

Hipermedia e Hipertexto

Introdução

Conceitos

Evolução

Tecnologias empregadas em Hipermedia

SGML

HTML

XML

Introdução

A relação entre o texto e seus leitores passou por diversas transformações ao longo da história. A distribuição, organização e estrutura dos textos encontrados nos livros em rolo da Antiguidade, nos manuscritos medievais e até mesmo nos impressos modernos e Contemporâneos, não são de forma alguma as mesmas que nos apresentam o **Hipertexto**. O objeto não é mais organizado em páginas, mas sim na tela do computador e o texto passa a ser eletrônico.

Na figura 1, temos um breve histórico da evolução da informação. Na Revolução da Linguagem temos a capacidade de passar as informações de maneira oral, porém extremamente ligada ao sujeito e, ainda, a sua disseminação de maneira oral dessa informação dificilmente mantém a correspondência com a sua original já que sem meios de registrá-la se torna dependente da memória de quem a possui e ainda passa muitas vezes pela interpretação de cada sujeito que a tem. Já na Revolução da Escrita, é possível saber ou registrar o que foi feito por pessoas em diferentes lugares e épocas. Existem agora registros escritos onde a originalidade da informação é mais fácil de ser mantida porém não muito simples de ser disseminada por poucas pessoas dominarem a escrita e pelo tempo despendido nesta tarefa. Com a Revolução da Imprensa, a automatização do processo de copiar textos escritos a informação começa a tomar um tom de universalidade, fazendo com que uma informação possa chegar a um número maior de pessoas ao mesmo tempo, trazendo conhecimento e informação. Com o advento da Internet estamos no auge da Revolução da Informação, é possível acessar praticamente todo o tipo de informação, livros, conhecimento de qualquer parte do mundo bastando apenas um computador ligado a rede mundial de computadores. Podemos acessar manuscritos medievais da Europa que estejam digitalizados, grandes enciclopédias que antes tinham um grande número de livros e que com o passar dos anos ficavam defasadas com alguns conceitos hoje estão acessíveis por um conjunto de mídias trazendo ainda além de informação escrita, imagens, vídeos, e sons para ajudar no aprendizado e na leitura, na comunidade científica temos uma grande disseminação de artigos publicados por Universidades o que nos permite acompanhar projetos em qualquer parte do planeta e uma grande gama de outras possibilidades.

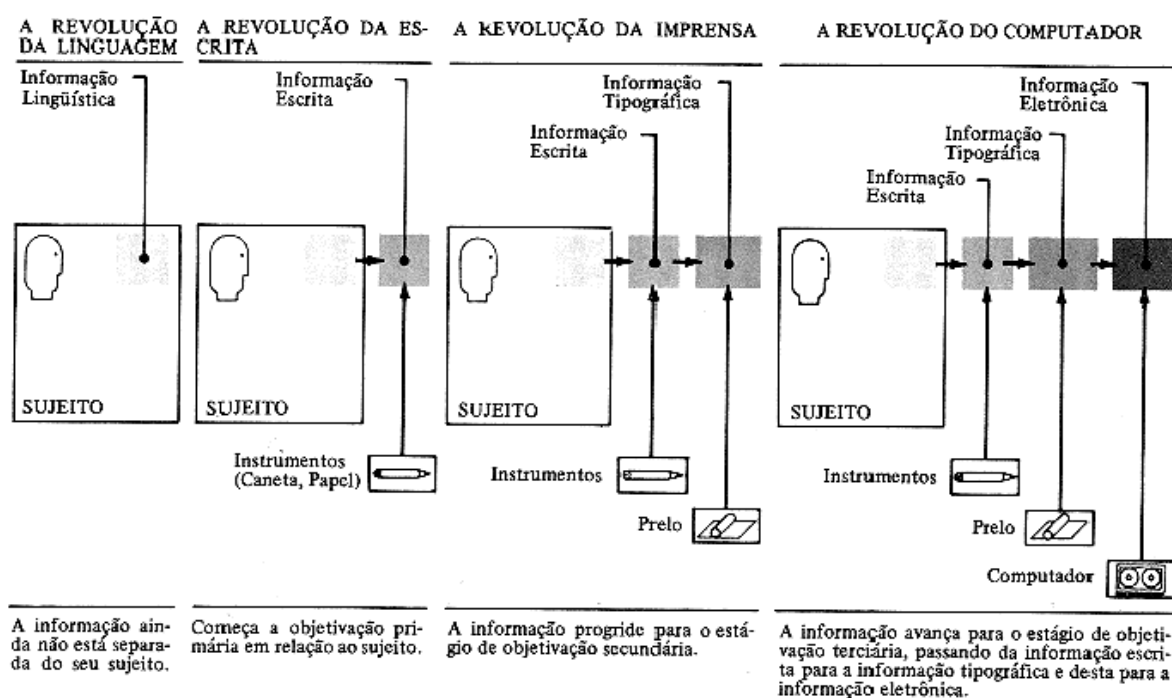


Figura 1 - Histórico da Evolução da Informação

Conceitos

Através da reunião de definições de diferentes autores sobre o termo hipertexto, pode-se dizer que o termo designa um processo de escrita/leitura não-linear e não hierarquizada e que permite o acesso ilimitado a outros textos de forma instantânea. Possibilita ainda que se realize uma trama, ou rede, de acessos sem seguir necessariamente, sequências ou regras. Na figura 2 e 3 temos uma representação das diferenças do hipertexto e do texto normal escrito e impresso.

Temos hoje o Hipertexto como base da Internet, tanto que a apresentação da grande maioria dos sites são baseadas em HTML (*HyperText Markup Language*) juntamente com outras tecnologias como XML e DHTML. Além da visualização de conteúdo textual em meio eletrônico temos ainda o conceito de Hipertexto e Multimídia juntos criando o que chamamos de **Hipermídia** trazendo novos recursos como imagens, vídeos e sons atrelados aos textos que ajudam a construir um tipo mais rico de conhecimento e informação.

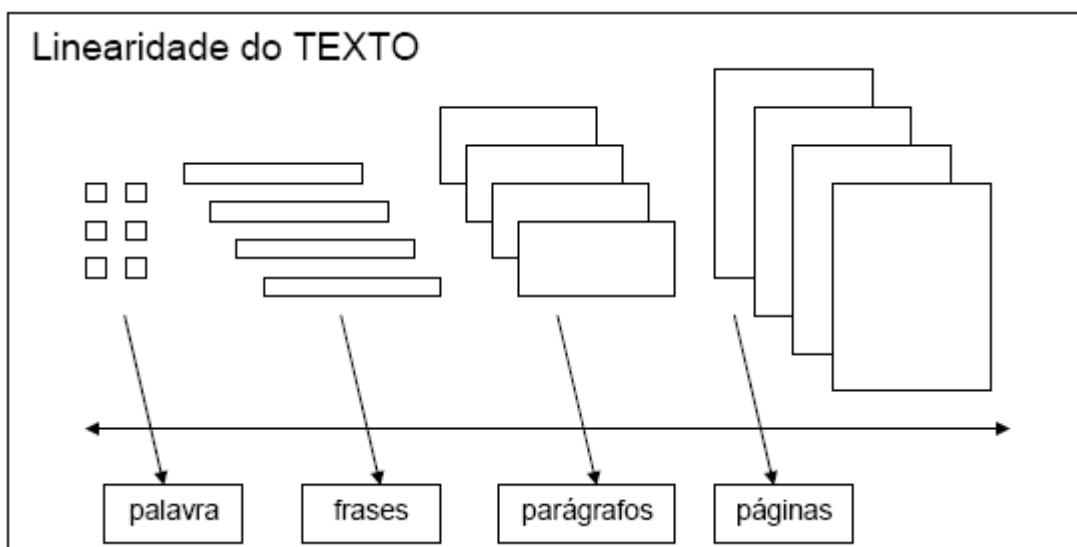


Figura 2 - Linearidade do Texto

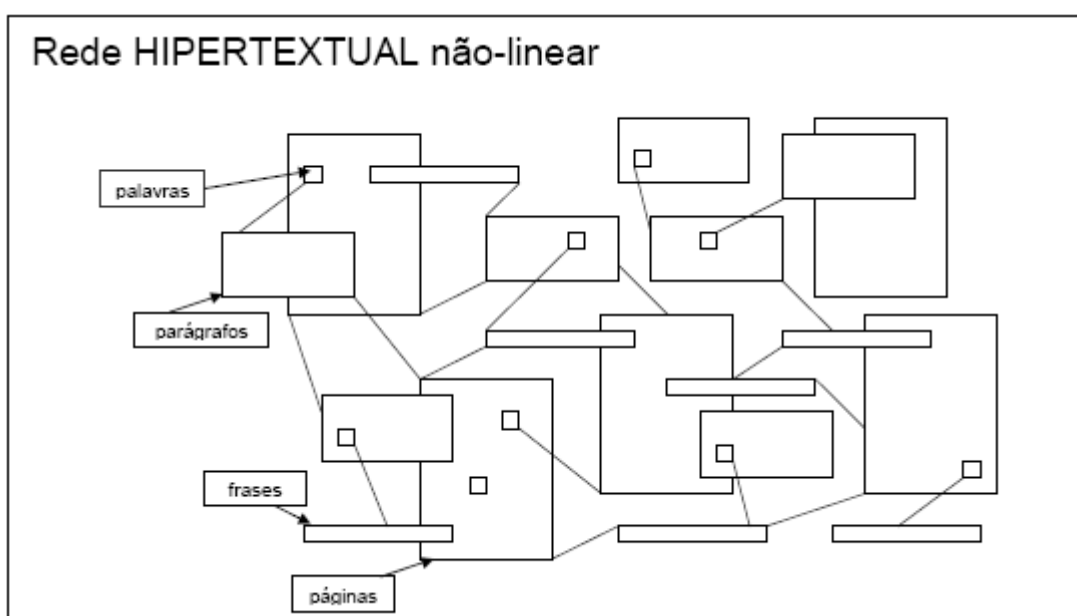


Figura 3 - Não-linearidade do Hipertexto

Tecnicamente, um hipertexto é um conjunto de nós ligados por conexões. Os nós podem ser palavras, páginas, imagens, gráficos ou partes de gráficos, seqüências sonoras, documentos complexos quem podem eles mesmos ser hipertextos.

Essa organização em nós, onde cada nó é uma informação com um formato diferente, reflete bem a não-linearidade e a falta de hierarquia da informação que o conceito de hipertexto prega já que o usuário tem a possibilidade de navegar entre os diversos nós sem que necessariamente haja uma ordem para isto. Essa navegação, são através de elos, conexões que ligam estes nós.

Estas conexões ou elos são definidos por uma das maiores autoridades a respeito. Estes elos são chamados de *links* e *hyperlinks*. O hipertexto pode ser

definido tamb m como um documento digital composto por diversos blocos de textos interconectados atrav s de *links*, que possibilitam o avan o da leitura de forma aleat ria. Na Web, cada endere o pode ser compreendido como um n  na rede, e os *links* podem remeter tanto para outros endere os, como para se  es ou outros n s dentro do pr prio endere o que se encontra como um grande portal de not cias onde na sua "capa" temos as principais manchetes com os *links* para a leitura das mat rias na sua  ntegra muitas vezes acompanhadas de fotos, v deos com entrevistas em cima daquela determinada mat ria ou gr ficos, por exemplo.

O termo hipertexto, entretanto, tem a sua origem nos anos 60 atrav s de Theodor H. Nelson e seu projeto Xanadu, cuja proposta era implementar uma rede de publica  es eletr nica, instant nea e universal – um verdadeiro sistema hipertexto, um universo documental. Conforme Nelson o hipertexto se constitui de escritas associadas n o-sequenciais, conex es poss veis de se seguir, oportunidades de leitura em diferentes dire  es. Isto pode ser representado muito bem nos dias de hoje atrav s dos conte dos veiculados atrav s da Internet, j  que temos endere os onde temos escritas n o-sequencias, conex es para outros endere os, ou conex es para dentro do pr prio endere o fazendo com que tenhamos oportunidade de leitura e acesso a informa  o em diferentes dire  es e tipo.

O hipertexto implica um texto composto de fragmentos de texto e os elos eletr nicos que os conectam entre si. Para ele, a express o **Hipermídia** estende a no  o de texto hipertextual ao incluir informa  o visual, sonora, anima  o e outras formas de informa  o.

Evolu  o

Em 1968, foi apresentado ao mercado o sistema de editora  o de texto Augment, considerado por muitos como o marco inicial do desenvolvimento do hipertexto. Desenvolvido no *Stanford Research Institute* come ava a implementar as id ias ditas por Nelson anos antes. O software tinha como recursos links entre diferentes arquivos, multiplas janelas controladas pelo usu rio entre outros. Vinha tamb m com uma interface mais amig vel homem-m quina e ainda tinha um suporte bastante interessante para trabalho colaborativo como compartilhamento de mensagens, mecanismos de mensagens eletr nicas e at  mesmo teleconfer ncia. Ainda no final desta d cada foi criada a ARPAnet que ser  detalhada posteriormente.

Em 1972, na *Carnegie Mellon University*, foi desenvolvido um sistema hipertexto distribuído (ZOG, mais tarde chamado de KMS – *Knowledge Management System*), onde n o havia uma separa  o r gida entre autores e leitores, de forma que as altera  es feitas por cada usu rio eram lidas por todos os outros. Ainda no final da d cada de 70 com o lan amentos dos primeiros PC's (Computadores Pessoais) da IBM e Apple e a inform tica invadindo escrit rios de grandes empresas surgiu o primeiro videodisco

hipermedia, uma espécie de hipertexto onde há links entre textos, sons e imagens, chamado Aspen Movie Map desenvolvido pelo MIT.

Na década de 80, temos o desenvolvimento de mídias óticas, videotexto, fibra ótica, a primeira tela sensível ao toque, o processador de texto do Macintosh da Apple com o paradigma do “*What You See Is What You Get*”, protocolos de comunicação e a própria base da Internet como conhecemos hoje. Com todos esses avanços, a melhoria das interfaces dos usuários com o computador, onde programas, menus, e janelas são acessíveis e manipuladas por um simples click no mouse permitiram o desenvolvimento de sistemas hipertexto comercialmente viáveis como:

- Guide – University of Canterbury – desenvolvido por Peter Brown, foi o primeiro sistema hipertexto para computadores pessoais;
- TIES (The Interactive Encyclopedia System, posteriormente chamado Hyperties) – University of Maryland – desenvolvido por Ben Schneiderman e usado em inúmeras aplicações, tais como exposições de museus, arqueologia, fotografia, manuais *on-line* etc;
- KMS (Sucessor do ZOG) – comercializado por *Knowledge Systems, INC.* para estações de trabalho Sun e Apollo;
- Notecards – Xerox Palo Alto Research Center – desenvolvido por Frank Halasz, Randy Trigg e Tom Moran, para suportar tarefas de leitura, categorização, interpretação e escrita de material técnico;
- Intermedia – Brown University – sistema para documentos multimídia desenvolvido pelo Institute for Research and Scholarship (Iris). Nesse sistema, os links pertenciam a webs (redes), de tal forma que, o usuário devia selecionar uma web específica. Os links eram, assim, dependentes de contexto;
- WE (Writing Environment) – University of North Carolina – projeto baseado em modelo de processos cognitivos envolvidos na autoria de textos, cujo objetivo era suportar todas as fases do processo de escrita de textos (conteúdo e estrutura);
- Hypercard – Apple Computer – Desenvolvido por Bill Atkinson, para computadores Macintosh.

No final da década, as primeiras teses de doutorado começam a ser defendidas e os primeiros eventos de grande porte específicos sobre o tema hipertexto começam a ser realizados como o workshop *Hypertext '87*.

A *World Wide Web*, é proposta em 1989 por Tim Berners-Lee. Sendo nada mais do que uma imensa biblioteca multimídia, ou seja, um grande conjunto de documentos hipertexto, com informações digitalizadas de textos, sons e imagens, conectados entre si e espalhadas por diversos computadores no mundo inteiro.

A década de 90 é marcada pelo desenvolvimento da linguagem de marcação extremamente difundida e sendo a base de todo o desenvolvimento de sites

que é o HTML (*Hypertext Markup Language*) e o protocolo de comunicação responsável pelas requisições via browser o http (HyperText Transfer Protocol) que possibilitaram a expansão da rede mundial de computadores e a disseminação de documentos hipertexto fazendo com que a Internet saísse dos meios acadêmicos e militares e se tornasse um instrumento de acesso mundial. Começaram nessa década ainda a utilização de mídias como CD-Rom com conteúdos para serem estudados, armazenados e pesquisados como grandes enciclopédias. Já na virada para o ano 2000 temos a Internet como um grande provedor de todo o tipo de serviços como bancos, conteúdo de ensino, compras on-line entre outros. E o panorama atual através de tecnologias como RSS, XML, Bancos de dados modernos e ágeis, linguagens de programação extremamente completas temos condições de ter ambientes de estudos todos a distância, temos sistemas de informação totalmente desenvolvidos em plataforma Web. Com o acesso facilitado a banda larga a cada ano que passa, temos ainda um conteúdo multimídia muito extenso sendo absorvido como vídeos, filmes, músicas, noticiários ao vivo pela Internet e muito mais que envolve a hiperMídia.

Tecnologias empregadas em HiperMídia

Um padrão para representação de documentos deve considerar, além da estrutura interna do mesmo, também a possibilidade de ligação entre vários documentos. Essas ligações podem ocorrer, por exemplo, em referências bibliográficas que levam aos documentos originais, ou ainda em palavras-chave, que são interligadas a documentos mais específicos. Quando o leitor ativa/processa uma ligação, ele é levado a outro ponto no mesmo documento ou a um novo documento.

A padronização do formato de representação de diferentes tipos de documentos facilita o intercâmbio das informações contidas nas instâncias de documentos específicos. Esta característica é extremamente importante, visto que o mercado de aplicações hiperMídia é atualmente composto por inúmeros fabricantes que utilizam tecnologia proprietária para a codificação e armazenamento de seus documentos.

Serão apresentados breves resumos de alguns padrões utilizados em documentos hipertextos e hiperMídia como SGML, HTML, XML e RSS.

SGML

O **Standard Generalized Markup Language (SGML)** é uma metalinguagem através da qual se podem definir linguagens de marcação para documentos. A SGML é uma descendente da Generalized Markup Language (GML) da IBM, desenvolvida na década de 60 por Charles Goldfarg, Edward Mosher e Raymond Lorie (cujas iniciais dos nomes próprios por acaso coincidem com GML). A SGML não deve ser confundida com a Geography Markup Language (GML) desenvolvida pelo consórcio Open GIS.

A SGML providencia uma variedade de sintaxes de marcação que podem ser usadas por várias aplicações. Ao alterar a Declaração SGML, deixa de ser necessário recorrer aos caracteres "<" e ">", apesar de serem o padrão.

A SGML foi inicialmente concebida para permitir a partilha de documentos que permitissem a leitura por máquina em projetos de grande dimensão governamentais e na indústria aeroespacial, que necessitam de permanecer legíveis por várias décadas -- um intervalo de tempo muito longo no que diz respeito à tecnologia de informação. Também tem sido usado extensivamente nas indústrias de impressão e publicação, mas a sua complexidade tem prevenido a sua difusão em outros campos.

Exemplo da sintaxe SGML:

```
<QUOTE TYPE="example">  
  typically something like <ITALICS>this</ITALICS>  
</QUOTE>
```

A SGML é uma norma ISO: "ISO 8879:1986 Information processing--Text and office systems--Standard Generalized Markup Language (SGML)". HTML e XML são ambas derivadas do SGML. Enquanto a HTML é uma aplicação da SGML, a XML é um perfil--um subconjunto específico da SGML, projetada para ser mais simples de se analisar gramaticalmente e de se processar do que SGML. Outra linguagem de marcação criada originalmente como uma aplicação da SGML é DocBook, desenhada para a edição de documentação técnica. DocBook é atualmente disponível como uma aplicação XML.

A XML é uma tentativa de simplificar SGML para aplicações de aplicação geral, tais como a Semantic Web¹. A XML tem sido usada num largo número de aplicações, incluindo as notáveis XHTML, RSS, XML-RPC e SOAP.

Existem também várias outras linguagens que são, pelo menos em parte, relacionadas com a SGML e XML, mas, por não poderem ser analisadas gramaticalmente, validadas ou processadas com outras ferramentas SGML e XML, não podem ser consideradas aplicações da SGML ou XML. Um exemplo é a Z Format, uma linguagem desenhada para composição e documentação.

As principais vantagens da utilização do SGML são:

- Possibilidade de representação da estrutura hierárquica dos elementos em um documento. Esta característica facilita quaisquer operações de processamento sobre os documentos;
- Flexibilidade dos documentos. Os tipos de elementos e relacionamento entre eles é definido pelo projetista da classe de documentos;

- Especificação formal dos elementos contidos no documento SGML. Os elementos são declarados de modo formal no DTD. O programa que utilizar o documento SGML processa esse documento em conjunto com a declaração, o que permite validá-lo;
- Conteúdo textual manipulado em qualquer editor e compreendido com facilidade.

Entretanto as principais desvantagens de SGML são:

- SGML não é poderoso o suficiente para lidar com documentos multimídia e hiperMídia, que podem possuir relações temporais ou ligações hipertextos complexas;
- SGML é difícil de utilizar.

HTML

HTML (*HyperText Markup Language*) tornou-se uma linguagem amplamente utilizada como formato para hiperdocumentos disponibilizados na WWW e reconhecida pelos browsers existentes. A partir da versão 2.0, HTML é formalmente definida por um DTD SGML, sendo parte do conjunto de especificações do World Wide Web Consortium (W3C). Assim, o DTD determina a maneira pela qual um conjunto de elementos deve estar estruturado para compor um documento, bem como a definição de links entre os elementos e a inclusão de outras mídias no corpo do documento.

A navegação em documentos HTML se faz através das seguintes operações básicas definidas pela tecnologia: seleção livre de relacionamento, complementada por operações de *backtracking* e controle de caminho percorrido (histórico de navegação). Tentando proporcionar maior interação aos documentos HTML, a especificação atual inclui o suporte a outras tecnologias, por exemplo, folhas de estilo e scripts.

A maioria dos documentos existentes hoje na WWW não está em completa conformidade com DTD SGML que define HTML. Este fato ocorre porque os browsers atuais processam documentos HTML com flexibilidade e sem utilizar todas as características de validação exigidas por SGML.

A simplicidade do padrão HTML implica uma série de limitações:

- HTML não é extensível. Um padrão extensível permite que os desenvolvedores de aplicações definam elementos específicos para situações diversas;
- HTML não possui estrutura semântica alguma, o que dificulta o processamento ou reuso do conteúdo dos documentos;
- HTML mistura estruturas lógicas do documento (título, parágrafos, etc) com elementos de apresentação (negritos, itálicos, alinhamentos, etc);

- Documentos HTML com ligações embutidas não permitem a separação dos dados sendo referenciados e de suas referências (links), dificultando a reutilização destes sem a herança obrigatória das relações;
- Em documentos HTML só podem ser estabelecidas ligações do tipo “ponto a ponto” unidirecionais, sempre expressando uma relação do tipo “go to”.

XML

XML (eXtensible Markup Language) é uma recomendação da W3C para gerar linguagens de marcação para necessidades especiais.

XML é um subtipo de SGML (*Standard Generalized Markup Language* - Linguagem Padronizada de Marcação Genérica) capaz de descrever diversos tipos de dados. Seu propósito principal é a facilidade de compartilhamento de informações através da Internet. Linguagens baseadas em XML: XHTML, RDF, SMIL, MathML, NCL, XBRL, XSIL e SVG. O SVG, por exemplo, é um formato gráfico vetorial (*Scalable Vectorial Graphics*).

Estimulado pela insatisfação com os formatos existentes (padronizados ou não), um grupo de empresas e organizações que se autodenominou World Wide Web Consortium (W3C) começou a trabalhar em meados da década de 1990 em uma linguagem de marcação que combinasse a flexibilidade da SGML com a simplicidade da HTML. O princípio do projeto era criar uma linguagem que pudesse ser lida por software, e integrar-se com as demais linguagens.

Sua filosofia seria incorporada por vários princípios importantes:

- Separação do conteúdo e da formatação;
- Legibilidade tanto por humanos quanto por máquinas;
- Possibilidade de criação de tags sem limitação;
- Criação de arquivos para validação de estrutura;
- Com seu uso pode-se interligar bancos de dados distintos;
- Simplicidade;
- XML concentra-se na estrutura da informação e não na sua aparência.

Um banco de dados pode, através de uma aplicação escrever em um arquivo xml, e um outro banco distinto pode ler então estes mesmos dados. Esse exemplo demonstra a sintaxe flexível do XML sendo usada para descrever uma receita de pão.

```
<?xml version="1.0" encoding="UTF-8"?>

<Receita nome="pão" tempo_de_preparo="5 minutos" tempo_de_cozimento="1 hora">
  <titulo>Pão simples</titulo>
  <ingrediente quantidade="3" unidade="xícaras">Farinha</ingrediente>
  <ingrediente quantidade="7" unidade="gramas">Fermento</ingrediente>
  <ingrediente quantidade="1.5" unidade="xícaras" estado="morna">Água</ingrediente>
  <ingrediente quantidade="1" unidade="colheres de chá">Sal</ingrediente>
  <Instruções>
    <passo>Misture todos os ingredientes, e dissolva bem.</passo>
    <passo>Cubra com um pano e deixe por uma hora em um local morno.</passo>
    <passo>Misture novamente, coloque numa bandeja e asse num forno.</passo>
  </Instruções>
</Receita>
```

Onde temos na 1ª linha

```
<Receita      nome="pão"      tempo_de_preparo="5      minutos"
tempo_de_cozimento="1 hora">
```

"Receita" é o nome principal para o seu documento. Note que a semelhança entre XML e HTML é grande, na 1ª linha abrimos a tag Receita e na última linha fechamos a mesma, como em HTML e assim se estende por todo o exemplo.

O W3C vêm investindo inúmeros esforços com o objetivo de garantir que o crescimento da WEB se dê de modo aberto e padronizado, por exemplo, ao definir recomendações que regulam a formalização da estrutura de documentos e do seu formato de apresentação. Especificamente, XML permite que projetistas de aplicações definam uma estrutura lógica para os documentos suportados. O objetivo é prover interoperabilidade entre aplicações e, assim, XML pode ser vista como a proposta do W3C de modelo de representação de dados para aplicações da Web. O uso de XML como formato de intercambio de dados fornece muitas vantagens. Em particular, do ponto de vista de projeto de aplicações na Internet em geral – e não apenas na Web, o uso de XML permite explorar as seguintes vantagens:

- Interoperabilidade entre aplicações: não é necessário que aplicações de terceiros tenham conhecimento da estrutura de uma base de dados em particular: é suficiente ter acesso a uma especificação XML, que define a sintaxe da informação para poder processar as instancias que apresentam seu conteúdo;
- Processamento distribuído: aplicações na Internet estão ficando cada vez mais comuns, e o uso de XML para transferência de informações entre os diversos módulos permite maior controle no processamento das informações pelos módulos;
- Processamento customizado: aplicações distintas ou módulos de aplicações distribuídas podem apresentar as informações recebidas de acordo com a preferência e/ou necessidade de seu usuário local;

- Pesquisa independente: o intercambio de documentos permite a busca de informações de interesse local, de maneira mais independente e eficiente por parte da aplicação;
- Trafico reduzido: o intercambio de documentos estruturados permite o processamento local da informação pela aplicação e assim, possibilita uma redução na quantidade de dados transmitidos pela rede.

Como os documentos XML não possuem recursos para especificar informações a respeito da apresentação de seu conteúdo, CSS (*Cascading Style Sheets*) e XSL (*Extended Style Language*) podem ser utilizadas para permitir que qualquer tipo de informação possa ser associado ao mesmo. Um exemplo é a criação de um documento HTML que contém informações de apresentação a partir do conteúdo de um documento XML, através do processamento de uma folha de estilo XSL.

