listas comuns de discussão.

Outros trabalhos relacionados, como [HKV08] e [TWL10], discutem ainda o conceito de grupos implícitos nos contextos de perfis de grupos, filtragem colaborativa e outras técnicas de personalização.

Embora os trabalhos anteriores sugiram e demonstrem distintas aplicações e benefícios consequentes da utilização de grupos implícitos, a aplicação do conceito é frequentemente descrita no âmbito de contextos, ambientes e objectivos muito específicos, e portanto, através de mecanismos moldados aos objectivos pretendidos em cada caso. De forma a expandir-se a abrangência do conceito de grupos implícitos e sua aplicação, propõem-se ao longo deste documento a concepção, e descreve-se a implementação, de uma ferramenta que permite a integração do conceito em distintas aplicações interactivas, baseadas na Internet.

A ferramenta designada por TIG combina e generaliza abordagens distintas, no que diz respeito à identificação e formação de grupos implícitos, nomeadamente baseando-se nas interações, características e interesses dos utilizadores. Adicionalmente, concretizou-se a integração da ferramenta na plataforma Moodle, tendo sido desenvolvidas novas funcionalidades que ilustram como os conceitos e mecanismos que oferece, podem ser utilizados no desenvolvimento de aplicações interactivas.

## 1.2 Descrição e Contexto do Trabalho

A existência de múltiplas aplicações e potenciais vantagens, associadas ao uso do conceito de grupos implícitos, no desenvolvimento de aplicações interactivas baseadas na Internet, torna interessante a criação de uma ferramenta de grupos implícitos, que permita a sua identificação, formação e gestão dinâmica. Esta ferramenta pretende-se modular e extensível, constituindo um componente reutilizável, que facilite a utilização e integração do referido conceito, em diversos tipos de aplicações, que envolvam comunidades de utilizadores.

No anterior contexto, os problemas associados à utilização de grupos implícitos em aplicações interactivas, podem agrupar-se em dois conjuntos distintos. O primeiro contendo as dificuldades inerentes à identificação e formação de grupos implícitos, e o segundo, referente à arquitectura e desenho da ferramenta pretendida.

O agrupamento implícito de utilizadores requer a existência de perfis, que representam cada utilizador inequivocamente, como um indivíduo único. Diferentes aplicações oferecem diferentes conjuntos de dados sobre os seus utilizadores. A quantidade dos dados e a especificidade dos mesmos pode variar, consoante o tipo de aplicação e dentro da mesma, de utilizador para utilizador, dependendo da estrutura dos perfis e dos campos obrigatórios a preencher. Além dos dados oferecidos por perfis convencionais de utilizadores, outros indicadores explícitos e implícitos [CLWB01, GK09, TMB09], relacionados com as acções que um utilizador efectua, podem fornecer pistas sobre as suas preferências, interesses e relações.

A problemática associada à identificação de grupos implícitos de utilizadores, no contexto desta dissertação, pode assim resumir-se à definição de mecanismos genéricos e flexíveis capazes de promover esta identificação com base numa quantidade limitada de dados, comum a grande parte das potenciais aplicações clientes. Por outro lado, para que



a ferramenta pretendida se torne deveras útil nos casos em que as aplicações possuam indicadores e dados adicionais, os referidos mecanismos devem permitir a sua própria extensibilidade e complementação, com vista à obtenção de soluções mais especializadas.

Numa segunda dimensão, surgem as questões associadas à concepção da ferramenta que engloba os mecanismos escolhidos. Devido à heterogeneidade das aplicações existentes na Internet, a nível de linguagens de programação e tecnologias que envolvem, a solução oferecida, deve basear-se e promover a interoperabilidade, com um conjunto aceitável de tecnologias, que respeitem os padrões aceites no desenvolvimento de aplicações web interactivas. Para tal, a sua arquitectura deve reflectir um desenho modular, assente em módulos ou componentes bem definidos, distribuídos por diferentes camadas de organização e abstracção, de forma a garantir a sua viabilidade, em termos de eficiência, manutenção e extensibilidade.

A ferramenta deve oferecer as suas funcionalidades, através de interfaces bem definidas e de fácil utilização. As técnicas e mecanismos utilizados na implementação dessas funcionalidades, devem ser adequados às complexidades temporais e espaciais previstas, assegurando o correcto funcionamento das mesmas.

As principais dimensões associadas à concepção da ferramenta de grupos implícitos, são discutidas no capítulo 3.

### 1.3 Solução Apresentada

De acordo com as problemáticas e contextos associados à identificação de grupos implícitos de utilizadores e a sua utilização em aplicações interactivas, discutidas nas anteriores secções, a solução apresentada baseou-se no desenvolvimento de uma ferramenta de grupos implícitos (TIG).

A ferramenta desenvolvida, suporta a identificação e gestão dinâmica de grupos implícitos de utilizadores. Os mecanismos e métodos utilizados pela ferramenta na identificação dos referidos grupos resultam da combinação de diferentes abordagens, anteriormente identificadas em trabalhos relacionados com o tema [CQL08, GK09, Pal10, RBDD<sup>+</sup>10]. As referidas abordagens foram generalizadas, modificadas e agregadas na ferramenta, originando novos mecanismos que permitem a geração de grupos com base num conjunto de dados, comum e disponível a diversas aplicações interactivas suportadas pela Internet.

A ferramenta TIG identifica e gere grupos implícitos de utilizadores, que partilham interesses, características e/ou interacções, através da observação de perfis estendidos gerados para cada utilizador. Os perfis estendidos de cada utilizador são construídos através da extensão dos perfis básicos e explicitamente definidos pelos próprios, nas aplicações cliente ou directamente na ferramenta.

A extensão dos perfis básicos dos utilizadores é efectuada, através da aplicação de ferramentas de análise e extracção de conteúdos [Reu11], que extraem dados implícitos

adicionais sobre os interesses dos utilizadores. Juntamente com o anterior processo baseado na análise de conteúdos, a ferramenta também recebe e processa eventos que fornecem informações sobre as interações existentes entre diferentes utilizadores, suportadas por mecanismos de comunicação internos ou externos à ferramenta.

Os dados reunidos no perfil estendido de cada utilizador, são posteriormente utilizados pelos processos de identificação de grupos implícitos implementados, para a geração de novos grupos, de acordo com um conjunto de critérios definidos. Os diferentes processos podem ainda ser configurados pelas aplicações cliente, de modo a se adaptarem a diferentes necessidades, de acordo com alguns parâmetros estabelecidos, essencialmente no que diz respeito ao nível de automatismo da identificação de grupos e da detecção de tipos de interesses. A configuração daqueles processos pode ser realizada através de um componente de configuração, o qual pode ser manipulado mediante qualquer uma das interfaces oferecidas.

A gestão dinâmica dos grupos identificados e formados é realizada com recurso a um sistema de eventos e um sistema de informação. O sistema de eventos permite à ferramenta ser notificada das entradas, saídas e mudanças de estado/perfis dos utilizadores e actualizar dinamicamente a constituição de cada grupo. O sistema de informação, por sua vez, é responsável por assegurar o armazenamento dos dados relativos aos grupos, aos utilizadores e às configurações da ferramenta.

A ferramenta baseia-se no padrão arquitectural MVC[KP88], implementado sobre a linguagem PHP, sendo que a interoperabilidade entre a ferramenta e as aplicações cliente é garantida através da utilização da API fornecida, que disponibiliza o conjunto das funcionalidades suportadas, isolando as aplicações externas dos detalhes de implementação.

## 1.4 Principais Contribuições

As principais contribuições, como resultado do trabalho desenvolvido, estão relacionadas com duas fases distintas na elaboração da dissertação. A primeira fase, associada à concepção e desenvolvimento da ferramenta TIG que disponibiliza o conceito de grupos implícitos, fornecendo uma solução genérica para a sua integração em aplicações web interactivas. A segunda, relativa à demonstração da utilidade da ferramenta e dos conceitos associados, através da sua aplicação e, mais concretamente, através da sua integração com a plataforma Moodle [Tru11].

Na primeira fase, partindo-se das diferentes conceptualizações existentes de grupos implícitos e suas dimensões associadas resultantes de contextos específicos, desenvolveu-se uma nova abordagem mais genérica que contribui para a aplicação deste tipo de grupos, num conjunto mais alargado de aplicações interactivas, baseadas na Internet. Esta nova abordagem conjuga a utilização de dados explicitamente definidos, pelos utilizadores dessas aplicações nos seus perfis, com dados implicitamente obtidos através de

técnicas de extracção e análise de conteúdos, na criação de representações digitais (perfis) mais completas, das identidades de cada utilizador e na criação de novos mecanismos. Mecanismos estes, capazes de identificar diferentes tipos de grupos implícitos de utilizadores.

Os anteriores processos constituem o núcleo da ferramenta, que permite criar e gerir os diferentes grupos implícitos de utilizadores identificados, de forma dinâmica e automática. Esta ferramenta concretiza a principal contribuição deste trabalho, permitindo que diversas aplicações integrem grupos implícitos de utilizadores na suas estruturas, aumentando a sua capacidade de personalização.

Na segunda fase, a ferramenta desenvolvida foi integrada na plataforma Moodle e utilizada na implementação de novas funcionalidades, sustentadas pelos conceitos e funcionalidades oferecidos, tendo-se verificado a sua aplicabilidade prática no contexto da plataforma Moodle. Novos caminhos de investigação e potenciais aplicações são sugeridas, assim como, são descritas diferentes formas de conjugar os dados resultantes da ferramenta com outras técnicas de personalização existentes, que podem ser úteis na maximização da utilidade dos grupos identificados.

Sumariamente, espera-se que o trabalho desenvolvido, contribua para consolidar a importância dos grupos implícitos e conceitos relacionados, como forma complementar de suporte à organização, interacção e comunicação em aplicações interactivas.

## 1.5 Organização da Dissertação

O restante documento, para além do actual capítulo introdutório, encontra-se organizado em seis capítulos complementares e uma secção com a bibliografia.

O segundo capítulo, intitulado "Trabalho Relacionado", descreve os conceitos relacionados com a ferramenta desenvolvida, através da discussão de outros trabalhos e estudos desenvolvidos em torno do conceito de grupos implícitos e temas associados. Nesse capítulo, apresenta-se primeiro uma conceptualização sociológica dos termos relevantes para o contexto da dissertação, seguida de uma discussão da bibliografia de referência em torno do conceito de grupos implícitos e sua identificação. O capítulo é fechado com a análise de diferentes ferramentas, aplicações e sistemas, que podem beneficiar da solução apresentada.

No terceiro capítulo, "Conceptualização da Ferramenta", são discutidas as opções de implementação relacionadas com os mecanismos de identificação de grupos implícitos e a arquitectura da respectiva ferramenta. No final do capítulo são ainda enumeradas as necessidades que a ferramenta criada pretende ajudar a colmatar.

O quarto capítulo, "TIG: Uma Ferramenta de Grupos Implícitos", descreve a ferramenta desenvolvida, nas suas múltiplas dimensões que incluem a sua arquitectura, o seu funcionamento e as funcionalidades oferecidas. A demonstração da viabilidade da integração da ferramenta, num cenário de aplicação concreto, é realizada no quinto capítulo, "Utilização da Ferramenta", onde se discute a implementação de um módulo de extensão

da plataforma Moodle e a sua integração com a ferramenta TIG. Ainda nesse capítulo, discute-se de que forma a ferramenta permite contemplar alguns indicadores e métricas adicionais que permitem aumentar a precisão e utilidade dos grupos identificados.

O sexto capítulo, "Avaliação de Resultados da Utilização", analisa os resultados de inquéritos, realizados a um conjunto reduzido de utilizadores que testaram a ferramenta, executados com vista à obtenção de uma percepção inicial da utilidade dos grupos implícitos identificados. Por último, no sétimo e derradeiro capítulo, "Conclusões e Trabalho Futuro", são apresentadas as conclusões resultantes da elaboração desta dissertação e revelados novos caminhos de investigação que conduzem a trabalho futuro.



# Trabalho relacionado

Este capítulo encontra-se dividido em duas secções principais. Na primeira secção, analisam-se os conceitos, metodologias e trabalhos relacionados com a construção de perfis de utilizadores, processos de identificação de grupos implícitos de utilizadores e enumeram-se algumas técnicas e ferramentas, úteis na aquisição de informação sobre um utilizador, seus interesses e características. Numa segunda secção, é observada a utilidade e importância dos grupos, no funcionamento de algumas aplicações interactivas, nomeadamente ao nível do software colaborativo e sites de redes sociais.

## 2.1 Grupos

Os grupos são estruturas muito utilizadas em múltiplas e distintas áreas, como a Matemática, Sociologia e a Computação. Como resultado da sua utilização diversificada, objectivos e contextos diferentes, as suas definições são múltiplas. O significado da palavra grupo, num dicionário comum [gr10], aparece frequentemente associado a um conjunto ou aglomerado. Uma generalização possível do termo grupo, pode assim consistir numa aglomeração de objectos ou pessoas, que possuem uma determinada ligação entre si, numa entidade única e bem definida, ou colecção.

Uma vez que a natureza e a especificidade das ligações ou interacções entre os membros de um determinado grupo, são determinantes na identificação do mesmo, o número e tamanho dos grupos que se podem formar varia consideravelmente.

Partindo de uma concepção generalizada, ao longo deste capítulo, serão analisados os processos de formação e funcionamento de grupos, adquirindo-se uma importante perspectiva sobre a necessidade e importância dos mesmos, no desenvolvimento de aplicações interactivas.

### 2.1.1 Grupos Sociais

Para as Ciências Sociais, um grupo social [Inf11] designa conjuntos de seres humanos que interagem de modo sistemático entre si, formando uma colectividade identificável e estruturada. Os indivíduos pertencentes a um determinado grupo, acordam sobre as expectativas e as obrigações decorrentes dos seus papéis e da sua filiação, definindo normas, interesses e valores com vista à prossecução de objectivos comuns. A dimensão de um grupo social pode variar, desde uma pequena associação até uma instituição de grande escala ou sociedade.

A sociedade humana, por exemplo, pode ser vista como um grande grupo e/ou uma grande diversidade de subgrupos de pessoas relacionadas entre si, através de relações persistentes, como o estatuto social, os papéis que desempenham e as redes sociais que formam. Numa nação/país, as pessoas possuem como factor comum a ocupação de um determinado território geográfico, estando sujeitas à mesma autoridade política e às expectativas resultantes da cultura dominante.

De acordo com a perspectiva sociológica [Har62, Lam08], um verdadeiro grupo deve exibir um certo grau de bem estar ou coesão social, sendo mais do que uma simples soma de indivíduos ou um conjunto de pessoas, presentes num certo local num dado instante. Na concepção e desenvolvimento de aplicações interactivas, nem sempre esta noção de coesão social é substanciada. Os elos de ligação, laços ou afinidades entre utilizadores num serviço ou aplicação, não implicam necessariamente a existência de interações, normas ou condutas a seguir. Por outro lado, os grupos formados no interior de um sistema, podem compreender várias motivações e objectivos, por vezes, relevantes apenas para o sistema em si e não para os seus utilizadores.

Embora existam algumas diferenças na conceptualização de grupo, todas as perspectivas partilham um núcleo semelhante, no que diz respeito ao processo de formação. É possível definir grupos baseados nas características de cada indivíduo. Estas características podem estar associadas à sua origem, naturalidade e a diversos factores sócio-culturais ou profissionais. As relações entre indivíduos são um exemplo óbvio de afinidades que podem ser traduzidas em grupos.

As redes sociais, sob a forma de aplicações e serviços, são o exemplo mais concreto da exploração dos grupos sociais e conceitos relacionados, para a construção de novas formas de interacção e comunicação.

#### 2.1.1.1 Redes Sociais

O termo "rede social"[Bar54] encontra-se associado, num sentido lato, ao campo de análise de redes sociais, embora seja normalmente empregue para descrever, num sentido mais estrito, os conhecidos *sites* de redes sociais ou comunidades virtuais. A análise de redes sociais observa as relações sociais à luz da teoria de redes [BE05], definindo as redes sociais como uma estrutura social, um conjunto de indivíduos e as relações entre

os próprios. Os indivíduos presentes numa rede são representados através de nós, enquanto as relações que formam entre si são identificadas através de arcos. Uma vez que uma rede pode conter um número amplamente variável de indivíduos ou utilizadores, as estruturas resultantes são grafos de estrutura e tamanho muito variáveis.

Podemos observar uma estreita relação entre uma rede social e o conceito de grupo social, definido anteriormente. Considerando-se uma rede social de escala suficientemente grande, observa-se que esta pode conter múltiplos grupos sociais, no seu interior. Para cada indivíduo, a rede formada pelas suas relações com outros indivíduos, representa um grupo social. Outros grupos sociais podem ainda ser obtidos, especificando ou generalizando-se, as relações mapeadas na rede e o universo de indivíduos a considerar.

Como acontece nos grupos sociais, as relações entre indivíduos podem ser de diferentes tipos, desde relações de amizade, financeiras ou sexuais, passando por interesses, ideais ou características comuns, ou associadas a objectivos, como a obtenção e partilha de conhecimento. Embora a complexidade possa parecer desencorajadora, diversos estudos nas várias áreas académicas evidenciam a importância deste tipo de redes, observando que o seu funcionamento pode ocorrer a diferentes níveis. Desde pequenas redes familiares a redes de dimensão global, os papéis que estas desempenham revelam-se decisivos na determinação da forma como os problemas são resolvidos, as organizações são geridas e os indivíduos se comportam [CLW05].

A diversidade de grupos e redes sociais existentes pode traduzir-se num número infinito de aplicações e serviços, pelo que a quantidade de *sites* de redes sociais e o ritmo do seu nascimento apenas confirma este potencial.

#### 2.1.1.2 *Sites* de Redes Sociais

O desenvolvimento dos sites de redes sociais, desde a sua introdução no final do século passado, levou a uma profunda reestruturação dos hábitos quotidianos de milhões de pessoas. Estabelecendo-se como uma nova forma de comunicação e interacção, vêm alterando de forma significativa a maneira como as aplicações e serviços web são pensados e desenvolvidos. Fruto da sua inegável importância, os conceitos por que se regem este tipo de aplicações e serviços, formam importantes temas de estudo em diversas áreas científicas. A figura 2.1 representa um mapa mundial dos sites de redes sociais e suas predominâncias geográficas.

O termo "rede social", quando aplicado à Internet, é utilizado para representar sites que privilegiam a formação de comunidades virtuais com interesses comuns, promovendo a interacção a diversos níveis entre os indivíduos dentro de cada grupo e para com os outros de diferentes grupos. Numa visão generalizada dos *sites* de redes sociais [BE07], Boyd e Ellison, definem-nos como serviços suportados na web, que permitem a cada utilizador construir um perfil público ou parcialmente público. Estes indivíduos podem, posteriormente, definir uma lista de outros utilizadores com quem partilham uma ligação, assim como ver e percorrer a sua lista de contactos ou quaisquer outras listas,

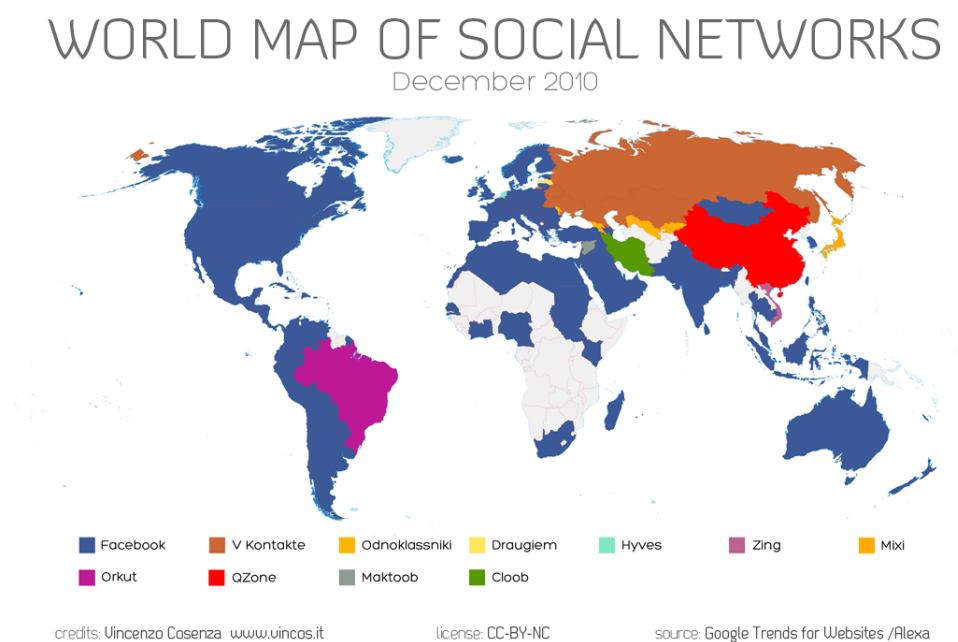


Figura 2.1: Mapa mundial dos sites de redes sociais em Dezembro de 2010, [fonte : GoogleTrends]

criadas por indivíduos registados no sistema.

Para Boyd e Ellison [BE07], o que torna estes sites únicos, não é a sua capacidade de permitir a indivíduos conhecer estranhos, mas sim a sua capacidade de permitir a cada utilizador, articular e expôr as suas redes sociais. O seu núcleo principal consiste em perfis de utilizadores visíveis, contendo listas de amigos, outros utilizadores do sistema. Na grande maioria destes sites, os utilizadores definem explicitamente a sua lista de contactos, formando grupos explícitos de contactos.

### 2.1.1.3 Grafo Social Implícito

Considerando a rede social definida pelas interações entre utilizadores e seus contactos e grupos de contactos, os autores de [RBDD<sup>+</sup>10] definem o grafo social implícito como um grafo orientado, cujos arcos estão associados a diferentes custos, determinados pela frequência, “frescura” e orientação das interações. Ao contrário do que acontece na formação de grafos sociais explícitos, em que os contactos ou indivíduos são adicionados explicitamente pelo utilizador, nesta abordagem o grafo é criado implicitamente, refletindo as interações dos vários utilizadores com as suas listas e grupos de contactos, através da computação dos referidos custos.

A frequência pretende medir a importância de determinados grupos para o utilizador, de tal modo que, os grupos com os quais este interage frequentemente são considerados mais importantes. A importância destes grupos é também considerada de forma



dinâmica, variando com o tempo, sendo quantificada através do custo “frescura” da interação. Relativamente à orientação, os autores consideram que as interações iniciadas por um determinado utilizador assumem maior importância do que aquelas em que participa mas não é responsável pelo seu início.

## 2.1.2 Grupos e Perfis de Utilizadores

Em muitas aplicações actuais, como os sites de redes sociais e outras comunidades virtuais, os utilizadores preenchem perfis, que consistem num conjunto de atributos, como a sua localização geográfica e os seus interesses. A informação disponibilizada por estes perfis, pode ser usada na criação de grupos de utilizadores, partilha de conteúdos ou sugestão de interações entre vários indivíduos [MVG10]. Nas seguintes subsecções, observa-se a importância e a utilidade dos perfis de utilizadores e das informações que armazenam.

### 2.1.2.1 Perfis de Utilizadores

Os sistemas e as aplicações que lidam com diferentes utilizadores e indivíduos, permitem aos seus utilizadores a criação de perfis individuais que os distinguem num contexto de operação.

Nos sites de redes sociais, em [BE07], verificamos que embora cada um dos sites, como o MySpace [Mys11], Facebook [Fac11] ou Cyworld [Com11], implemente funcionalidades diversas, o núcleo comum consiste em perfis de utilizador visíveis que contêm uma lista de “amigos”, outros utilizadores do sistema. Quando um utilizador se regista no *site*, é solicitado a responder a diversas questões através do preenchimento de formulários, caracterizando-se perante o sistema através de dados como a idade, a localização, os interesses, a descrição do próprio e a sua lista de contactos. A maioria das aplicações permite a adição de conteúdos multimédia a estes perfis: algumas permitem uma personalização dos mesmos, enquanto que outras, como o Facebook, permitem a adição de módulos ou aplicações que incrementam estes perfis. Nos sites de redes sociais, observa-se também que estes perfis adquirem uma importância decisiva, permitindo aos utilizadores usar as listas de contactos a que têm acesso, como hiperligações, para navegar no sistema, consultando sucessivamente páginas e perfis de outros utilizadores.

O perfil de um utilizador armazena um conjunto de dados pessoais associados a um utilizador específico. Os dados são uma representação digital da identidade do utilizador e podem ser obtidos explicitamente, implicitamente ou de forma híbrida [SA09]. Quando utilizados para armazenar a descrição das características e preferências do indivíduo, os perfis podem sustentar a criação de sistemas adaptáveis de personalização da interação entre utilizadores e computadores.

Algumas páginas web são personalizadas de acordo com as características de um indivíduo. Diferentes técnicas, metodologias e algoritmos de personalização [AT05a, AM07, STE01] podem utilizar os dados presentes num ou vários perfis semelhantes, para

adaptar o sistema às características de cada utilizador. Exemplos de tais técnicas estão incluídos nos sistemas de recomendações [AT05b] e respectivas abordagens (baseadas em conteúdo, baseadas em filtragem colaborativa e híbridas), oferecidas, por exemplo, em serviços de comércio electrónico e *sites* como a Amazon [Ama11].

No aspecto mais relevante para este trabalho, a utilização de perfis pode também constituir uma fonte valiosa de informações na criação de grupos de utilizadores, identificando indivíduos "semelhantes", nos seus interesses, características e relações. Em [TMB09] os perfis de utilizador são combinados com dados implícitos, recolhidos a partir das informações nos computadores pessoais de cada utilizador e utilizados na determinação de diferentes grupos. Os grupos formados são úteis na identificação de informações relevantes para os seus utilizadores, personalizando as suas pesquisas.

A quantidade e diversidade de dados existentes num perfil, pode variar de acordo com a sua função e peso na aplicação, sendo importante salientar que a veracidade e quantidade de dados definidos explicitamente por cada utilizador, influencia a eficácia e precisão das técnicas de personalização e formação de grupos.

### 2.1.2.2 Personalização

As técnicas de personalização recolhem dados sobre um indivíduo, com o objectivo de identificar e apresentar informação relevante ao mesmo. O uso da personalização pode estender-se a várias áreas, tais como, a saúde, a educação, a televisão e encontra-se também associado às aplicações na Internet. Os dados podem ser obtidos através da análise de perfis ou grupos de indivíduos, dos seus comportamentos e através da colaboração de múltiplos agentes e conjuntos de dados.

A personalização pode também ocorrer explicitamente, implicitamente ou de forma híbrida. No caso explícito, o sistema é modificado pelo utilizador através das funcionalidades oferecidas pelo próprio sistema. Na abordagem implícita, o sistema promove as mudanças automática e dinamicamente, através da análise das informações recolhidas e consequentes da análise do comportamento do utilizador, de forma transparente sem esforço adicional para o mesmo. Um sistema híbrido combina as vantagens das anteriores abordagens.

No artigo [TMB09] foi analisada a eficácia da utilização combinada de dados pessoais de um indivíduo e de dados relativos a outros grupos de indivíduos, com vista à melhoria da personalização de consultas na Internet. Os resultados obtidos demonstraram que a utilização de grupos melhora a personalização, para diferentes tipos de grupos, especialmente para grupos explicitamente definidos e melhorando procuras, efectuadas por um utilizador, relacionadas com o interesse que define o grupo, ou seja, o interesse à volta do qual este se formou.

Os algoritmos e diferentes técnicas usadas na recolha e análise de dados, durante a personalização, permitem a definição de uma identidade para cada indivíduo, tornando-se extremamente interessantes na identificação de grupos de utilizadores, ao mesmo

tempo que beneficiam dos mesmos para personalizar a informação oferecida aos seus membros. De acordo com esta perspectiva, as seguintes secções apresentam diferentes técnicas e metodologias usadas em sistemas de recomendação, com foco particular na importância dos perfis e grupos de utilizadores para estes mesmos sistemas e consequentemente, para as tecnologias de personalização em geral.

### 2.1.2.3 Sistemas de Recomendação

Os sistemas de recomendação têm como objectivo essencial encontrar itens adequados às preferências dos utilizadores, filtrando a informação apresentada aos mesmos. Além de itens como filmes, programas televisivos, músicas, páginas web ou literatura científica, os sistemas de recomendação são também utilizados na recomendação de elementos sociais como pessoas, eventos e grupos que apresentem uma certa probabilidade de serem interessantes para os utilizadores.

Diversas técnicas, algoritmos e métricas foram desenvolvidos, originando diferentes sistemas de recomendação [OCKR01, MKR04, KKMB05, MSC<sup>+</sup>06, CNN<sup>+</sup>10]. Adomavicius e Tuzhilin fornecem uma perspectiva global sobre os sistemas de recomendação e suas características em [AT05b]. Os autores reduzem a formulação da problemática da recomendação, ao problema de estimar as classificações atribuídas por indivíduos a itens, sobre os quais ainda não possuem conhecimento.

Jameson em [Jam04], analisa os desafios colocados pelos actuais sistemas de recomendação. Herlocker et al. por outro lado, apresentam, em [HKTR04], uma análise das estratégias que permitem avaliar diferentes sistemas de recomendação.

Apesar da imensa variedade, estes sistemas podem ser classificados, de acordo com a forma como produzem recomendações, em três categorias principais:

- **recomendações baseadas em conteúdos:** em que são recomendados ao utilizador, itens semelhantes aos preferidos pelo mesmo no passado;
- **recomendações baseadas em filtragem colaborativa:** em que os itens recomendados ao utilizador resultam das preferências e gostos anteriores de indivíduos pertencentes ao seu ambiente social;
- **recomendações baseadas em abordagens híbridas:** em que os itens recomendados resultam da combinação dos métodos anteriores.

Todas as anteriores abordagens necessitam de informação sobre os utilizadores, suas preferências, gostos, características e outros atributos. Os perfis de utilizadores e as técnicas de recolha de dados que permitem construir os mesmos, constituem assim, componentes essenciais na estimativa/previsão das classificações e preferências que os utilizadores atribuem a itens que não conhecem e invariavelmente, das recomendações que lhes são oferecidas.

#### 2.1.2.4 Estratégias de Construção de Perfis

Múltiplos conjuntos e tipos de dados sobre os utilizadores podem contribuir para a formação dos seus perfis. A construção de perfis individuais e perfis de grupos de utilizadores, no interior dos sistemas de recomendação, recorre a diversos métodos associados ao termo *Profiling*. Um exemplo de um desses métodos é a análise do comportamento do utilizador, durante a sua utilização de diferentes aplicações e sistemas na Internet.

A análise do comportamento de um utilizador pode ser realizada através de *Web Usage Mining* [PPPS03], capaz de identificar padrões de utilização associados a utilizadores e seus grupos. No artigo [HKV08], Hu et al. a propósito dos sistemas de recomendação, identificam um dos seus objectivos principais, como a melhoria da experiência do utilizador (cliente) através da personalização de recomendações. Esta personalização pode ser conseguida através da análise do seu comportamento passado, como o histórico das suas compras, hábitos de navegação, padrões de procuras e cliques, em sites e aplicações.

No referido trabalho, os autores identificam propriedades únicas em conjuntos de dados implícitos, relativos ao comportamento dos utilizadores, obtendo indicações sobre as suas preferências e animosidades, associadas a diferentes níveis de confiança. O processo origina um modelo especialmente adequado ao uso em sistemas de recomendações, capaz de se expandir linearmente com a dimensão dos dados, testado com sucesso, por exemplo, na recomendação de programas televisivos.

As vantagens do modelo e dos algoritmos propostos por Hu et al. [HKV08] e outros semelhantes, assentam na redução do esforço exigido aos utilizadores, uma vez que, aprovado o processo de recolha de dados sobre o seu comportamento, torna-se desnecessário o preenchimento de extensos questionários e campos, usados na identificação dos seus interesses. Não obstante, essa vantagem, por si só, não invalida a utilização de dados explicitamente definidos, uma vez que a sua precisão é normalmente mais elevada, passando a situação óptima pela combinação de ambos os tipos de dados, num modelo flexível e eficiente.

Os indicadores que podem ser utilizados na caracterização de um utilizador, sejam explícitos ou implícitos, podem variar consoante as aplicações e sistemas considerados, as suas funcionalidades, ambientes e objectivos. Claypool et al. descrevem alguns desses indicadores, comparando as suas utilidades em [CLWB01]. A escolha dos indicadores e o modo da sua utilização na construção de perfis são descritos em [SA09].

Além da construção dos perfis individuais de utilizadores, abordada até aqui, alguns sistemas de recomendação necessitam de mecanismos para identificar grupos de utilizadores e construir os seus perfis [BF10, BCC<sup>+</sup>09, LB10]. A geração de recomendações para grupos de utilizadores, ao invés de recomendações individuais, está normalmente associada aos sistemas de recomendações para grupos [OCKR01, MSC<sup>+</sup>06, BC11]. Neste tipo particular de sistemas de recomendação, a identificação de grupos e a definição de perfis de grupos assumem um papel fulcral.

A extracção de atributos descritivos de um grupo de utilizadores é uma técnica designada por *Group Profiling* [Hil08]. Analisando o funcionamento e os resultados obtidos por diferentes variantes desta técnica, podemos inferir sobre quais os atributos mais importantes na caracterização de grupos de utilizadores. Em [TWL10, SKB<sup>+</sup>10], são descritas algumas dessas variantes (agregação, diferenciação e a diferenciação egocêntrica) e comparam-se os resultados da sua aplicação, no contexto dos sistemas de recomendação e *sites* de redes sociais.

No estudo efectuado [TWL10], os autores observam que nos *sites* de redes sociais, os utilizadores muitas vezes escolhem os grupos a que se juntam, de acordo com procuras realizadas sobre palavras chave, que reflectem sobretudo os seus interesses. Este facto, associado ao poder descritivo das técnicas referidas, sugere que estas podem ser usadas na construção de melhores sistemas de recomendação/personalização.

### 2.1.3 Identificação de Grupos

À semelhança do que acontece nas técnicas de personalização, os dados essenciais na determinação de grupos de indivíduos, podem ser obtidos de diferentes formas. Um indivíduo pode definir explicitamente as suas características e interesses no seu perfil de utilizador. Estes dados, por si só, podem ser utilizados na inferência de outros dados relevantes sobre o indivíduo, implicitamente definidos, através de relações entre conceitos (*tags* ou conjuntos de palavras que identificam interesses, entidades e eventos) e o uso de ontologias. A análise do comportamento, interacções e fluxos de dados relacionados com um utilizador, no interior e exterior de um sistema, representam também possíveis formas de recolha de informações importantes na definição de um indivíduo e dos grupos a que pertence.

A forma como se processa a adesão de um indivíduo a um grupo, permite-nos distinguir dois tipos de grupos, explícitos e implícitos. Em [TMB09], os grupos explícitos são definidos como aqueles cujo processo de adesão dos utilizadores ao grupo é realizado explicitamente pelos próprios membros. Por oposição, nos grupos implícitos, a adesão é inferida através da actividade dos indivíduos, pertencentes ao universo de potenciais membros.

No contexto das redes sociais, [RBDD<sup>+</sup>10] oferece uma definição distinta do conceito de grupo implícito, considerando que cada um destes grupos é uma combinação única, de um ou mais contactos sociais com quem um utilizador interage. Neste caso, a identificação de grupos é sempre realizada de acordo com o grafo social implícito (2.1.1.3) definido e centrado no utilizador em questão. Cada grupo implícito contém apenas contactos sociais presentes na lista desse utilizador. Apesar de os grupos resultantes serem exclusivamente formados por contactos conhecidos, a identificação de cada grupo permanece implícita, uma vez que não existe informação específica no sentido de se agruparem os contactos. Para Palsma em [Pal10], além dos grupos formados explicitamente

pelos utilizadores dos sites de redes sociais, estes podem ter relações baseadas em interesses comuns, implicitamente detectados através das curtas mensagens de texto que produzem no Twitter [Twi11].

No modelo definido por Morgado et al. em [MCCC10], a associação de utilizadores a grupos implícitos pode variar com o tempo, sendo baseada nos utilizadores presentes no sistema, suas características e regras previamente definidas pela aplicação, no instante da criação de cada grupo. As regras são definidas como conjunções de atributos, que ao serem válidas para determinado utilizador, permitem que o mesmo seja adicionado automaticamente ao grupo. A verificação dos utilizadores para os quais as regras são satisfeitas, é levada a cabo por um mecanismo próprio cuja responsabilidade é procurar constantemente novos utilizadores aptos.

#### 2.1.4 Grupos Implícitos

Uma vez conceptualizados dois tipos distintos de grupos, de acordo com o seu processo de identificação, procede-se, nesta secção, à descrição dos diferentes algoritmos e metodologias capazes de identificar grupos implícitos de utilizadores. As principais dimensões consideradas estão separadas em subsecções, de acordo com as metodologias usadas.

##### 2.1.4.1 Relações e *Tie-Strength*

Analisar as diferentes relações entre utilizadores, especialmente no contexto das redes sociais, é uma das formas possíveis para se identificarem grupos implícitos de utilizadores. Gilbert e Karahalios [GK09], definem um modelo capaz de prever a força das relações entre utilizadores, abrindo caminho para a categorização das diferentes relações existentes e consequentes grupos implícitos.

Gilbert e Karahalios exploram a força das ligações (*tie-strength*) como conceito fundamental na distinção das relações entre indivíduos. Este conceito permite-lhes atribuir uma classificação a cada relação, estimando a “força” das mesmas, distinguindo-as dos dois habituais tipos de relações existentes nos sites de redes sociais, amigos e desconhecidos, caracterizando todo um conjunto de relações posicionadas entre estes dois extremos.

O conceito de *tie-strength* foi inicialmente proposto, em 1973, no artigo "The Strength of Weak Ties"[Gra73], realizado por Mark Granovetter. Nas últimas décadas, outras investigações e autores contribuíram para a evolução do conceito, especialmente, tentando definir as suas diferentes dimensões e seus pesos no cálculo de métricas baseadas no conceito. O modelo proposto por Gilbert e Karahalios, baseia-se em sete dimensões distintas: intimidade, intensidade, duração, distância social, serviços, apoio emocional e estrutura. Cada dimensão consiste em uma ou mais variáveis premonitórias da força da relação. Alguns exemplos ilustrativos destas variáveis, por exemplo, tal como foram utilizadas pelos autores no contexto do Facebook, são a data da última interacção existente, o número de amigos, o número de grupos comuns e o número de aplicações e hiperligações

partilhadas. O indicador de *tie-strenght* é modelado como a combinação linear das variáveis premonitórias, juntamente com termos para a dimensão das interações e a estrutura da rede social. Nos resultados apresentados, o modelo prevê, em média, o valor de *tie-strenght* dentro de um décimo do seu valor real, numa escala contínua entre 0 e 1.

O modelo proposto e seus resultados, como afirmam os autores, possuem algumas limitações. As variáveis premonitórias, utilizadas na estimativa das forças das relações, estão especificamente associadas ao ambiente de teste e recolha de dados considerado, o Facebook, não podendo ser exportadas para outras aplicações distintas. Os pesos associados a cada dimensão, no entanto, possuem essa versatilidade. Outra limitação observada é a não utilização de dados complementares, distintos dos obtidos através do *browser* e do Facebook, com potencial para influenciarem os resultados obtidos.

Apesar das limitações enumeradas, os autores também sugerem potenciais aplicações para o modelo criado, como a sua utilização conjunta com elementos de *design*, controlos de privacidade e priorização da informação, em sites de redes sociais.

O modelo, proposto no âmbito da identificação de grupos implícitos, possui o potencial para permitir a identificação de diferentes grupos implícitos, em aplicações cujos utilizadores partilham relações sociais semelhantes.

#### 2.1.4.2 Interação

Roth et al. [RBDD<sup>+</sup>10] observam que a maioria dos utilizadores de aplicações web, raramente caracteriza os seus contactos (outros utilizadores) definindo grupos, como “família” e “colegas de trabalho”. No entanto, em função das suas interações com esses contactos, acabam por agrupar os mesmos implicitamente, formando grupos implícitos de contactos.

Em [RBDD<sup>+</sup>10], os autores definem um grafo social implícito (2.1.1.3), utilizando-o depois como base para a introdução de uma métrica baseada na interação, que lhes permite estimar a afinidade de cada utilizador com os seus contactos e grupos de contactos. A métrica criada é depois analisada e utilizada na implementação de um algoritmo de sugestão de “amigos”, capaz de gerar um novo grupo de “amigos”, com base num pequeno subconjunto dos contactos pré-existentes de cada utilizador, ao qual são adicionados novos contactos, de acordo com as suas afinidades estimadas face a esse subconjunto.

As afinidades, estimadas através da métrica criada, estão relacionadas com o conceito de *tie-strenght*, analisado na secção anterior, distinguindo-se ligeiramente do mesmo, na medida em que as interações consideradas, neste caso, não se limitam a interações entre dois utilizadores distintos, sendo adicionalmente exploradas as interações com grupos de utilizadores.

O grafo social implícito utilizado, tem como base a rede de correio electrónico associada ao uso do Google Mail [Goo11a], sendo cada nó composto por um endereço de email. Os arcos são formados através do envio ou recepção de mensagens de email. No caso de

mensagens enviadas a múltiplos destinatários, apenas se considera um único arco que representa essa interação com um grupo de contactos. Para cada utilizador, são relevantes todos os arcos que possuam ligação com o nó que o representa, ou seja, o grafo centrado no próprio, representando cada arco desse grafo, um grupo implícito.

Cada grupo implícito é visto com uma combinação única de um ou mais contactos, com os quais um utilizador interagiu. Os arcos (grupos implícitos) no referido grafo, possuem uma orientação e um custo associado. A orientação é determinada observando-se o sentido da mensagem, enviada ou recebida, em relação ao utilizador considerado. O custo ou peso, é determinado pela “frescura” e frequência das interações, correspondentes a mensagens de email entre o utilizador e o grupo. Roth et al. [RBDD<sup>+</sup>10] propõem uma métrica para calcular os referidos custos, designada *Interactions Rank*, afirmando que o custo de cada arco é um indicador importante da força da relação entre o utilizador e um grupo específico, o que lhes permite distinguir quais os grupos implícitos mais importantes para cada utilizador.

A anterior solução proposta, para a identificação de grupos implícitos, centra-se exclusivamente no utilizador e nos seus dados locais, conferindo privacidade ao mesmo e aos seus contactos. A privacidade é garantida, não se utilizando o conteúdo das interações (mensagens de email) e ignorando potenciais relações entre os contactos de um utilizador, relações essas que podem ser desconhecidas do mesmo, devendo por isso, ser mantidas privadas.

A utilidade da construção de um grafo social implícito e da métrica associada, foi comprovada pelos autores [RBDD<sup>+</sup>10], indirectamente, através do sucesso e precisão do algoritmo de sugestão de amigos e suas aplicações. O referido algoritmo foi implementado em duas funcionalidades oferecidas em Gmail Labs, “Don’t forget Bob!” e “Got the wrong Bob?”. Os autores antevêm também a possibilidade dos algoritmos e abordagens, discutidas nesta secção, serem aplicados a qualquer rede social ou aplicação baseada na interação entre utilizadores. Exemplos de tais aplicações são redes implícitas de partilha de fotos e documentos, aplicações de mensagens instantâneas e blogs.

Futuras investigações são indicadas como necessárias, principalmente no que diz respeito ao impacte desta solução no funcionamento de sistemas de recomendações e na melhoria da partilha de conteúdos entre vários utilizadores, na Internet.

### 2.1.4.3 Mensagens de Texto

Uma abordagem distinta, na identificação de grupos implícitos, é a utilização das mensagens de texto que compõem os perfis dos utilizadores, na inferência de relações semânticas entre diferentes perfis, úteis na descrição das características de grupos e relações [Pal10].

Palsma [Pal10] constata que a descoberta deste tipo de relações é difícil, quando se tenta simplesmente estabelecer correspondências entre palavras semelhantes em mensagens de texto, abordagem que designa por "abordagem ingénua". De forma a atenuar



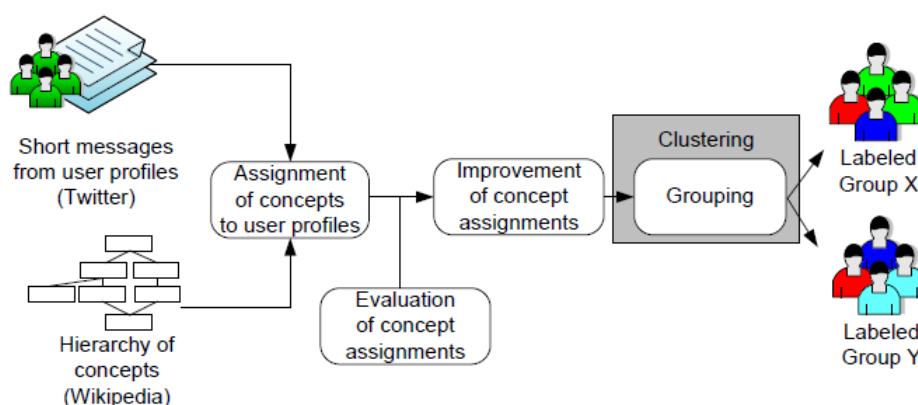


Figura 2.2: Vista geral do processo de identificação de grupos, [retirado de [Pal10]]

este problema, introduz na solução proposta, uma estrutura hierárquica de conceitos, obtida através do sistema de categorização da Wikipédia. O uso desta estrutura permite-lhe descobrir grupos e relações semânticas entre perfis de utilizadores do Twitter, a níveis superiores de abstracção. Os conceitos são posteriormente atribuídos a cada perfil, quando termos relacionados com os mesmos, ocorrem numa mensagem de texto associada ao perfil.

Na avaliação manual dos resultados obtidos através deste processo de classificação, o autor observa que cerca de 37,4% das atribuições de conceitos a perfis são correctas (caracterizam precisamente os utilizadores associados aos perfis), resultando num *F-Score*<sup>1</sup> de 0.54. Para aperfeiçoar este desempenho recorre depois, ao uso de um conjunto de métodos supervisionados de aprendizagem, que analisam dados e identificam padrões, ou *Support Vector Machines* (SVMs). Com a utilização destes novos métodos, obtém uma melhoria dos resultados do processo de classificação, que agora atinge os 69.6% de atribuições correctas e um novo *F-Score* de 0.68.

Com base nos anteriores pressupostos, o processo de identificação de grupos implícitos (consultar figura 2.2), consiste em agrupamentos de dados estatísticos sobre a ocorrência de conceitos em perfis de utilizadores (*clustering*), possuindo o próprio processo, uma natureza implícita, na medida em que os utilizadores não especificam directamente os seus interesses.

O trabalho realizado por Palsma [Pal10], recorrendo a técnicas de *text mining* e algoritmos de *clustering* hierárquicos constituem, simultaneamente, uma alternativa e um possível complemento, aos métodos observados neste capítulo, úteis na identificação de grupos implícitos de utilizadores.

Embora a estratégia utilizada permita associar perfis e seus utilizadores a múltiplos

<sup>1</sup>F-Score ou F-measure: em estatística, esta medida representa a média ponderada entre a precisão e cobertura de um dado processo analisado, num valor entre 0 e 1, sendo 1 o valor óptimo e 0 o oposto. No contexto da Recuperação da Informação (RI), esta medida é normalmente utilizada para medir o desempenho de pesquisas e classificações de documentos.

grupos baseados em conceitos, a utilidade dos grupos formados não é comprovada. A investigação efectuada levanta ainda importantes questões associadas ao uso de uma solução baseada em *text mining*, especialmente no que toca à temporalidade das mensagens, uma vez que, se os utilizadores não efectuarem uma actualização constante do seu perfil, as mensagens presentes nos mesmos podem não reflectir interesses actuais. A data em que uma mensagem foi publicada pode ser um factor importante, na avaliação da utilidade da informação que é possível extrair da mesma.

### 2.1.5 Grupos Implícitos e Redes de Computadores

Na actual secção, é descrito um possível processo de identificação e criação de grupos implícitos, no contexto das redes de computadores, especificamente das redes *peer-to-peer*, que complementa as metodologias enumeradas anteriormente e oferece uma diferente perspectiva da utilidade deste tipo de grupos, na disseminação de informação para múltiplos utilizadores.

Cutting et al. descrevem ao longo do seu trabalho [CLQ06], um sistema de mensagens que recorre ao conceito de grupos implícitos. O sistema apresentado consiste numa forma de distribuição de mensagens *many-to-many*, onde os utilizadores responsáveis pela publicação das mensagens, não necessitam de enumerar os recipientes e onde os últimos, não precisam de se juntar a grupos explícitos. O funcionamento do sistema apenas requer, por parte dos utilizadores que publicam mensagens, que estes especifiquem as características da audiência pretendida. O próprio sistema é capaz de, dinamicamente, entregar as mensagens a todos os membros que satisfaçam as condições especificadas, no instante da publicação.

Dentro do sistema P2P idealizado, todos os utilizadores (*peers*) são considerados iguais, independentemente de, num determinado momento, poderem adquirir funções distintas como produtores ou consumidores de mensagens. Cada utilizador presente na rede, mantém um conjunto de palavras chave que descrevem as suas características ou atributos, dos quais são exemplos, as suas capacidades, serviços ou interesses. Os autores não discutem a forma como estas palavras chave são realmente atribuídas a cada utilizador assumindo, no entanto, que a sua distribuição seguirá um padrão normalmente encontrado em aplicações baseadas na Internet, uma distribuição Zipfian [Zip32, New06] (distribuição de probabilidades de ocorrência de palavras e itens que se rege pela lei de Zipf, também conhecida por distribuição de Yule).

Os grupos implícitos são especificados como expressões booleanas de palavras chave. O conjunto formado por todos os utilizadores, cujos atributos satisfaçam a expressão que define esse grupo, é o conjunto dos seus membros. Um exemplo fornecido em [CLQ06], ilustra esta especificação, observando que, dada uma expressão alvo:

$$(surgeon|doctor)\&dysphonia \quad (2.1)$$

um utilizador com os atributos *doctor* e *dysphonia* satisfaz a expressão, enquanto outro,

com apenas o atributo *doctor*, não pertence aos membros deste grupo implícito.

A linguagem de expressões utilizada permite a criação de expressões alvo através da combinação dos operadores lógicos, conjunção e disjunção. As expressões obtidas são também comutativas e associativas, respeitando uma gramática formal bem definida, não podendo especificar a ausência de um atributo, isto é, limitando-se apenas a especificar quais os atributos que um recipiente deve declarar.

O conjunto dos atributos de cada utilizador é codificado no seu endereço lógico, através de estruturas de dados probabilísticas utilizadas para testar se um determinado elemento pertence a um dado conjunto, *Bloom Filters* [Blo70]. Quando um utilizador pretende disseminar uma mensagem, utiliza este mesmo método de endereçamento, para obter os endereços de todos membros do grupo alvo, usando depois um algoritmo de encaminhamento geométrico para distribuir a mensagem pretendida, através dos utilizadores vizinhos.

A utilização de grupos implícitos na disseminação de informação, numa rede de computadores, é caracterizada como uma possível solução eficiente, quando o criador de uma mensagem conhece as características do público alvo, em vez dos seus "nomes", constituindo uma abordagem potencialmente útil para ligar produtores e consumidores. Ao contrário das *mailing lists*, onde um utilizador necessita de se registar explicitamente para receber informação, estando muitas vezes apenas interessado num subconjunto de dados associados ao tema retratado, a abordagem sugerida [CLQ06] reduz o esforço do recipiente e o número de mensagens irrelevantes que recebe.

Por outro lado, ao considerar os *sites*, blogs e fóruns de mensagens, onde as mensagens são publicadas em formatos publicamente acessíveis, os autores [CLQ06] verificam que a utilização de grupos implícitos pode contribuir para a redução do esforço dos utilizadores, que necessitam de efectuar procuras e filtrar os dados publicados, ignorando as informações irrelevantes.

### 2.1.6 Análise de Conteúdo

A análise de conteúdo [Ber71] é uma metodologia, usada no estudo do conteúdo das comunicações e textos, definida principalmente no âmbito das ciências sociais e humanas, sendo no entanto utilizada, em muitas outras áreas científicas, entre as quais a ciência da computação. Através de uma perspectiva quantitativa, estes métodos permitem a construção de bases de conhecimento, analisando numericamente a frequência de ocorrência de determinados termos, construções e referências num dado texto.

Nesta secção, são mencionados dois conjuntos de ferramentas úteis na análise de conteúdos, em aplicações e serviços baseados na Internet, conteúdos esses que podem fornecer dados adicionais sobre características e interesses dos utilizadores, aumentando o lote de informação disponível para a criação de perfis e facilitando os processos de personalização e identificação de grupos implícitos, observados anteriormente. Apesar de

existirem as mais variadas ofertas ao nível deste tipo de ferramentas, a maioria compreende um subconjunto das funcionalidades, que se encontram nas duas ferramentas a seguir abordadas. Por esta razão, acrescida do facto de se tratar de software disponível gratuitamente, é justificada a escolha das ferramentas analisadas.

#### 2.1.6.1 AlchemyAPI

A ferramenta AlchemyAPI [Orc11], como refere o seu *website*, fornece aos programadores de aplicações baseadas na Internet, um conjunto de ferramentas para a análise de conteúdo e anotações através de meta-dados, nomeadamente, extracção rotulada de entidades, extracção de frases e termos, categorização de documentos, detecção de linguagem e *content scraping* estruturado, entre outras funcionalidades.

A sua utilização permite expor a riqueza semântica escondida nos conteúdos analisados. A tecnologia envolvida no funcionamento deste serviço, baseia-se na aplicação de processamento estatístico de linguagem natural e de algoritmos de *machine learning*, para a análise do conteúdo e extracção de meta-dados semânticos. Alguns exemplos dos meta-dados, possivelmente extraídos, são informações sobre pessoas, locais, companhias e tópicos. Os meta-dados podem ser representados num dos vários formatos, XML, JSON, RDF ou em microformatos rel-tag [Kc06].

O serviço descrito proporciona REST API *endpoints* capazes de realizar análise de conteúdos em páginas web acessíveis na Internet, HTML, texto e imagens digitalizadas de documentos. Além do acesso directo através de uma interface HTTP REST, a interface de programação pode também ser integrada nas principais linguagens de programação, como Java, C/C++, C#, Perl, PHP, Python e Ruby, através do respectivo sistema de apoio ao desenvolvimento (SDK).

O conjunto completo de ferramentas que constituem este serviço, está disponível gratuitamente, encontrando-se limitado a cerca de trinta mil chamadas, por dia à API, para usos não comerciais.

#### 2.1.6.2 OpenCalais

Semelhante ao anterior serviço, o ambiente OpenCalais [Reu11] é um conjunto de ferramentas, que permite a integração de funcionalidades semânticas, num blogue, sistema de gestão de conteúdos ou em qualquer outro tipo de aplicação web. O seu núcleo é composto pelo serviço *OpenCalais Web Service*, que automaticamente cria valiosos meta-dados semânticos para o conteúdo pretendido, com base em processamento de linguagem natural, *machine learning* e outros métodos.

O processo de análise de um documento permite-lhe identificar as entidades incorporadas no mesmo, devolvendo depois, além dessas entidades, factos e eventos escondidos no interior do texto sob a forma de *tags* reutilizáveis em qualquer outra aplicação, conforme se ilustra na figura 2.3.

O vasto conjunto de ferramentas oferecidas pelo OpenCalais está bem documentado

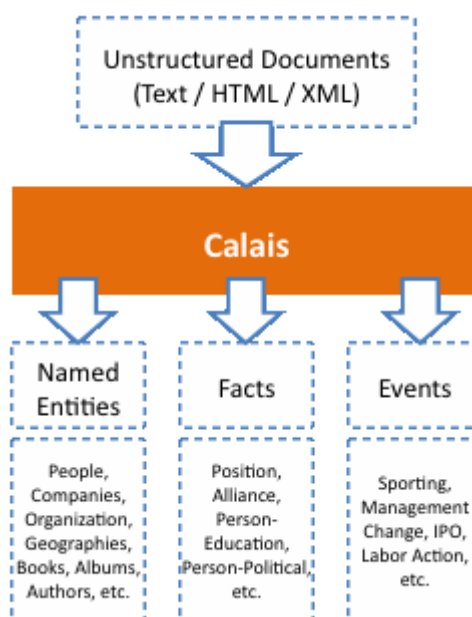


Figura 2.3: Vista do processo de extração de meta-dados [retirado de [Reu11]]

e exemplificado no seu *website*. Exceptuando OpenCalais Web Service já mencionado, outras funcionalidades importantes são listadas de seguida, acompanhadas de uma breve descrição:

- Calais Linked Data : utiliza a incorporação do Calais na *Linking Open Data (LOD) Cloud*, permitindo a exposição, partilha e aquisição de dados adicionais sobre as entidades extraídas de documentos, assim como a sua devolução ao utilizador, no modelo de meta-dados RDF, através de URIs (*Uniform Resource Identifiers*);
- Calais Marmoset : permite a um site ou aplicação fornecer automaticamente metadados micro-formatados a Yahoo [Yah11b] e outros *crawlers*<sup>2</sup> automatizados de metadados;
- Calais Yahoo! Pipes : permite aos utilizadores de Yahoo! Pipes, enriquecer os seus *feeds* RSS personalizados com meta-dados semânticos;
- Calais Tagaroo : uma extensão para o Wordpress [Aut11] que melhora as capacidades destes blogues, com sugestões automáticas de *tags*, pontuações de relevância, categorização de documentos e rápida incorporação de imagens disponíveis em Flickr [Yah11a];
- Gnosis : uma extensão para Firefox (3.5) e Internet Explorer, que automaticamente analisa o conteúdo lido no *browser*, fornecendo uma variedade de ferramentas para

<sup>2</sup>*crawler* : um programa que navega pela Internet de forma metódica e automatizada; em geral, começa com uma lista de URLs para visitar e, à medida que visita esses endereços, identifica todas as hiperligações na página adicionando-as a uma lista de hiperligações a visitar; tais hiperligações são visitadas recursivamente de acordo com um conjunto de políticas definidas.

explorar informação sobre indivíduos, companhias, locais e outros assuntos, associados com aquilo que o utilizador está a consultar.

O ambiente OpenCalais encontra-se disponível numa versão gratuita, incluindo o conjunto completo de ferramentas e funcionalidades, para uso comercial e não comercial.

## 2.2 Grupos e Aplicações Interactivas

A nova vaga de aplicações desenhadas para a Internet vem redefinir os níveis de interactividade oferecidos aos utilizadores. Aplicações como os sites de redes sociais, GooglePlus [Goo11b], Flickr [Yah11a] e outras aplicações que se sustentam no conceito de *Social Media* [KH10], utilizam técnicas escaláveis e acessíveis de publicação de conteúdos, transformando a comunicação em diálogos interactivos.

À medida que estas aplicações abarcam um grande número de utilizadores, a utilização de grupos surge naturalmente, como uma forma de organizar e estruturar os utilizadores, dentro dos sistemas e aplicações. Devido à enorme quantidade de aplicações interactivas existentes, na pesquisa efectuada, optou-se por analisar exemplos de dois subconjuntos destas aplicações e sistemas: software colaborativo e sites de redes sociais.

As aplicações e sistemas escolhidos abrangem duas vertentes distintas, relacionadas com o uso profissional e o lazer, sendo que algumas delas, são utilizadas simultaneamente com os dois propósitos pelos seus utilizadores. As aplicações escolhidas são também aquelas que se baseiam fundamentalmente, nos utilizadores, seus perfis e grupos, servindo de casos de estudo e potenciais ambientes de teste para os conceitos e métodos abordados, durante esta dissertação.

### 2.2.1 Software Colaborativo

Software colaborativo [BAVR08], *groupware* ou *social software* são designações genéricas para caracterizar um software desenhado para apoiar utilizadores que partilhem uma tarefa comum, ajudando-os a atingir os seus objectivos. Sendo um conceito amplo, pode estar associado a sistemas de correio electrónico assíncrono, agendas corporativas, aplicações de *chat*, sistemas de gestão de conteúdos e sistemas que envolvam comunidades, entre muitos outros tipos de sistemas ou aplicações, que incluam quaisquer grupos de trabalho.

Ao contrário do termo mais geral de software social, aplicado a sistemas usados num contexto afastado do trabalho, este tipo de software pretende criar um ambiente de trabalho colaborativo. O ambiente criado tem a vantagem de apoiar o trabalho de indivíduos ou grupos cooperantes, independentemente da sua localização geográfica, suportando uma qualquer forma de organização humana, onde a colaboração seja um requisito.

Devido à sua própria natureza estar associada à utilização de grupos de utilizadores, este tipo de software pode beneficiar do conceito de grupos implícitos, discutido ao longo deste trabalho.

### 2.2.1.1 Moodle

A plataforma Moodle [Tru11] baseia-se em software livre, estando vocacionada para a gestão de cursos baseados na Internet (*Course Management System*), também conhecido como um sistema de apoio à aprendizagem num ambiente virtual (*e-learning*). Orientada para o apoio a educação, esta ferramenta permite uma fácil criação de *websites* dinâmicos de apoio ao ensino, apesar de poder ser usada, também, em processos de formação, desenvolvimento ou outros contextos.

Baseada na linguagem PHP, a arquitectura da plataforma Moodle apresenta uma estrutura modular que potencia o desenvolvimento de extensões, relevantes no aumento das suas funcionalidades. A sua arquitectura é constituída por um núcleo de vários componentes principais, um conjunto de bibliotecas para usos mais específicos e um conjunto de módulos e extensões.

Os componentes principais comportam uma camada de abstracção de base de dados, um sistema de controlo de permissões e papéis desempenhados pelos utilizadores, uma biblioteca para a criação de formulários e uma interface de programação (API) para gerir vários tipos de ficheiros. O esquema da base de dados, juntamente com três bibliotecas (*moodlelib.php*, *weblib.php* e *datalib.php*) que contêm as principais funções do sistema e o ficheiro de configuração *config.php*, finalizam o conjunto dos componentes fundamentais da infraestrutura.

Entre as várias bibliotecas mais específicas, encontram-se funcionalidades de autenticação, gestão de sessões, processamento de correio electrónico e um sistema de grupos. Os grupos são geridos e definidos explicitamente, pelos administradores e professores no interior do sistema. A actual funcionalidade do sistema de grupos permite a atribuição de alunos e professores a um ou mais grupos, utilizados ao nível das disciplinas ou actividades existentes no interior dessas mesmas disciplinas. As extensões ou módulos desenvolvidos podem também utilizar as funcionalidades nativas do sistema de grupos, recorrendo a uma interface de programação específica para o efeito, que fornece as funções necessárias (localizada em *modulelib.php*).

Na mais recente distribuição do Moodle [Tru11], versão 2.0, foi implementada uma nova versão de grupos globais de utilizadores (*Site-wide groups*), que podem ser usados num processo de inscrição manual ou automaticamente sincronizada, dos seus utilizadores em disciplinas, facilitando um processo, de outro modo, exaustivo e penoso.

Como é referido na própria documentação na Internet, possíveis melhorias do sistema de grupo incluem a criação de grupos de utilizadores baseados em contextos, reutilizáveis nos diferentes níveis do sistema, tornando o sistema mais flexível e adequado, por exemplo, a ambientes universitários, onde os grupos podem variar consideravelmente, consoante as disciplinas e actividades.

### 2.2.1.2 Alfresco

O sistema Alfresco [Sof11] é um sistema de gestão de conteúdos, desenvolvido em Java e direccionado para ambientes empresariais. Encontra-se disponível em duas versões distintas, sendo a primeira uma versão gratuita de código aberto e a segunda, uma versão comercial, mais completa em termos de funcionalidades e suporte técnico.

O objectivo deste software é fornecer uma plataforma, baseada na Internet, para a gestão de conteúdos, registos, imagens e desenvolvimento de conteúdo colaborativo. As tecnologias utilizadas são inúmeras, assim como as funcionalidades oferecidas. No contexto deste trabalho, no entanto, o facto considerado mais importante é a sua utilização de grupos de utilizadores, como base organizativa.

Os grupos de utilizadores do sistema Alfresco, assim como no anterior software (Module), são definidos explicitamente pelos utilizadores com os privilégios adequados. Os grupos criados permitem aos seus membros acederem a determinados recursos e conteúdos, formarem equipas destacadas para tarefas/projectos e beneficiarem das várias funcionalidades oferecidas pelo sistema, entre as quais se destacam:

- Gestão Documental
- Gestão de Conteúdos Web
- Colaboração
- Plataforma de Conteúdos e Repositório
- Serviços de Interoperabilidade de Gestão de Conteúdos
- Gestão de Registos
- Gestão de Imagens

Existem diferentes ambientes de desenvolvimento, capazes de fornecer as ferramentas adequadas ao desenvolvimento de aplicações, serviços e novos módulos, associados ao sistema Alfresco. Encontram-se disponíveis sistema de apoio ao desenvolvimento, sistemas de controle de versão e interfaces de programação remotas, que permitem criar, modificar e reutilizar as funcionalidades oferecidas, incluindo o sistema de grupos, acessível através de um serviço REST API Group Service.

## 2.2.2 Sites de Redes Sociais

Sendo o exemplo mais abrangente, deste tipo de serviços e sites, o Facebook é brevemente abordado na subsecção seguinte. Uma conceptualização genérica, do funcionamento dos referidos sites, foi apresentada na secção 2.1.1.2, tendo por base as investigações efectuadas por Boyd e Ellison [BE07, Boy11].



### 2.2.2.1 Facebook

O sistema Facebook [Fac11] é suportado por um serviço e um *website*, pertencentes à categoria dos sites de redes sociais (brevemente descritos na secção 2.1.1.2). Como foi também observado anteriormente, o seu funcionamento baseia-se na possibilidade dos utilizadores construírem perfis pessoais, podendo depois adicionar outros utilizadores como "amigos", com os quais podem interagir de diversas formas, por exemplo, trocando mensagens. Adicionalmente, os utilizadores podem juntar-se explicitamente a grupos de interesses, organizados por local de trabalho, estabelecimento de ensino frequentado e outras características.

Além do funcionamento semelhante a outros ambientes do mesmo género, o sistema Facebook oferece a capacidade aos utilizadores, de utilizarem múltiplas extensões sociais, como por exemplo o botão Like, nos seus *websites*, aplicações web e aplicações móveis. Permite também a construção de aplicações, directamente integráveis no funcionamento base do sistema, disponibilizando funcionalidades nativas reutilizáveis, através de um conjunto de interfaces e sistemas de apoio ao desenvolvimento.

Existem quatro conceitos fundamentais na plataforma [Fac11]. A interface Graph API é o núcleo da plataforma, permitindo ler e escrever dados no sistema e ter acesso ao grafo social, que representa as entidades como pessoas, fotos, eventos e páginas e as ligações entre elas. O processo de autenticação do Facebook permite a uma aplicação interagir com a Graph API como se fosse um utilizador normal do sistema, constituindo um mecanismo único que pode ser usado com todos os tipos de aplicações suportadas. O terceiro conceito está relacionado com a importância das extensões sociais que oferece, que permitem a personalização dos conteúdos consultados pelos utilizadores. Por último, o protocolo Open Graph Protocol permite aos utilizadores, integrarem as suas páginas web no grafo social construído pelo Facebook, adquirindo as funcionalidades adicionais conferidas pela plataforma.

Em Outubro de 2010, os criadores do Facebook anunciaram novas funcionalidades e futuros objectivos para a componente de grupos (Facebook Groups)[Par10]. A componente redesenhada tem, como propósito, permitir a socialização de utilizadores, em torno de um tópico ou comunidade. Os novos grupos podem ser vistos como representações mais aproximadas dos grupos, com que cada utilizador interage quotidianamente.

As novas funcionalidades implementadas incluem notificações de grupo, conferindo aos administradores dos grupos a possibilidade de enviarem directamente mensagens para os membros, assim que qualquer objecto ou mensagem, seja adicionado ao conteúdo do grupo. Outra funcionalidade acrescentada é o *chat* de grupo, uma sessão de *chat* para cada grupo, onde os membros podem participar. Por último, cada grupo fornece, também, a possibilidade de cada membro contribuir com documentos básicos de texto, que possuem um histórico de modificações, possivelmente consultado por qualquer utilizador que seja membro.

Mark Zuckerberg e Chris Cox [Par10], relativamente aos novos grupos no Facebook,

delinearam alguns objectivos futuros para a referida componente, pretendendo desenvolver métodos capazes de medirem as interações de cada utilizador com os seus grupos, de forma a determinar a sua importância. A determinação destas forças das relações, ao longo do tempo, servirá de base para oferecer novas funcionalidades, como filtros de amigos baseados nos grupos de que estes fazem parte, ou outras funcionalidades que minimizem ou invalidem a necessidade de os utilizadores definirem explicitamente grupos de afiliações (ou grupos de contactos), nas suas listas de amigos.

#### 2.2.2.2 Twitter

O sistema Twitter [Twi11] é suportado por um *site* de redes sociais que fornece aos seus utilizadores um serviço de *microblogging*, que consiste na capacidade destes poderem enviar e receber pequenas mensagens de texto, com um limite de 140 caracteres, conhecidas como *tweets*. Estas mensagens podem ser enviadas através do referido site, via SMS ou através de outros sites e aplicações que utilizem a sua interface de programação (API).

A funcionalidade do Twitter baseia-se na utilização da rede de contactos de cada utilizador. Quando um utilizador envia uma mensagem para a sua página principal (perfil), o sistema distribui automaticamente, quase em tempo real, essa mensagem para os utilizadores que optaram por segui-lo, adicionando-o à sua lista de contactos. Por omissão, as mensagens enviadas por cada utilizador estão publicamente visíveis. No entanto, o criador pode optar por restringir a sua distribuição, aos seus seguidores.

A interface de programação fornecida pelo Twitter, consiste em três partes distintas: duas interfaces baseadas em REST [Fie00, FST02] e uma terceira, mais recente, designada Streaming API. Os métodos disponibilizados permitem a outras aplicações aceder aos dados do serviço, incluindo informações sobre o seu estado, cronogramas de actualização e dados dos utilizadores.

A utilização dos dados, presentes nos perfis dos utilizadores, pode contribuir para identificar grupos implícitos, como nos anteriores casos estudados. Os grupos identificados poderiam, entre outras funcionalidades, ser usados na sugestão de novos contactos a seguir, de acordo com os interesses e características partilhados.

## 2.3 Resumo dos Aspectos Essenciais

Ao longo do capítulo, foram abordados vários conceitos relacionados com grupos de utilizadores, no contexto das aplicações interactivas baseadas na Internet. Os conceitos, técnicas e metodologias apresentados, visaram promover a compreensão dos factores relevantes para a identificação, gestão e utilização de grupos de utilizadores, particularmente grupos implícitos.

A importância dos grupos de indivíduos na organização social e na modelação da sociedade, assume um papel determinante na maneira como as pessoas vivem, interagem

e se relacionam. Partindo deste facto, optou-se por abordar inicialmente a conceptualização do termo a nível social e humano, observando-se a sua ligação com os conceitos de grupo social, rede social e respectivas características.

A base anterior forneceu a sustentação para a compreensão das motivações e expectativas associadas aos membros de grupos de utilizadores num ambiente virtual, com foco especial nos emergentes *sites* de redes sociais e no software colaborativo. De acordo com esta linha de investigação, estudaram-se as representações digitais das identidades dos utilizadores, através da construção de perfis.

A busca de conhecimento sobre a identificação e recolha dos dados que permita caracterizar indivíduos e utilizadores, conduziu a pesquisa às áreas temáticas associadas a tecnologias de personalização. Os trabalhos desenvolvidos neste campo contribuíram com valiosa informação, permitindo vislumbrar inúmeras técnicas e metodologias úteis para a construção de perfis de utilizadores. No conjunto global das técnicas de personalização, as aproximações utilizadas nos sistemas de recomendação e as suas ramificações, assumiram uma ênfase especial, orientando a investigação.

Entre as metodologias importantes para a caracterização de utilizadores, destacaram-se a análise dos seus conteúdos e recursos, a análise dos seus comportamentos e a análise das suas interacções e relações. Estas diferentes metodologias que recorrem, por sua vez, a diferentes técnicas de recolha, técnicas de processamento e tipos de dados, são essenciais na agregação da informação que permite, posteriormente, identificar grupos implícitos de utilizadores.

Os trabalhos anteriormente desenvolvidos por múltiplos autores, em torno do conceito de grupos implícitos, complementaram a informação reunida com as diferentes estratégias desenvolvidas para a identificação de grupos implícitos, estratégias essas, adequadas a diferentes contextos e cenários de aplicação.

A recolha bibliográfica executada, relacionada com os temas deste capítulo, permitiu a identificação de diferentes soluções para a problemática da dissertação, juntamente com as respectivas vantagens e desvantagens. O conjunto de conhecimentos adquiridos e sintetizados neste capítulo, permitiram a conceptualização da ferramenta pretendida, apresentada e discutida nos seguintes capítulos, definindo os aspectos essenciais que moldam os conceitos inerentes ao seu desenvolvimento.





## Conceptualização da Ferramenta

O trabalho de síntese desenvolvido em torno do conceito de grupos implícitos e descrito no capítulo anterior, indica e antevê possíveis benefícios resultantes da utilização do conceito em aplicações distintas, como sites de redes sociais [Pal10], serviços de correio electrónico [RBDD<sup>+</sup>10] e redes de partilha de conteúdo ponto-a-ponto [CLQ06, CQL08].

Embora a abrangência do conceito englobe aplicações distintas, a generalidade das soluções apresentadas difere na metodologia usada na identificação de grupos implícitos, centrando-se em contextos aplicativos e objectivos específicos, aos quais o conceito é moldado.

Com o principal intuito de generalizar a aplicação do conceito discutido e facilitar a sua integração em aplicações interactivas distintas, desenvolveu-se uma ferramenta capaz de identificar e gerir dinamicamente grupos implícitos de utilizadores.

Nas secções seguintes deste capítulo, descrevem-se os principais pressupostos e objectivos da ferramenta desenvolvida, assim como os conceitos envolvidos na sua idealização.

### 3.1 Enquadramento

Na maioria das aplicações interactivas baseadas em comunidades de utilizadores, das quais são exemplos os sites de redes sociais como o Facebook [Fac11], Twitter [Twi11] e os sistemas de gestão e partilha de conteúdos como a plataforma Moodle [Tru11] e Alfresco [Sof11], a afiliação de utilizadores a grupos é realizada explicitamente pelos próprios ou por utilizadores aos quais é atribuído esse papel específico. A escolha dos grupos relevantes é realizada por cada utilizador, tendo em conta diversos factores, incluindo os seus objectivos pessoais, as relações sociais com os restantes membros, o tema do grupo e

outros objectivos decorrentes das funcionalidades particulares da aplicação considerada.

Uma vez que a identificação de grupos implícitos é realizada com base na informação existente sobre cada utilizador, é importante definir este processo/mecanismo de identificação da forma mais genérica possível, para se obter uma ferramenta flexível e útil a um maior número de aplicações e utilizadores. Devido à diversidade das aplicações existentes e aos diferentes níveis de especificidade da informação que contém sobre os seus utilizadores, é importante que a informação analisada e utilizada na identificação de grupos implícitos seja comum.

Tendo em conta os anteriores factores, a ferramenta desenvolvida deve basear-se na análise de informação textual que descreve o utilizador, seus interesses e características. Este tipo de informação comum, encontra-se frequentemente disponível nos campos de perfis de utilizadores (descrição de interesses, profissão, data de nascimento, etc..) e contida em recursos/conteúdos associados a esses mesmos utilizadores, tais como páginas web favoritas ou os meta-dados que classificam os conteúdos multimédia reproduzidos.

A dimensão temporal assume também um papel relevante na identificação de grupos implícitos. À medida que os utilizadores alteram os seus interesses, características e objectivos ao longo do tempo, torna-se essencial captar essas mudanças e reajustar dinamicamente os grupos identificados e seus respectivos membros, garantindo coerência. Por outro lado, as próprias interações entre utilizadores, ao longo do tempo, podem fornecer indicadores importantes de proximidade e afinidade entre os mesmos, devendo ser consideradas na identificação de grupos.

A utilidade dos grupos identificados depende da precisão nos processos/mecanismos de análise conteúdos e aferição dos interesses associados a cada utilizador. A utilidade destes mesmos grupos pode também variar consoante a aplicação cliente, seus objectivos, contextos e ambientes. A maximização da utilidade dos grupos identificados pode ser alcançada através da utilização de métricas que considerem, por exemplo, o número de utilizadores por grupo e a junção de grupos com interesses directamente relacionados. Todavia, para se manter a flexibilidade da ferramenta é necessário que esta seja neutra em relação a este tipo de métricas, oferecendo funcionalidades que permitam que estas sejam definidas e/ou sob controlo da aplicação cliente.

Por último, é interessante considerar o nível de automatismo do processo de identificação de grupos implícitos e o seu impacte nas aplicações que recorram à ferramenta. A privacidade dos dados existentes nos perfis dos utilizadores varia consoante a aplicação, pelo que é importante fornecer várias opções às aplicações cliente, no que diz respeito ao controlo do processo de análise de perfis e formação de grupos de utilizadores. Ainda assim, a utilização de ferramentas de análise de conteúdos oferece uma boa base para a implementação de um processo de identificação automático e configurável. A possibilidade de grupos implícitos serem detectados e formados automaticamente aumenta o seu potencial no desenvolvimento de funcionalidades e aplicações interactivas.

## 3.2 Objectivos

De acordo com os requisitos e as condicionantes referidas na secção anterior, enumeram-se os principais objectivos definidos para a ferramenta desenvolvida:

- apoiar a investigação do conceito de grupos implícitos, fornecendo uma ferramenta que explore o conceito;
- facilitar a integração do conceito de grupos implícitos em aplicações interactivas diversas, baseadas na Internet;
- definir um processo de identificação de grupos implícitos flexível e útil a diversas aplicações e contextos aplicativos;
- explorar processos e mecanismos automáticos de identificação e gestão de grupos implícitos;
- oferecer a possibilidade de diferentes aplicações partilharem a mesma ferramenta e contribuírem com os seus utilizadores para os grupos identificados, promovendo a interacção e partilha de conhecimento entre utilizadores de aplicações distintas, mas com interesses e/ou características comuns;
- desenvolver uma ferramenta neutra em relação a métricas que visem maximizar a utilidade dos grupos identificados, oferecendo no entanto mecanismos que facilitem a definição destas nas aplicações cliente, consoante seus contextos e objectivos;
- garantir a extensibilidade da ferramenta desenvolvida através de uma arquitectura modular.

De forma complementar ao trabalho realizado e à ferramenta desenvolvida, definem-se também as seguintes actividades:

- avaliar a utilidade dos grupos identificados pelo processo definido, através da realização de inquéritos de satisfação a utilizadores de teste;
- testar a ferramenta desenvolvida, num contexto aplicativo específico, através da sua integração na plataforma Moodle.

## 3.3 Conceitos

Nesta secção, contextualizam-se os principais conceitos envolvidos no desenvolvimento da ferramenta e no alcance dos objectivos definidos na secção 3.2. Os conceitos definidos, nas seguintes subsecções, introduzem os aspectos essenciais que definem as metodologias e os mecanismos utilizados.

### 3.3.1 Grupos Implícitos

No contexto da ferramenta desenvolvida, considera-se como grupo implícito de indivíduos ou utilizadores, o grupo definido por um conjunto de utilizadores que partilhem uma mesma característica, interesse ou qualquer outro tipo de relação/interacção. Como referem Teevan et al. e Morgado et al. em [TMB09, MCCC10] respectivamente, considera-se que este tipo de grupos pode ser formado dinamicamente e implicitamente, sem intervenção ou acção explícita dos membros nesse sentido.

A identificação e formação de grupos é realizada pela ferramenta, com base nas informações presentes nos perfis estendidos de cada utilizador.

### 3.3.2 Perfis Estendidos

Os perfis estendidos são construídos por extensão dos perfis básicos de cada utilizador, com dados inferidos através da análise de conteúdos e das interacções estabelecidas por cada utilizador.

Os perfis básicos, encontrados habitualmente em aplicações interactivas, fornecem por si só informação útil que descreve características como a faixa etária e a ocupação dos seus proprietários. Apesar deste facto, alguns dos campos existentes em perfis, dos quais são ilustrativos campos de descrição do utilizador e outros campos baseados em texto, podem conter implicitamente dados sobre as preferências e interesses dos seus utilizadores.

Estes interesses e preferências podem ser inferidos através da análise do texto contido em campos de perfis e outros documentos relacionados com o utilizador. Na ferramenta desenvolvida, a extracção de interesses é realizada por um componente responsável pela análise de conteúdos textuais.

No que diz respeito às interacções entre utilizadores, a metodologia usada na extensão dos perfis baseia-se na recolha e análise de três propriedades resultantes da troca de mensagens, concretamente, a "frescura", a frequência e a direcção/sentido das mesmas. A figura 3.1 ilustra o processo de obtenção de perfis estendidos.

### 3.3.3 Análise de Conteúdos

A análise de conteúdos relacionados com os utilizadores é da responsabilidade de um componente próprio na arquitectura da ferramenta. Este componente pode ser facilmente modificado ou substituído por outro componente que desempenhe funções similares.

Na actual implementação, o referido componente recorre a um serviço externo OpenCalais Web Service [Reu11]. O serviço OpenCalais é capaz de extrair meta-dados semânticos dos conteúdos fornecidos, identificando categorias de documentos, *tags* sociais e entidades presentes. Os meta-dados extraídos reflectem os temas, assuntos e entidades contidos nos documentos analisados, pelo que, implicitamente, indicam potenciais interesses dos utilizadores aos quais os documentos estão associados.



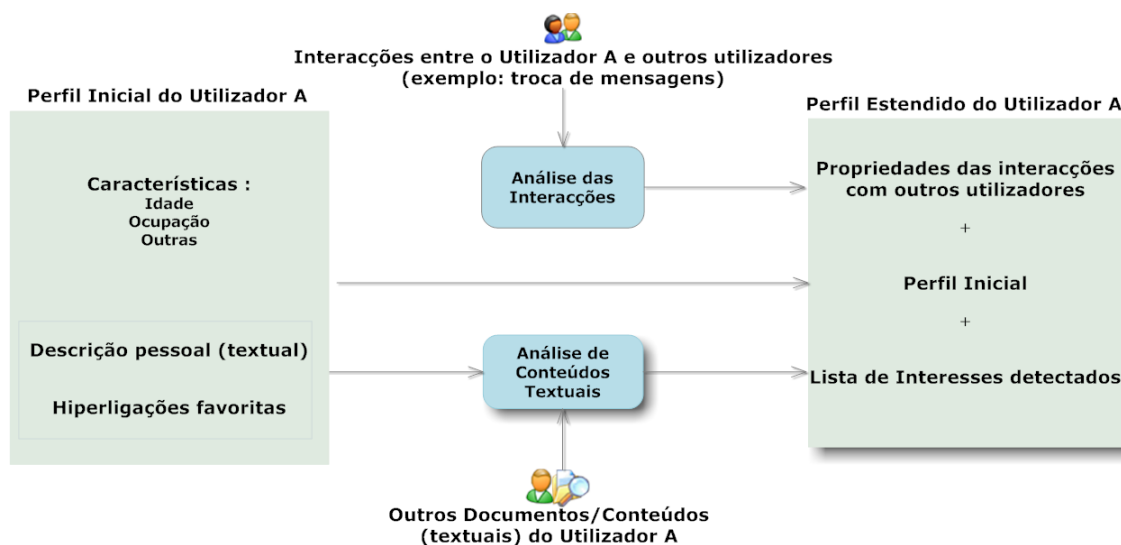


Figura 3.1: Vista generalizada da geração de perfis estendidos.

A escolha de OpenCalais como serviço externo introduz algumas limitações relacionadas com as funcionalidades que disponibiliza. O conteúdo analisável pelo componente que implementa o serviço baseia-se exclusivamente em conteúdo textual de língua inglesa. A diversidade e as relações entre as entidades, categorias de documentos e tags sociais identificadas, têm como base a ontologia própria do serviço. Informações adicionais sobre as funcionalidades oferecidas por este serviço podem ser encontradas na secção 2.1.6.2.

É importante referir que o conteúdo de um documento nem sempre está relacionado com o interesse que um indivíduo tem nesse mesmo documento. O motivo da relevância que um determinado utilizador atribui a certo documento pode basear-se em aspectos pessoais, dificilmente previsíveis. Este factor de incerteza e imprecisão, na aferição dos interesses de cada utilizador, pode ser minimizado através da configuração dos processos de identificação e gestão de grupos na ferramenta ou através da implementação de métricas complementares na aplicação cliente. Os processos de identificação e gestão de grupos são introduzidos nas secções seguintes.

### 3.3.4 Processo de Identificação

Os perfis estendidos gerados para cada utilizador constituem a base do processo de identificação de grupos implícitos de utilizadores. Múltiplos grupos de utilizadores podem ser identificados e categorizados em três tipos distintos, de acordo com os respectivos critérios utilizados na sua identificação : grupos baseados em interações, grupos baseados em interesses e grupos baseados em características.

Os diferentes tipos de grupos identificados, podem posteriormente ser combinados na obtenção de subgrupos de utilizadores que satisfaçam certas expressões booleanas de interesses, características e interações. A combinação dos diferentes grupos confere

versatilidade ao desenvolvimento de funcionalidades interactivas e dinâmicas de personalização.

#### 3.3.4.1 Grupos baseados em Interações

Segundo este critério, cada utilizador possui um grupo de contactos (outros utilizadores) com quem interage. Os grupos referidos são formados através da análise da direcção, frequência e frescura das interações entre os utilizadores no interior da ferramenta e/ou aplicação cliente, numa aproximação genérica da solução apresentada em [RBDD<sup>+</sup>10]. A análise das propriedades das interações entre utilizadores permite definir custos associados a cada interacção e obter a relevância de cada contacto para o utilizador considerado.

As interações entre os utilizadores, através da ferramenta, são suportadas pelo sistema de mensagens disponibilizado. As interações que ocorram externamente à ferramenta (por exemplo, através de sistemas de mensagens das aplicações cliente) podem contribuir para a formação deste tipo de grupos, através do uso de funcionalidades específicas oferecidas pela interface de programação disponibilizada. A descrição detalhada destas funcionalidades é apresentada nos capítulos 4 e 5, referentes à arquitectura e utilização da ferramenta.

#### 3.3.4.2 Grupos baseados em Características

Os grupos identificados segundo este critério são constituídos por utilizadores que partilham uma mesma característica. A identificação é realizada através da análise dos valores contidos em campos comuns, tais como campos de data de nascimento, sexo e ocupação. Esta identificação apenas é possível caso seja conhecido o domínio de valores possíveis para cada característica.

Uma vez que a formatação dos dados presentes nos campos mencionados pode variar consoante a aplicação, é possível definir os valores possíveis para cada um destes campos, através da interface de administração e configuração da ferramenta ou API fornecida.

#### 3.3.4.3 Grupos baseados em Interesses

Este critério permite o agrupamento de utilizadores que possuem interesses comuns. Os interesses afectos a cada utilizador são identificados pelo componente de análise de conteúdos, através do processamento dos dados existentes em campos de descrição do utilizador ou qualquer outro conteúdo/recurso baseado em texto e relevante para o utilizador. Artigos de notícias ou *websites* favoritos, são exemplos de recursos que podem ser utilizados na identificação de interesses de um utilizador.

Para se garantir a flexibilidade da ferramenta, assim como a utilidade dos grupos identificados, o processo de identificação de grupos de interesses pode ser realizado com diferentes níveis de automatismo.

Cada nível de automatismo oferece, às aplicações cliente, um certo grau de controlo sobre os interesses identificados e a sua atribuição aos respectivos utilizadores. Embora os interesses detectados possam ser directamente associados a utilizadores e contribuir para a formação automática de grupos de interesses, torna-se importante oferecer às aplicações uma maneira de filtrar os interesses detectados, definindo quais são realmente relevantes para os seus utilizadores, contextos aplicativos e para a criação de grupos. As opções fornecidas às aplicações cliente e administradores da ferramenta, no que diz respeito ao automatismo do processo abordado, são descritas detalhadamente no próximo capítulo, na secção 4.2.2.

### 3.3.5 Gestão de grupos

Os interesses, características e interações, que cada utilizador exhibe, variam ao longo do tempo, implicando que após a conclusão do processo de identificação e consequente formação de grupos, seja necessário um processo de gestão que mantenha a coerência dos perfis e grupos existentes.

Alterações efectuadas pelos utilizadores nos campos dos seus perfis, desde que comunicadas pela aplicação cliente à ferramenta, desencadeiam um conjunto de acções que visam manter a integridade e coerência dos interesses e características de cada utilizador e respectivos grupos.

Embora a ferramenta ofereça um mecanismo explícito que permite a cada utilizador abandonar um determinado grupo, existe também um processo de gestão automático configurável e independente da ordem explícita dos utilizadores. Este processo regista continuamente as interações que cada utilizador realiza com os seus grupos, lendo ou enviando mensagens para o grupo. O registo efectuado permite que seja possível estabelecer limites mínimos de interacção para que um utilizador continue a ser considerado membro de um grupo.

A solução anterior é especialmente benéfica para os grupos de interesses, uma vez que como foi discutido anteriormente, nem sempre os grupos identificados podem ser úteis para os utilizadores. Estes utilizadores podem assim ser removidos automaticamente, caso não utilizem com a frequência exigida e definida os grupos atribuídos.

A gestão de grupos compreende também outras potenciais restrições que podem ser utilizadas no controlo do número de grupos activos e na consequente gestão dos recursos da ferramenta. O número mínimo de utilizadores para que um grupo seja considerado activo é passível de configuração. Um grupo activo permite o envio e consulta de mensagens e recursos, enquanto um grupo inactivo funciona apenas como uma simples lista de utilizadores associados, os quais não podem aceder ou utilizar o espaço do grupo.

Os grupos baseados em interações representam um caso particular para o processo de gestão de grupos. As funcionalidades disponibilizadas sobre estes grupos incluem o registo das interações entre utilizadores distintos e a definição dos pesos associados às propriedades analisadas de uma interacção. Devido à natureza destes grupos e à sua

relação com o tempo, a gestão deste tipo de grupos é simples, consistindo apenas na ordenação dos utilizadores com base na relevância calculada após cada nova interacção.

Independentemente dos mecanismos introduzidos anteriormente, é importante salientar que mecanismos explícitos de atribuição e remoção de interesses, utilizadores e grupos estão disponíveis para uso das aplicações cliente. Estas funcionalidades, disponibilizadas pelas interfaces desenvolvidas, podem ser utilizadas pelas aplicações cliente na gestão directa dos grupos formados e seus membros.

### 3.4 Resumo dos Aspectos Essenciais

Ao longo deste capítulo foram identificadas e discutidas dimensões relevantes associadas ao conceito de grupos implícitos e aos mecanismos utilizados na concepção de uma ferramenta de suporte ao conceito. Resumidamente, consideram-se essenciais para a concepção e implementação de uma ferramenta de grupos implícitos, os seguintes aspectos:

- o tipo de informação/conteúdo utilizado na identificação dos interesses dos utilizadores e consequentes grupos implícitos baseados em interesses;
- privacidade dos utilizadores;
- análise da interacção entre utilizadores como forma de identificar grupos implícitos;
- dimensão temporal/evolução dos interesses, características e interacções dos utilizadores;
- neutralidade da ferramenta e seus mecanismos em relação a métricas de maximização da utilidade dos grupos identificados;
- nível de automatismo e controlo dos processos/mecanismos de atribuição de interesses aos utilizadores e da geração de grupos implícitos.

Os aspectos enumerados, anteriormente, influenciaram as estratégias de implementação associadas aos mecanismos utilizados na identificação e gestão de grupos implícitos, discutidos de seguida, no capítulo 4.

# 4

## TIG : Uma Ferramenta de Grupos Implícitos

No decorrer deste capítulo é descrita a arquitectura definida e utilizada no desenvolvimento da ferramenta de grupos implícitos, TIG (Tool for Implicit Groups). Numa fase posterior são discutidos os detalhes relacionados com a implementação de TIG e com as funcionalidades oferecidas através de diferentes interfaces.

### 4.1 Arquitectura

O desenvolvimento desta ferramenta de grupos implícitos baseia-se no padrão arquitectural de software MVC (Model-View-Controller) [KP88], tendo sido associado ao uso da linguagem PHP e à plataforma de desenvolvimento CakePHP [CSF11].

A escolha do padrão MVC confere à ferramenta uma estrutura modular e bem definida, impondo uma clara separação entre as diferentes camadas da aplicação. A separação destas camadas contribui para a extensibilidade, reutilização e organização das funcionalidades oferecidas [CG08]. A próxima subsecção detalha o padrão seleccionado e a forma como é usado para disponibilizar as funcionalidades implementadas a outras aplicações interactivas, baseadas na web.

#### 4.1.1 Padrão Arquitectural MVC

À semelhança das tradicionais arquitecturas cliente-servidor, baseadas em múltiplas camadas, o padrão MVC permite isolar diferentes camadas dentro de uma aplicação.

A diferença mais significativa para um modelo clássico de três camadas (apresentação, lógica e dados), reside no facto de o MVC não se comportar como uma arquitectura linear, ignorando o pressuposto que proíbe que a camada de apresentação (interface/cliente), comunique directamente com a camada de dados. Enquanto que no modelo de três camadas, toda a comunicação passa obrigatoriamente pela camada intermédia (lógica da negócio), a arquitectura triangular do modelo MVC permite que a camada de dados interaja directamente com a camada de apresentação (ver figura 4.1).

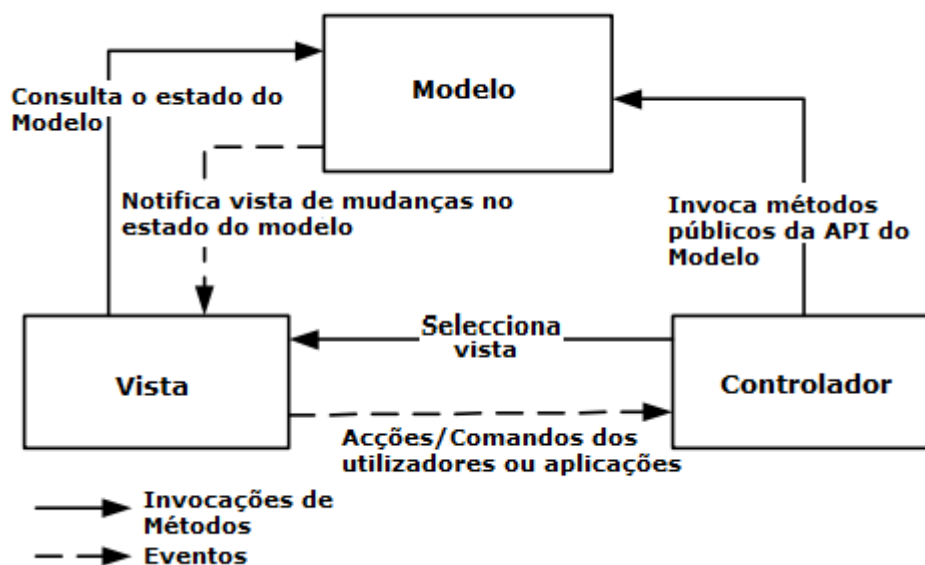


Figura 4.1: Ilustração do padrão arquitetural de software MVC

No caso concreto da ferramenta desenvolvida sobre a plataforma CakePHP, a utilização do padrão MVC separa a aplicação em três partes distintas: modelos (1), controladores (2) e vistas (*views*) (3).

1. Os modelos gerem o comportamento e os dados do domínio aplicacional, respondendo a pedidos de vistas sobre a informação do seu estado e a instruções dos controladores, que alteram esse mesmo estado. Além de fornecem uma camada de abstracção de dados à ferramenta desenvolvida, possuem a capacidade de alertar as vistas quando se verifica alguma mudança no seu estado.
2. Os controladores recebem os pedidos provenientes das vistas e iniciam uma resposta, instruindo os modelos das acções a realizar, de acordo com os *inputs* recebidos.
3. As vistas formam a interface com as aplicações, encaminhando pedidos para os controladores e devolvendo o resultado das operações efectuadas aos clientes.

No caso particular da plataforma CakePHP e da ferramenta desenvolvida, um pedido

básico à arquitectura MVC é ilustrado na figura 4.2. A título exemplificativo, assuma-se

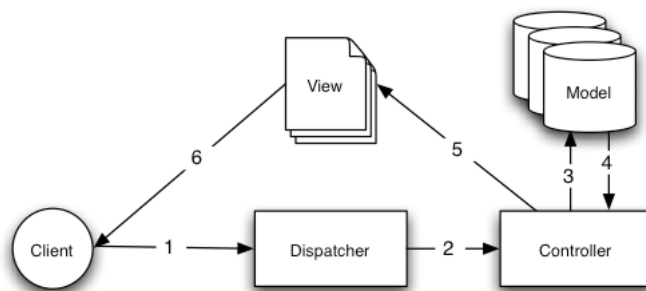


Figura 4.2: Pedido básico MVC, <http://book.cakephp.org/view/10/Understanding-Model-View-Controller>

que uma aplicação cliente pretende consultar os grupos implícitos existentes em determinado momento:

- a aplicação cliente executa um pedido para o URL `http://tig.com/groups/index`, ao servidor web onde está alojada a ferramenta;
- o *dispatcher* analisa o URL pedido (`/groups/index`) e transfere o pedido para o controlador apropriado;
- o controlador executa a lógica específica da aplicação; por exemplo, verificando as credencias do cliente e o seu acesso aos dados requisitados;
- o controlador usa depois os modelos para garantir o acesso aos dados da aplicação; neste exemplo, o controlador utilizaria o modelo Grupo para obter os grupos disponíveis;
- depois do controlador executar o processamento necessário sobre os dados adquiridos, passa-os para uma vista; a vista recebe os dados e procede à sua apresentação ao cliente.

Historicamente, as aplicações web desenvolvidas sobre o modelo MVC, utilizam vistas que recorrem a interfaces HTML. Embora o HTML continue a ser a interface dominante para aplicações web, novas opções para as vistas continuam a emergir. Estas opções incluem a linguagem Flash, linguagens de anotação como XHTML, XML/XSL e *Web Services*.

Devido à necessidade de se oferecer uma interface versátil a aplicações cliente que pretendam utilizar a ferramenta, optou-se por utilizar vistas que permitem a formatação dos dados das respostas nas linguagens XML e JSON. As linguagens anteriores são utilizadas em conjunto com as funcionalidades oferecidas pela plataforma CakePHP para desenvolvimento de *web services*, na disponibilização de uma API estilo REST. As aplicações

cliente podem aceder aos dados pretendidos e introduzir novos dados na ferramenta, utilizando o protocolo HTTP e as operações GET e POST, com os parâmetros especificados. A especificação da API REST oferecida pode ser consultada na secção 4.3.3.

Os modelos, controladores e outros componentes respeitantes à arquitectura da ferramenta são descritos nas seguintes subsecções. A figura 4.3 ilustra a arquitectura da ferramenta desenvolvida.

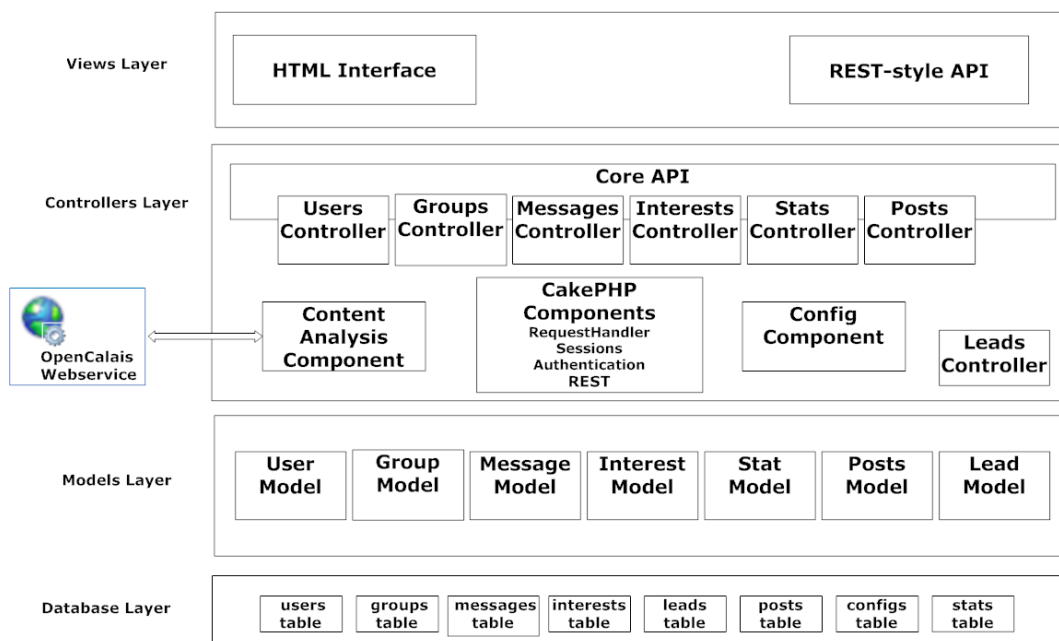


Figura 4.3: Visão das diferentes camadas arquiteturais da ferramenta TIG.

#### 4.1.1.1 Modelos e Entidades

Na plataforma de desenvolvimento escolhida, os modelos representam, normalmente, tabelas de uma base de dados, servindo na modelação das entidades persistentes da aplicação.

Os modelos e tabelas definidos na implementação da ferramenta respeitam as convenções<sup>1</sup> impostas pela plataforma.

O desenvolvimento de aplicações em CakePHP assume que as relações entre as diferentes entidades, não são definidas ao nível da base de dados, mas sim ao nível dos modelos.

A figura 4.4 ilustra os modelos existentes e as suas relações na notação do CakePHP. As relações `belongsTo`, presentes na referida figura, implicam uma relação `hasMany` no sentido oposto e vice-versa, enquanto a relação `hasOne` implica uma relação `belongsTo` no sentido inverso. Exemplificando as implicações anteriores, uma Mensagem pertence a

<sup>1</sup>As convenções impostas relativamente à definição de tabelas da base de dados e aos modelos associados podem ser consultadas em <http://book.cakephp.org/view/903/Model-and-Database-Conventions>.



um único Utilizador (fonte) mas um Utilizador pode enviar várias Mensagens. O mesmo acontece para os Grupos que podem conter várias Mensagens embora cada Mensagem distinta só possa estar associada a um único Grupo.

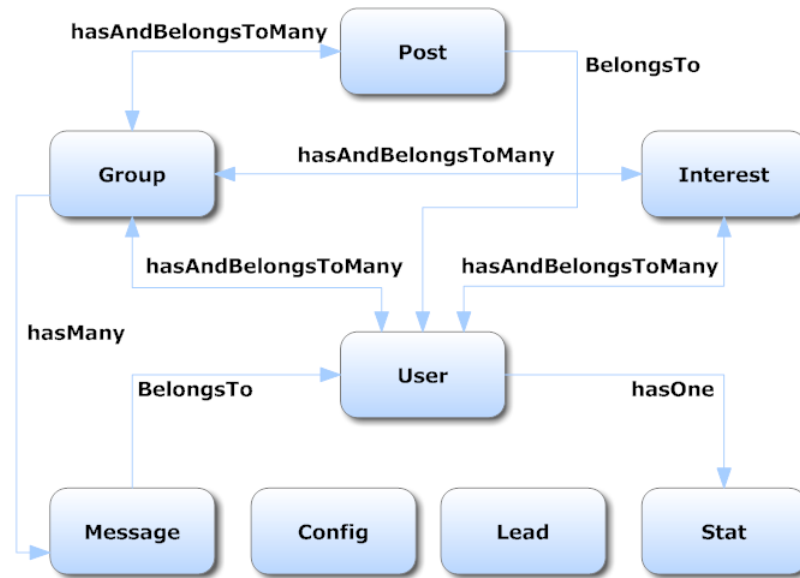


Figura 4.4: Relações entre modelos na ferramenta TIG (notação CakePHP)

Os modelos principais da ferramenta TIG estão associados às entidades User, Group, Interest, Stat, Message e Post.

- **Modelo User:** um utilizador possui, entre outros atributos, um identificador único e um conjunto de atributos que define o seu perfil, exemplos desses atributos incluem o nome de utilizador, a descrição do utilizador e a sua lista de interesses.
- **Modelo Group:** um grupo implícito ou explícito possui os atributos, nome, descrição e tipo; o tipo do grupo varia conforme o critério utilizado na sua identificação (secção 3.3.4) e é representado por um único carácter.
- **Modelo Interest:** a entidade Interest representa qualquer palavra que classifique um potencial objecto de interesse de um utilizador. O universo de interesses está directamente relacionado com o componente de análise de conteúdo escolhido e a ontologia que utiliza; os seus atributos consistem num identificador único e na sua designação;
- **Modelo Stat:** a entidade Stat engloba os dados estatísticos recolhidos através da análise das interacções entre utilizadores. Os seus atributos incluem uma lista de identificadores de utilizadores e as respectivas propriedades das interacções com cada um deles (frequência, direcção e "frescura"). Este modelo é centrado no utilizador, cada utilizador na plataforma possui uma entidade/modelo deste tipo.

- **Modelo Message:** uma entidade Message possui, entre outros atributos, um conteúdo, um recipiente e uma fonte. Todos os utilizadores ou grupos existentes na aplicação formam possíveis recipientes. As fontes, por outro lado, pertencem exclusivamente ao universo de utilizadores.
- **Modelo Post:** Posts representa um tipo específico de mensagem que apenas pode ser dirigida a grupos. Cada Post pode, no entanto, estar associado a múltiplos grupos simultaneamente. Os atributos do modelo são semelhantes aos atributos de uma entidade Message.

Outros modelos existentes incluem as entidades Lead e Config.

- **Modelo Lead:** representa um log de acções pedidas/realizadas por um cliente da ferramenta;
- **Modelo Config:** representa uma possível configuração da ferramenta. Os atributos deste modelo incluem o tipo de automatismo do processo de identificação de grupos de interesses, os pesos atribuídos a cada propriedade de interacção entre utilizadores e outras parametrizações, relacionadas com o componente de análise de conteúdos e os mecanismos de gestão de grupos.

A cada modelo está associado um conjunto de primitivas que permitem, modificar, actualizar e consultar os dados que representa. Estas primitivas podem ser invocadas, exclusivamente, por métodos definidos nos controladores.

#### 4.1.1.2 Controladores

Os controladores são utilizados na gestão da lógica da ferramenta. Cada controlador gere a lógica de uma parte da aplicação e está normalmente associado, directamente, a um único modelo. Um único controlador pode, no entanto, executar alterações/consultas em diferentes modelos através das relações entre o modelo que controla directamente e os restantes associados ao mesmo.

Na arquitectura de TIG, cada modelo possui um controlador associado. Assim para cada um dos modelos, referidos na anterior secção 4.1.1.1, existe um controlador, cuja designação é formada pela concatenação da forma plural do nome do modelo com o sufixo “*Controller*” (Exemplo: *UsersController*).

Um controlador pode incluir um qualquer número de métodos, funções ou acções, conforme a designação utilizada no ambiente CakePHP. Estas acções são métodos do controlador utilizados no preenchimento de vistas. Uma acção corresponde a um método único do controlador.

As aplicações externas e utilizadores individuais podem utilizar as funcionalidades da ferramenta através de pedidos de URL's. Cada pedido é traduzido pelo *dispatcher* do CakePHP na execução de uma acção que produz como resultado uma vista ou conjunto

de dados devolvidos ao cliente. Um componente de autenticação partilhado pelos diferentes controladores é capaz de autenticar a origem do pedido, permitindo que as acções determinem as vistas apropriadas para os diferentes clientes.

Os detalhes sobre a lógica utilizada nos diferentes controladores e processos da ferramenta, são discutidos na secção 4.2 deste capítulo.

#### 4.1.1.3 Componentes

Os componentes representam pacotes de lógica que podem ser partilhados entre controladores. A ferramenta desenvolvida utiliza vários componentes pertencentes, por omissão, à plataforma CakePHP: componente de sessão, componente de autenticação e componente de tratamento de pedidos.

- **Componente de Sessão:** fornece a capacidade de tornar os dados dos clientes persistentes, entre pedidos de páginas distintos. Actua como um *wrapper* para a variável `$_SESSION`, facilitando o acesso a métodos e funções relacionados com a mesma;
- **Componente de Autenticação:** fornece um sistema de autenticação de utilizadores, incluindo uma variedade de processos, tais como *callbacks* de controladores e listas de controlo de acessos. Permite definir as acções, métodos e respectivos controladores que podem ser acedidos por cada utilizador, papel de utilizador ou aplicação cliente;
- **Componente de Tratamento de Pedidos:** permite inspeccionar os pedidos dos utilizadores ou aplicações cliente. Informa a ferramenta sobre os tipos de conteúdos incluídos num pedido e a informação que requisita.

Para além dos componentes oferecidos pela plataforma de desenvolvimento, existem dois componentes adicionais e essenciais no funcionamento de TIG: componente de análise de conteúdos e componente de configuração da ferramenta.

- **Componente de Análise de Conteúdos:** responsável pela análise de documentos textuais e pela extracção dos interesses que reflectem os seus conteúdos. Este componente pode ser substituído facilmente por outro componente com funções semelhantes, devido à estrutura modular da ferramenta. A actual implementação do componente utiliza um cliente REST para o *webservice* externo OpenCalais, permitindo à ferramenta usufruir das suas funcionalidades. Outro componente semelhante, baseado no conjunto de ferramentas AlchemyAPI [Orc11] é também fornecido pela ferramenta.
- **Componente de Configuração da Ferramenta:** disponibiliza aos vários controladores da ferramenta, os parâmetros e as configurações necessários para o funcionamento pretendido. As diferentes configurações podem ser definidas pelos utilizadores ou aplicações cliente, afectando os processos de gestão e identificação de

grupos implícitos de utilizadores. Cada uma das possíveis opções de configuração é descrita na secção correspondente ao processo que afecta, ao longo da secção seguinte (4.2).

## 4.2 Implementação de um Protótipo

Nesta secção, são apresentados e descritos os principais mecanismos que promovem a identificação, formação e gestão de grupos implícitos de utilizadores.

### 4.2.1 Perfis Estendidos

O perfil estendido de cada utilizador consiste no conjunto dos dados contidos nas instâncias dos seus modelos User, Stat e Interest. O modelo User representa os dados do perfil básico de um utilizador, incluindo as suas características, descrições e recursos associados. O modelo Stat representa os dados estatísticos relativos às suas interacções com outros utilizadores. O modelo Interest representa cada interesse associado ao utilizador considerado.

Cada um destes modelos possui um conjunto de primitivas que permite alterar e consultar os dados que modela. Estas primitivas são invocadas pelos controladores de acordo com as acções desencadeadas pelos utilizadores ou aplicações cliente.

A plataforma de desenvolvimento utilizada, oferece um conjunto básico de primitivas comuns aos vários modelos definidos, encapsulando as operações necessárias para manipular as entidades modeladas e os dados contidos nas respectivas tabelas da base de dados.

A lista de primitivas oferecida por omissão, pela plataforma CakePHP, abrange três conjuntos de operações essenciais sobre modelos, permitindo consultar, guardar ou remover os dados modelados. A lista completa de primitivas oferecidas, pode ser consultada na documentação da plataforma CakePHP <sup>2</sup>.

### 4.2.2 Identificação de Grupos baseados em Interesses

A figura 4.5 ilustra de forma simplificada as principais entidades, componentes e controladores, envolvidos na identificação e gestão de grupos implícitos baseados em interesses.

No primeiro passo a), indicado na figura 4.5, os dados do perfil estendido de um utilizador são submetidos para o componente de análise de conteúdos. O componente referido é responsável por extrair ricos meta-dados semânticos do conteúdo fornecido, identificando categorias de documentos e *tags* sociais que indicam interesses dos utilizadores. Esta acção é despoletada automaticamente, cada vez que existe uma mudança no perfil de um utilizador, através da captura do respectivo evento de alteração do Modelo Utilizador associado.

<sup>2</sup><http://book.cakephp.org/view/1000/Models>

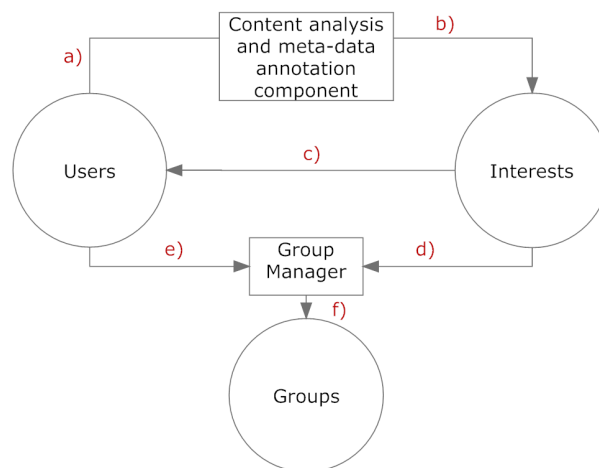


Figura 4.5: Processo de identificação de grupos implícitos de interesses e entidades envolvidas

Em b) após o processo de extração, cada nova *tag* identificada origina uma entidade Interest única. Durante o passo c), o conjunto de todos os interesses identificados é automaticamente atribuído ao utilizador considerado (modo automático), estendendo o seu perfil básico inicial, ou é oferecido como uma lista de sugestões, a partir da qual o utilizador pode seleccionar os interesses que considera mais relevantes (modo assistido).

Os seguintes passos d) e e) ilustram o conhecimento que os controladores possuem sobre os vários utilizadores e respectivos interesses, através das relações existentes entre os modelos que controlam. Com este conhecimento de todas as entidades disponíveis, no passo f), a ferramenta é capaz de criar grupos com qualquer número de interesses relacionados, atribuindo os utilizadores a esses mesmos grupos, de acordo com os seus perfis estendidos actualizados.

Considere-se a lista de potenciais interesses de um determinado utilizador, contida no seu perfil estendido e representada pela variável de nome *Suggestions*. Os seguintes blocos de pseudocódigo ilustram os procedimentos realizados:

Listagem 4.1: Pseudocódigo: Atribuição de interesses automática

```

//UsersController method.
foreach Suggestions as Interest{
  if(InterestRelevanceIsAcceptable)
    select(Interest); //Action.
}
  
```

Cada interesse é formado por um par (nome, relevância). Por omissão, a relevância associada a cada interesse é a relevância relativa do interesse para o assunto geral do documento analisado, fornecida pelo *webservice* OpenCalais (componente de análise de conteúdo). O processo automático pode ser configurado através da definição da relevância mínima permitida (valor entre 0 e 2 com uma casa decimal), usada para filtrar

os interesses que são atribuídos a cada utilizador, no componente de configuração da ferramenta.

A especificidade dos interesses, pode também ser controlada através da escolha dos diferentes tipos de interesses que podem ser identificados. Os tipos possíveis reflectem as funcionalidades do componente de análise de conteúdo, baseado no *webservice* OpenCalais:

- Entidades : este tipo está associado à identificação de locais, endereços, organizações e pessoas entre muitas outras entidades. A lista completa de entidades, passíveis de identificação, pode ser consultada na documentação do OpenCalais [Reu11].
- *Tags* sociais: identifica termos/conceitos mapeados nas bases de conhecimento do OpenCalais e da Wikipédia.
- Categorias: identifica categorias de documentos cobertas pela taxonomia de categorização do OpenCalais.

Listagem 4.2: Pseudocódigo: Selecção de um interesse

```
//InterestsController select action.  
if (exists(Interest)) {  
    InterestGroup = getGroupForThis(Interest);  
    addThisUserTo(InterestGroup);  
    deleteInterestFrom(Suggestions);  
}  
else {  
    create(Interest);  
    create(InterestGroup);  
    addThisUserTo(InterestGroup);  
    deleteInterestFrom(Suggestions);  
}
```

A acção de selecção de um interesse pode, adicionalmente, ser executada por iniciativa do próprio utilizador ou aplicação cliente, caso seja seleccionada a opção de atribuição de interesses assistida. Nesse caso, a selecção de um interesse é conseguida através da utilização da API disponibilizada (consultar 4.3.3). Os diferentes modos de atribuição de interesses e tipos de interesses identificáveis, podem ser definidos através da interface de administração da ferramenta.

### 4.2.3 Identificação de Grupos baseados em Características

Os grupos de utilizadores baseados em características são identificados através de uma simples comparação dos valores contidos em determinados campos dos perfis de cada utilizador. Por omissão, a ferramenta analisa dois campos distintos, relacionados com

a data de nascimento e a ocupação de um utilizador. No entanto, qualquer campo cujo domínio de valores possíveis seja conhecido, pode ser utilizado na identificação e criação deste tipo de grupos.

A escolha dos dois campos oferecidos por omissão, está relacionada com duas observações. A primeira, a observação de que ambos os campos acima mencionados são bastante comuns em aplicações interactivas e respectivos perfis de utilizadores. A segunda está relacionada com a importância que esta informação, sobre as características de cada utilizador, assume quando conjugada com os grupos de interesses de cada utilizador. A combinação destes dois tipos de grupos oferece vantagens para as aplicações cliente, permitindo restringir o acesso a conteúdos e funcionalidades, utilizar *marketing* direccionado, identificar possíveis correlações entre as faixas etárias e ocupações profissionais dos utilizadores e os interesses que possuem, ou mesmo, desenvolver sistemas de recomendações.

Após a definição dos campos e valores possíveis, utilizados na formação de grupos de características, a ferramenta cria um grupo para cada valor possível de cada característica. A seguinte função demonstra a criação de grupos de idades:

Listagem 4.3: Função createAgeGroups: Criação de grupos de características

```
var $agegroups =
    array( '01-18' , '18-24' , '25-35' , '36-50' , '51-69' , '70-?' );

//GroupsController Method.
function createAgeGroups() {
    if ($this->enableAutoGroups) {
        foreach ($this->agegroups as $agegroup) {
            $this->Group->createGroup($agegroup, NULL);
        }
    }
}
```

Depois de criados os grupos, cada utilizador é automaticamente atribuído ao grupo correspondente ao valor que possui no campo associado. Qualquer mudança efectuada pelo utilizador no seu perfil, por exemplo, alteração da sua profissão, desencadeia um evento que provoca uma nova atribuição do utilizador ao grupo correcto. No exemplo utilizado, a adição e remoção de utilizadores a grupos de idades, é realizada através dos métodos *joinAgeGroups(..)* e *leaveAgeGroups(..)*, presentes no controlador de utilizadores.

A figura 4.6, ilustra a criação de grupos de características em TIG:

#### 4.2.4 Identificação de Grupos baseados em Interações

Cada utilizador registado na ferramenta possui um grupo de utilizadores com os quais interage, através da troca de mensagens. Estas mensagens podem ser trocadas através

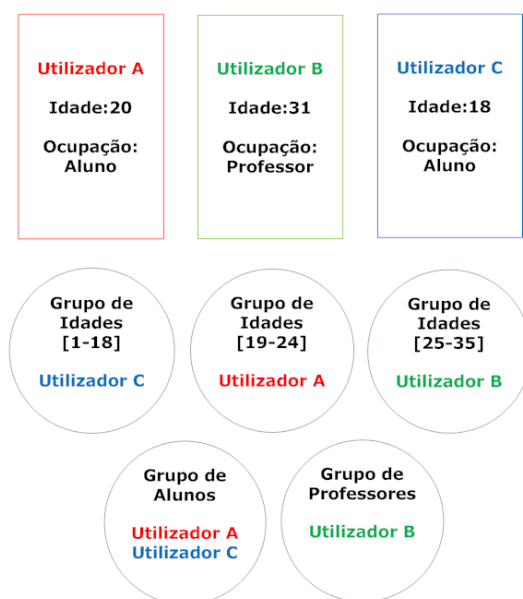


Figura 4.6: Grupos de características formados através dos perfis dos utilizadores A, B e C.

do sistema de mensagens privadas, fornecido pela ferramenta e implementado no controlador de mensagens, ou qualquer outro sistema externo de mensagens associado às aplicações cliente.

Quando um utilizador envia/recebe uma mensagem de outro utilizador, são actualizadas no seu Modelo Stat, as propriedades direcção, frequência e frescura, resultantes da mensagem trocada com esse utilizador. O Modelo Stat consiste assim numa estrutura de dados que mantém para cada utilizador, com quem o utilizador considerado interage, as seguintes propriedades:

- **frequência:** indica o número de interacções efectuadas (número de mensagens trocadas) com determinado utilizador;
- **direcção das mensagens trocadas:** dada uma certa frequência, contabiliza as interacções feitas em cada um dos diferentes sentidos (Mensagem enviada/Mensagem recebida);
- **frescura:** indica há quanto tempo ocorreu a última interacção entre dois utilizadores.

A cada uma destas propriedades pode ser atribuído um qualquer peso pertencente ao conjunto dos números reais. Os administradores e aplicações cliente podem alterar estes pesos através da respectiva interface de administração. Cada uma destas propriedades contribui para o estabelecimento de um *ranking* (valor numérico) que indica a importância de outro utilizador para o utilizador em questão. Este *ranking* é calculado para um utilizador A que interage com um utilizador B, por omissão, da seguinte forma:



$$ranking = frequencyAB \times w_1 + frequencyBA \times w_2 + freshness \times w_3$$

Na fórmula anterior, *frequencyAB* e *frequencyBA* representam respectivamente, o número de mensagens enviadas pelo utilizador A ao utilizador B e o número de mensagens enviadas pelo utilizador B ao utilizador A. A variável *freshness* representa o quociente da divisão de 1 pelo número de dias (calculado com duas casas decimais) passados desde a última interacção entre os utilizadores considerados e  $w_1$ ,  $w_2$  e  $w_3$ , os pesos referidos anteriormente.

A utilização de uma fórmula diferente, no cálculo da relevância das interacções, pode ser obtida através da modificação das funções implementadas no Modelo Stat (*stat.php*).

As aplicações cliente podem informar a ferramenta sobre a ocorrência de uma interacção, entre dois utilizadores distintos, através da primitiva *updateStats(..)* ou da utilização do sistema de mensagens oferecido, o qual procede automaticamente à recolha das propriedades pretendidas.

Listagem 4.4: Pseudocódigo: Actualização de propriedades de interacção

```
//MessagesController method.
function sendPM($to_id){
    //...send message logic...
    // $msg : array of message attributes , including sender and
    // recipient ids and date.
    if(messageSent){
        $this->Message->Sender->Stat->updateStats($msg);
        $this->Message->Recipient->Stat->updateStats($msg);
    }
}
```

#### 4.2.5 Gestão Dinâmica de Grupos

A gestão dos grupos de utilizadores gerados é realizada dinamicamente, através da captura de eventos de mudança dos vários modelos existentes na ferramenta. A utilização do padrão arquitectural MVC garante que o resultado de cada pedido, efectuado por um utilizador ou aplicação cliente, se encontra devidamente actualizado, por via da sua arquitectura triangular e das relações existentes entre os diferentes modelos existentes na ferramenta.

Devido à utilização do referido padrão de desenho, torna-se desnecessária a existência de um componente próprio para gerir assincronamente os grupos identificados, estando a lógica de gestão de grupos contida nos vários métodos e acções responsáveis por definir os grupos e seus membros.

As seguintes opções, relativas à gestão dos grupos existentes e seus membros, podem ser configuradas através da interface de administração da ferramenta:

- **manutenção de membros:** ao activar-se esta opção, pode-se simultaneamente definir um limite mínimo de interações (consulta/publicação de mensagens/conteúdos) que um utilizador deve efectuar perante um grupo, num determinado período de tempo, para continuar a ser considerado membro do mesmo; esta opção tem particular importância, quando é utilizado o processo automático de identificação de grupos de interesses, uma vez que alguns dos interesses e grupos atribuídos a cada utilizador, podem não ser úteis para os mesmos;
- **limite mínimo de membros:** esta opção indica o número mínimo de utilizadores necessários para um grupo ser considerado activo; um grupo activo permite a consulta e a publicação de mensagens e conteúdos; um grupo inactivo representa apenas um conjunto de utilizadores relacionados entre si.

Além das opções anteriores são ainda fornecidas primitivas que permitem aos administradores e aplicações cliente da ferramenta, criar e destruir manualmente grupos de qualquer tipo, assim como definir os membros de cada um desses grupos. Além de ajustes aos grupos gerados e respectivos membros, as primitivas anteriores permitem também a coexistência de grupos explícitos com os grupos implícitos identificados pela ferramenta. A lista de primitivas disponíveis é apresentada numa secção seguinte (4.3.3).

## 4.3 Interfaces Disponibilizadas

As funcionalidades da ferramenta consistem num conjunto de métodos implementados ao nível dos controladores. Este conjunto de métodos forma o núcleo de funções que permite a utilização e manipulação da ferramenta (Core API). Sobre este núcleo de funções foram implementadas, ao nível das vistas, dois tipos distintos de interfaces, como ilustra a figura 4.7.

A primeira interface, baseada em HTML, pode ser utilizada por utilizadores e administradores, para aceder directamente à ferramenta via *web browser*, de modo a testar e configurar as suas funcionalidades (ver 4.3.2). A segunda interface, consiste numa API estilo REST que pode ser utilizada pelas aplicações cliente que pretendam integrar a ferramenta TIG e as funcionalidades oferecidas (ver 4.3.3).

Em ambos os casos anteriores, as primitivas oferecidas baseiam-se no conjunto de funções que constitui o núcleo da ferramenta. A lista de funções que forma a Core API é apresentada na secção seguinte.

### 4.3.1 Core API

As tabelas 4.1, 4.2, 4.3,4.4 e 4.5 apresentam a lista de funções essenciais da ferramenta desenvolvida e uma breve descrição das suas funcionalidades.



Figura 4.7: Diferentes interfaces da ferramenta TIG

Controlador	Primitiva	Descrição
Users	home(\$id)	Devolve toda a informação associada a um utilizador. A informação devolvida inclui os grupos, interesses, mensagens e <i>posts</i> associados ao utilizador autenticado.
Users	import_bookmarks(\$id)	Permite adicionar um url favorito ao perfil estendido de um utilizador.
Users	addUserToGroup(\$userID, \$groupID)	Permite adicionar um utilizador a um determinado grupo.
Users	delUserFromGroup(\$userID,\$groupID)	Permite remover um utilizador de um determinado grupo.
Users	add(\$id,\$userdata)	Adiciona um novo utilizador à ferramenta.
Users	delete(\$id)	Remove um utilizador e toda a informação associada ao mesmo.
Users	edit(\$id,\$userdata)	Permite a edição do perfil básico de um utilizador.
Users	index()	Devolve todos os utilizadores existentes na ferramenta e respectiva informação.
Users	clearSuggestions(\$id)	Apaga a lista de interesses sugeridos a um utilizador.
Users	admin_settings(\$options)	Permite definir as opções de configuração da ferramenta.

Tabela 4.1: Funções da Core API implementadas no *UsersController*.

Controlador	Primitiva	Descrição
Groups	add(\$id,\$groupdata)	Adiciona um novo grupo de qualquer tipo à ferramenta.
Groups	delete(\$id)	Remove um grupo e toda a informação associada ao mesmo.
Groups	edit(\$id,\$groupdata)	Permite a edição de um grupo.
Groups	view(\$id)	Devolve a informação associada a um determinado grupo.
Groups	index()	Devolve todos os grupos existentes na ferramenta e respectiva informação.

Tabela 4.2: Funções da Core API implementadas no *GroupsController*.

Controlador	Primitiva	Descrição
Interests	add(\$id,\$interestdata)	Adiciona um novo interesse de qualquer tipo à ferramenta.
Interests	delete(\$id)	Remove um interesse e toda a informação associada ao mesmo.
Interests	edit(\$id,\$interestdata)	Permite a edição de um interesse.
Interests	view(\$id)	Devolve a informação associada a um determinado interesse.
Interests	index()	Devolve todos os interesses existentes na ferramenta e respectiva informação.
Interests	select(\$name,\$userID)	Atribui a um utilizador, o interesse com nome \$name, juntando-o automaticamente ao grupo(s) associado(s) ao interesse. Se o referido interesse especificado ainda não existir, tanto o interesse como o respectivo grupo são criados automaticamente.

Tabela 4.3: Funções da Core API implementadas no *InterestsController*.

Controlador	Primitiva	Descrição
Messages	sendPM(\$from_id,\$to_id,\$toGroup,\$msgdata)	Permite enviar uma mensagem para um determinado utilizador ou grupo.
Messages	inbox(\$userID)	Devolve a lista de mensagens recebidas por um utilizador.
Messages	outbox(\$userID)	Devolve a lista de mensagens enviadas por um utilizador.
Messages	view(\$id)	Devolve uma mensagem.
Messages	delete(\$id,\$userID)	Marca para remoção uma mensagem associada a um utilizador. Uma mensagem é efectivamente apagada da base de dados da ferramenta, quando ambos, receptor e fonte, marcam a mensagem para remoção.
Messages	edit(\$id,\$msgdata)	Permite a edição do conteúdo de uma mensagem.

Tabela 4.4: Funções da Core API implementadas no *MessagesController*.

Controlador	Primitiva	Descrição
Posts	delete(\$id)	Elimina um <i>post</i> .
Posts	edit(\$id,\$postdata)	Permite a edição do conteúdo de um <i>post</i> .
Posts	view(\$id)	Devolve o conteúdo de um <i>post</i> .
Posts	addToGroups(\$postdata,\$filters)	Cria e associa automaticamente um <i>post</i> a grupos de interesse de acordo com os filtros definidos pelo criador.
Stats	updateStats(\$from_id,\$to_id)	Informa a ferramenta sobre a ocorrência de uma interacção entre dois utilizadores (interacção externa à ferramenta).

Tabela 4.5: Funções da Core API implementadas no *PostsController* e no *StatsController*.

### 4.3.2 Interface do Utilizador

Esta interface da ferramenta é composta por três componentes principais : espaço de utilizador (a), espaço de grupo (b) e o espaço de administração. A figura 4.8 ilustra e relaciona os dois primeiros componentes. A figura 4.9 ilustra as opções de configuração disponíveis no terceiro componente, o espaço de administração.

#### 4.3.2.1 Espaço de Utilizador

No espaço de utilizador (a), cada utilizador pode executar três acções importantes, que permitem à ferramenta gerar o seu perfil estendido.

The screenshot displays the 'Implicit Groups' web application interface, divided into two main sections: (a) User space and (b) Group space.

**Section (a) - User space:** This section shows the user profile for 'user03'. It includes a navigation menu with 'Home', 'User Guide', and 'Admin'. On the left, there are buttons for 'Edit', 'Import Bookmark', 'Messages', and 'Logout'. The profile details include:
 

- Profile:** user03
- Username:** user03
- Email:** user03@hotmail.com
- Description:** I love sports mainly soccer and tennis. I am a student of computer science at FCT - UNL and my favourite subjects are database and web development related (PHP, CSS, Ajax).
- Occupation:** Student
- Birthdate:** 1982-06-26

 Below the profile, there are sections for 'My Interests' (listing Software engineering, Computer programming, FCT, Sports, Technology\_Internet) and 'My Groups' (listing Computer programming group, FCT group, Sports group, Software engineering group, PHP group, Technology\_Internet group). A 'Suggested Interests' section lists Computing, Software design patterns, and PHP programming. A red box highlights the 'Import Bookmark' button and the 'Import Bookmark' form, which includes a text area for a URL and an 'Import' button. A red box also highlights the 'Group space access' link.

**Section (b) - Group space:** This section shows the details for a group with 'Group Id 44'. It includes:
 

- Name:** PHP group
- Description:** description PHP group

 Below this, there are sections for 'Group Messages', 'Related Users', and 'Related Interests'. The 'Related Users' table lists:
 

Username	Email
user03	user03@hotmail.com
user04	user04@someemail.com

 The 'Related Interests' section lists 'PHP'.


Figura 4.8: Vista geral da GUI: Espaço de Utilizador (a), Espaço de Grupo (b)

A primeira acção, Edit, permite ao utilizador definir explicitamente o seu perfil básico, preenchendo os campos apropriados no formulário disponível, incluindo a sua data de nascimento, ocupação e um campo de texto com a sua descrição. Após a referida acção estar completa, a ferramenta é capaz de atribuir, automaticamente, o utilizador aos grupos de características, com base nos valores dos campos de ocupação e data de nascimento que preencheu. O campo de descrição de utilizador, também preenchido, é utilizado para inferir interesses do utilizador e contribui para o perfil estendido.

A segunda acção disponível, Import Bookmark, permite ao utilizador fornecer à ferramenta, endereços para os seus artigos favoritos ou mais relevantes, notícias ou qualquer outro conteúdo *online* baseado em texto. Tais conteúdos são também usados pelo componente de análise de conteúdos na aferição de interesses dos utilizadores, estendendo os seus perfis básicos.

A acção final no espaço de utilizador é a acção Messages. Como o nome indica, esta acção está relacionada com o sistema de mensagens privadas incluído na ferramenta e fornece ao utilizador a capacidade de enviar mensagens privadas a outros utilizadores ou grupos existentes na ferramenta. As propriedades de interacção, resultantes das mensagens enviadas através deste sistema, são também usadas na extensão do perfil de utilizador e na identificação/formação dos grupos baseados em interacções, como referido na secção 4.2.4.

Adicionalmente ao menu de acções, o espaço de utilizador (a), fornece a cada utilizador uma vista geral do estado actual do seu perfil, interesses sugeridos, associados e todos os grupos a que esse utilizador pertence. A selecção de qualquer um dos grupos a que pertence redirecciona o utilizador para o espaço de grupo (b) correspondente.



**TIG Settings**

**Profile Settings**

Occupations definition : please insert in the form : Occupation1,Occupation2,OccupationN

Teacher,Student,Other

Age Groups definition : please insert in the form : 01-18,18-24,25-35,36-50,51-69,70-?

01-18,18-24,25-35,36-50,5

**Group Settings**

Interest Discovery Mode

Social

Interest-Based Group Assignment Mode

Assisted

Content-Analysis Component

Opencalais

**Interaction Settings**

Interaction-Based Metric Weights : please specify weight values respectively to frequency, freshness and direction in the form : 1.00,0.85,0.50

0.50,1.00,0.50

Interaction-Based Metric Enable Direction

on

Figura 4.9: Espaço de Administração : exemplo de configurações possíveis

#### 4.3.2.2 Espaço de Grupo

No espaço de grupo (b), os utilizadores podem aceder a toda a informação relacionada com um grupo em particular, incluindo mensagens, conteúdo partilhado e outras informações sobre os membros do grupo. As mensagens de grupo podem ser enviadas, por qualquer membro do grupo, para o espaço de grupo correspondente e consultadas por todos os membros desse grupo. Cada mensagem de grupo pode conter HTML ou texto simples.

#### 4.3.2.3 Espaço de Administração

O último componente da interface de utilizador da ferramenta é o espaço de administração. Este espaço está reservado para os utilizadores com papel de administração e pode ser utilizado com dois propósitos principais. O primeiro está relacionado com a configuração do comportamento da ferramenta, enquanto o segundo se relaciona com a consulta e edição de todas as informações relacionadas com os utilizadores e grupos existentes.

Entre as várias opções possíveis, um administrador pode definir uma lista de possíveis ocupações, as faixas etárias utilizadas para formar grupos de utilizadores, o modo

de atribuição de utilizadores a grupos (Automático ou Assistido) e os pesos atribuídos a cada propriedade de interacção (direcção, frequência e frescura). A figura 4.9 ilustra parte desta interface de configuração. As referidas opções pretendem aumentar a flexibilidade da ferramenta, na sua integração em ambientes, cenários e aplicações distintos.

Toda a informação relacionada com os utilizadores registados, grupos disponíveis e afiliação de utilizadores, pode ser acedida através do espaço de administração. Os administradores podem também usar este espaço, para definir explicitamente/manualmente grupos, interesses e atribuir utilizadores a grupos existentes. Estes comandos explícitos podem ser utilizados para complementar ou substituir os processos automáticos fornecidos pela ferramenta, caso seja necessário.

### 4.3.3 API REST

A ferramenta TIG oferece uma API estilo REST, constituída por um conjunto de primitivas. Esta API permite que outras aplicações baseadas na web, utilizem e integrem as funcionalidades da ferramenta.

A autenticação dos clientes da API, baseia-se em autenticação básica por HTTP. Cada cliente deve autenticar-se através de um cabeçalho HTTP, com o seguinte formato:

Listagem 4.5: Cabeçalho HTTP de autorização

```
Authorization: TRUEREST username=john&password=xxx&apikey=247b5a2f7s
```

A autenticação básica pressupõe que os nomes de utilizador e palavras passe são passados em limpo no referido cabeçalho. Caso seja necessária segurança adicional, como a utilização de conexões seguras usando SSL, os administradores da ferramenta podem instalar facilmente essa funcionalidade, através da utilização do componente de segurança fornecido pela plataforma de desenvolvimento CakePHP:

Listagem 4.6: Utilizar conexões SSL

```
var $components = array('Security');  
function beforeFilter() {  
    $this->Security->requireSecure();  
}
```

O exemplo anterior mostra como as conexões seguras SSL podem ser activadas em qualquer controlador, carregando-se o componente de segurança e adicionando-se a chamada ao método `requireSecure()` no método `beforeFilter()` do controlador pretendido. Adicionando-se a referida chamada ao método `beforeFilter()` do controlador `app_controller`, todos os métodos de todos os controladores passam a requerer conexões SSL.

As respostas a pedidos efectuados através da API, consistem num conjunto de dados no formato XML ou opcionalmente, no formato JSON. O formato de dados retornado é

seleccionado através da especificação da extensão correspondente, no pedido dirigido ao endereço do *web service*:

Listagem 4.7: Exemplo de pedido à API

```
//Web service base url example:
http://example.tig.api.domain.com/implicit/

//Generic API call format:
webServiceBaseUrl/controller_name/api_method.reply_extension?arg1=value1
or
webServiceBaseUrl/controller_name/arg1/arg2.reply_extension
or
webServiceBaseUrl/controller_name/api_method.reply_extension

//Example API call to import_bookmark method:
webServiceBaseUrl/users/import_bookmarks.xml?bookmark=http://someurl.com
```

Cada primitiva oferecida pela API REST está directamente associada a uma função da Core API. Os pedidos HTTP efectuados podem utilizar os métodos GET ou POST, consoante a primitiva pretendida. A chamada da primitiva `import_bookmark`, no exemplo anterior, reflecte o uso de HTTP GET. Na construção do pedido HTTP, todos os argumentos devem estar no formato URLencode.

A lista das primitivas oferecidas por cada controlador, assim como a descrição das respectivas funcionalidades, é apresentada nas seguintes subsecções. Informações e exemplos práticos adicionais, sobre a utilização da API REST, podem ser encontrados nas secções 5.3 e 5.4 do capítulo 5, referente à utilização e integração da ferramenta.

#### 4.3.3.1 Utilizadores

As seguintes primitivas oferecidas pela API REST, baseiam-se nas funções da Core API com o mesmo nome, implementadas no *UsersController*.

Listagem 4.8: addUserToGroup

```
- HTTP method: GET
- Arguments:
  (int) userID (mandatory)
  (int) groupID (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/users/addUserToGroup/2/43.xml
```

Adiciona o utilizador identificado por `$userID`, ao grupo com identificador `$groupID`. Permite o controlo explícito do processo de atribuição de utilizadores, a qualquer tipo de grupo.



## Listagem 4.9: delUserFromGroup

```

- HTTP method: GET
- Arguments:
  (int) userID (mandatory)
  (int) interestID (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/users/delUserFromGroup/2/43.xml

```

Remove o utilizador identificado por \$userID do grupo associado ao interesse identificado por \$interestID.

## Listagem 4.10: edit

```

- HTTP method: POST
- Arguments:
  (int) id (mandatory)
  (serialized array) userdata (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/users/edit.xml

```

Permite a alteração dos dados contidos no perfil de um utilizador, com identificador \$id.

## Listagem 4.11: add

```

- HTTP method: POST
- Arguments:
  (serialized array) userdata (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/users/add.xml

```

Adiciona à ferramenta um novo utilizador e devolve o identificador do utilizador criado.

## Listagem 4.12: delete

```

- HTTP method: GET
- Arguments:
  (int) id (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/users/delete.xml?id=2

```

Remove da ferramenta o utilizador com identificador \$id.

## Listagem 4.13: import\_bookmark

```
- HTTP method: GET
- Arguments:
  (string) bookmark (mandatory)
- Example call:
  webServiceBaseUrl/users/import_bookmark.xml?bookmark=http://someurl.com
```

Permite importar um endereço (url) favorito para o perfil estendido de um utilizador e consequentemente, a detecção dos interesses que classificam o documento associado ao endereço fornecido. A resposta a este pedido devolve os interesses identificados para o documento fornecido.

#### Listagem 4.14: home

```
- HTTP method: GET
- Arguments: none
- Example call:
  webServiceBaseUrl/users/home.xml
```

Devolve toda a informação associada ao utilizador autenticado no cabeçalho HTTP. A informação devolvida inclui grupos, interesses, mensagens e os *posts* associados ao utilizador autenticado.

#### Listagem 4.15: index

```
- HTTP method: GET
- Arguments: none
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/users.xml
```

Devolve todos os utilizadores da ferramenta.

#### Listagem 4.16: clearSuggestions

```
- HTTP method: GET
- Arguments: none
- Example call:
  webServiceBaseUrl/users/clearSuggestions.xml
```

Elimina a lista de interesses sugeridos ao utilizador autenticado.

#### Listagem 4.17: settings

```
- HTTP method: POST
- Arguments:
  (serialized array) options (mandatory)
- Admin privileges required for authenticated user
```

```

- Example call:
  webServiceBaseUrl/admin/users/settings.xml

```

Permite a um utilizador com privilégios de administrador definir as opções de configuração da ferramenta.

#### 4.3.3.2 Interesses

As seguintes primitivas oferecidas pela API REST, baseiam-se nas funções da Core API com o mesmo nome, implementadas no *InterestsController*.

##### Listagem 4.18: view

```

- HTTP method: GET
- Arguments:
  (int) interest_id (mandatory)
- Example call:
  webServiceBaseUrl/interests/58.xml

```

Devolve o interesse identificado por \$interest\_id e marca a mensagem como lida/visualizada.

##### Listagem 4.19: add

```

- HTTP method: GET
- Arguments:
  (string) interest_name (mandatory)
  all empty spaces must be replaced with underscores
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/interests/add.xml?interest_name

```

Adiciona um novo interesse e cria o respectivo grupo de interesse. A resposta a este pedido inclui os identificadores do novo interesse e grupo criados.

##### Listagem 4.20: delete

```

- HTTP method: GET
- Arguments:
  (int) interestID (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/interests/delete/20.xml

```

Remove o interesse com identificador \$id e o respectivo grupo associado. A resposta inclui um indicador do sucesso da operação de remoção.

## Listagem 4.21: select

```
- HTTP method: GET
- Arguments:
  (string) interest_name (mandatory)
  all empty spaces must be replaced with underscores
  (int) relevance (optional) (relevance integer values: low=1 or high=2)
- Example call:
  webServiceBaseUrl/interests/select/Soccer/2
```

De acordo com o utilizador autenticado no pedido, atribui ao mesmo o interesse com nome \$name, juntando-o automaticamente ao grupo com o referido interesse. Se o referido interesse especificado ainda não existir, tanto o interesse como o respectivo grupo são criados.

## Listagem 4.22: index

```
- HTTP method: GET
- Arguments: none
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/interests.xml
```

A resposta a este pedido contém toda a informação associada aos interesses existentes na ferramenta, incluindo respectivos grupos e utilizadores relacionados.

### 4.3.3.3 Grupos

As seguintes primitivas oferecidas pela API REST, baseiam-se nas funções da Core API com o mesmo nome, implementadas no *GroupsController*.

## Listagem 4.23: add

```
- HTTP method: POST
- Arguments:
  (serialized array) groupdata (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/groups/add.xml
```

Adiciona um novo grupo com os atributos especificados no pedido. Pode ser utilizado para criar qualquer tipo de grupo, excepto grupos baseados em interacções.

## Listagem 4.24: delete

```
- HTTP method: GET
- Arguments:
  (int) group_id (mandatory)
```

```

- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/groups/delete/10.xml

```

Remove o grupo identificado por \$id. Retorna o estado da operação.

#### Listagem 4.25: index

```

- HTTP method: GET
- Arguments: none
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/groups.xml

```

A resposta a este pedido contém toda a informação associada aos grupos existentes na ferramenta, incluindo a lista de membros, *posts*, mensagens e interesses relacionados com cada grupo.

#### Listagem 4.26: view

```

- HTTP method: GET
- Arguments:
  (int) group_id (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/groups/58.xml

```

A resposta a este pedido contém toda a informação associada ao grupo identificado por \$group\_id.

#### 4.3.3.4 Mensagens

As seguintes primitivas oferecidas pela API REST, baseiam-se nas funções da Core API com o mesmo nome, implementadas no *MessagesController*.

#### Listagem 4.27: sendPM

```

- HTTP method: POST
- Arguments:
  (serialized array) message_attributes (mandatory)
- Example call:
  webServiceBaseUrl/messages/sendPM.xml

```

Permite o envio de uma mensagem privada para um utilizador ou grupo.

#### Listagem 4.28: inbox

```
- HTTP method: GET
- Arguments: none
- Example call:
  webserviceBaseUrl/messages/inbox.xml
```

Retorna a lista de mensagens recebidas pelo utilizador autenticado.

#### Listagem 4.29: outbox

```
- HTTP method: GET
- Arguments: none
- Example call:
  webserviceBaseUrl/messages/outbox.xml
```

Retorna a lista de mensagens enviadas pelo utilizador autenticado no pedido.

#### Listagem 4.30: view

```
- HTTP method: GET
- Arguments:
  (int) message_id (mandatory)
- User can only view his related messages
- Admin users can view all messages
- Example call:
  webserviceBaseUrl/messages/12.xml
```

Devolve a mensagem identificada por \$message\_id e marca a mensagem como lida/visualizada.

#### Listagem 4.31: delete

```
- HTTP method: GET
- Arguments:
  (int) message_id (mandatory)
- User can only mark for deletion his related messages
- Example call:
  webserviceBaseUrl/messages/delete/10.xml
```

De acordo com o utilizador considerado, marca para remoção a mensagem associada ao mesmo. Uma mensagem é efectivamente apagada da base de dados da ferramenta, quando ambos, receptor e fonte, marcam a mensagem para remoção.

#### Listagem 4.32: edit

```
- HTTP method: POST
- Arguments:
  (int) message_id (mandatory)
  (serialized array) msgdata (mandatory)
```

```

- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/messages/edit.xml

```

Permite a modificação do conteúdo da mensagem com o identificador especificado.

#### 4.3.3.5 Posts

As seguintes primitivas oferecidas pela API REST, baseiam-se nas funções da Core API com o mesmo nome, implementadas no *PostsController*.

##### Listagem 4.33: addToGroups

```

- HTTP method: POST
- Arguments:
  (serialized array) postdata (mandatory)
- Example call:
  webServiceBaseUrl/posts/addToGroups.xml

```

Disseminação de um *post* para grupos de interesse. O utilizador ou aplicação responsável por criar o *post*, pode especificar um conjunto de filtros que definem quais os utilizadores e grupos que poderão aceder ao mesmo.

##### Listagem 4.34: view

```

- HTTP method: GET
- Arguments:
  (int) post_id (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/posts/2.xml

```

Devolve o post identificado por \$id.

##### Listagem 4.35: edit

```

- HTTP method: GET
- Admin privileges required
- Arguments:
  (int) post_id (mandatory)
  (serialized array) postdata (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/posts/edit.xml

```

Permite a alteração do conteúdo do *post* identificado por \$id.

## Listagem 4.36: delete

```
- HTTP method: GET
- Admin privileges required
- Arguments:
  (int) post_id (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/posts/delete/10.xml
```

Remove o post identificado por \$id.

#### 4.3.3.6 Stats

A seguinte primitiva oferecida pela API REST, baseia-se na funções da Core API com o mesmo nome, implementada no *StatsController*.

## Listagem 4.37: updateStats

```
- HTTP method: GET
- Arguments:
  (int) source_id (mandatory)
  (int) destination_id (mandatory)
  (string) destination_username (mandatory)
- Admin privileges required for authenticated user
- Example call:
  webServiceBaseUrl/stats/updateStats/10/11/user11.xml
```

Permite a indicação da ocorrência de uma interação entre dois utilizadores.

## 4.4 Conclusão

Ao longo deste capítulo, descreveu-se a arquitectura utilizada na implementação da ferramenta de suporte a grupos implícitos (TIG), conceptualizada no capítulo 3.

O padrão arquitectural definido, facilitou a obtenção de uma ferramenta modular e extensível, separando a aplicação em diferentes camadas (dados, lógica e apresentação) com componentes independentes e facilmente substituíveis.

Os mecanismos desenvolvidos e implementados, associados a diferentes processos de identificação, geração e gestão de grupos implícitos, pretendem facilitar a integração do conceito de grupos implícitos em diferentes aplicações, permitindo várias opções de configuração. Estas configurações permitidas visam também aumentar a flexibilidade e utilidade da ferramenta implementada, alargando a sua adaptabilidade a diferentes cenários aplicativos.

A integração da ferramenta numa plataforma colaborativa específica é concretizada



no próximo capítulo, provando a sua capacidade de suportar o desenvolvimento de novas funcionalidades, baseadas no conceito de grupos implícitos e nos mecanismos oferecidos.

Uma versão *online* do protótipo operacional desenvolvido, encontra-se disponível no endereço "[www.seawoman.eu/cpm/implicit/cakephp-cakephp-f6748d4/implicit/](http://www.seawoman.eu/cpm/implicit/cakephp-cakephp-f6748d4/implicit/)".





# Utilização e Integração da Ferramenta

Ao longo deste capítulo são descritas possíveis aplicações e utilizações para a ferramenta desenvolvida (TIG). Um cenário aplicacional é apresentado e concretizado, através da integração da ferramenta na plataforma Moodle. No fim do capítulo são ainda discutidas possíveis formas de complementar as funcionalidades oferecidas pela ferramenta TIG, com a implementação de métricas que visem aumentar a utilidade dos grupos identificados e gerados.

## 5.1 Cenários de Aplicação Gerais

O conceito de grupos implícitos, os mecanismos utilizados na concepção e as funcionalidades oferecidas, projectam várias utilizações possíveis para a ferramenta TIG. Os grupos implícitos identificados formam uma base de informação, passível de ser utilizada na realização de diversas funcionalidades e análises. Os seguintes itens, descrevem algumas das utilizações possíveis, para além da óbvia formação de grupos de utilizadores:

- **Partilha e Disseminação de Informação:** a organização de utilizadores em diversos grupos baseados nos seus interesses, características e interações fornece um meio adequado à partilha e disseminação de informação. Sistemas e aplicações podem optar por utilizar os grupos identificados dinamicamente para disseminar conteúdos relacionados com diversos interesses, grupos de idades e profissões, reduzindo o esforço exigido aos utilizadores, na procura de informação relevante de acordo com os seus interesses e contextos. Os utilizadores, por outro lado, podem utilizar os grupos aos quais pertencem para identificar utilizadores previamente desconhecidos, com os quais podem partilhar informação sobre gostos, interesses,

características e contextos comuns;

- **Marketing dirigido:** os grupos identificados e gerados pela ferramenta TIG, agrupam os utilizadores com as mesmas características, interesses e os utilizadores que partilham interacções. Estes grupos podem ser utilizados em marketing e publicidade, permitindo a empresas e aplicações dirigir campanhas de forma eficiente, desde que conheçam *a priori*, os atributos do público alvo. A combinação de diferentes grupos e seus membros, permite identificar os utilizadores alvo (recipientes) que satisfazem expressões booleanas de características, interesses e relações;
- **Estudos associados a grupos de utilizadores, correlações e padrões existentes:** os grupos de interesses, características e interacções podem também ser utilizados na identificação de correlações e padrões de interesses, características e interacções, em comunidades de utilizadores com um número significativo de membros. Uma vez que uma mesma instância da ferramenta pode ser partilhada por diferentes aplicações cliente, espera-se que a sua utilização contribua para a geração de grupos de dimensão significativa e útil para os referidos estudos e análises. Um exemplo deste tipo de análise é a descoberta de características associadas a interesses e preferências. Por exemplo, dado um conjunto de utilizadores interessados num determinado género musical, os grupos a que estes pertencem podem ser utilizados para descobrir características relacionadas com essa preferência musical, como as faixas etárias predominantes para o género considerado;
- **Estudo de tendências relacionadas com os interesses dos utilizadores, ao longo do tempo:** outros estudos que podem beneficiar dos grupos identificados são aqueles que identificam tendências de interesses e preferências ao longo do tempo. Uma vez que a dimensão temporal assume um papel importante na ferramenta TIG, espera-se que a evolução dos interesses dos utilizadores ao longo do tempo possa fornecer dados importantes sobre a emergência e regressão de tendências;
- **Sistemas de recomendações e filtragem colaborativa:** os dados fornecidos pela utilização de TIG, conjugados com métricas e algoritmos associados a técnicas de filtragem colaborativa, podem contribuir para a percepção da utilidade do conceito de grupos implícitos como componente no desenvolvimento de sistemas de recomendações, particularmente sistemas automatizados que não impliquem a interferência dos utilizadores (consultar secção 5.5.3.2). Além dos interesses dos utilizadores, espera-se que os dados contidos nos seus perfis estendidos, nomeadamente resultantes das interacções com outros utilizadores e da análise das suas características, ofereçam dados valiosos para a alimentação deste tipo de sistemas;
- **Funcionalidades de suporte e promoção da interacção entre utilizadores:** a combinação de mecanismos de interacção com os diferentes tipos de grupos identificados pela ferramenta TIG, pode ser útil no desenvolvimento de interfaces e funcionalidades “inteligentes” que sugiram, promovam e facilitem a interacção e aproximação

de utilizadores desconhecidos. Utilizadores desconhecidos que possuam interesses e características comuns, por exemplo, podem beneficiar da partilha de informação, ideias e recursos associados a essas relações implícitas. Um exemplo de uma funcionalidade que pode usufruir desta combinação é a criação de núcleos, associações e grupos de trabalho, relacionados com um determinado assunto;

- **Processos automáticos de análise e categorização de documentos / personalização da informação:** devido à aproximação utilizada (baseada na análise de conteúdos) na identificação dos interesses de cada utilizador e respectivos grupos de interesses, a ferramenta TIG pode também ser utilizada para organizar e classificar documentos textuais de acordo com o seu conteúdo. Principalmente no que diz respeito aos grupos implícitos de interesses, além da agregação de utilizadores com interesses comuns, os grupos mencionados podem ser utilizados como depósitos de recursos e informações relacionadas com esses interesses. Os membros de cada grupo de interesse podem beneficiar da personalização da informação, filtrada de acordo com as suas preferências. Saliente-se que a manutenção de grupos de interesses, criados e geridos automaticamente, pode contribuir para o desenvolvimento de interfaces com usabilidade acrescida, devido ao facto de possuírem informação sobre os indivíduos que as utilizam.

## 5.2 Um Cenário Aplicacional

Considere-se o contexto associado a um Campus universitário, partilhado por docentes, estudantes, investigadores e funcionários administrativos, assim como as diversas actividades relacionadas com os departamentos, cursos, núcleos de investigação e núcleos de interesses existentes.

No anterior contexto, a ferramenta desenvolvida, TIG, pode ser utilizada para identificar dinamicamente diversos grupos implícitos de indivíduos, consoante os interesses, características e interações que partilham. No cenário proposto, estes grupos podem ser utilizados para facilitar a divulgação de informação relevante para os seus membros, por exemplo, permitindo a partilha de artigos, notícias e ideias entre alunos e professores interessados num mesmo assunto.

A plataforma Moodle [Tru11] assenta em software de código aberto baseado na Internet, usado na gestão de cursos e respectivos conteúdos, é também reconhecida como uma plataforma que suporta o ensino e aprendizagem em ambientes virtuais (*e-learning*).

A escolha da plataforma Moodle, como um exemplo concreto da integração da ferramenta de grupos implícitos desenvolvida, foi determinada pela sua adequação ao cenário de aplicação considerado, pela abertura do seu código e pelo conjunto completo de bibliotecas e funcionalidades que assistem o processo de desenvolvimento. A documentação fornecida é bastante completa, oferecendo vários tutoriais e exemplos práticos, úteis na integração pretendida.

Na plataforma Moodle, os utilizadores podem assumir múltiplos papéis, sendo no entanto os mais relevantes, os papéis Estudante e Professor. No tradicional ambiente universitário os professores são responsáveis por gerir e administrar cursos e os estudantes estão inscritos em um ou mais cursos. Embora o Moodle ofereça várias funcionalidades e a capacidade de gerir cursos e os seus conteúdos, carece de outros aspectos ou componentes, relevantes para o dia-a-dia académico, tais como a gestão de actividades de investigação, muitas vezes partilhadas por professores e alunos.

A extensão do Moodle através da integração de TIG, permite criar rapidamente funcionalidades que auxiliam actividades e projectos de investigação, relacionados com diversos temas. Através da extensão implementada, qualquer utilizador pode partilhar automaticamente um artigo com todos os outros utilizadores interessados no sistema. Concretamente, foi desenvolvido um módulo de extensão da plataforma Moodle que permite aos seus utilizadores, acederem aos grupos identificados e utilizarem formulários de partilha de notícias e artigos, disponíveis na Internet.

Cada utilizador pode recorrer aos formulários definidos para importar os seus recursos favoritos ou recursos que considera relevantes. O módulo de extensão acede às funcionalidades da ferramenta TIG através da API REST disponibilizada, permitindo que os recursos submetidos sejam utilizados na extensão do perfil dos utilizadores que iniciam a partilha e na identificação de outros utilizadores interessados. Adicionalmente, os utilizadores podem especificar filtros que restringem o acesso aos recursos partilhados.

## 5.3 Integração TIG-Moodle

A integração da ferramenta desenvolvida no ambiente de execução de qualquer aplicação baseada na web, pode ser conseguida através da API estilo REST, apresentada anteriormente. No caso concreto da plataforma Moodle, desenvolveu-se um módulo de extensão da plataforma, no qual está incluído um cliente para a API.

O módulo desenvolvido utiliza as funcionalidades oferecidas pela ferramenta TIG, na implementação das funcionalidades pretendidas e introduzidas no cenário apresentado. As funcionalidades implementadas, resultam da combinação das primitivas TIG com outras funcionalidades oferecidas pelas bibliotecas internas do Moodle.

A figura 5.1 ilustra a integração da ferramenta TIG na plataforma Moodle :

### 5.3.1 Registo de Utilizadores

Os utilizadores de diferentes aplicações web podem registar-se directamente na ferramenta TIG através da Interface HTML disponível. O registo encontra-se disponível através da acção *register*, apresentada na página inicial da ferramenta.

Um utilizador/cliente registado possui um identificador único em TIG. Este identificador que está associado às credenciais definidas durante o registo, permite à ferramenta

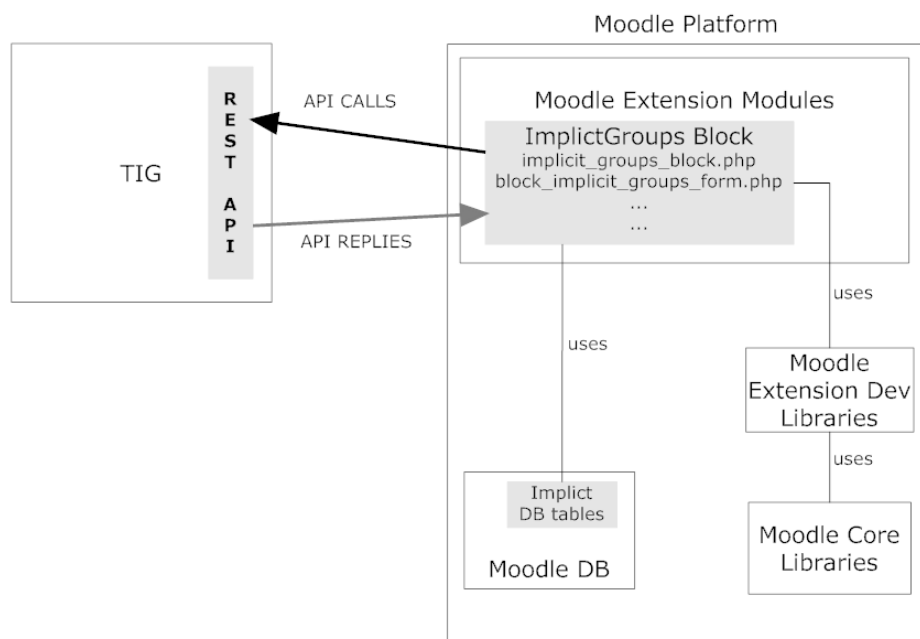


Figura 5.1: Ilustração da integração de TIG na plataforma Moodle.

identificar a origem dos pedidos efectuados à API REST, determinando quais os utilizadores afectados pelas operações desencadeadas.

A escolha deste método de registo directo, usando a interface HTML da ferramenta, possui várias vantagens para utilizadores e aplicações:

- manutenção do direito de escolha e da privacidade de cada utilizador;
- uma única conta de utilizador TIG (credenciais obtidas) permite a utilização das funcionalidades da API através de múltiplas e distintas aplicações;
- quando aplicações distintas partilham uma mesma instância da ferramenta TIG, evita a criação de múltiplas contas associadas a um mesmo utilizador que pertença a mais do que uma dessas aplicações.

Embora este método possua as referidas vantagens, exige que cada utilizador proceda individualmente ao registo. No caso de aplicações em que os direitos dos seus utilizadores estejam claramente definidos e contemplem a utilização consentida dos dados presentes nos seus perfis, este esforço adicional pedido aos utilizadores pode ser evitado das seguintes formas:

1. utilização da interface HTML de administração da ferramenta, na definição manual de contas de utilizadores;
2. utilização de uma conta TIG com privilégios de administrativos, para aceder a funcionalidades adicionais da API, como a primitiva que permite adicionar utilizadores a pedido;

The screenshot shows the Moodle Local administration interface. At the top, it says 'Moodle Local' and 'You are logged in as Admin User (Logout)'. Below that, there's a section titled 'Implicit Moodle'. Inside this section, there's a sub-section 'Implicit Groups Settings: admin'. This sub-section contains a form with two input fields: 'Username' with the value 'admin' and 'Password' with masked characters '.....'. There is also an 'Unmask' checkbox. Below the form are two buttons: 'Save changes' and 'Cancel'. At the bottom of the page, there is a link for 'Moodle Docs for this page' and another 'You are logged in as Admin User (Logout)' message, along with a 'Home' button.

Figura 5.2: Definição das credenciais de um utilizador (extensão Moodle)

3. utilização de outros métodos e aplicações de gestão de base de dados, para executar o mapeamento de tabelas de utilizadores das aplicações cliente na ferramenta TIG. Por exemplo, através de processos de exportação/importação de dados e utilização de *scripts* SQL, permitidos pelo *mySQLAdmin*.

No caso da integração na plataforma Moodle, cada utilizador tem de se registar individualmente na ferramenta TIG. Após a conclusão do processo de registo, cada utilizador possui um nome de utilizador e uma palavra-passe, as credenciais necessárias para utilização da ferramenta. Quando estas credenciais são introduzidas num pedido efectuado à API, a ferramenta é capaz de autenticar o utilizador, completar o pedido e devolver a resposta apropriada.

O módulo de extensão do Moodle desenvolvido utiliza uma tabela adicional na base de dados da plataforma Moodle, para manter as correspondências entre os identificadores de utilizadores do Moodle e de TIG. Campos suplementares da tabela guardam adicionalmente, para cada utilizador, as credenciais introduzidas no formulário próprio de configuração do módulo de extensão (ver figura 5.2).

### 5.3.2 Pedidos e Respostas

Os pedidos efectuados à TIG API, pelo módulo de extensão, são realizados através de um cliente REST, implementado com recurso à biblioteca PHP, *Client URL Library*. O bloco de código 5.1, ilustra o referido cliente.

Listagem 5.1: Código: cliente REST na linguagem PHP

```
<?php
//../rest_client/util.php

function callResource($url, $credentials, $method, $data = null)
{
    $url    = $GLOBALS['config']['apiUrl'] . $url;
```



```

$handle = curl_init();

$headers = array(
    'Accept: text/xml',
    'Content-Type: text/xml',
    'Authorization: '. 'TRUEREST username='.$credentials->name. '&class='↵
        . 'Customer'. '&password='.$credentials->password. '&apikey=↵
        testapikey '
);

$jheaders = array(
    'Accept: text/xml',
    'Content-Type: application/x-www-form-urlencoded',
    'Authorization: '. 'TRUEREST username='.$GLOBALS['config']['username'↵
        ]. '&class='.'Customer'. '&password='.$GLOBALS['config']['↵
        password']. '&apikey=testapikey '
);

curl_setopt($handle, CURLOPT_URL, $url);
curl_setopt($handle, CURLOPT_HTTPHEADER, $headers);
curl_setopt($handle, CURLOPT_RETURNTRANSFER, true);
curl_setopt($handle, CURLOPT_SSL_VERIFYHOST, false);
curl_setopt($handle, CURLOPT_SSL_VERIFYPEER, false);

switch ($method) {
    case 'POST':
        $postfields = urlencodeArray($data);
        curl_setopt($handle, CURLOPT_POST, true);
        curl_setopt($handle, CURLOPT_HTTPHEADER, $jheaders);
        curl_setopt($handle, CURLOPT_POSTFIELDS, $postfields);
        break;
    case 'PUT':
        curl_setopt($handle, CURLOPT_PUT, true);
        $file_handle = fopen($data, 'r');
        curl_setopt($handle, CURLOPT_INFILE, $file_handle);
        break;
    case 'DELETE':
        curl_setopt($handle, CURLOPT_CUSTOMREQUEST, 'DELETE');
        break;
}

$response = curl_exec($handle);
$code = curl_getinfo($handle, CURLINFO_HTTP_CODE);
curl_close($handle);

return array(
    'code' => $code,
    'data' => $response,
);
}

```

A função *callResources* em 5.1, recebe quatro argumentos distintos :

1. *\$url*: o endereço do recurso, chamada à TIG API.
2. *\$credentials*: conjunto de credenciais do utilizador associado ao pedido.
3. *\$method*: o método HTTP definido para a respectiva chamada.
4. *\$data*: os dados a enviar por POST (opcional).

No endereço da chamada pretendida é indicado o formato de resposta esperado. As respostas devolvidas para cada pedido, em XML ou JSON, apresentam uma estrutura comum, embora os dados que contêm variem consoante a respectiva chamada. O bloco 5.2, representa genericamente, a resposta XML devolvida a um pedido efectuado pelo cliente REST implementado.

Listagem 5.2: Código: resposta a um pedido de *../groups/view.xml*

```
<groups_response>
  <meta>
    <status>ok</status>
    <feedback/>
    <request>
      <http_host>localhost</http_host>
      <server_addr>127.0.0.1</server_addr>
      <remote_addr>127.0.0.1</remote_addr>
      <server_protocol>HTTP/1.1</server_protocol>
      <request_method>GET</request_method>
      <request_uri>/cakephp-1.3.6/cakephp-cakephp-f6748d4/implicit/groups/70.<
        xml</request_uri>
      <request_time>1313424674</request_time>
    </request>
    <credentials>
      <class>Customer</class>
      <apikey>testapikey</apikey>
      <username>admin</username>
    </credentials>
    <time_epoch>1313424674</time_epoch>
    <time_local>Mon, 15 Aug 2011 17:11:14 +0100</time_local>
    <version>0.3</version>
  </meta>
  <data>
    ..... Dados relativos ao grupo que se pretende consultar...
  </data>
</groups_response>
```

Em qualquer resposta, as *tags* “meta” englobam informações sobre o pedido efectuado, como a sua origem e o URI requisitado. As *tags* “data” contêm toda a informação específica de resposta ao pedido efectuado, neste caso, informações sobre o grupo que se

pretende consultar, incluindo os utilizadores, interesses, mensagens e *posts* relacionados com o mesmo (ver 5.3).

Listagem 5.3: Código: dados XML contidos na resposta à consulta de um determinado grupo implícito.

```

<data>
  <group>
    <Group>
      <group_id>70</group_id>
      <name>Physics group</name>
      <description>description Physics group</description>
      <type>I</type>
    </Group>
  </Message/>
  <User>
    <item>
      <user_id>1</user_id>
      <username>user01</username>
      <email>user01@somemail.com</email>
      <description>..This user text description..</description>
      <birthdate>1984-06-17</birthdate>
      <occupation>1</occupation>
      <created>2011-03-21 01:54:52</created>
      <modified>2011-07-13 04:29:36</modified>
      <GroupsUser>
        <id>310</id>
        <user_id>1</user_id>
        <group_id>70</group_id>
      </GroupsUser>
    </item>
    <item>
      <user_id>2</user_id>
      <username>admin</username>
      <email>admin@somemail.com</email>
      <description>..This user text description..</description>
      <birthdate>1984-06-18</birthdate>
      <occupation>1</occupation>
      <created>2011-03-21 23:58:13</created>
      <modified>2011-07-17 18:54:05</modified>
      <GroupsUser>
        <id>312</id>
        <user_id>2</user_id>
        <group_id>70</group_id>
      </GroupsUser>
    </item>
  </User>
  <Interest>
    <item>

```

```

    <interest_id>53</interest_id>
    <name>Physics</name>
    <InterestsGroup>
      <id>58</id>
      <interest_id>53</interest_id>
      <group_id>70</group_id>
    </InterestsGroup>
  </item>
</Interest>
<Post>
  <item>
    <post_id>17</post_id>
    <sender_id>2</sender_id>
    <type>B</type>
    <title>Bookmark</title>
    <content>no_content</content>
    <url>http://www.sciencedaily.com/releases/2011/07/110714072416.htm</url>
    <filter>NULL</filter>
    <created>2011-07-14 15:32:31</created>
    <modified>2011-07-14 15:32:31</modified>
    <GroupsPost>
      <id>10</id>
      <group_id>70</group_id>
      <post_id>17</post_id>
    </GroupsPost>
  </item>
</Post>
</group>
</data>

```

Após a recepção da resposta, as aplicações cliente podem usar os dados recebidos como base para a implementação de diversas funcionalidades ou verificar o sucesso da operação requisitada.

No caso do módulo de extensão Moodle, as bibliotecas SimpleXML e JavaScript Object Notation são utilizadas na conversão das respostas XML e JSON em *arrays* associativos PHP. Outras linguagens podem recorrer a métodos semelhantes ou optar por manipular directamente os documentos XML e JSON recebidos.

A secção seguinte detalha a extensão Moodle desenvolvida, exemplificando como os dados recebidos são utilizados no desenvolvimento de funcionalidades adicionais, úteis aos utilizadores da plataforma Moodle.

## 5.4 Aplicação Desenvolvida

A aplicação desenvolvida consiste num módulo de extensão da plataforma Moodle. O módulo criado é compatível com as versões 1.9.x da plataforma. As próximas secções detalham o módulo implementado, assim como as bibliotecas internas fornecidas pelo

Moodle e utilizadas no desenvolvimento das funcionalidades pretendidas.

### 5.4.1 Módulo de Extensão

O Moodle permite o alargamento das funcionalidades oferecidas, através da implementação de diferentes tipos de módulos de extensão ou *plugins*. Dos diferentes tipos de módulos possíveis, optou-se por implementar um novo bloco Moodle.

A API de blocos do Moodle, permite o desenvolvimento de blocos configuráveis e multiusos que exibem HTML. A plataforma oferece, por omissão, vários destes blocos com os mais variados propósitos. Alguns exemplos destes blocos podem ser visualizados na figura 5.3.

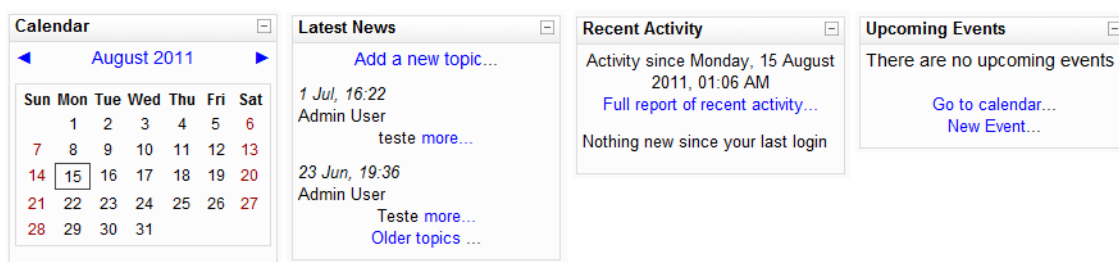


Figura 5.3: Diversos blocos existentes na plataforma Moodle

Um novo bloco consiste num conjunto de ficheiros indispensáveis, que devem ser incluídos na pasta *blocks* presente na pasta raiz do Moodle, "moodle/blocks/newblock-name/". O bloco de grupos implícitos desenvolvido possui a seguinte estrutura, na pasta "moodle/blocks/block\_implicit\_groups/":

- **block\_implicit\_groups.php**: contém a classe que define um novo bloco, estendendo a classe nativa *block\_base*. Corresponde ao ficheiro necessário e essencial na definição de qualquer novo bloco, no Moodle.
- **block\_implicit\_groups\_bookmarks\_form.php**: Define a estrutura do formulário utilizado na partilha de artigos e notícias.
- **block\_implicit\_groups\_form.php**: Define a estrutura do formulário utilizado na recolha das credenciais TIG de cada utilizador.
- **ibookmarks.php**: responsável pela instanciação e exibição do formulário de partilha, definido no ficheiro *block\_implicit\_groups\_bookmarks\_form.php*.
- **issettings.php**: responsável pela instanciação e exibição do formulário de credenciais, definido no ficheiro *block\_implicit\_groups\_form.php*.
- **iinterest.php**: permite a selecção explícita de qualquer interesse, pertencente à lista de sugestões de um utilizador.

- **isuggestions.php**: permite apagar a lista de interesses sugeridos a determinado utilizador.
- **view\_groups.php**: permite a visualização da informação associada a determinado grupo de interesse.
- **lib.php**: conjunto diverso de funções úteis, reutilizadas em diferentes partes da aplicação. Inclui, entre outras, funções de acesso à base de dados Moodle e controlo do cliente REST.
- **/rest\_client/util.php**: implementação do cliente REST utilizado na comunicação com a TIG API.
- **/rest\_client/config.php**: dados de parametrização do cliente REST.
- **/db/..**: directório que contém os ficheiros associados à criação e utilização de tabelas adicionais, presentes na base de dados da plataforma Moodle.
- **/pix/..**: directório que contém as imagens associadas à interface do bloco.

O bloco desenvolvido pode ser instalado numa distribuição compatível da plataforma Moodle, colocando-se a pasta associada ao bloco na pasta "moodle/blocks/". O bloco será automaticamente instalado, assim que for visitada a secção de administração do Moodle. Informações adicionais sobre o progresso da instalação, podem ser consultadas na zona de notificações, na secção de administração do Moodle.

### 5.4.2 Bibliotecas Moodle

A maioria dos ficheiros que compõem o bloco criado recorre a bibliotecas internas do Moodle e a funcionalidades que estas oferecem. As principais bibliotecas nativas utilizadas são:

- **formlib.php**: biblioteca de classes para a criação de formulários no Moodle, baseada em PEAR QuickForms.
- **weblib.php**: biblioteca de funções e constantes utilizadas na geração de HTML.
- **datalib.php**: biblioteca de funções de acesso à base de dados Moodle.
- **blocklib.php**: biblioteca de funções associadas a blocos Moodle.

### 5.4.3 Funcionalidades Implementadas

O módulo de extensão (bloco) desenvolvido, permite que os utilizadores do Moodle utilizem os grupos identificados pela ferramenta TIG, na partilha de artigos, notícias e mensagens relacionadas com os seus interesses.

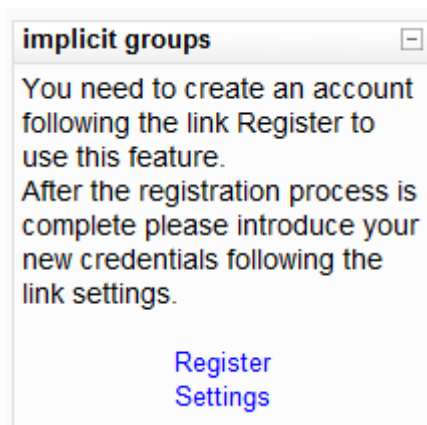


Figura 5.4: Visualização do bloco apresentado inicialmente a cada utilizador.

Um utilizador pode interagir com a extensão desenvolvida, através do bloco que forma a sua interface. A figura 5.4 ilustra o bloco de grupos implícitos apresentado, inicialmente, a cada utilizador.

Após seguir e executar as instruções apresentadas, um utilizador pode começar a utilizar as funcionalidades oferecidas pelo bloco. Depois de introduzidas as credenciais que permitem autenticar um utilizador perante a ferramenta TIG, o bloco apresenta quatro secções distintas, como se pode observar na figura 5.5.

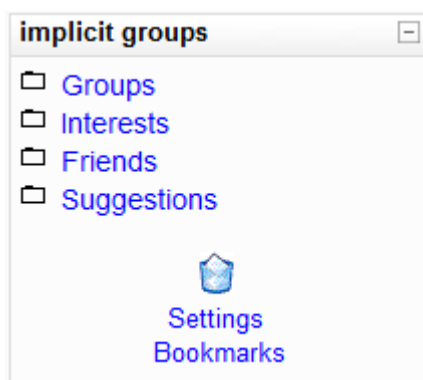


Figura 5.5: Visualização do bloco após a introdução de credenciais de um utilizador.

Cada uma das secções (*Groups*, *Interests*, *Friends* e *Suggestions*), fornece acesso a diferentes informações ou funcionalidades. Para além das quatro secções enumeradas, existem ainda duas secções que podem ser acedidas através dos respectivos links, no rodapé do bloco (*Settings* e *Bookmarks*).

Os dados contidos em cada uma das quatro secções principais são fornecidos pela a resposta ao pedido realizado à primitiva "users/home.xml" de TIG, através da sua API REST. Os dados devolvidos correspondem ao utilizador autenticado no pedido, por via

das credenciais introduzidas anteriormente. A chamada efectuada no ficheiro de definição do bloco desenvolvido, pode ser consultada na listagem seguinte:

Listagem 5.4: Código: extracto do ficheiro `block_implicit_groups.php` contendo a chamada da primitiva `../users/home.xml`.

```
function get_content() {
    ...
$response = $this->getData("users/home.xml", $credentials);

$currentuser = new stdClass;
$currentuser->iid = (int) $response->data->user->User->user_id;
$currentuser->mid = $USER->id;

update_user_map($currentuser);

$this->implicitMenuItems[0]['topname'] = "Groups";
$this->implicitMenuItems[0]['Groups'] = $response->data->user->Group[0];
$this->implicitMenuItems[1]['topname'] = "Interests";
$this->implicitMenuItems[1]['Interests'] = $response->data->user->Interest[0];
$this->implicitMenuItems[2]['topname'] = "Friends";
$this->implicitMenuItems[2]['Friends'] = $response->data->user->Stat[0];
$this->implicitMenuItems[3]['topname'] = "Suggestions";
$this->implicitMenuItems[3]['Suggestions'] = $response->data->user->User[0]-><->
    suggestions;

foreach ($this->implicitMenuItems as $stop){
    $this->buildGroupTree($stop);
}
    ...
}
```

#### 5.4.3.1 Groups

Seleccionando a secção de grupos, um utilizador pode consultar todos os grupos dos quais é membro. A selecção de um dos grupos listados, redirecciona o utilizador para o espaço de grupo respectivo, onde pode consultar as mensagens, artigos e notícias associadas ao grupo, assim como informações sobre os restantes membros.

Quando um utilizador selecciona o nome de utilizador de outro membro do grupo, obtém acesso ao perfil Moodle desse utilizador, caso este exista. Uma vez que a ferramenta TIG pode ser partilhada por várias aplicações cliente, existe a possibilidade de que um utilizador de determinado grupo seja utilizador de outra aplicação que não o Moodle. Neste último caso, o utilizador externo ao Moodle é assinalado com um ícone que evidencia a incapacidade de se promover uma interacção com o mesmo, através do sistema de mensagens da plataforma Moodle.



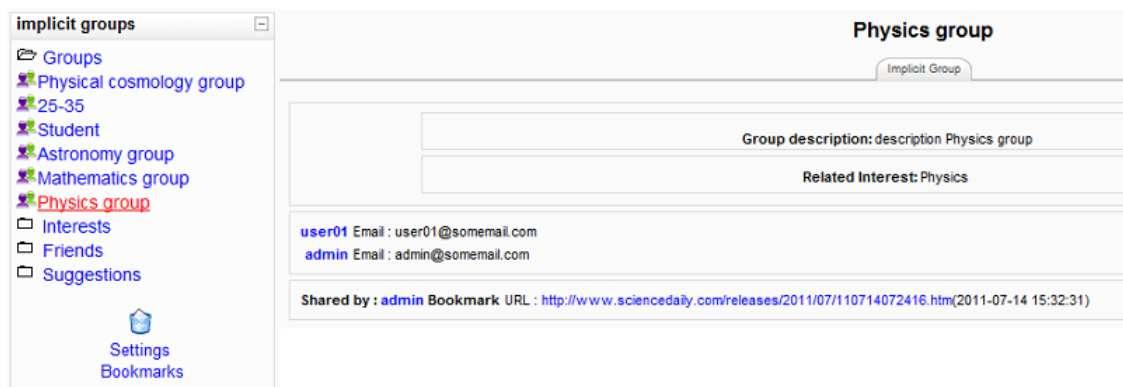


Figura 5.6: Vista da secção de grupos e do espaço de um grupo.

A visualização do espaço de um grupo é promovida pelo ficheiro *view\_groups.php*. Quando um grupo é seleccionado, o *script view\_groups.php* recebe como argumento o identificador do grupo em TIG e utiliza-o para executar uma chamada à primitiva REST "groups/group\_id.xml". Na sequência do processo, os dados contidos na resposta ao pedido efectuado são apresentados ao utilizador, através das bibliotecas HTML do Moodle (ver figura 5.6).

#### 5.4.3.2 *Interests*

A secção de interesses é meramente informativa, apresentando uma simples lista de todos os interesses associados ao utilizador considerado. Esta lista permite que o utilizador tome conhecimento dos vários interesses que lhe são atribuídos ao longo do tempo, incluindo aqueles que seleccionou utilizando a secção de sugestões.

#### 5.4.3.3 *Friends*

A lista de amigos de um utilizador (figura 5.7) mostra outros utilizadores com os quais existiu interacção (troca de mensagens via TIG ou via Moodle). A lista encontra-se ordenada por relevância descendente. A relevância consiste num valor numérico, calculado de acordo com as propriedades de interacção previamente abordadas (em 4.2.4) e os respectivos pesos, definidos através da interface de administração de TIG.

#### 5.4.3.4 *Suggestions*

A secção de sugestões apresenta um conjunto de interesses, implicitamente detectados através da análise de documentos, conteúdos e campos de perfil associados a um utilizador.

Esta funcionalidade apenas está disponível quando o processo de atribuição de interesses a cada utilizador, funciona no modo assistido. Apesar do referido modo fornecer

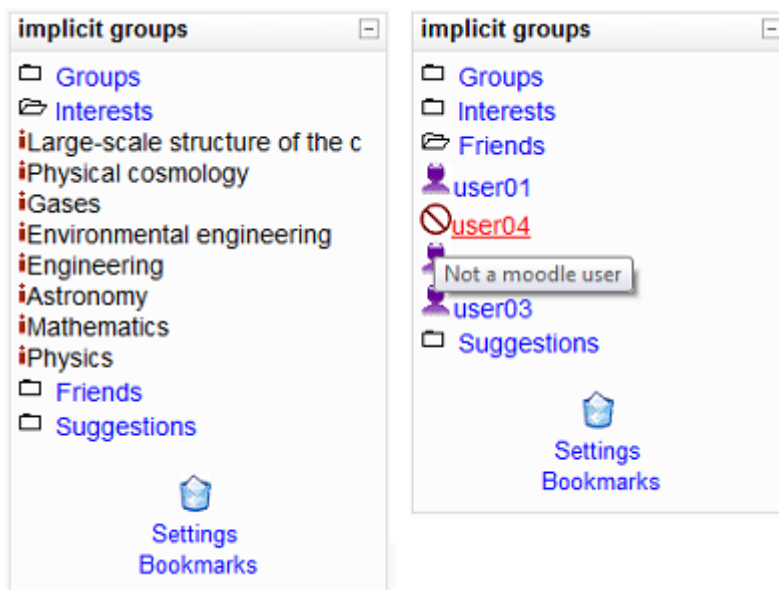


Figura 5.7: Vista das secções de interesses e amigos.

a capacidade de um utilizador filtrar explicitamente os interesses que vão sendo detetados ao longo do tempo, os grupos identificados são considerados implícitos, uma vez que não existe uma acção concreta de afiliação a esses grupos.

Através do ícone posicionado imediatamente a seguir à secção de sugestões, o utilizador pode também eliminar as sugestões existentes. Esta funcionalidade recorre a uma chamada a "interests/clearSuggestions.xml".

#### 5.4.3.5 *Settings*

As credenciais de acesso à ferramenta TIG podem ser introduzidas, por cada utilizador, na secção *settings*. Estas credenciais consistem num conjunto definido durante a acção de registo na ferramenta, formado pelo nome de utilizador e palavra-passe.

A introdução das credenciais permite ao módulo de extensão desenvolvido, recorrer às funcionalidades da API REST, associadas a cada utilizador. A secção *settings* pode ser alcançada seguindo-se a hiperligação existente na parte inferior do bloco criado.

#### 5.4.3.6 *Bookmarks*

A secção *bookmarks* fornece o acesso às principais funcionalidades da extensão Moodle, permitindo a importação e partilha de artigos e notícias, relevantes para os utilizadores. Uma visualização desta secção é ilustrada na figura 5.8.

O utilizador pode introduzir o endereço de um recurso web no campo de texto, identificado pela etiqueta *Bookmark*. Caso o utilizador opte por não seleccionar a opção de partilha, o recurso importado apenas contribuirá para a identificação dos seus interesses. No caso contrário, o recurso é também partilhado com outros utilizadores interessados,

Import/Share Bookmark: admin

Bookmark (url)

Share Bookmark  select to share this bookmark with other interested users (if they exist)

Sharing filters

Recipients

Custom Filter

Occupation  Age Group(s)   
18-24  
25-35  
36-50  
51-69  
70-?

Only Friends

[Moodle Docs for this page](#)

You are logged in as [Admin User](#) ([Logout](#))

Figura 5.8: Vista da secção de *bookmarks*.

de acordo com filtros opcionalmente definidos pelo utilizador que inicia a partilha.

Quando um recurso é submetido através do formulário, a sua classificação é realizada automaticamente de acordo com o seu conteúdo. Posteriormente, esta classificação é utilizada na identificação dos grupos de interesses que beneficiam do novo recurso. Caso o utilizador não especifique nenhum filtro adicional, o recurso será acessível a todos os utilizadores que estejam associados, no mínimo, a um dos interesses que o classificam. Os utilizadores interessados podem aceder aos recursos partilhados através dos espaços de cada grupo implícito (figura 5.6).

O utilizador responsável pela partilha de um recurso pode restringir os utilizadores com acesso a esse recurso, especificando filtros adicionais. Estes filtros são implementados através da combinação dos grupos existentes na ferramenta. Os grupos baseados em interesses podem ser combinados com grupos baseados em características e grupos baseados em interações, na restrição do conjunto de utilizadores alvo da partilha (recipientes).

A título exemplificativo, um utilizador com papel de Professor na plataforma Moodle, pode utilizar os filtros disponíveis para partilhar um determinado artigo com todos os outros utilizadores interessados, que sejam Professores. Além da utilização da “característica” ocupação como filtro, a funcionalidade implementada permite ainda a utilização de faixas etárias e listas de amigos, como forma de restringir os beneficiários da partilha.

## 5.5 Utilização de Métricas

A ferramenta de grupos implícitos desenvolvida, baseia-se em mecanismos e processos genéricos de identificação e atribuição de grupos implícitos, que podem ser ajustados às aplicações cliente, seus objectivos e contextos.

Os três diferentes critérios utilizados na identificação de grupos possuem diferentes potenciais, no que diz respeito à definição e utilização de métricas que aumentem a utilidade dos grupos identificados para os utilizadores e aplicações.

### 5.5.1 Grupos baseados em características

Os grupos baseados em características, sendo directamente inferidos através de valores explicitamente definidos pelos próprios utilizadores nos campos dos seus perfis, reflectem grupos de utilizadores com características comuns. Embora estes grupos possuam maior ou menor utilidade no contexto da aplicação cliente, dificilmente beneficiam de métricas adicionais na sua identificação e criação. Esta observação deriva do facto de as características não dependerem de indicadores implícitos variáveis de utilizador para utilizador, como juízos de valor, opiniões pessoais e comportamentos. De facto alguns destes grupos, como os baseados em faixas etárias, podem ser utilizados para se estimar e prever alguns desses indicadores, como os comportamentos e gostos associados a diferentes idades.

### 5.5.2 Grupos baseados em interacções

No caso dos grupos formados com base nas interacções entre utilizadores, o potencial para a utilização de métricas adicionais aumenta. Note-se que a ferramenta TIG apenas considera a troca de mensagens privadas entre utilizadores, como indicador do nível de interacção existente. Neste caso particular, as aplicações cliente podem permitir outro tipo de interacções que podem ser contabilizadas e analisadas apenas do seu lado, como por exemplo, comentários a *posts* e a utilização de salas de *chat* em tempo real. A diversidade das aplicações e mecanismos que utilizam para promover a interacção entre utilizadores e a importância que cada um desses mecanismos pode possuir para cada aplicação é extremamente variável. Além da diversidade de mecanismos, existe ainda a diversidade associada às possíveis dimensões analisadas e aos seus pesos, como sugere e identifica [GK09]. Estes factos tornam inviável a contabilização e a análise dos referidos mecanismos numa ferramenta como a desenvolvida, essencialmente devido a factores de *performance*, dificuldade de implementação e ao nível de abrangência que se pretende.

Como alternativa para a implementação de métricas adicionais, que permitam calcular índices de interacção mais abrangentes, as aplicações cliente podem combinar os resultados devolvidos por TIG, relativos às interacções contabilizadas através da utilização de sistemas de mensagens privados, com outros valores inferidos e calculados pelas próprias. Esta combinação possui várias vantagens, uma vez que a proximidade dos

utilizadores às aplicações cliente é menor e os objectivos e contextos específicos de cada aplicação, podem ser satisfeitos de forma mais concreta e eficiente.

Caso as aplicações cliente optem apenas por contabilizar as interacções através do mecanismo oferecido por TIG e sua API, podem sempre, no entanto, optar por definir os pesos das propriedades/dimensões contabilizadas, através da interface de administração da ferramenta (ver 4.3.2.3) e de acordo com as suas preferências.

### 5.5.3 Grupos baseados em interesses

Os grupos baseados em interesses identificados e gerados pela ferramenta desenvolvida, baseiam-se na análise dos conteúdos relacionados com os utilizadores. A problemática associada à descoberta dos interesses associados a determinado utilizador é bastante comum nas áreas dos sistemas de recomendações e filtragem colaborativa, diferindo consoante o nível de automatismo dos processos de identificação pretendidos e consoante as acções exigidas aos utilizadores.

De forma semelhante, os grupos implícitos de interesses podem beneficiar dos mecanismos, processos e métricas, utilizados para estimar e descobrir os interesses que cada utilizador possui, com precisão superior e tendo em conta outros recursos para além dos conteúdos associados aos utilizadores. Como nos grupos baseados em interacções, no entanto, a imposição deste tipo de métricas do lado da ferramenta, limita consideravelmente a utilidade da mesma, por exemplo, limitando o nível de automatismo de identificação de grupos que pode ser oferecido, quando se exigem constantes classificações de informação aos utilizadores. Devido a este facto, este tipo de métricas e mecanismos relacionados com a identificação de interesses, devem ser implementados e definidos nas aplicações cliente, tendo em conta os objectivos pretendidos e de acordo com os indicadores de interesses disponíveis.

Com base nos anteriores factores, a implementação de métricas e mecanismos adicionais que maximizem a precisão e utilidade dos interesses e respectivos grupos identificados por TIG, deve ser ponderada de acordo com condicionantes distintas, abordadas nas seguintes subsecções.

#### 5.5.3.1 Automatismo do processo de identificação de grupos de interesses

Uma solução frequentemente utilizada nos sistemas de recomendações, para inferir os interesses dos utilizadores, prende-se com o desenvolvimento de interfaces que exigem aos mesmos a classificação explícita de pedaços de informação. Estas classificações, que indicam a relevância atribuída por cada utilizador a certos pedaços de informação, são posteriormente utilizadas na identificação de novos interesses, contribuindo para a geração de recomendações.

Em [CLWB01], Claypool et al. identificam uma limitação essencial que este tipo de soluções impõe a processos automatizados de identificação de interesses, interrompendo o fluxo normal de visualização e consulta de documentos, com pedidos recorrentes de

classificação aos utilizadores. Na alternativa sugerida em [CLWB01], os sistemas podem recorrer a interfaces que analisem indicadores implícitos de interesses, utilizados na classificação de informação sem a intervenção dos utilizadores.

Alguns dos indicadores implícitos de interesses identificados por Claypool et al., correspondendo a acções que os utilizadores executam quando interagem com uma página web, como os movimentos do rato, os cliques efectuados, o *scrolling* e o tempo despendido na consulta, provaram possuir uma forte correlação com interesses explícitos dos utilizadores. Entre os indicadores estudados, o tempo despendido na consulta de uma página, a quantidade de *scrolling* efectuado e a combinação de ambos, podem ser utilizados na previsão de interesses explícitos dos utilizadores. Estes indicadores podem complementar ou filtrar os interesses detectados por TIG e contribuir para a identificação de grupos implícitos adicionais.

A conjugação deste tipo de indicadores com a ferramenta TIG pode ser obtida através das funcionalidades oferecidas pela API e da configuração do nível de automatismo pretendido. Concretamente, uma vez seleccionado o processo de identificação automático de grupos de interesses, em TIG, as aplicações cliente podem analisar os referidos indicadores e atribuir os interesses detectados a cada utilizador, através da chamada "interests/select" da API. Adicionalmente, as aplicações cliente podem optar por criar manualmente grupos de interesses e atribuir os utilizadores aos grupos criados, através das funcionalidades da API associadas a administradores da ferramenta, qualquer que seja o processo de identificação de grupos de interesses definido.

Considere-se o seguinte exemplo, utilizando o indicador de tempo despendido na consulta de uma página web. Uma aplicação cliente pode contabilizar o tempo despendido por cada utilizador na consulta de uma qualquer secção ou página da mesma e fornecer o endereço da página à ferramenta TIG, que devolve os interesses contidos na mesma (passo dispensável caso os interesses sejam já conhecidos pela própria aplicação). De seguida, os interesses detectados de acordo com os indicadores analisados, podem contribuir para a extensão do perfil de cada utilizador e conseqüentemente, utilizados na atribuição dos mesmos aos grupos de interesses correspondentes.

Para além dos indicadores implícitos que favorecem os processos automáticos e reduzem o esforço exigido aos utilizadores, indicadores explícitos e métodos habituais de classificação por parte dos utilizadores, podem também ser implementados nas aplicações cliente e conjugados com a ferramenta TIG. A escolha dos processos manuais ou assistidos de identificação de grupos de interesses em TIG, permite a implementação destes mecanismos onde as acções explícitas dos utilizadores podem ser utilizadas para filtrar e precisar os interesses que se consideram relevantes. O processo assistido de identificação de grupos de interesses em TIG, por exemplo, utiliza por omissão um mecanismo deste género, pedindo aos utilizadores para seleccionarem de entre uma lista de sugestões de interesses, aqueles que consideram mais relevantes.

Alguns exemplos de indicadores explícitos/implícitos que podem ser considerados na maximização da utilidade dos interesses seleccionados e grupos identificados por TIG,

são também abordados em [CLWB01] e incluem :

- classificações explícitas de utilizadores/indicação de preferências;
- acções de marcação explícita, como gravar, apagar ou adicionar aos favoritos, um determinado documento/recurso;
- indicadores de manipulação como cortar, colar, modificar ou procurar num documento por uma determinada palavra;
- indicadores de navegação, por exemplo, analisando-se as hiperligações seguidas;
- indicadores de repetição associados à frequência de consulta/modificação de documentos;
- indicadores externos relacionados com a resposta física e emocional dos utilizadores (batimento cardíaco, movimento dos olhos, temperatura, expressões faciais, etc.);
- indicadores negativos, representando acções que ao não serem realizadas, podem indicar o desgosto ou desinteresse por algo.

### 5.5.3.2 Aproximação baseada em conteúdo e filtragem colaborativa

A abordagem seguida pela ferramenta TIG na identificação de interesses, centra-se na análise dos conteúdos relacionados aos utilizadores. Como complemento da abordagem escolhida, o uso de técnicas e algoritmos de filtragem colaborativa pode ser considerado. Herlocker et al. identificam em [HKBR99] três vantagens importantes que a filtragem colaborativa acrescenta a técnicas baseadas somente em conteúdos:

1. suporte para a filtragem de itens, cujo conteúdo não possa ser facilmente analisado por componentes de análise de conteúdos, como videos, músicas, ideias e sentimentos;
2. suporte para a filtragem de itens, cujo conteúdo não possa ser facilmente analisado por processos automáticos de análise de conteúdos;
3. suporte para recomendações inesperadas, ou seja, itens interessantes para os utilizadores mas cujo conteúdo é diferente do esperado. Herlocker et al., indicam que este tipo de recomendações é bastante comum no domínio do cinema, sugerindo filmes que os utilizadores nunca considerariam de outra forma.

Como ilustra o trabalho desenvolvido, por exemplo em [HKBR99], a filtragem colaborativa possui um enorme potencial na melhoria de ferramentas utilizadas na filtragem de informação, uma das possíveis utilizações para o conceito de grupos implícitos e para a ferramenta TIG. Aqueles autores concluem, no entanto que, para se maximizar este

potencial, a combinação com aproximações e metodologias baseadas em conteúdos é essencial. Isto deve-se ao facto de, por si só, a filtragem colaborativa não ser adequada à localização de informação associada a uma necessidade, mas sim eficiente na predição de itens que satisfazem os interesses e gostos dos utilizadores.

Por outro lado, devido à natureza dos diferentes tipos de grupos implícitos identificados, as próprias técnicas de filtragem colaborativa podem beneficiar das informações associadas a esses grupos. Quando combinados os grupos de características com os grupos de interesses, é possível estudar possíveis correlações existentes entre as características e os interesses dos utilizadores. Estas correlações, caso sejam identificadas, podem ser úteis na geração de novas recomendações.

Consoante os objectivos aplicacionais existentes, diferentes metodologias e algoritmos relacionados com filtragem colaborativa, podem ser conjugados com as funcionalidades oferecidas pela ferramenta TIG. Quer utilizando o componente de análise de conteúdos incluído em TIG e simultaneamente capaz de classificar esses mesmos conteúdos, quer através da combinação dos dados reflectidos nos grupos identificados, as aplicações cliente podem optar por utilizar o conceito de grupos implícitos, no desenvolvimento de sistemas de filtragem colaborativa e sistemas de recomendações que satisfaçam diferentes necessidades.

### 5.5.3.3 Objectivos aplicacionais, *performance* e precisão

A ferramenta desenvolvida pretende valorizar o conceito de grupos implícitos generalizando-o, de forma a permitir a sua integração num leque mais abrangente de aplicações interactivas, baseadas na Internet.

Devido às opções escolhidas e relacionadas com a conceptualização da ferramenta (ver capítulo 3), a precisão dos interesses e respectivos grupos identificados pode não ser suficiente para determinadas aplicações, uma vez que os mecanismos utilizados na atribuição de interesses a um utilizador, não conseguem prever se o mesmo considera um documento interessante por outra razão que não o seu conteúdo. Os objectivos aplicacionais, a *performance* pretendida/exigida e a precisão necessária, constituem importantes condicionantes para a escolha das metodologias, mecanismos, métricas e a estratégia da sua implementação.

Embora a ferramenta tente agregar processos suficientemente genéricos, em último caso, a decisão da implementação de metodologias de maximização da utilidade dos grupos gerados, pertence exclusivamente às aplicações cliente. O padrão arquitectural utilizado facilita a extensão directa da ferramenta TIG, assim como a extensão das funcionalidades que oferece através da sua API, caso as condicionantes de *performance*, precisão e os objectivos aplicacionais assim o justifiquem.

Alguns dos indicadores de interesses abordados anteriormente, podem ser implementados sem sacrifício da *performance*, eficiência ou dos automatismos pretendidos, estendendo-se tanto a ferramenta como as aplicações cliente. Por outro lado, alguns



desses indicadores, quando implementados directamente na ferramenta, implicam sacrifícios de generalidade e eficiência, entre outros. A escolha do local de implementação associado a métricas e funcionalidades adicionais, constitui assim um problema que deve ser analisado conscientemente, de acordo com os objectivos concretos de cada aplicação e com compreensão das implicações que cada metodologia acarreta.

## 5.6 Conclusão

A integração da ferramenta TIG concretizada na plataforma Moodle foi ilustrada neste capítulo. A utilidade da ferramenta no desenvolvimento de novas funcionalidades na plataforma Moodle foi comprovada para o cenário idealizado, relativo a um Campus universitário.

Embora a adaptabilidade da ferramenta a diferentes cenários, apenas possa ser comprovada com validações adicionais, a generalidade e flexibilidade dos mecanismos/processos concebidos e presentes na ferramenta, como por exemplo, as capacidades para permitir diversas opções de configuração e a identificação de interesses com níveis de especificidade variáveis, prevêem que a mesma possa ser aplicada noutros contextos.

Tendo em vista a obtenção de alguns indicadores preliminares, sobre o potencial do conceito de grupos implícitos e as funcionalidade oferecidas pela ferramenta, no próximo capítulo são apresentados os testes realizados com um conjunto limitado de utilizadores, assim como os respectivos inquéritos que revelam a satisfação e as expectativas dos mesmos.



# 6

## Avaliação de Resultados da Utilização

Após a conclusão da implementação do primeiro protótipo da ferramenta TIG, com o objectivo de recolher dados relacionados com as opiniões e expectativas dos utilizadores acerca da utilidade dos grupos identificados e das funcionalidades oferecidas, foram conduzidos breves testes preliminares com utilizadores voluntários. Os testes realizados foram posteriormente complementados com inquéritos sobre a satisfação dos utilizadores teste.

Nas próximas secções, é discutida a relevância dos testes e inquéritos efectuados, nas opções tomadas ao longo do desenvolvimento da ferramenta, assim como as limitações e vantagens identificadas, como consequência dos resultados obtidos. Salienta-se que os referidos testes não pretenderam comprovar ou verificar formalmente a utilidade da ferramenta criada, mas sim, extrair indicadores relevantes sobre o modo de utilização e a satisfação dos utilizadores, no contexto dos grupos implicitamente identificados pela ferramenta.

### 6.1 Testes de Utilizadores e Inquéritos de Satisfação

Os testes de utilizadores foram realizados através da Internet e da interface HTML oferecida, apresentada no capítulo 4, tendo sido iniciados na última semana de Junho de 2011 e concluídos um mês depois. O inquérito de satisfação foi lançado na segunda semana de Julho, tendo sido recebidas 22 respostas válidas até à data de 31 de Agosto de 2011.

Do total de 22 utilizadores que testaram a ferramenta, seguindo o guião estabelecido e respondendo depois ao inquérito disponibilizado, 18 indicaram o sexo masculino e os restantes 4, o sexo feminino. As profissões e áreas profissionais dos inquiridos concentraram-se no contexto académico, sendo 12 dos inquiridos estudantes e 6 professores, associados a diferentes níveis de ensino. Os restantes 4 inquiridos representaram outras áreas profissionais.

A figura 6.1 ilustra a distribuição dos utilizadores e inquiridos por diferentes faixas etárias.

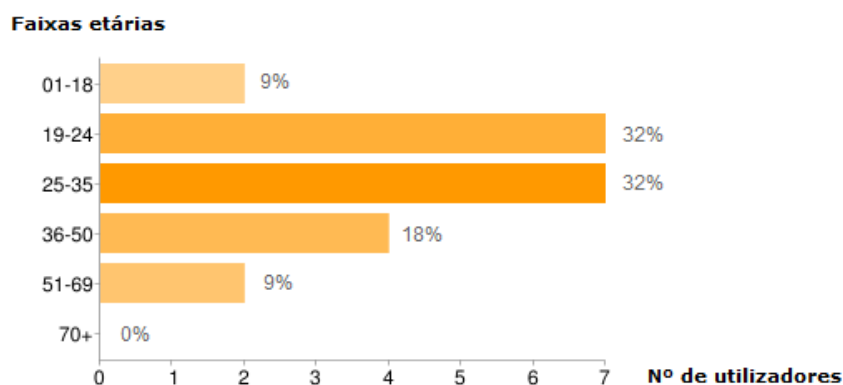


Figura 6.1: Utilizadores/inquiridos e respectivas faixas etárias

A configuração utilizada no teste da ferramenta, assim como a metodologia seguida são apresentadas na próxima subsecção. As questões colocadas durante o inquérito são enumeradas na subsecção 6.1.2.

### 6.1.1 Testes de Utilizadores e Metodologia

A utilização do protótipo TIG, durante a fase de testes, foi suportada pelo guião de utilização fornecido aos utilizadores, guião esse, responsável por apresentar as funcionalidades indispensáveis à sua manipulação.

Após cada utilizador efectuar um registo convencional no protótipo, estabelecendo o seu nome de utilizador e a sua palavra-chave de acesso, foi pedido o preenchimento de um perfil básico contendo, como campos essenciais, a sua idade de nascimento, a sua ocupação actual e uma breve descrição pessoal.

Adicionalmente ao preenchimento do perfil, foi pedido, a cada utilizador, a introdução mínima de três endereços web, relativos a artigos/notícias escritos na língua inglesa e relacionados com os seus temas favoritos. A restrição linguística imposta resulta das funcionalidades oferecidas pelo OpenCalais, o serviço utilizado na implementação do componente de análise de conteúdo textual.

De acordo com o nível do automatismo do processo de identificação de grupos de interesses, definido anteriormente como "assistido", a introdução dos referidos endereços

gerou, para cada utilizador, um conjunto de sugestões, ou seja, uma lista de interesses representados por *tags*. A partir da lista anterior, cada utilizador seleccionou os interesses que considerou relevantes, estendendo o seu perfil básico. Para cada interesse distinto seleccionado por um qualquer utilizador, um novo grupo baseado no mesmo foi criado dinamicamente pelo sistema. Os interesses contidos nos perfis estendidos de cada utilizador definiram os membros de cada grupo formado.

De forma complementar ao processo descrito anteriormente, os restantes campos definidos no perfil básico e a troca de mensagens entre utilizadores, contribuíram também para a geração automática de grupos, de acordo com o critério baseado em características (idade de nascimento, ocupação), o critério baseado em interacção (mensagens trocadas entre utilizadores) e novamente o critério baseado em interesses (descrição pessoal).

A observação da distribuição dos utilizadores pelos diferentes grupos formados e a relação entre interesses sugeridos e seleccionados permitiu efectuar importantes ajustes nos mecanismos concebidos e identificar as suas limitações chave. Os resultados e as considerações resultantes destas observações são apresentados de seguida (secção 6.1.1.1).

#### 6.1.1.1 Análise de Resultados

A análise dos *logs* produzidos durante a fase de testes, para cada utilizador, juntamente com os grupos identificados e seus membros, permitiu observar dados importantes.

Um total de 85 interesses distintos foram seleccionados entre 289 sugestões, originando o mesmo número (85) de grupos de interesses. A taxa global de interesses seleccionados estabeleceu-se assim em 29,41%. A baixa percentagem de selecções relativas às sugestões pode ser justificada por diversos factores.

Após a análise cuidada dos endereços fornecidos pelos utilizadores e utilizados na extracção de interesses, verificou-se que os utilizadores revelaram uma tendência para fornecer endereços, cujos conteúdos reflectiam interesses semelhantes. Este facto originou diversas sugestões repetidas, referentes aos mesmos interesses (por vezes já seleccionados), contribuindo para a reduzida percentagem obtida.

Por outro lado, alguns dos endereços fornecidos pelos utilizadores, dos quais são exemplos,

- "<http://www.materialstoday.com/view/20256/>";
- "<http://www.wizards.com/Magic/Digital/DuelsOfThePlaneswalkers.aspx>".

não continham grande informação textual para ser analisada, pelo que os interesses extraídos, nestes casos, reflectiram publicidade e outros conteúdos textuais presentes no código fonte. Outro dos problemas verificados relacionou-se com a existência de componentes Flash, responsáveis por oferecer a informação textual em alguns *websites*, tornando essa informação omissa do código fonte analisado.

As limitações identificadas, associadas à estruturação e definição das páginas processadas, são inerentes ao componente de análise de conteúdo utilizado e ao conjunto de

Grupo de Interesse	Nº de membros
Entertainment_Culture group	7
Technology_Internet group	5
Computer programming group	4
FCT group	4
Sports group	3

Tabela 6.1: Os 5 grupos implícitos de interesses com maior número de membros.

ferramentas em que se baseia (OpenCalais). Por esta razão, tornou-se essencial optar por uma arquitectura modular no desenvolvimento da ferramenta, facto que torna possível a substituição do componente por outros, mais adequados aos conteúdos que as aplicações cliente pretendam processar. Outra possibilidade para minimizar as limitações apresentadas, está associada ao uso de ferramentas capazes de extrair o texto de objectos Flash (*flash decompilers*), como o KineticFusion [Ltd11].

O número de grupos de interesse a que cada utilizador foi atribuído (ou interesses seleccionados por cada um deles), variou entre 0 (um único utilizador) e 10 (um único utilizador), sendo que em média cada utilizador foi atribuído a 4,86 grupos diferentes. Para além dos grupos de interesse, todos os utilizadores foram atribuídos automaticamente a 2 grupos distintos, consoante a sua ocupação e a sua data de nascimento. Os grupos de interacção centrados em cada utilizador, foram também criados para os utilizadores envolvidos em trocas de mensagens, 11 dos 22 (50% do total de utilizadores).

Não foi possível identificar qualquer relação entre a quantidade de endereços fornecidos e o número de interesses seleccionados, por utilizador. De tal facto, conclui-se que, na experiência realizada, a quantidade de endereços analisada é menos importante que a riqueza textual do seu conteúdo. Esta assumpção foi verificada no conjunto de teste, especialmente no caso dos utilizadores que forneceram artigos académicos ricos em conteúdo textual, obtendo conseqüentemente um número mais elevado de sugestões. No entanto, apenas a elaboração de testes mais substanciais e significativos estatisticamente, pode comprovar a ilação retirada.

A distribuição dos utilizadores pelos grupos gerados demonstrou que grande parte dos grupos continham apenas um único membro. Dos 85 grupos identificados, apenas 13 continham 2 ou mais membros (15,29%). Esta observação vai de encontro ao esperado, devido ao reduzido conjunto de utilizadores de teste.

Outro resultado esperado está relacionado com a especificidade/generalidade dos interesses, associada aos grupos implícitos com maior número de membros (consultar tabela 6.1). Os dados mostram que os interesses que definem os grupos com mais membros são bastante gerais, ou resultantes de contextos sociais comuns a grande número de utilizadores teste. Mais uma vez, testes de maior dimensão são necessários para comprovar esta assumpção.

Embora não seja possível inferir, através dos testes realizados, se a especificidade dos

interesses é adequada à formação de grupos realmente úteis, a ferramenta desenvolvida prevê diferentes cenários, fornecendo diferentes tipos de interesses que podem ser detectados. Os tipos oferecidos, por ordem de especificidade crescente, são : categoria, entidade e *tag* social. O tipo *tag* social foi o escolhido durante a realização dos testes. Todos os tipos resultam das funcionalidades oferecidas pelo OpenCalais [Reu11].

### 6.1.2 Inquérito de Satisfação

Após a conclusão dos testes mencionados na secção anterior, cada utilizador respondeu a um inquérito sobre a sua satisfação. As questões colocadas a cada utilizador e focadas exclusivamente nos grupos implícitos de interesse identificados, foram as seguintes:

1. **Do ponto de vista de utilizador, considera interessante a criação de grupos baseados nos seus interesses ? Se respondeu não à questão anterior, indique as razões.**  
(Respostas possíveis : sim ou não).
2. **Considera que os interesses sugeridos reflectem os conteúdos introduzidos?**  
(Resposta numa escala de 1 a 5: nunca, poucas vezes, às vezes, muitas vezes e sempre);
3. **Como analisa a especificidade dos interesses sugeridos e respectivos grupos?**  
(Respostas possíveis : interesses demasiado específicos, interesses com especificidade adequada ou interesses demasiado gerais).
4. **Como classifica a utilidade potencial dos grupos criados?**  
(Resposta numa escala de 1 a 5: nenhuma utilidade, pouca utilidade, alguma utilidade, muito útil e extremamente útil);
5. **Considera que o tipo de grupos criados, baseados nos seus interesses, favorece a comunicação e partilha de informação?**  
(Resposta numa escala de 1 a 5: nunca, poucas vezes, às vezes, muitas vezes e sempre).

Para além das respostas às questões anteriores, cada utilizador indicou ainda a sua faixa etária, profissão/área profissional e opcionalmente deixou sugestões, opiniões ou comentários sobre a sua experiência.

Os resultados para as questões anteriores, podem ser consultados na seguinte subsecção (6.1.2.1), juntamente com a respectiva análise.

#### 6.1.2.1 Análise de Resultados

Os inquéritos realizados a cada utilizador pretenderam reunir dados relativos à satisfação e expectativas dos utilizadores, resultantes do contacto com o conceito de grupos implícitos e com a ferramenta desenvolvida.

Do ponto de vista de utilizador, considera interessante a criação de grupos baseados nos seus interesses ?



Se respondeu não à questão anterior, indique as razões.

Apenas não costumo utilizar e considero que o serem interessantes ou não pode variar de acordo com as pessoas que poderão estar nos mesmos grupos que eu. prefiro escolher os grupos

Figura 6.2: Resultados obtidos para a questão 1

Os resultados referentes à primeira questão colocada, demonstraram que a quase totalidade dos utilizadores considera relevante a criação de grupos baseados nos seus interesses. Como mostra a figura 6.2, apenas em dois casos a resposta foi negativa. O primeiro utilizador indicou a preferência por métodos explícitos de selecção, como justificação para a sua resposta negativa, embora neste caso o método implícito ou explícito de adesão, não estivesse em causa. O segundo utilizador associou o seu interesse por grupos aos membros que o compõem, afirmando ainda que não costuma utilizar grupos baseados em interesses.

Considera que os interesses sugeridos reflectem os conteúdos introduzidos ?

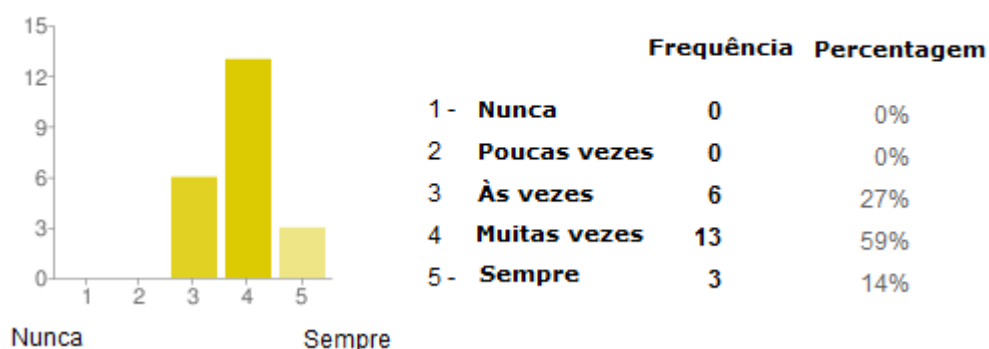


Figura 6.3: Resultados obtidos para a questão 2

Da análise dos resultados da segunda questão colocada (figura 6.3), resulta que a maioria dos utilizadores (59%), considerou que os interesses sugeridos reflectem frequentemente (4 em 5) os conteúdos introduzidos. A não existência de respostas para os níveis 1 e 2, indica também que mesmo os utilizadores que não consideram úteis os grupos de



interesses, tiveram sugestões válidas.

Outro facto interessante resulta do cruzamento desta questão com o facto de um utilizador, durante o teste, não ter sido atribuído a qualquer grupo. A não existência de respostas associadas ao valor 1 (nunca), permite conjecturar que esse utilizador recebeu sugestões válidas, mas optou por não seleccioná-las.

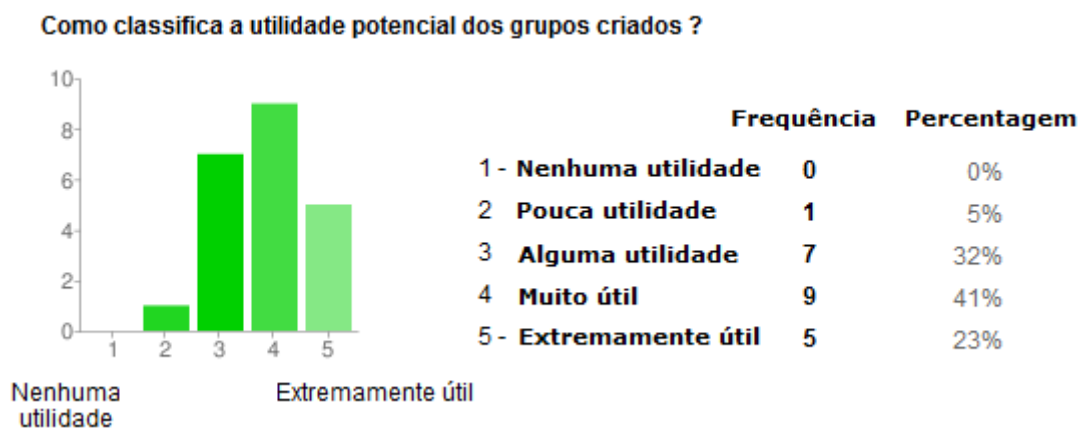


Figura 6.4: Resultados obtidos para a questão 3

A figura 6.4 apresenta os resultados derivados da terceira questão, revelando uma clara concentração das respostas nos níveis 3, 4 e 5. O nível 4 apresentou o número mais elevado de respostas (41%). Estes resultados indicam que os utilizadores possuem boas expectativas em relação à utilização futura do conceito, em aplicações baseadas na Internet.

**Como analisa a especificidade dos interesses sugeridos e respectivos grupos?**

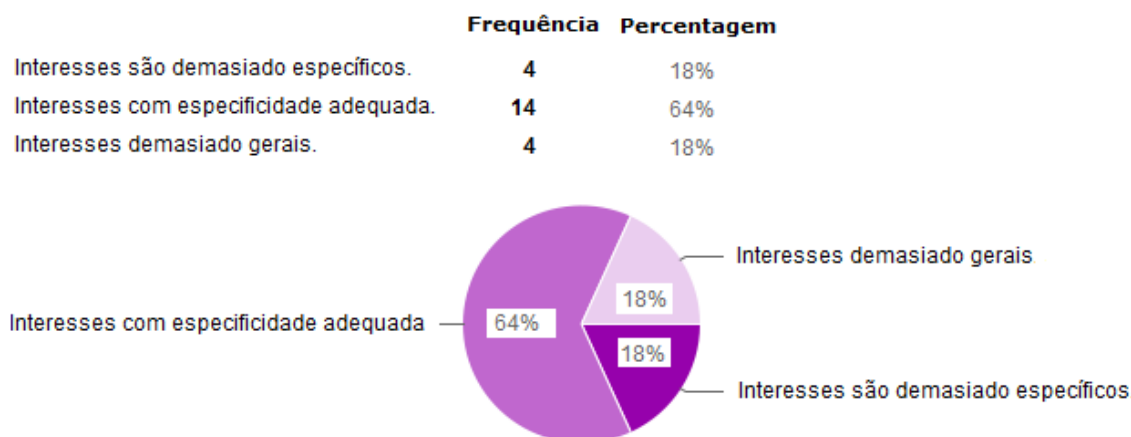


Figura 6.5: Resultados obtidos para a questão 4

As respostas dadas relativas à quarta questão colocada (ver figura 6.5), indicam que

64% dos utilizadores consideram adequada a especificidade dos interesses detectados. Dos restantes 36%, metade considerou os interesses demasiado específicos, enquanto a outra metade considerou o oposto. Este resultado reforça a necessidade do desenvolvimento de mecanismos que controlem a especificidade dos interesses detectados e sugeridos, como a possibilidade de controlar o tipo de interesses identificados, já discutida anteriormente (secção 6.1.1.1).

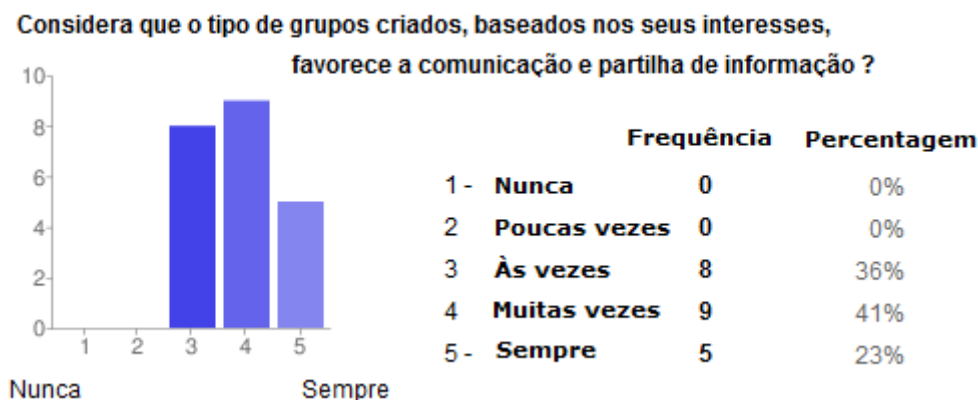


Figura 6.6: Resultados obtidos para a questão 5

Na última questão colocada, os resultados (ver figura 6.6) reflectem as opiniões positivas dos utilizadores, no que diz respeito ao impacte do tipo de grupos analisado, na promoção da interacção entre utilizadores e na partilha de ideias e conhecimentos. Alguns dos comentários deixados pelos utilizadores complementam também estes resultados, salientando a sua utilidade (ver figura 6.7).

#### Comentários/Opiniões

Gostei da facilidade de uso. Pena só funcionar com documentos em Inglês. Gostaria apenas de salientar o mérito desta ideia, parece-me bastante útil no contexto educativo e na promoção da partilha de conhecimento entre pessoas com os mesmos interesses. Ideia útil, pena so funcionar com artigos e notícias em lingua inglesa. Nos artigos relacionados a biotecnologia que foram os que utilizei, funcionou muito bem, pois sugeri os interesses que estavam relacionados com as notícias indicadas. Apenas queria deixar um aviso, visto que na página do perfil algumas caixas estão desalinhadas. Só queria a ...

Figura 6.7: Comentários e opiniões dos utilizadores inquiridos.

## 6.2 Conclusões

Os resultados dos testes de utilizadores efectuados revelaram indicadores importantes para a concepção da ferramenta de grupos implícitos, apesar do limitado número de

utilizadores teste. A qualidade da informação utilizada na identificação de interesses dos utilizadores, revelou-se determinante na geração de perfis estendidos que caracterizam correctamente os seus proprietários.

Os resultados decorrentes da distribuição dos utilizadores pelos grupos de interesse criados indicam que a criação de grupos baseados em interesses mais gerais pode beneficiar pequenas comunidades de utilizadores, enquanto a utilização de interesses mais específicos, funcionará melhor em comunidades de grandes dimensões.

As limitações identificadas, relacionadas com o componente de análise de conteúdos implementado, contribuíram para o ajuste arquitectural da ferramenta, reforçando a sua modularidade e flexibilidade.

Por outro lado, o resultado global dos inquéritos de satisfação, indicou uma aceitação positiva dos utilizadores em relação ao conceito de grupos implícitos baseados em interesses, deixando boas perspectivas para testes futuros de maior dimensão e significado.





## Conclusões e Trabalho Futuro

O trabalho desenvolvido anteriormente, em torno do conceito de grupos implícitos, demonstrou a sua aplicação e utilidade em cenários e ambientes específicos, antecipando o seu potencial no processo de desenvolvimento de funcionalidades de personalização em aplicações interactivas.

O trabalho realizado ao longo desta dissertação, o qual envolveu diversas fases, centrou-se na generalização do conceito de grupos implícitos, através da concepção e implementação de uma ferramenta extensível e flexível capaz de identificar e gerir grupos implícitos, que visa permitir a integração do referido conceito em diferentes aplicações interactivas e facilitar o desenvolvimento de funcionalidades de personalização nessas mesmas aplicações.

Após uma primeira fase relacionada com a necessária recolha bibliográfica e com o estudo das várias dimensões associadas ao conceito de grupos implícitos, conceptualizou-se uma ferramenta de suporte ao referido conceito, especificando-se as suas funcionalidades e definindo-se as abordagens e mecanismos necessárias à concretização dos objectivos estabelecidos.

Numa segunda fase, de concepção, foi implementada a ferramenta idealizada, a qual resulta na principal contribuição desta dissertação. O protótipo operacional concebido combina técnicas de análise dos conteúdos associados aos utilizadores e a análise das interacções que envolvem os mesmos ao longo do tempo, na recolha de informação implícita que os descreve. A informação recolhida é depois utilizada na extensão de perfis básicos definidos explicitamente pelos próprios utilizadores.

Os perfis estendidos de cada utilizador são, por sua vez, utilizados por vários mecanismos automáticos e dinâmicos, na identificação e gestão de grupos implícitos de utilizadores, que exibem interesses, características e interacções comuns. A arquitectura modular da ferramenta, resultante do padrão arquitectural definido, garante a sua extensibilidade, enquanto os mecanismos genéricos e configuráveis concebidos visam aumentar a sua flexibilidade e utilidade.

Numa terceira fase, de validação, foi idealizado um cenário de aplicação específico, através do qual foi concretizada a integração da ferramenta desenvolvida na plataforma colaborativa Moodle. A integração consistiu no desenvolvimento de um módulo de extensão da referida plataforma, o qual utiliza a ferramenta criada e a respectiva interface de programação, para oferecer novas funcionalidades aos utilizadores deste sistema colaborativo. Estas funcionalidades podem ser usadas por docentes e estudantes, no contexto de um Campus universitário, para partilhar facilmente artigos e notícias de investigação entre si.

Adicionalmente, a ferramenta concebida foi também utilizada numa avaliação preliminar do modo de utilização e da satisfação dos utilizadores, em relação aos grupos implícitos de interesses identificados. Esta avaliação consistiu em duas fases distintas. A primeira compreendeu um conjunto restrito de utilizadores de teste, os quais utilizaram a ferramenta desenvolvida testando as suas funcionalidades. Após os testes mencionados, numa segunda fase, cada utilizador foi inquirido sobre a satisfação e expectativas decorrentes da sua experiência.

Os resultados da avaliação preliminar indicam que as interacções entre utilizadores, numa dada aplicação e com a mesma, podem beneficiar da identificação e gestão de grupos implícitos. Do trabalho realizado resulta a convicção, de que os grupos identificados podem ser utilizados como estruturas dinâmicas de suporte à implementação de funcionalidades que se adaptam às actividades frequentes dos utilizadores, uma vez que consideram a evolução das suas características, interesses e relações ao longo do tempo. Do trabalho elaborado resultou também um artigo aceite para publicação [MRP11].

Dos estudos e avaliações efectuados emergem novas linhas de investigação, relacionadas com a obtenção de um processo automatizado de identificação de grupos implícitos, suficientemente flexível para se adaptar a diferentes aplicações e comunidades de utilizadores. Apesar da ferramenta desenvolvida ter dado os primeiros passos nesse sentido, oferecendo essa possibilidade, os mecanismos concebidos, neste domínio, podem ser refinados com vista à obtenção de uma maior eficiência e utilidade, por exemplo, através da utilização conjugada de diferentes componentes de análise de conteúdos.

Ainda no que diz respeito à identificação automática de grupos implícitos de utilizadores, trabalho futuro justifica a realização de testes com um conjunto de utilizadores mais alargado, a fim de se concluir sobre a viabilidade desta abordagem, na aquisição e organização de dados em ambientes dinâmicos de grande escala. A comprovação da flexibilidade e adaptabilidade da ferramenta exige também a sua aplicação em cenários e contextos diversificados.

# Bibliografia

- [AM07] Sarabjot Singh Anand e Bamshad Mobasher. Introduction to intelligent techniques for web personalization. *ACM Trans. Internet Technol.*, 7, October 2007.
- [Ama11] Amazon. Amazon. <http://www.amazon.com/>, 1996-2011.
- [AT05a] Gediminas Adomavicius e Alexander Tuzhilin. Personalization technologies: a process-oriented perspective. *Commun. ACM*, 48:83–90, October 2005.
- [AT05b] Gediminas Adomavicius e Alexander Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, 17(6):734–749, 2005.
- [Aut11] Automattic. Wordpress. <http://wordpress.com/>, 2011.
- [Bar54] J. A. Barnes. Class and committees in a norwegian island parish. *Human Relations*, 7(1):39–58, February 1954.
- [BAVR08] R. O. Briggs, P. Antunes, De G. J. Vreede, e A. S. Read. *Groupware: Design, Implementation, and Use*. Lecture Notes in Computer Science. Springer, December 2008.
- [BC11] Ludovico Boratto e Salvatore Carta. State-of-the-art in group recommendation and new approaches for automatic identification of groups. In Alessandro Soro, Eloisa Vargiu, Giuliano Armano, e Gavino Paddeu, editores, *Information Retrieval and Mining in Distributed Environments*, volume 324 of *Studies in Computational Intelligence*, pág. 1–20. Springer Berlin / Heidelberg, 2011.
- [BCC<sup>+</sup>09] Ludovico Boratto, Salvatore Carta, Alessandro Chessa, Maurizio Agelli, e M. Laura Clemente. Group recommendation with automatic identification of users communities. In *Proceedings of the 2009 IEEE/WIC/ACM International*

- Joint Conference on Web Intelligence and Intelligent Agent Technology - Volume 03, WI-IAT '09*, pág. 547–550, Washington, DC, USA, 2009. IEEE Computer Society.
- [BD11] Inc BigTent Design. Bigtent. <http://www.bigtent.com/>, 2011.
- [BE05] Ulrik Brandes e Thomas Erlebach. *Network Analysis Methodological Foundations*. Springer Berlin / Heidelberg, 2005.
- [BE07] Danah M. Boyd e Nicole B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):article 11, October 2007.
- [Ber71] Bernard Berelson. *Content Analysis in Communication Research*. Glencoe, Ill: Free Press, 1971.
- [BF10] Shlomo Berkovsky e Jill Freyne. Group-based recipe recommendations: analysis of data aggregation strategies. In *Proceedings of the fourth ACM conference on Recommender systems, RecSys '10*, pág. 111–118, New York, NY, USA, 2010. ACM.
- [Blo70] Burton H. Bloom. Space/time trade-offs in hash coding with allowable errors. *Commun. ACM*, 13:422–426, July 1970.
- [Boy11] Danah Boyd. Bibliography of research on social network sites. <http://www.danah.org/researchBibs/sns.php>, 2011.
- [CG08] Edward Curry e Paul Grace. Flexible self-management using the model-view-controller pattern. *IEEE Softw.*, 25:84–90, May 2008.
- [CLQ06] Daniel Cutting, Björn Landfeldt, e Aaron Quigley. Implicit group messaging in peer-to-peer networks. Relatório técnico, University of Sydney, School of Information Technologies, University of Sydney, 2006.
- [CLW05] Rob Cross, Jeanne Liedtka, e Leigh Weiss. A practical guide to social networks. *Harvard Business Review*, 83(3):124–132, 150, 2005.
- [CLWB01] Mark Claypool, Phong Le, Makoto Waseda, e David Brown. Implicit interest indicators. In *In Intelligent User Interfaces*, pág. 33–40. ACM Press, 2001.
- [CNN<sup>+</sup>10] Jilin Chen, Rowan Nairn, Les Nelson, Michael Bernstein, e Ed Chi. Short and tweet: experiments on recommending content from information streams. In *Proceedings of the 28th international conference on Human factors in computing systems, CHI '10*, pág. 1185–1194, New York, NY, USA, 2010. ACM.
- [Com11] SK Communications. Cyworld. <http://us.cyworld.com/>, 2011.



- [CQL08] Daniel Cutting, Aaron Quigley, e Björn Landfeldt. Spice: Scalable p2p implicit group messaging. *Comput. Commun.*, 31(3):437–451, 2008.
- [CSF11] Inc Cake Software Foundation. Cakephp. <http://cakephp.org/>, 2005-2011.
- [Fac11] Facebook. Facebook. <http://www.facebook.com/>, 2011.
- [Fie00] Roy T. Fielding. *Architectural Styles and the Design of Network-based Software Architectures*. Tese de Doutorado, University of California, Irvine, 2000.
- [FST02] Roy T. Fielding, Day Software, e Richard N. Taylor. Principled design of the modern web architecture. *ACM Transactions on Internet Technology*, 2:115–150, 2002.
- [GK09] Eric Gilbert e Karrie Karahalios. Predicting tie strength with social media. In *Proceedings of the 27th international conference on Human factors in computing systems, CHI '09*, pág. 211–220, New York, NY, USA, 2009. ACM.
- [Goo11a] Google. Google mail. <http://mail.google.com/>, 2011.
- [Goo11b] Google. Googleplus. <http://plus.google.com/>, 2011.
- ["gr10] "group". Oxford dictionaries. Oxford University Press, [http://oxforddictionaries.com/view/entry/m\\_en\\_gb0353180](http://oxforddictionaries.com/view/entry/m_en_gb0353180), 2010. [consultado em 3-10-2010].
- [Gra73] M.S. Granovetter. The strength of weak ties. *The American Journal of Sociology*, 78(6):1360–1380, 1973.
- [Har62] A. Paul Hare. *Handbook of small group research*. Free Press of Glencoe, 1962.
- [Hil08] Mireille Hildebrandt. Defining profiling: A new type of knowledge? In Mireille Hildebrandt e Serge Gutwirth, editores, *Profiling the European Citizen*, pág. 17–45. Springer Netherlands, 2008.
- [HKBR99] Jonathan L. Herlocker, Joseph A. Konstan, Al Borchers, e John Riedl. An algorithmic framework for performing collaborative filtering. In *SIGIR '99: Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, pág. 230–237, New York, NY, USA, 1999. ACM.
- [HKTR04] Jonathan L. Herlocker, Joseph A. Konstan, Loren G. Terveen, e John T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, 22:5–53, January 2004.
- [HKV08] Yifan Hu, Yehuda Koren, e Chris Volinsky. Collaborative filtering for implicit feedback datasets. In *Proceedings of the 2008 Eighth IEEE International*

- Conference on Data Mining*, pág. 263–272, Washington, DC, USA, 2008. IEEE Computer Society.
- [Inf11] Infopédia. grupo social. [http://www.infopedia.pt/\\$grupo-social](http://www.infopedia.pt/$grupo-social), 2003-2011. Consult. 2011-01-10.
- [Jam04] Anthony Jameson. More than the sum of its members: challenges for group recommender systems. In *Proceedings of the working conference on Advanced visual interfaces, AVI '04*, pág. 48–54, New York, NY, USA, 2004. ACM.
- [Kc06] Rohit Khare e Tantek Çelik. Microformats: a pragmatic path to the semantic web. In *Proceedings of the 15th international conference on World Wide Web, WWW '06*, pág. 865–866, New York, NY, USA, 2006. ACM.
- [KH10] Andreas M. Kaplan e Michael Haenlein. Users of the world, unite! the challenges and opportunities of social media. *Business Horizons*, 53(1):59 – 68, 2010.
- [KKMB05] J.A. Konstan, N. Kapoor, S.M. McNee, e J.T. Butler. Techlens: Exploring the use of recommenders to support users of digital libraries. 2005.
- [KP88] Glenn E. Krasner e Stephen T. Pope. A cookbook for using the model-view controller user interface paradigm in smalltalk-80. *J. Object Oriented Program.*, 1(3):26–49, 1988.
- [Lam08] Airi Lampinen. We are all here! : Multiple groups on a social network site. Tese de Mestrado, University of Helsinki, 2008.
- [LB10] Michele Satta Ludovico Boratto, Salvatore Carta. Groups identification and individual recommendations in group recommendation algorithms. In *ACM RecSys 2010 Workshop on the Practical Use of Recommender Systems, Algorithms and Technologies*, Barcelona, September 2010.
- [Ltd11] Kinesis Software Ltd. Kineticfusion. <http://www.kinesissoftware.com/>, 2011.
- [MCCC09] C. Morgado, J.C. Cunha, J. Custodio, e N. Correia. A group-based model for dynamic communities. In *Computational Aspects of Social Networks, 2009. CASON '09. International Conference on*, pág. 161 –164, June 2009.
- [MCCC10] C. Morgado, Jose C. Cunha, N. Correia, e J. F. Custódio. Group-based interactions for multiuser applications. *Scalable Computing Practice and Experience*, 11(1):63–71, 03 2010.
- [Mee11] Meetup. Meetup. <http://www.meetup.com/>, 2011.
- [MKR04] Bradley N. Miller, Joseph A. Konstan, e John Riedl. Pocketlens: Toward a personal recommender system. *ACM Trans. Inf. Syst.*, 22:437–476, July 2004.

- [MRP11] Jose C. Cunha Miguel Rangel Pais, Carmen Morgado. Implicit groups in web-based interactive applications. In *Proceedings of the Third International Conference on Computational Aspects of Social Networks*. Aceite para publicação, 2011.
- [MSC<sup>+</sup>06] Kevin McCarthy, Maria Salamó, Lorcan Coyle, Lorraine McGinty, Barry Smyth, e Paddy Nixon. Group recommender systems: a critiquing based approach. In *Proceedings of the 11th international conference on Intelligent user interfaces, IUI '06*, pág. 267–269, New York, NY, USA, 2006. ACM.
- [MVGD10] Alan Mislove, Bimal Viswanath, Krishna P. Gummadi, e Peter Druschel. You are who you know: inferring user profiles in online social networks. In *Proceedings of the third ACM international conference on Web search and data mining, WSDM '10*, pág. 251–260, New York, NY, USA, 2010. ACM.
- [Mys11] Myspace. Myspace. <http://www.myspace.com/>, 2011.
- [New06] M. E. J. Newman. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5):323–351, Maio 2006.
- [OCKR01] Mark O'Connor, Dan Cosley, Joseph A. Konstan, e John Riedl. PolyLens: a recommender system for groups of users. In *ECSCW'01: Proceedings of the seventh conference on European Conference on Computer Supported Cooperative Work*, pág. 199–218, Norwell, MA, USA, 2001. Kluwer Academic Publishers.
- [Orc11] Orchestr8. Alchemyapi. <http://www.alchemyapi.com/>, 2011.
- [Pal10] T. Palsma. Discovering groups using short messages from social network profiles. Tese de Mestrado, Univ. of Twente (The Netherlands), July 2010.
- [Par10] Ben Parr. Mark zuckerberg on the "biggest problem in social networking"[interview]. <http://mashable.com/2010/10/06/mark-zuckerberg-on-the-biggest-problem-in-social-networking/>, October 2010.
- [PPPS03] Dimitrios Pierrakos, Georgios Paliouras, Christos Papatheodorou, e Constantine D. Spyropoulos. Web usage mining as a tool for personalization: A survey. *User Modeling and User-Adapted Interaction*, 13:311–372, November 2003.
- [RBDD<sup>+</sup>10] Maayan Roth, Assaf Ben-David, David Deutscher, Guy Flysher, Ilan Horn, Ari Leichtberg, Naty Leiser, Yossi Matias, e Ron Merom. Suggesting friends using the implicit social graph. In *KDD '10: Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pág. 233–242, New York, NY, USA, 2010. ACM.
- [Reu11] Thomson Reuters. Opencalais. <http://www.opencalais.com/>, 2011.

- [SA09] Silvia Schiaffino e Analía Amandi. Artificial intelligence. chapter Intelligent user profiling, pág. 193–216. Springer-Verlag, Berlin, Heidelberg, 2009.
- [SKB<sup>+</sup>10] Christophe Senot, Dimitre Kostadinov, Makram Bouzid, Jérôme Picault, Armen Aghasaryan, e Cédric Bernier. Analysis of strategies for building group profiles. In Paul De Bra, Alfred Kobsa, e David Chin, editores, *User Modeling, Adaptation, and Personalization*, volume 6075 of *Lecture Notes in Computer Science*, pág. 40–51. Springer Berlin / Heidelberg, 2010.
- [Sof11] Alfresco Software. Alfresco. <http://www.alfresco.com/>, 2011.
- [STE01] Suwimol Sae-Tang e Vatcharaporn Esichaikul. Web personalization techniques for e-commerce. In *Proceedings of the 6th International Computer Science Conference on Active Media Technology, AMT '01*, pág. 36–44, London, UK, 2001. Springer-Verlag.
- [TDH10] Jaime Teevan, Susan T. Dumais, e Eric Horvitz. Potential for personalization. *ACM Trans. Comput.-Hum. Interact.*, 17:4:1–4:31, April 2010.
- [TMB09] Jaime Teevan, Meredith Ringel Morris, e Steve Bush. Discovering and using groups to improve personalized search. In *WSDM '09: Proceedings of the Second ACM International Conference on Web Search and Data Mining*, pág. 15–24, New York, NY, USA, 2009. ACM.
- [Tru11] Moodle Trust. Moodle. <http://moodle.org/>, 2011.
- [Twi11] Twitter. Twitter. <http://twitter.com/>, 2011.
- [TWL10] Lei Tang, Xufei Wang, e Huan Liu. Understanding emerging social structures - a group profiling approach. Relatório Técnico TR-10-002, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, 2010.
- [Uni11] International Telecommunication Union. The world in 2010: Ict facts and figures. <http://www.itu.int/ITU-D/ict/material/FactsFigures2010.pdf>, 2011.
- [Yah11a] Yahoo! Flickr. <http://www.flickr.com/>, 2011.
- [Yah11b] Yahoo! Yahoo!. <http://www.yahoo.com/>, 2011.
- [Zip32] G. K. Zipf. Selective studies and the principle of relative frequency in language, 1932.