



Independentemente dos seus conhecimentos atuais e de modo a que todos possam acompanhar, vamos começar do zero e começar por apresentar a Web Semântica no contexto de evolução da *World Wide Web* (WWW ou simplesmente Web) como a conhecemos hoje.

## Contexto

[Sir Tim Berners-Lee](#) e outros inventaram a Web em 1989. Nessa altura, e sob o ponto de vista do utilizador, a tecnologia-chave foi/é a **hiperligação**. Num determinado documento (e.g. este *post*), uma hiperligação permite através de um simples *click* que o utilizador consulte (quase) imediatamente o documento identificado nessa hiperligação (e.g. [informação sobre o inventor da Web](#)), independentemente da sua localização física (e.g. noutra máquina, noutra rede informática, noutra cidade, noutro país ou até mesmo noutro continente).

A Web original promove a adoção contínua e generalizada desta tecnologia na elaboração de documentos, contribuindo dessa forma para que todos os documentos fiquem interligados entre si. Por esta razão, a Web original é também vulgarmente denominada de Web 1.0 ou Web dos Documentos.

Na Web 1.0 as atividades dos utilizadores resumem-se basicamente a (i) pesquisar documentos de interesse, recorrendo a motores de pesquisa (e.g. [Google](#), [Bing](#)) e (ii) a consultar os documentos disponibilizados na Web tipicamente por organizações e utilizadores especializados. Contudo, o aparecimento de aplicações como fóruns, blogs (e.g. [Blogger](#)), redes sociais (e.g. [Facebook](#), [LinkedIn](#), [Twitter](#)) e Wikis (e.g. [Wikipedia](#)), entre outras, conduziu a uma alteração significativa nas atividades desempenhadas pelos utilizadores.



Consequentemente, todos os utilizadores passaram eles próprios a gerar conteúdos Web, desde anotações, comentários, classificações sobre (parte d') os documentos já existentes até à criação de novos documentos. O resultado desta evolução é *comumente* designado de Web 2.0 ou Web Social.

Ainda no contexto da Web Social, a expressão “*Enterprise 2.0*” refere-se ao empacotamento e aplicação das tecnologias advindas da Web 2.0 às organizações, nomeadamente ao uso de plataformas de software social dentro das organizações, ou entre organizações e os seus parceiros ou clientes.

Um problema intensificado por esta evolução relaciona-se com a quantidade, cada vez maior, de informação com que nós, humanos, somos confrontados e temos necessidade de lidar todos os dias (desde e-mails profissionais e pessoais a promoções, notícias, comentários das redes sociais, etc.).

Neste contexto, acabamos muitas vezes por despendir imenso tempo, por exemplo, (i) a tentar distinguir entre o que é relevante e menos relevante ou até não relevante de todo; entre o que deve merecer a nossa atenção imediata e menos imediata; entre o que é prioritário e menos prioritário; (ii)

a executar tarefas repetitivas que podiam/deviam ser (facilmente) automatizadas (e.g. dada uma lista extensa de algo, como restaurantes ou pessoas, verificar quais é que satisfazem um pré-determinado conjunto de critérios). Como resultado, o tempo disponível para a realização das atividades (realmente) importantes torna-se muitas vezes diminuto ou insuficiente.

Esta evolução veio também acentuar o problema da fragmentação da informação, tanto na Web como nas organizações. De uma forma simples, este problema reside na dificuldade de alguém obter uma perspetiva global, integrada e atualizada da informação existente sobre um dado objeto (e.g. uma pessoa, um projeto, um automóvel).



Vejamos dois exemplos: (i) sobre uma pessoa existe informação disponível, entre outros locais, na sua página pessoal, nas redes sociais onde está registada, nos seus blogs e na Autoridade Tributária; (ii) sobre um automóvel existe informação disponível, por um lado, nos vários sistemas informáticos usados na empresa fabricante como suporte às mais diversas atividades desde a sua concepção, passando pelo seu fabrico, gestão de qualidade, encomendas, vendas, marketing até ao serviço pós-venda e, por outro lado, pelos sistemas informáticos dos distribuidores e (re)vendedores, redes sociais, etc.

Na origem desta dificuldade estão essencialmente dois fatores. Por um lado, a diversidade existente entre as pessoas, as suas necessidades e práticas que conduzem ao uso de diferentes e variados serviços (e.g. as aplicações usadas por quem concebe um automóvel satisfazem requisitos e necessidades diferentes das aplicações usadas por quem é responsável pelo marketing ou pelas vendas desse mesmo automóvel). Por outro lado, o facto desses serviços atuarem como silos independentes, cada um com a sua terminologia e nível de detalhe adequado aos utilizadores-alvo e, portanto, cada um com o seu próprio formato/esquema de dados, a sua API, etc.

Como resultado, a informação de um dado objeto reside e fica dispersa por vários serviços e, muitas vezes até, em duplicado potenciando, assim, a sua desatualização. Por exemplo, quando uma pessoa atualiza o seu estado no **Facebook** o mesmo não é automaticamente atualizado nas restantes redes sociais onde essa mesma pessoa está registada. A solução para este problema passa basicamente por uma maior interoperabilidade e integração de dados entre as várias aplicações. Contudo, esta solução tende a ser um processo moroso e muito dispendioso.



Assim, torna-se necessário que sejamos ajudados por tecnologias e ferramentas que promovem, facilitam, agilizam e automatizam atividades como (i) a pesquisa e processamento da informação, (ii) a reutilização dos dados independentemente dos documentos onde esses dados se



encontram e, se possível, (iii) a inferência de nova informação com base na informação existente.

## Web 3.0 - Ideias e Objetivos

A Web 3.0, também conhecida como Web Semântica ou Web dos Dados, surge com o intuito de dar resposta aos principais desafios colocados pela Web Social, nomeadamente os descritos anteriormente.

E como é que a Web Semântica pretende fazer isso? De forma resumida e simples, a ideia é para além de se ligar e partilhar documentos conforme a perspectiva da Web original, passar-se também a identificar, ligar e partilhar os dados constantes nesses mesmos documentos. Tal, pretende possibilitar que os utilizadores acessem aos dados de uma forma transparente e independente da aplicação ou documento onde os dados residem.

Dito de outra forma, tal como a Web original abstraiu o acesso aos documentos das dificuldades relacionadas com as camadas (físicas e lógicas) de armazenamento e de interligação de redes de computadores, a Web Semântica pretende resolver as dificuldades advindas da interoperabilidade e integração de dados entre aplicações, bem como da capacidade de um computador ser capaz de compreender o significado desses dados e, por conseguinte, de processá-los adequadamente.

De acordo com [Sir Tim Berners-Lee](#), se a Web original fez com que todos os documentos online pareçam um enorme livro, a Web Semântica irá fazer com que todos os dados do mundo pareçam uma enorme base de dados. Nesta perspectiva, os utilizadores terão a possibilidade de formular inquéritos bastante mais complexos do que aqueles que realizam hoje em dia e obter uma resposta concreta a esse inquérito.



Por exemplo, inquirir quais são os restaurantes de comida oriental no Algarve cujo preço médio de refeição/pessoa é inferior a 25€ e cujo chefe já tenha recebido pelo menos uma distinção internacional e obter como resultado, de facto, uma lista de restaurantes que satisfazem os critérios pretendidos.

Infelizmente, hoje em dia, o resultado típico a um inquérito deste género é obter uma lista de restaurantes e de documentos que contêm as palavras-chave do inquérito, cabendo ao utilizador a tarefa de verificar um a um quais são os restaurantes dessa lista que satisfazem os critérios pretendidos. Felizmente a Web Semântica não é uma Web separada, mas sim uma extensão da Web atual, onde a informação tem um significado bem definido, permitindo dessa forma uma melhor cooperação entre computadores e pessoas, nomeadamente no que concerne à automatização de tarefas como a descrita anteriormente.

Para alcançar este objetivo, é fundamental que os dados constantes nos documentos sejam (ou possam ser) representados num modelo de dados comum e cuja semântica (i.e. significado e/ou compreensão) seja aceite e partilhada. Neste sentido, e sob um ponto de vista técnico, a Web Semântica

assenta principalmente em três recomendações da [World Wide Web Consortium \(W3C\)](#):

- [RDF - Resource Description Framework](#): modelo de dados para representação de informação, nomeadamente da Web Semântica;
- [SPARQL - Simple Protocol and RDF Query Language](#): linguagem para realização de inquéritos formais sobre dados representados em RDF;
- [OWL - Web Ontology Language](#): linguagem para definição de esquemas de dados ou para representação de conhecimento complexo, podendo esse conhecimento ser sobre qualquer coisa desde que a possamos identificar (teoricamente tudo pode ser identificado), grupos de coisas, e as relações entre essas coisas.

Estas recomendações e outras tecnologias serão abordadas nos próximos *posts*.

Esperamos que tenha gostado e deixe a sua opinião sobre o tema.



**Manuel Rocha**