



Comércio Electrónico e Contratação Electrónica

Mestrado em Direito em Informática

Escola de Engenharia

Objetivos

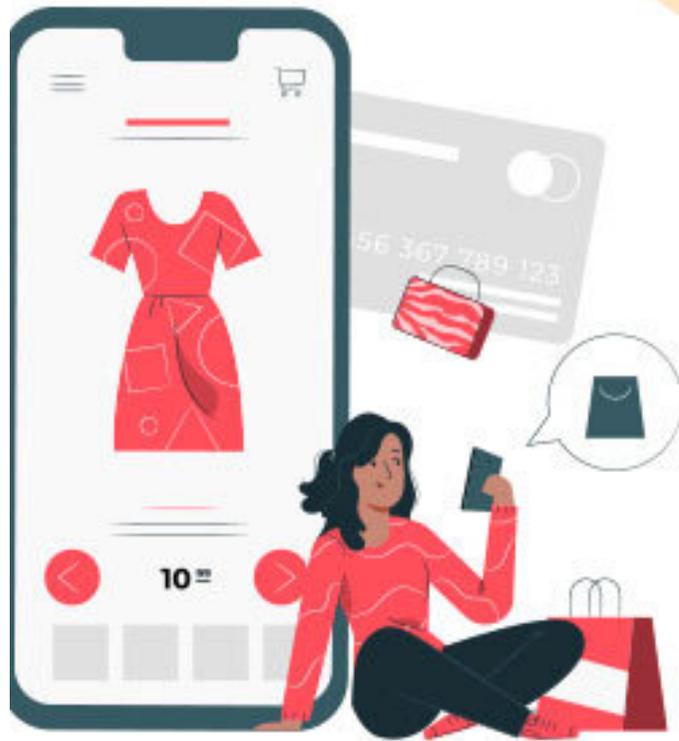
- Comércio e Comércio Eletrónico
- Os 3 Es (E-Commerce, E-Business, E-Governance)
- Mercados
- Cadeias de valor
- Custos das transações

Objetivos

- Vantagens e Desvantagens do Comércio Eletrónico
- 1ª e 2ª onda do Comércio Eletrónico
- Categorias

eCommerce Versus Traditional Commerce

Which is Better for Your Business?



eCommerce Versus



**Comércio e Comércio
Eletrónico**

Traditional Commerce

Comércio

The exchange or buying and selling of commodities; esp. the exchange of merchandise, on a large scale, between different places or communities; extended trade or traffic.

E- Commerce

Electronic commerce (e-commerce) is a general term for any type of business, or commercial transaction that involves the transfer of information across the Internet. This covers a range of different types of businesses from consumer-based retail sites, like Amazon.com, through auction and music sites like eBay or MP3.com, to business exchanges trading goods or services between corporations.

Definições

Electronic commerce is the use of electronic communication to do business.

E-commerce is not about technology. It is not a new business. E-commerce is a method for companies to create and operate their business in new and efficient ways.

Definições

Algumas pessoas usam o termo e-business para caracterizar todas as categorias de e-commerce.

e-Business:

- *The transformation of key business processes through the use of Internet technologies (IBM)*

E- commerce

E-Commerce or Electronics Commerce is a methodology of modern business which addresses the need of business organizations, vendors and customers to reduce cost and improve the quality of goods and services while increasing the speed of delivery. E-commerce refers to paperless exchange of business information using following ways.

- Electronic Data Exchange (EDI)
- Electronic Mail (e-mail)
- Electronic Bulletin Boards
- Electronic Fund Transfer (EFT)
- Other Network-based technologies

Comércio Vs E-Commerce

Heavy dependency on information exchange from person to person.	Information sharing is made easy via electronic communication channels making little dependency on person to person information exchange.
Communication/ transaction are done in synchronous way. Manual intervention is required for each communication or transaction.	Communication or transaction can be done in asynchronous way. Electronics system automatically handles when to pass communication to required person or do the transactions.
It is difficult to establish and maintain standard practices in traditional commerce.	A uniform strategy can be easily established and maintain in e-commerce.
Communications of business depends upon individual skills.	In e-Commerce or Electronic Market, there is no human intervention.
Unavailability of a uniform platform as traditional commerce depends heavily on personal communication.	E-Commerce website provides user a platform where all information is available at one place.
No uniform platform for information sharing as it depends heavily on personal communication.	E-Commerce provides a universal platform to support commercial / business activities across the globe.

Trocas de Informação

Tipos de Transações

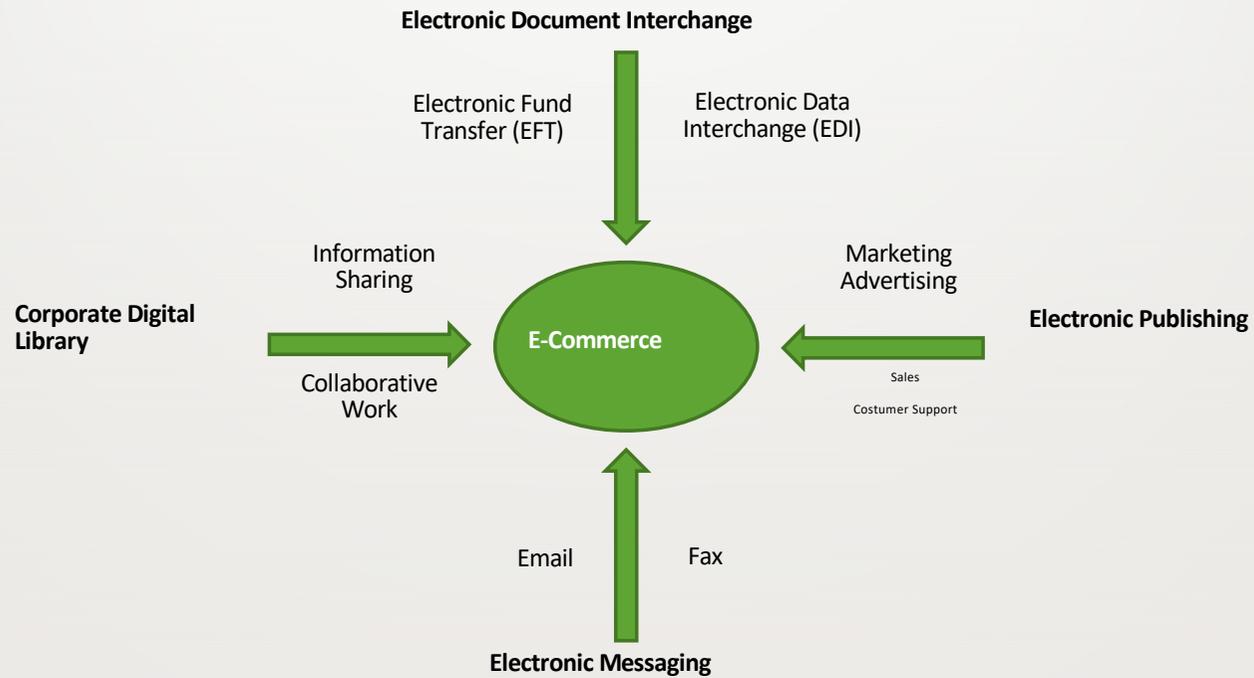
Tipos de Estratégias

Intervenção Humana

Informação Acessível

Vendas Internacionais

E-Commerce





Principais Características do E-commerce

- Pagamentos eletrônicos
- Disponibilidade 24x7 h
- Informação / Marketing
- Vendas automatizadas
- Assistência 24h
- Gestão de inventários
- Comunicação com os clientes

Vantagens do Comércio Electrónico

- Aumentar as vendas diminuindo os custos
- Oportunidades de negócio
- Aumenta a velocidade e a precisão da informação

+Vantagens

- 24 horas por dia
- Nível de detalhe pode ser escolhido pelo utilizador
- Entrega imediata dos produtos digitais
- Reembolso de impostos e menores custos (Internet)
- Produtos e serviços disponíveis em áreas remotas (e-learning)

Vantagens para as organizações

- Internacionalização / globalização a baixo custo
- Redução de custos operacionais
- Imagem da empresa
- Melhoria do serviço ao consumidor
- *Business Process* mais eficiente e mais rápido
- Sem papel
- Aumento da produtividade
- *“just-in-time”*

Vantagens para o consumidor

- 24 x 7 h
- Mais opções e entrega mais rápida do produto
- Maior comparação e maior facilidade em escolher o mais barato ou a melhor opção
- Possibilidade de colocar *review comments*
- Leilões virtuais
- Informação legível mais disponível
- Maior competitividade entre as organizações

Vantagens para a sociedade

- Menor tráfego rodoviário
- Menor poluição
- Redução dos preços e aumento do acesso aos produtos
- Acesso a produtos em áreas rurais
- Os governos podem disponibilizar serviços eletrónicos, por exemplo na Saúde, educação, serviços sociais, mais baratos e mais eficientes.

+ Desvantagens

- Problemas com alguns produtos (e.g., pericíveis) apesar de haver super e hipermercados a vender na Net
- A evolução tecnológica
- Alguns produtos têm de ser vendidos em quantidade

+ Desvantagens

- Investimento inicial
- Dificuldades em integrar bases de dados e sistemas de processamento transaccional nas soluções
- Obstáculos legais e culturais (cartões de crédito, resistência à mudança, legislação ainda não desenvolvida de forma clara).
- Perfil de transporte (alguns produtos são difíceis de embalar e transportar)

Desvantagens técnicas

- Falta de segurança, fiabilidade ou standards devido a fracas implementações
- Em alguns países, a Internet é lenta
- Pode haver a necessidade de um software específico
- Dificuldade em integrar os sistemas com as bases de dados ou aplicações existentes
- Incompatibilidades de software e/ou hardware

Desvantagens Não Técnicas

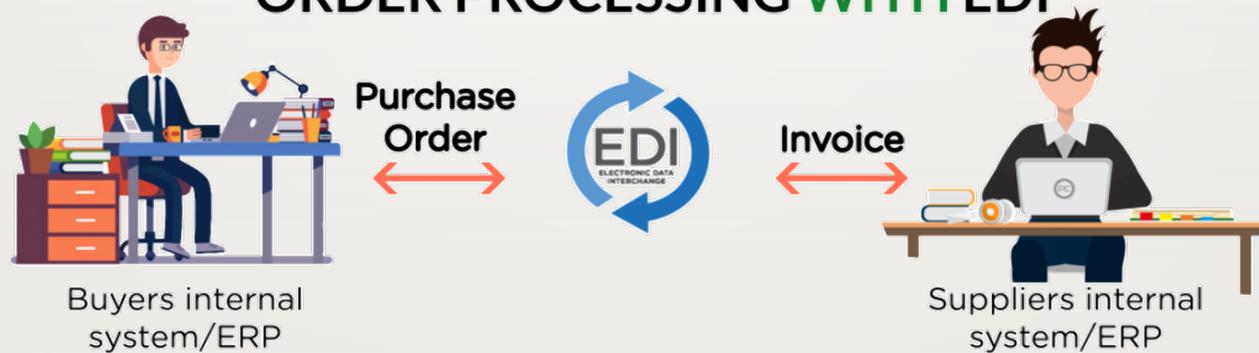
- Custo inicial
- Falta de experiência
- Resistência dos utilizadores
- Segurança / privacidade no pagamento
- Dificuldade em avaliar o produto por imagens / fotografias
- As aplicações estão em constante evolução
- Acesso à internet caro e nem sempre disponível em regiões mais remotas

Electronic Data Interchange (EDI)

ORDER PROCESSING **WITHOUT** EDI



ORDER PROCESSING **WITH** EDI



Adaptado de <https://www.horizoncore.com/blog/adopt-the-electronic-data-interchange-edi-and-be-successful/>.

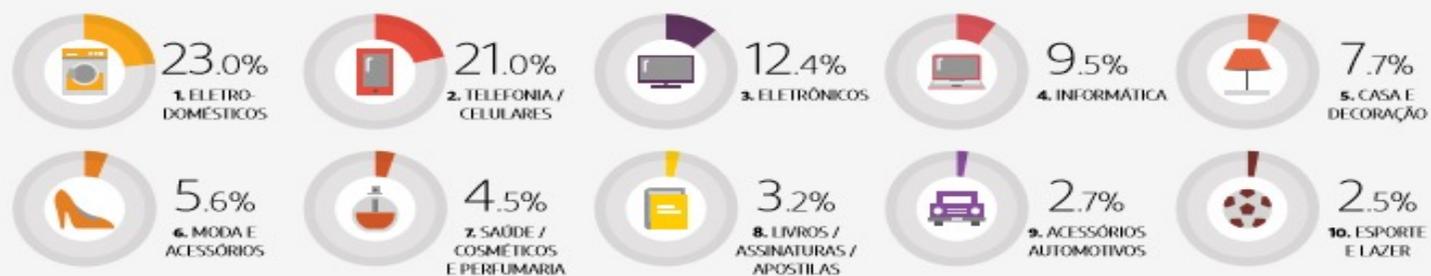
Categorias Mais Vendidas

EM VOLUME DE PEDIDOS



FONTE: EBIT INFORMAÇÃO

EM VOLUME FINANCEIRO



FONTE: EBIT INFORMAÇÃO

Em Suma

- ✓ Definições de Comércio Tradicional e E-commerce
- ✓ Diferentes Formas de troca de informação no E-commerce
- ✓ Principais Características do E-commerce
- ✓ Vantagens do E-commerce (Organizações, Consumidor, Sociedade)
- ✓ Desvantagens do E-commerce (Técnicas e Não Técnicas)
- ✓ Importância do *Electronic Data Interchange* (EDI)
- ✓ Categorias Atualmente Mais Vendidas através do E-commerce

Combinções

- Comércio tradicional
 - Chamados “produtos por impulso” (joias, antiguidades) que precisam de toque, cheiro, examinação prévia...
 - Pequenas compras (pagamento de pequenas quantias)

A 1ª Onda

- 1990s a 2003
- dot-com boom com um investimento estimado em 100 biliões de dólares (1990s – 2000)
- The dot-com bust occurred in 2000
- Anos tristes 2000 – 2003 (investimento estimado em 200 biliões de dólares)

1ª Onda

- Fenómeno U.S.
- Páginas Web em Inglês
- Internet lenta e barata
- Uso de códigos de barra e scanners para rastreamento (tracking)
- Email
- Publicidade era a principal receita



A 2ª Onda

- 2003 – (nova vida)
- Empresas como a Amazon.com (livros) e eBay.com (leilões) sobreviveram e começaram a apresentar lucros
- Aumento contínuo nas vendas todos os anos

2ª Onda

- Muitos países e muitas línguas
- Internet mais rápida
- Uso de RFID em vez de barcode
- Fingerprint readers e retina scanners (biometria) usados no rastreamento
- Email é uma parte integrante do negócio

2ª Onda

- É uma parte integral do marketing e das estratégias para contactar clientes
- Substituiu as formas tradicionais de “avisar”
- Problemas com as línguas e as moedas

The Three Waves of Electronic Commerce



First Wave: Traditional EDI

Ordering
Shipping
Invoicing
Inventory
Established
partners only



Second Wave: Electronic Commerce

Elements in 1st wave plus:
Online shopping
Online payments
Increased Information
Sharing
New partners allowed
Interactive Websites

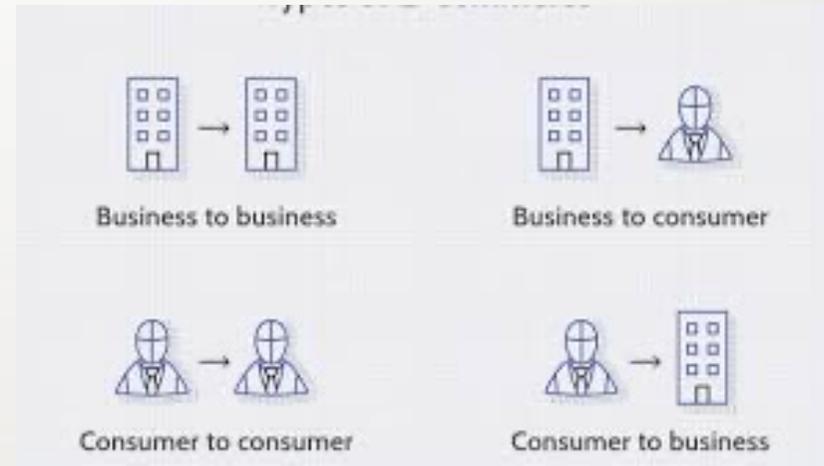


Third Wave: Electronic Society

Elements in 2nd wave
plus:
Cashless transactions
Transaction integrity
Intelligent Agents
Continuous testing
Wireless capabilities

Categorías

- Foram definidas cinco categorías:
 - Business to Consumer (B2C)
 - Business to Business (B2B) (e-procurement)
 - Business processes that support buying and selling activities
 - Consumer-to-consumer (C2C)
 - Business-to-government (B2G)



B2C e-commerce

- Descrição
 - Negócios / empresas vendem produtos ou serviços a clientes individuais
- Exemplo
 - www.walmart.com



B2B e-commerce

- Descrição
 - Negócios vendem produtos ou serviços a outros negócios
- Exemplo
 - www.grainger.com

Business Processes that Support Buy/Sell Activities

- Descrição:

Empresas ou organizações mantêm e usam informação para identificar, avaliar clientes, fornecedores e colaborador (para apoiar vendas, compras, contratação, planejamento, etc...)

Exemplo:

A Dell partilha informação de vendas com os seus fornecedores que a usam para planear a produção.

C2C e-commerce

- Descrição
 - Os participantes num mercado online pode comprar ou vender bens uns aos outros
- Exemplo:
 - www.ebay.com



B2G e-commerce

- Descrição
 - Empresas vendem bens ou serviços ao Estado ou a organismos públicos

Exemplo

- A empresa de informática que vende computadores pode desenvolver as seguintes actividades na Internet:
 - Vender computadores a pessoas(**B2C**)
 - Comprar matéria prima ou peças a fornecedores(**B2B**)
 - Contratar pessoas, gerir fornecedores, emitir avisos, etc. (**Business processes**)
 - Vender computadores a escolas (**B2G**)
 - No eBay, as pessoas compram e vendem os computadores (**C2C**)



Outras Categorias

- Consumer to Business (C2B)
- Government to Government (G2G)
- Government to Employee (G2E)
- Government to Business (G2B)
- Government to Citizen (G2C)
- Citizen to Government (C2G)

What is E-Business?

“E-business is the practice of performing and coordinating critical business processes such as designing products, obtaining supplies, manufacturing, selling, fulfilling orders, and providing services through the extensive use of computer and communication technologies and computerized data.”

(Alter, 2002)



What is E-Business

“E-business ... incorpora tudo o que tem a ver com a aplicação da informação e das tecnologias da comunicação (ICT) na condução de processos de negócio entre organizações ou envolvendo computadores / computação”

(Huff et al. (2000) cited in Alter, 2002)



What is E-Business

“E-business incorpora e-commerce mas também cobre os processos internos tais como produção, inventários, desenvolvimento do produto, gestão de riscos, finanças, gestão do conhecimento, e recursos humanos.”

(Bartels (2000) cited in Alter, 2002)



Mercados

- Um mercado é um lugar onde os compradores contactam os vendedores usando um meio de troca (e.g., a moeda)
- A integração vertical permite a algumas empresas prescindir do mercado



Custo da Transação

- Os custos de transação são os custos indirectos de um vendedor ou de um comprador num determinado negócio
- A integração vertical reduz esses custos
- O Comércio electrónico também



Características da Transação

- Atomicidade (tudo ou nada)
- Consistência (integridade)
- Isolamento (independência)
- Durabilidade (persistência)



Pagamentos

- Credit Card
- Debit Card
- Smart Card
- E-Money
- Electronic Fund Transfer (EFT)



Cartão de crédito

Step	Description
Step 1	Bank issues and activates a credit card to customer on his/her request.
Step 2	Customer presents credit card information to merchant site or to merchant from whom he/she want to purchase a product/service.
Step 3	Merchant validates customer's identity by asking for approval from card brand company.
Step 4	Card brand company authenticates the credit card and paid the transaction by credit. Merchant keeps the sales slip.
Step 5	Merchant submits the sales slip to acquirer banks and gets the service chargers paid to him/her.
Step 6	Acquirer bank requests the card brand company to clear the credit amount and gets the payment.
Step 6	Now card brand company asks to clear amount from the issuer bank and amount gets transferred to card brand company.

Segurança

- Confidencialidade
- Integridade
- Disponibilidade
- Autenticidade
- Não repúdio
- Encriptação.
- Auditabilidade



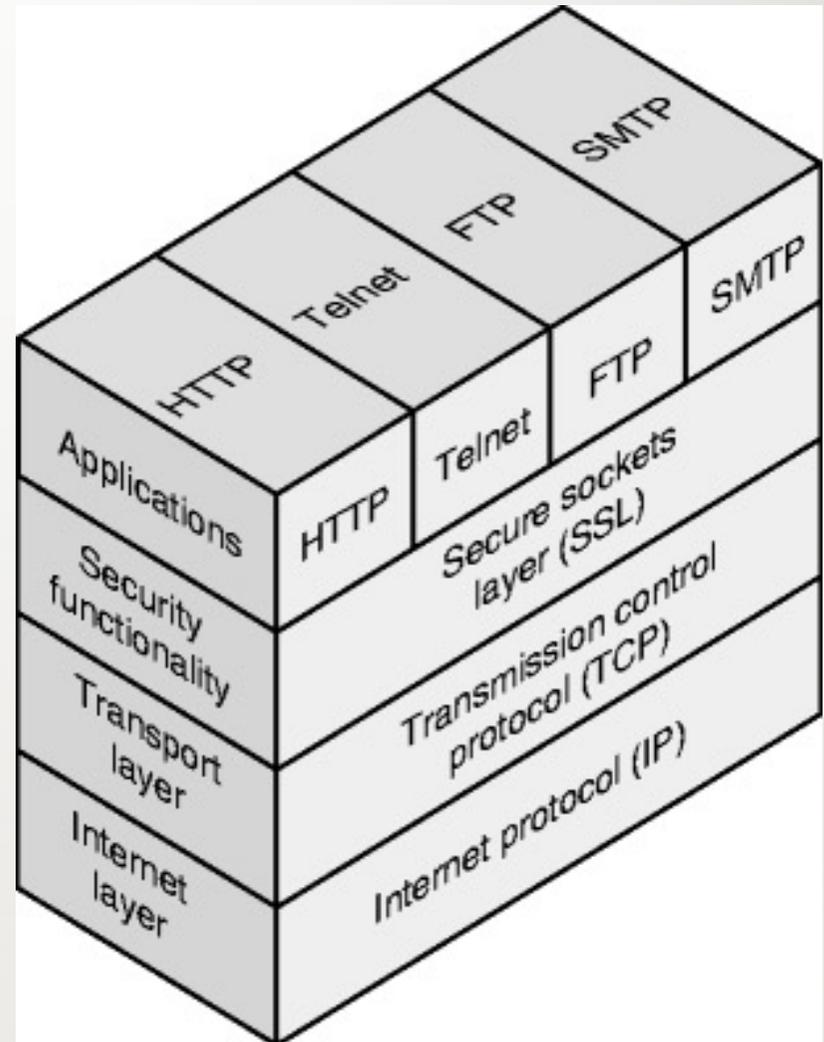
Medidas de segurança

- Encriptação
- Assinatura digital
- Certificados de segurança

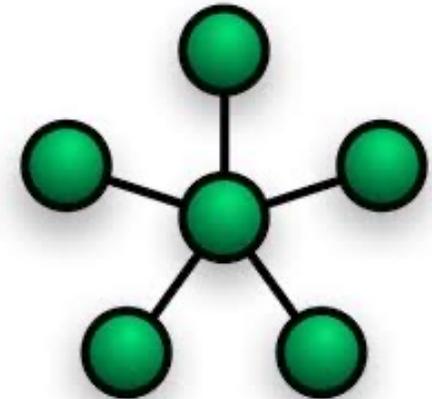


Protocolos de segurança

- Secure Socket Layer (SSL)
- Secure Hypertext Transfer Protocol (SHTTP)
- Secure Electronic Transaction



Estruturas económicas em rede



- Algumas empresas trabalham numa estrutura económica que não é nem mercado nem vertical
- Essas empresas têm alianças estratégicas, partilham objectivos e estratégias.
- Essas alianças podem ser formadas na Internet - Empresas virtuais
 - Equipas podem ser dissolvidas quando deixam de ser necessárias
 - Novas equipas são criadas sempre que necessárias

AGENTES?

Cadeias de valor

- Uma cadeia de valor é uma forma de organizar actividades de negócio para desenhar, produzir, promover, vender, entregar e manter os produtos ou serviços vendidos





Internacionalização

- Questões de confiança
- Questões linguísticas
- Questões culturais
- Questões de infraestruturas

Confiança

- Qualquer um pode criar um website
- Os criadores podem ficar anónimos
- Os consumidores não confiam principalmente quando têm de dar dados pessoais e número de cartão de crédito:
- Devem ser desenvolvidos métodos para estabelecer relações de confiança com os consumidores



Língua

- O uso da língua nativa é importante
- A mesma língua têm diferentes vocabulários (Portugal - Brasil – Angola)
- Traduzir é caro e difícil



Cultura

- Cultura / tradição e costumes
- A cultura pode ser regional
- Importante para a propaganda
- Icons (significados diferentes, às vezes inconvenientes)



Conclusão

- Comércio Electrónico não é um negócio novo mas uma nova forma de fazer negócio
- Os custos transacionais devem ser reduzidos
- As questões da confiança, língua, cultura e das infraestruturas foram levantadas

Objectivos

- Usabilidade
- Princípios básicos de um website
- 5 princípios para aferir a credibilidade de um website
- Um website para o comércio electrónico

Princípios Básicos



Princípios básicos

- O que é uma interface?



Princípios básicos

- O que é uma boa interface web?
 - Deduzir as funcionalidades
 - Items importantes sempre disponíveis
 - Icons expressivos
 - Objectivos e arranjos lógicos
 - Site interessante, colorido, não pode irritar
 - Audiência (idade, cultura, tema)
 - Tamanho do ecran
 - Smartphone, tablet ou PC
 - Identificação (logos, texto,...)



Princípios básicos

- Apresentação não linear
- Página principal simples
- Navegação simples
- Imagens pequenas para carregamento rápido
- Efeitos visuais chamativos

Princípios básicos

- A WWW é caracterizada por:
 - Informação não linear
 - As páginas podem ser vistas em periféricos diferentes, sistemas operativos diferentes, browsers diferentes
 - Velocidade de ligação à Internet diferentes
 - Audiência numerosa e heterogénea
- Estas características devem ser consideradas na construção do website





Princípios básicos

- Poucos clicks
- O website deve funcionar com ligações lentas e com browsers mais recentes
- Simplicidade

SWOT Analysis

- A análise SWOT serve para analisar e avaliar oportunidades de negócio
- SWOT significa:
 - **S**trengths (Forças)
 - **W**eaknesses (Fraquezas)
 - **O**pportunities (Oportunidades)
 - **T**hreats (Ameaças)



TOWS

- Pontos fortes vs Oportunidades: estratégia ofensiva / desenvolvimento das vantagens competitivas.
- Pontos fortes vs Ameaças: estratégia de confronto para modificação do ambiente a favor da empresa.
- Pontos fracos vs Oportunidades: estratégia de reforço para poder aproveitar melhor as oportunidades.
- Pontos fracos vs Ameaças: estratégia defensiva com possíveis modificações profundas para proteger a empresa.
- Como é que os pontos fracos se podem transformar em pontos fortes?

TOWS Matrix for PowerPoint



Usabilidade

Usability = quality attribute that assesses how easy user interfaces are to use.

Learnability

Efficiency

Memorability

Errors

Satisfaction

Usability = methods for improving the usefulness during design phase

Usabilidade

- Exemplo:
 - Um website não fornece um endereço de email
 - Um website fornece um endereço de email mas não há pessoal para ler e responder aos emails atempadamente

Usabilidade

- Um visitante deve ter acesso à:
 - História da organização
 - Objectivos e missão
 - Informação sobre produtos e serviços
 - Uma forma de comunicação
- Cada visitante é um potencial comprador
 - É difícil atender a todas as vontades / necessidades do utilizador



Usabilidade

- Conhecer os produtos ou serviços
- Comprar produtos e serviços
- Obter informação sobre a organização
- Identificar e contactar responsáveis
- Obter informações financeiras para possíveis investimentos

Usabilidade

Devem ser considerados

- O nível de expectativa e experiência do utilizador
- O canal de comunicação e a largura de banda
- O web browser usado
- Os add-ins disponíveis para o browser

Usabilidade / Flexibilidade

- Versões com ou sem frames
- Versões só com texto
- Versões com gráficos mais “leves”
- O utilizador deve poder escolher a largura de banda para ver videos ou ouvir audio
- O utilizador pode escolher se quer visualizar ou “fazer download” de docs
- Docs em formatos diferentes



Usabilidade

Confiança e Lealdade

satisfação e lealdade

problemas de integração com backoffice

email é lento e pouco funcional (assíncrono)



Usabilidade

Avaliar a usabilidade e aumentar a usabilidade com pequenas alterações

Aumentar o tempo de resposta mantendo a qualidade (síncrono e assíncrono) – quando isso não é possível manter o utilizador sobre aviso que tem de esperar e que o sistema não bloqueou!

Avaliação de Web Sites

- A Internet oferece uma fonte muito rica de informação, mas
 - Nem tudo está na Internet!
 - Muitos websites têm informação errada, confusa e incompleta



Avaliar Web Sites

- Existem inúmeros critérios para avaliar a credibilidade de fontes de informação disponíveis na Internet



Avaliar Web Sites

- Critérios:
 - Propriedade
 - Correção
 - Objectividade
 - Actualidade
 - Cobertura

Para construir um Web Site

- Requisitos de Software e hardware



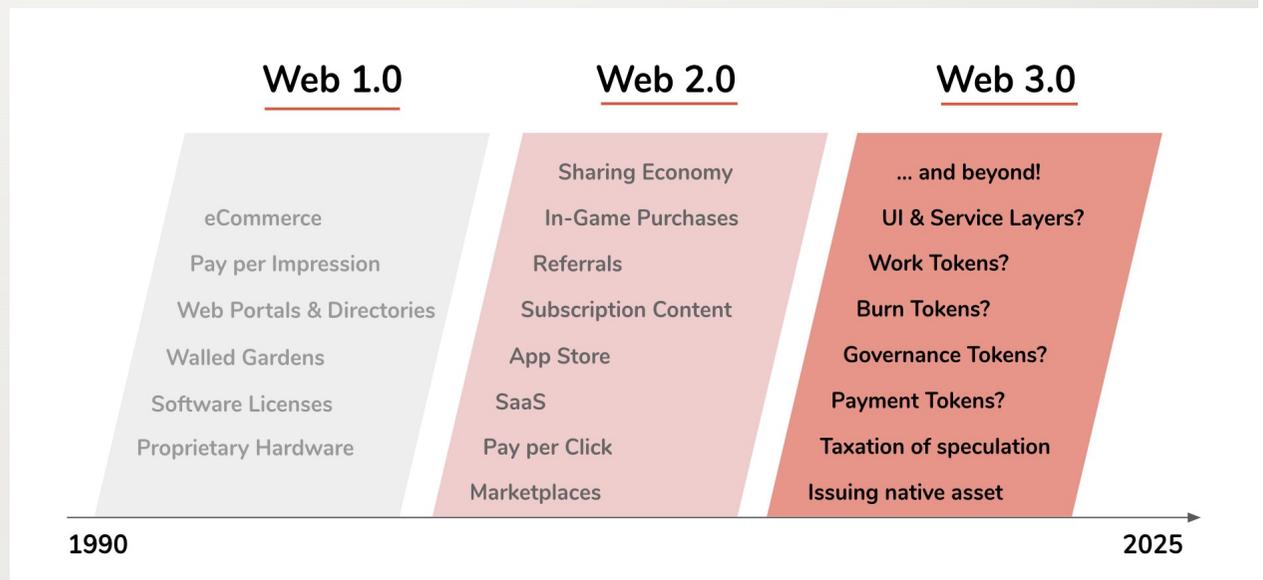
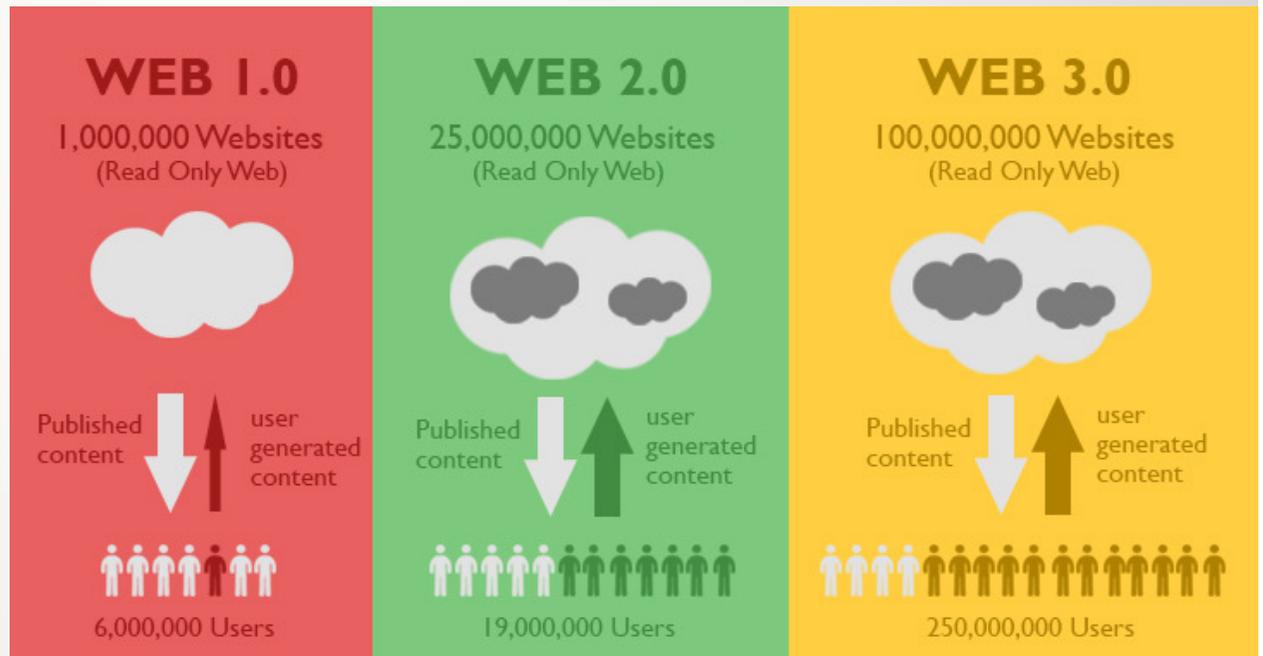
O Web Site

- Estimar
 - O número de visitantes
 - O número de páginas visitadas por visitante
 - O tamanho médio e máximo de cada página
 - O número máximo de visitantes em simultâneo
- Estes valores estimados irão condicionar o hardware e o software

o Web Site

Conteúdos

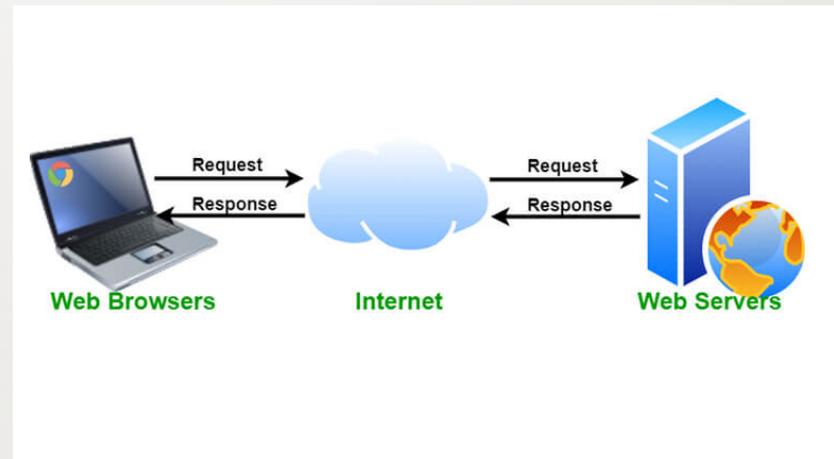
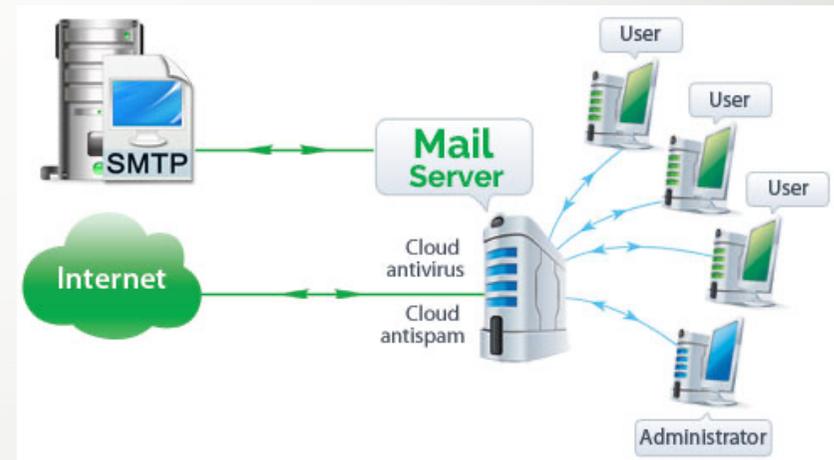
- Uma página estática não muda (web 1.0)
- Uma página dinâmica é criada por um programa (web 2.0)
- Web 3.0



O Web Site

Servidores

- O servidor
 - O software do Servidor
 - Os servidores Web
 - A integração com outras aplicações (agentes?)





O Web Site

Servidores

- **Servidores de E-mail** para enviar e receber email
- **Servidores de bases de dados** com sistemas gestores de bases de dados

Maximum Availability Architecture (MAA)

Low-Cost, Integrated, Fully Active, High ROI

Production Site

RAC

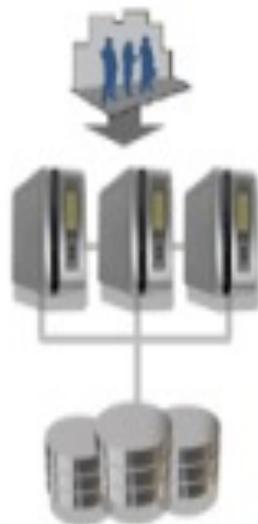
- Scalability
- Server HA

Flashback

- Human error correction

Edition-based Redefinition, Data Guard, GoldenGate

- Minimal downtime maintenance, upgrades, and migrations



ASM

- Volume Management

RMAN & Fast Recovery Area

- On-disk backups

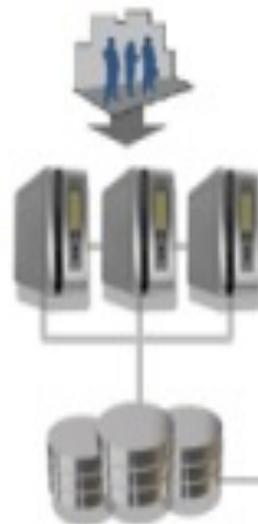
Active Replica

Active Data Guard

- Data Protection, DR
- Query Offload

GoldenGate

- Active-active
- Heterogeneous



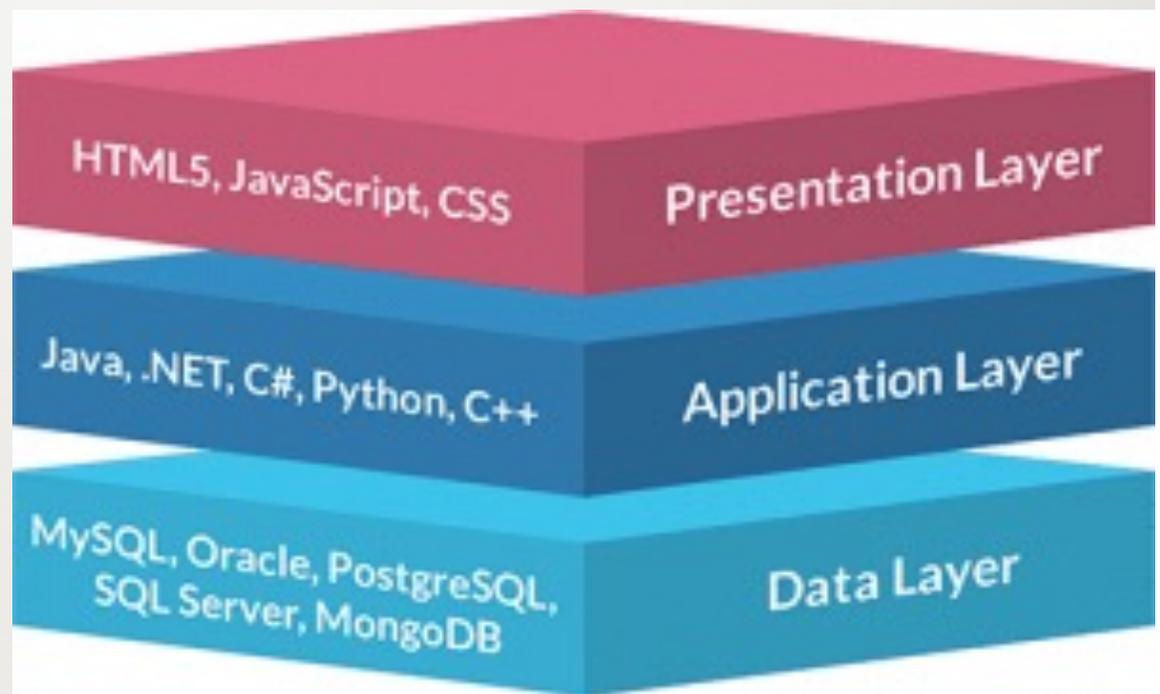
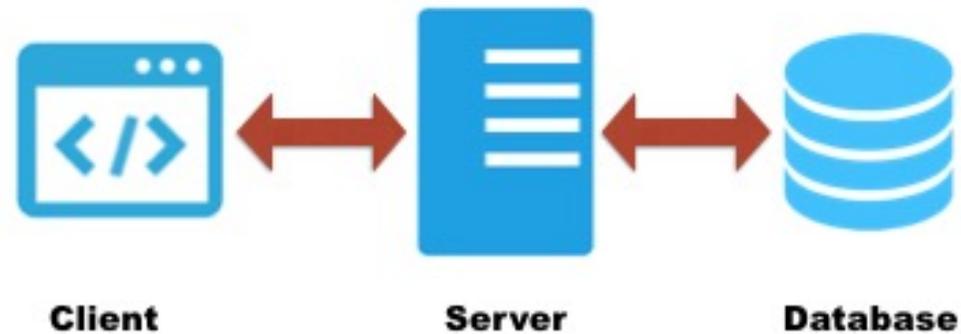
Oracle Secure Backup

- Backup to tape / cloud

O Web Site

Arquitecturas 2/3 camadas

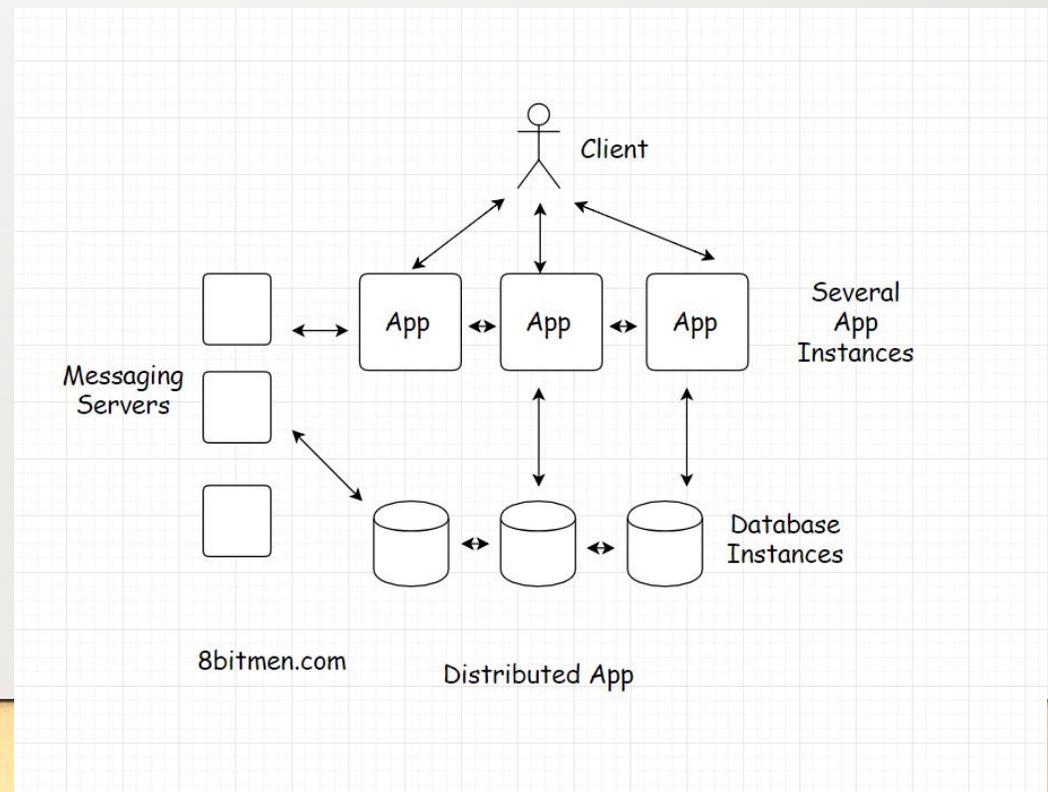
- Cliente – servidor (2 camadas)
- Cliente – Servidor – Outros sistemas (3 camadas)



O Web Site

Arquitecturas dos Web Servers

- Farms (podem ir até centenas ou milhares de computadores)
- Arquitectura centralizada (poucos computadores muito rápidos)
- Arquitectura distribuída / descentralizada (muitos computadores normais)





O Web Site

Balanceamento?

Redundância?

Backup?

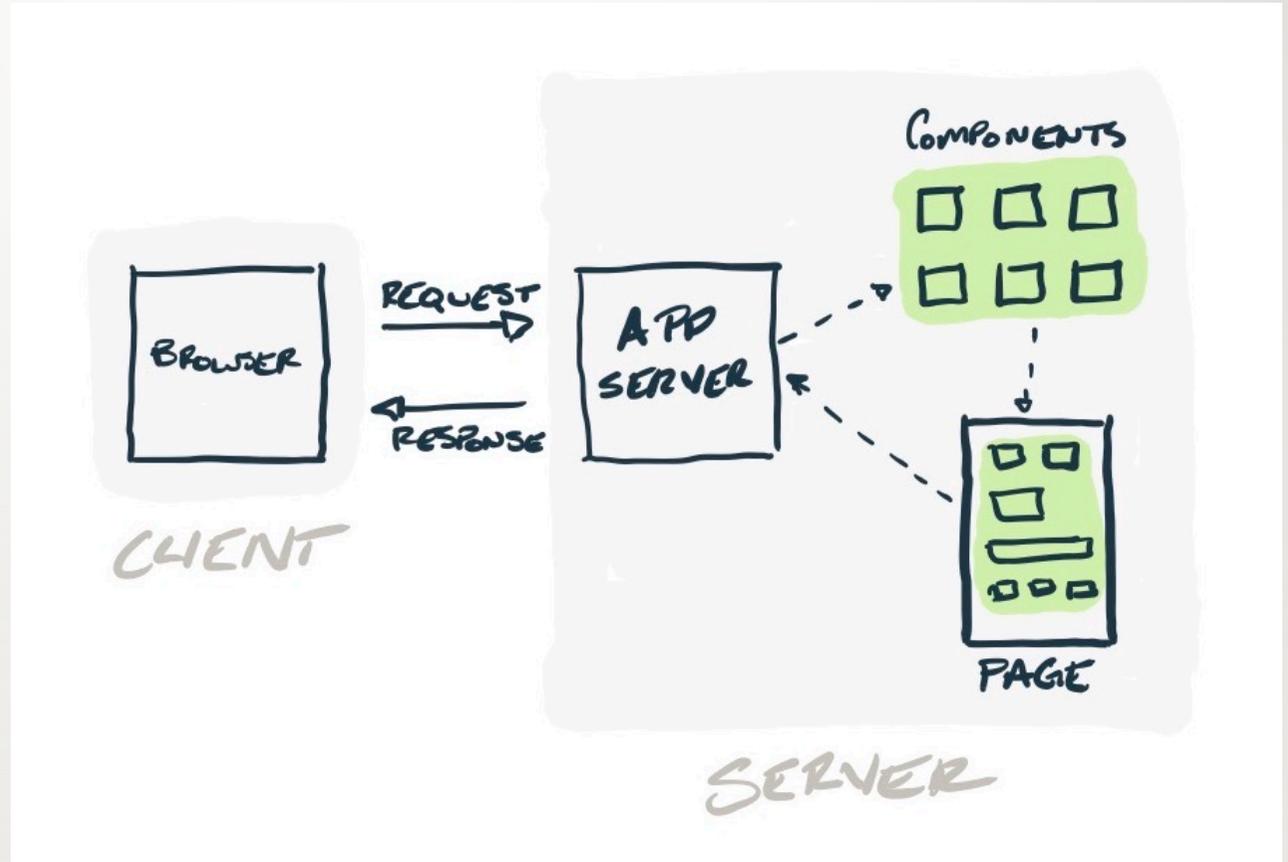
Disponibilidade permanente?

Failover?

O Web Site

Tecnologias Client-side

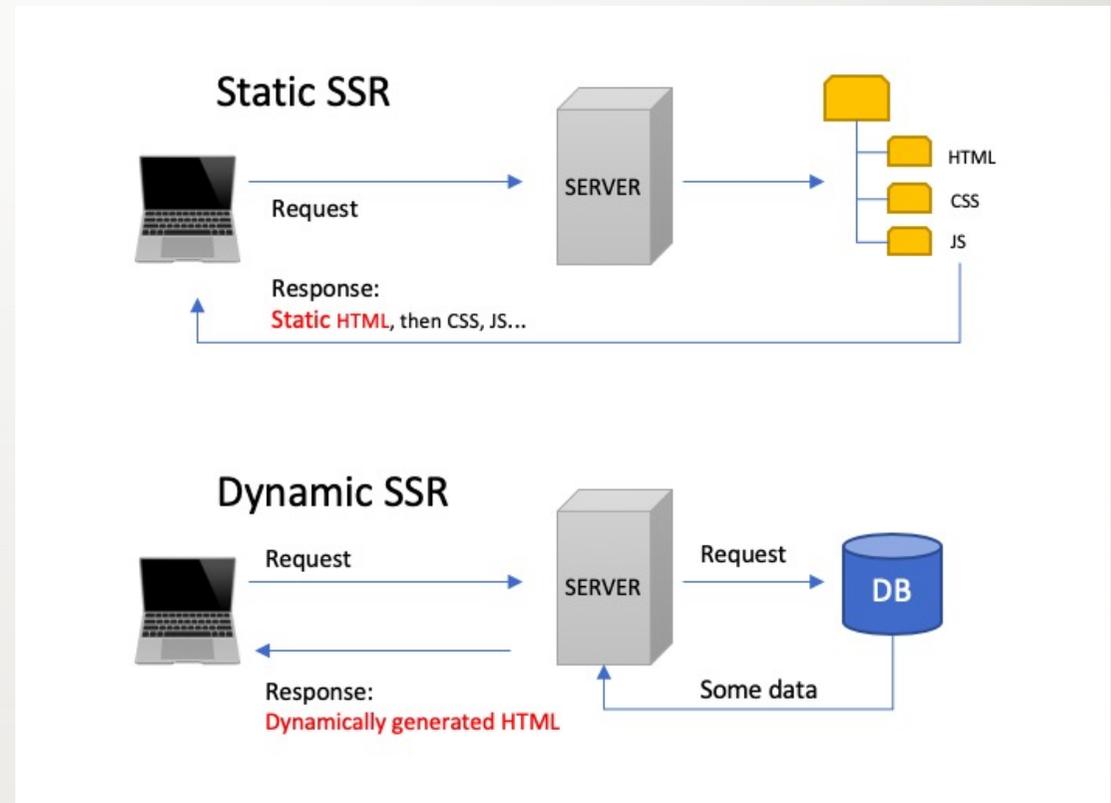
- HTML
- XML
- JavaScript
- VBScript
- Java
- Applets



O Web Site

Tecnologias Server-side

- Perl/CGI
- JSP
- PHP
- Microsoft ASP/ASPX





O Web Site

- Self-hosting ou outsourcing?



O Web Site

Hosts

- Self-hosting
- Hosting service providers
- Internet Service Providers (ISPs), Commerce Service Providers (CSPs), Managed Service Providers (MSPs) ou Application Service Providers (ASPs)

O Web Site

Software de Comércio Electrónico (variáveis)

- Tamanho da empresa
- Orçamento

Funcionalidades (mínimas):

catálogo

carrinho de compras

processamento de transações

O Web Site

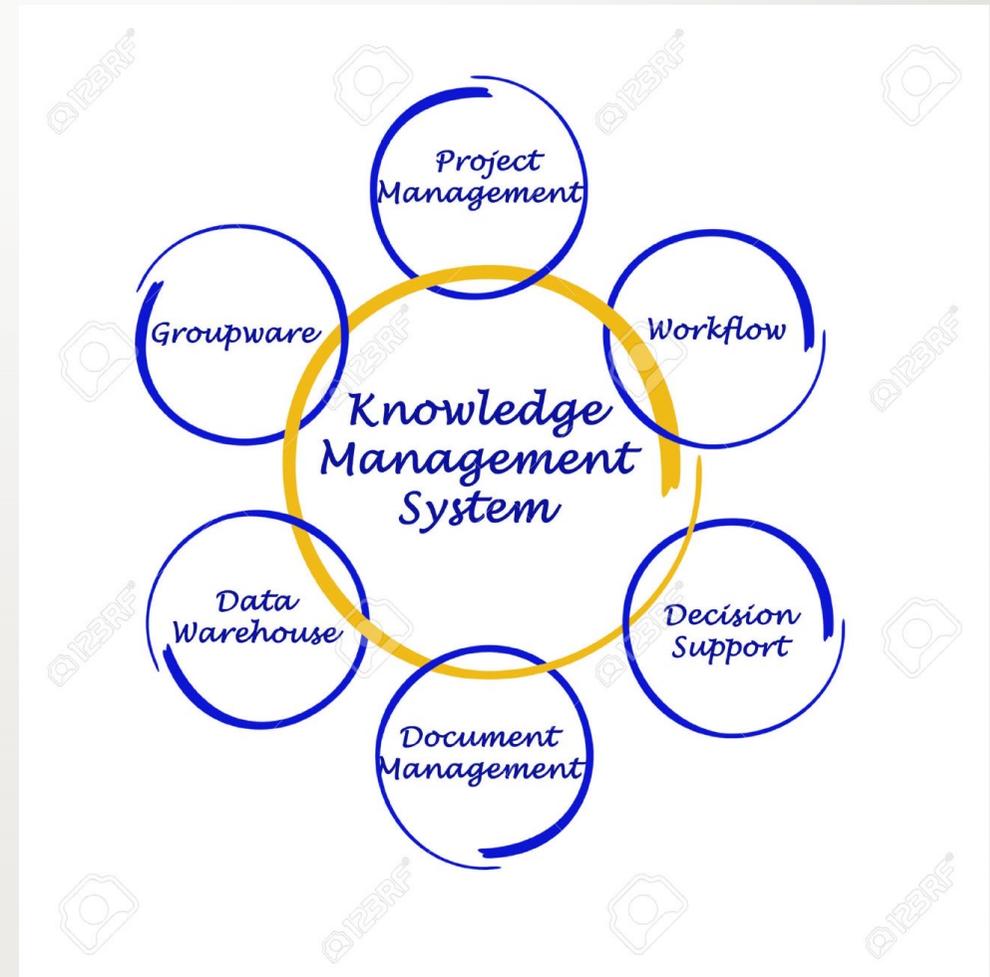
Funcionalidades adicionais

- Middleware para ligar a outros sistemas (stocks, encomendas, contabilidade e finanças)
- Bases de dados
- Gestão da produção (planeamento) (Supply Chain Manager)
- Gestão comercial (Customer Relationship Manager)
- Gestão de conteúdos (Content Management System)
- Gestão de conhecimento (Knowledge Management System)
- Web services

O Web Site

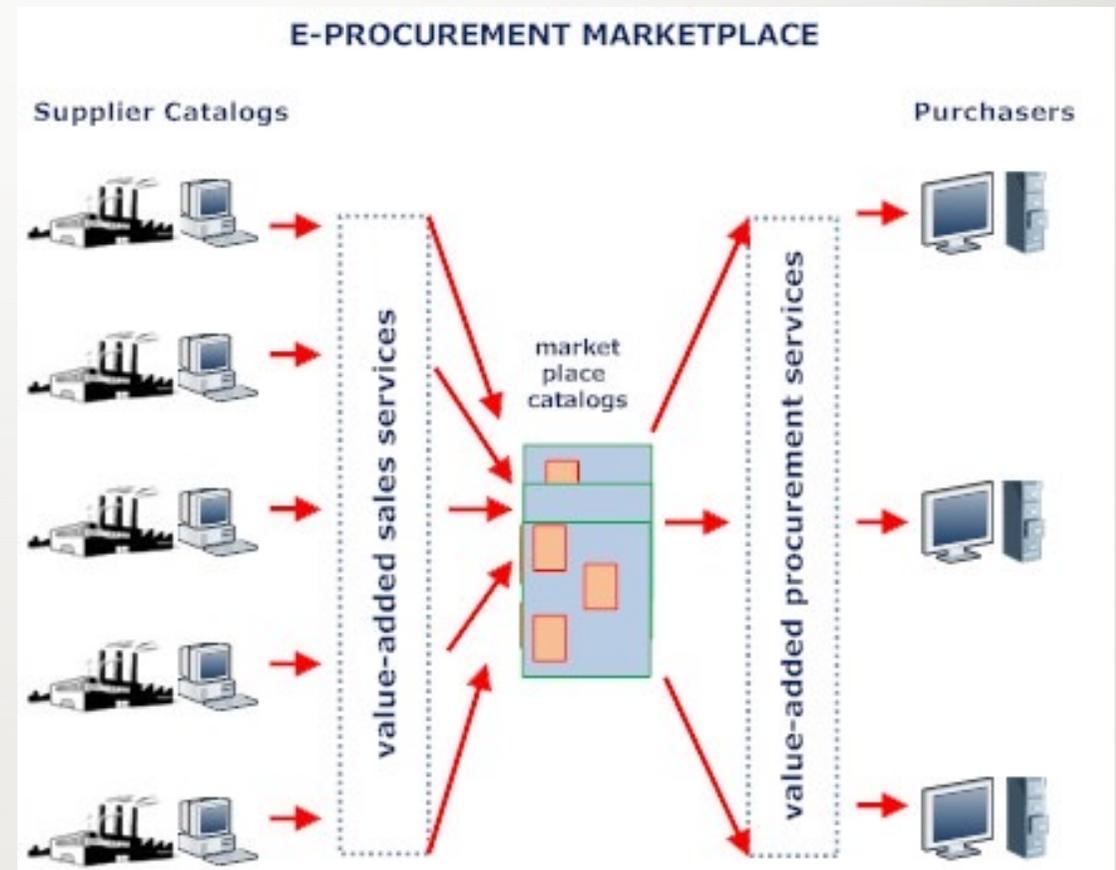
Knowledge Management Software

- Colectar e organizar informação
- Partilhar informação
- Realça a habilidade do utilizador em colaborar
- Retem conhecimento para futuras operações (aprendizagem)



E-procurement

- E-Procurement é compra de bens e serviços, por meio da Internet sendo que o e-procurement pode ser implementado através de um processo manual ou usando um software de gestão do tipo ferramenta de ERP.





E-procurement

- Negociação,
- Concurso
- Leilão,
- gestão de fornecedores,
- Gestão de catálogos,
- Integração de Ordens de Compra,
- Status do pedido
- Guia de transporte
- ,eFactura
- ePagamento
- gestão de contratos.

E-procurement

- Indent management is the workflow involved in the preparation of tenders (concurso)
- This part of the value chain is optional, with individual procuring departments defining their indenting process.
- In work procurement, administrative approval and technical sanction are obtained in electronic format.
- In goods procurement, indent generation activity is done online.



E-procurement

- Elements of e-procurement include
 - request for information,
 - request for proposal,
 - request for quotation,
 - RFX (the previous three together),
- and eRFX (software for managing RFX projects).



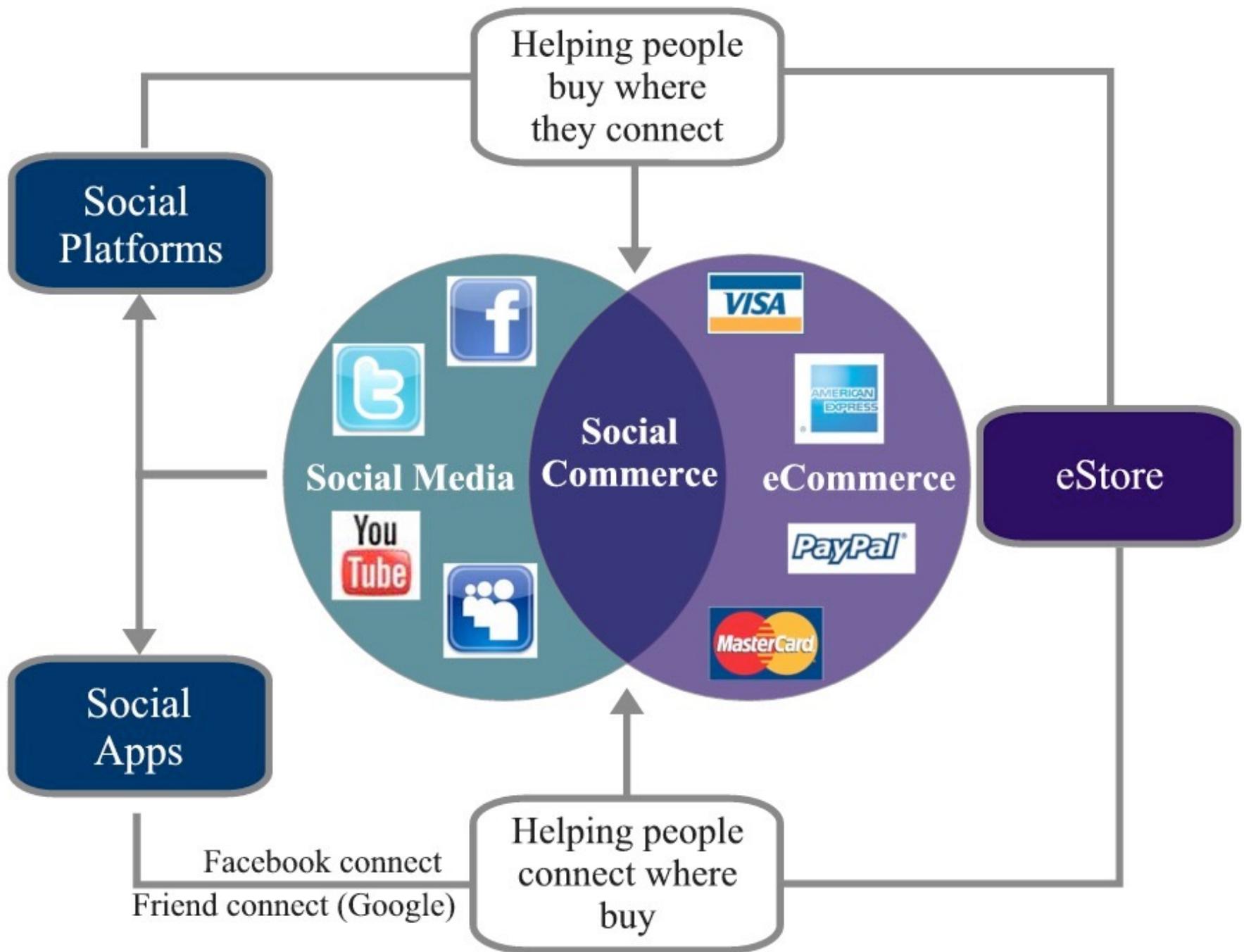
- Public sector organizations use e-procurement for contracts to achieve benefits such as increased efficiency and cost savings (faster and cheaper) in government procurement and improved transparency (to reduce corruption) in procurement services.
- E-procurement in the public sector has seen rapid growth in recent years.
- E-procurement in the public sector is emerging internationally.





S-Commerce

Social Commerce: A subset of electronic commerce that involves using social media, online media that supports social interaction and user contributions, to assist in the online buying and selling of products and services.



Comércio social



O S-Commerce ou (social commerce), é o nome dado à influência das pessoas no comércio eletrônico. Alguns estudos mostram que 80% da população segue a opinião de outras pessoas (amigos/familiares), os outros 20% levam em consideração o que está na net

Social commerce fuels real business growth



Tamara discovers
an event on Eventbrite



She shares
it with friends



Friends see
Tamara's post
and click on
the event



They share the
event with more
friends



Friends
buy tickets

BY THE NUMBERS



90%
of all purchases are
subject to social influence.
WIRED MAGAZINE (UK)



150 million
people engage with Facebook
on external websites each month.
FACEBOOK 2011



1 million +
sites have implemented
Facebook's Social Layer.
TECHCRUNCH



90%
of consumers trust recommend-
ations from people they know.
THE NIELSEN COMPANY



\$30 billion
predicted revenues for social
commerce market in 2015.
BOOZ & COMPANY



4x
Facebook friends are four
times more valuable than
Twitter followers.
CHOMPON.COM



3.2
average unique visitors
per share
SPINBACK.COM



67%
spend more online after
recommendations.
BAZAARVOICE.COM



Comércio Social

- Compra social
- Classificações e comentários
- Recomendações e referências
- Fóruns e comunidades
- Social Media optimization
- Anúncios sociais e aplicativos



Recommendation Systems

There is an extensive class of Web applications that involve predicting user responses to options. Such a facility is called a recommendation system (**Jeffrey D. Ullman, Stanford Univ.**)



two good examples of recommendation systems are:

- 1. Offering news articles to on-line newspaper readers, based on a prediction of reader interests.
- 2. Offering customers of an on-line retailer suggestions about what they might like to buy, based on their past history of purchases and/or product searches.



Utility matrices

Recommendation systems deal with users and items. A utility matrix offers known information about the degree to which a user likes an item. Normally, most entries are unknown, and the essential problem of recommending items to users is predicting the values of the unknown entries based on the values of the known entries.

Two Classes of Recommendation Systems

These systems attempt to predict a user's response to an item by discovering similar items and the response of the user to those.

One class of recommendation system is content-based; it measures similarity by looking for common features of the items.

A second class of recommendation system uses collaborative filtering; these measure similarity of users by their item preferences and/or measure similarity of items by the users who like them.



Item Profiles

These consist of features of items. Different kinds of items have different features on which content-based similarity can be based. Features of documents are typically important or unusual words. Products have attributes such as screen size for a television. Media such as movies have a genre and details such as actor or performer. Tags can also be used as features if they can be acquired from interested users.



User Profiles

A content-based collaborative filtering system can construct profiles for users by measuring the frequency with which features appear in the items the user likes. We can then estimate the degree to which a user will like an item by the closeness of the item's profile to the user's profile.



Classification of Items

An alternative to constructing a user profile is to build a classifier for each user, e.g., a decision tree. The row of the utility matrix for that user becomes the training data, and the classifier must predict the response of the user to all items, whether or not the row had an entry for that item.

Similarity of Rows and Columns of the Utility Matrix

Collaborative filtering algorithms must measure the similarity of rows and/or columns of the utility matrix.

Jaccard distance is appropriate when the matrix consists only of 1's and blanks (for “not rated”). Cosine distance works for more general values in the utility matrix. It is often useful to normalize the utility matrix by subtracting the average value (either by row, by column, or both) before measuring the cosine distance.



Clustering Users and Items

Since the utility matrix tends to be mostly blanks, distance measures such as Jaccard or cosine often have too little data with which to compare two rows or two columns. A preliminary step or steps, in which similarity is used to cluster users and/or items into small groups with strong similarity, can help provide more common components with which to compare rows or columns.

UV-Decomposition

One way of predicting the blank values in a utility matrix is to find two long, thin matrices U and V , whose product is an approximation to the given utility matrix. Since the matrix product UV gives values for all user-item pairs, that value can be used to predict the value of a blank in the utility matrix. The intuitive reason this method makes sense is that often there are a relatively small number of issues (that number is the “thin” dimension of U and V) that determine whether or not a user likes an item.



Root-Mean-Square Error

A good measure of how close the product UV is to the given utility matrix is the RMSE (root-mean-square error). The RMSE is computed by averaging the square of the differences between UV and the utility matrix, in those elements where the utility matrix is nonblank. The square root of this average is the RMSE.

Computing U and V

One way of finding a good choice for U and V in a UV-decomposition is to start with arbitrary matrices U and V . Repeatedly adjust one of the elements of U or V to minimize the RMSE between the product UV and the given utility matrix. The process converges to a local optimum, although to have a good chance of obtaining a global optimum we must either repeat the process from many starting matrices, or search from the starting point in many different ways.