



Instituto Politécnico de Tomar
Escola Superior de Tecnologia de Tomar
Departamento de Engenharia Informática
Unidade Curricular de Introdução à Tecnologia

A Internet pré-Hipermédia

José M. Palma R. Ramos
docente

Resumo

Entre 1969, ano da sua criação, e 1984 (durante 25 anos, portanto), a Internet pouco mais foi que uma obscura invenção tecnológica de muito restrita utilidade. Porém, entre 1984 e 1989, o desenvolvimento da micro-informática, das arquiteturas cliente-servidor e de outras tecnologias, vão multiplicar por 100, em apenas 5 anos, tudo o que se relaciona com a Internet: o número de computadores interligados, o número de pessoas envolvidas, o volume da informação disponível, o volume do tráfego de dados, etc. Deste sobreaquecimento do fenómeno Internet, resultaram novas atitudes sociais, as comunidades virtuais, de gente ávida de novas tecnologias para partilhar informação entre si, o mais rapidamente possível, o mais facilmente possível e ao menor custo possível. É deste ambiente que, a partir de 1989, surge a Web, como um sistema de hipermédia, efectivamente realizado e rapidamente distribuído à escala planetária, via Internet.

1. Introdução

Em 1967, a ARPA (Advanced Research Projects Agency) lidera um projecto para uma ampla rede de computadores interligados. Em 1969, surge a ARPANET (de onde irá nascer mais tarde a *Internet*), na sua primeira configuração de quatro computadores interligados, concretizando quatro nós (figura 1): o *SRI* (*Stanford Research Institute*), a *UCLA* (*University of California Los Angeles*), a *UCSB* (*University of California Santa Barbara*) e a *Universidade de Utah*.

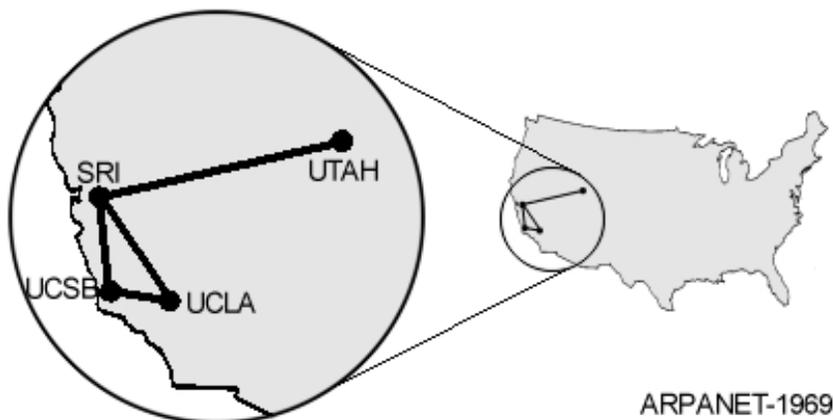


figura 1 – embrião da Internet: a primeira configuração da ARPANET, em 1969, com quatro nós.

Foram precisos 15 anos para que, em 1984, se ultrapassassem os 1.000 computadores interligados. Em 1989 (em apenas mais 5 anos), ultrapassaram-se os 100.000 computadores na *Internet*. Estava atingida a massa crítica que iria modificar tudo.

2. Arquitecturas Cliente-Servidor na Internet

No mundo das redes de computadores, antes de 1983, imperavam as *arquitecturas mainframe*. Uma *rede mainframe* consistia num grande computador central ao qual se ligavam vários terminais passivos. Todas as aplicações e todos os dados estavam num enorme e dispendioso computador central. Este esperava comandos de teclado enviados pelos terminais e executava-os, retornando a visualização dos resultados (figura 2-a). Este modo de funcionamento centralizado implicava na partilha temporal (*time-sharing*) da capacidade de processamento do computador central.

Por muito rápido que fosse o computador central, o aumento do número de terminais provocava uma lentidão extremamente limitativa. Além disso, como todos os utilizadores da rede tinham os seus dados no computador central, por maiores que fossem as capacidades de armazenamento, as perdas de tempo no acesso aos dados eram críticas. Quando uma *rede mainframe* estava conectada a outras, através da *Internet*, tudo se agravava, pois a conexão era mantida à custa de uma parte do processamento e de uma parte da capacidade de armazenamento de dados do computador principal (figura 2-b) o que reduzia ainda mais a eficiência partilhada pelos utilizadores.

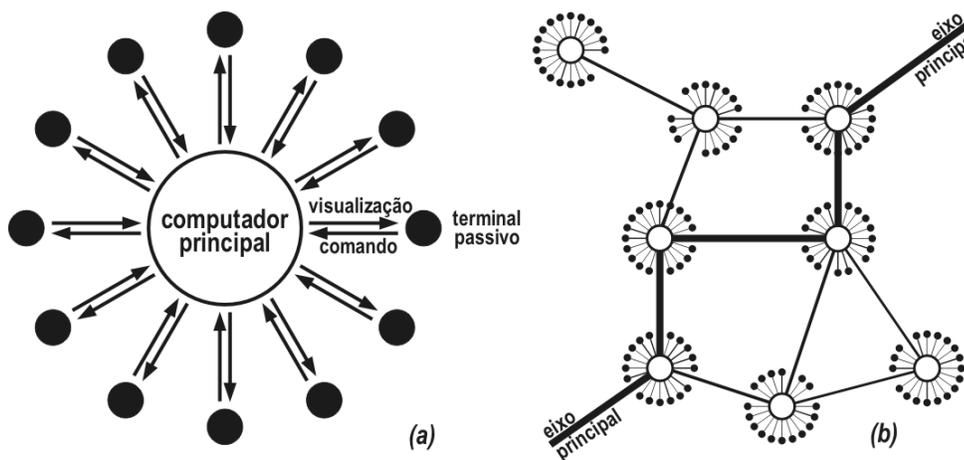


figura 2 – A arquitetura mainframe: a) todo o processamento e armazenamento de dados encontram-se no computador central, operado através de terminais passivos; b) a rede é sustentada à custa da sobre-utilização do computador central.

O conceito de *cliente-servidor* descreve uma nova relação entre duas aplicações computacionais: a aplicação *cliente* executa um pedido de *serviço* à aplicação *servidora*, que satisfaz esse pedido. Sendo adequado a aplicações em execução no mesmo computador, este

conceito é mais relevante na ideia de diversos computadores interligados em rede, proporcionando o “diálogo” síncrono entre aplicações computacionais, espacialmente dispersas.

Assim, o *modelo cliente-servidor*, divulgado nos anos 80, tornou-se na ideia central, *arquitetural*, das actuais redes de computadores, em que qualquer aplicação *cliente*, instalada num computador, solicita à respectiva aplicação *servidora*, instalada noutra computador, a transferência (em qualquer dos sentidos) de “pacotes de dados”. Cada relação *cliente-servidor* é regulada por um *protocolo* estabelecido entre as duas aplicações (figura 3-a). Porém, este *protocolo* é regulado por outro *protocolo* (de nível mais baixo) que determina o funcionamento geral da rede e a identificação de cada computador nela presente. O actual *protocolo* básico da Internet, o *TCP/IP* (*Transmission Control Protocol/Internet Protocol*), foi instalado em 1983 para suportar *arquitecturas cliente-servidor* para as mais variadas finalidades de transferência de dados, incluindo o *protocolo FTP* (*File Transfer Protocol*).

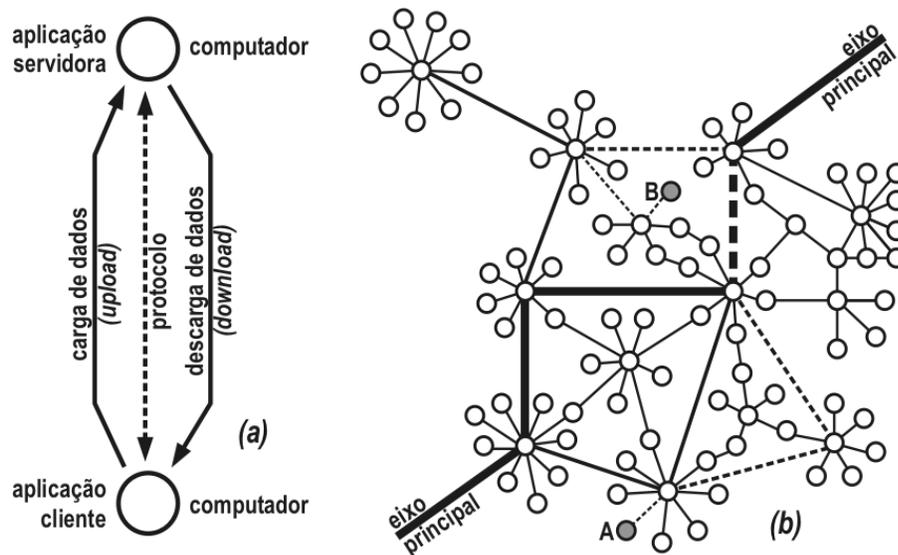


figura 3 – *Arquitetura cliente-servidor*: a) o protocolo entre as aplicações cliente e servidora transfere dados entre dois computadores; b) a rede é sustentada por todos os computadores, havendo redundância conectiva; a tracejado, apresenta-se o percurso conectivo de um protocolo estabelecido entre os computadores A e B.

Deste modo, qualquer computador de uma rede usa as suas próprias aplicações (sejam elas *clientes* ou *servidoras*) e processa e armazena os seus próprios dados, ocupando outros computadores apenas para obter ou para prestar determinado serviço (figura 3-b). Ao descentralizar o processamento e o armazenamento, aumentando a velocidade e a economia, a introdução de *arquitecturas cliente-servidor* trouxe consigo uma maior *facilidade de utilização* e maiores *flexibilidade, interoperabilidade e escalabilidade* das redes, o que promoveu o rápido crescimento da Internet entre 1984 e 1989.

3. Comunidades na Internet (1984-1989)

A primeira finalidade da *Internet* foi a transmissão de dados digitais entre computadores. Neste contexto inicial, o primeiro serviço da *Internet* foi algo semelhante ao que hoje designamos por *correio electrónico*.

Até 1984, os cerca de 1.000 computadores da *Internet* eram, na sua grande maioria, máquinas de grande porte, muito caras e complicadas. Por esse motivo, os utilizadores da *Internet* reduziam-se à escassa elite da comunidade computacional, sediada em poucas grandes universidades ou alguns centros de investigação avançada. Esse ano de 1984 marca a *revolução da micro-informática* e das *arquitecturas cliente-servidor*, com a popularização dos primeiros *computadores pessoais*, ou *micro-computadores*, com *processadores de 32 bits*. Muito mais acessíveis e de mais fácil operação, alguns destes sistemas já ultrapassavam, em capacidade de armazenamento e velocidade de processamento, uma grande parte das grandes máquinas que constituíam, até então, a *ossatura* da *Internet*. Esta súbita viragem permitiu o acesso em massa de universidades e instituições científicas ao mundo da computação. Antes, limitados pelos constrangimentos do *time-sharing* (acesso remoto a grandes computadores, em tempo partilhado e por operação através de terminais passivos), ao poder adquirir os seus próprios sistemas, passaram, assim, a dispor, entre portas, de capacidade de processamento e de armazenamento em tempo ilimitado.

Este cenário alterou também o *padrão utilitário* do *conceito de computador*. Antes, quase exclusivamente, ferramenta de cálculo pesado e ininterrupto, estava a adquirir uma nova utilidade como gerador e armazenador de informação digital. A possibilidade de ligar vários computadores em *rede local (LAN - Local Area network)*, permitia aos membros de cada grupo local, produzir, reeditar, enviar e receber cópias de documentos digitais, especialmente de texto, por processos semelhantes ao actual *correio electrónico* e pelo *protocolo FTP*.

A *Internet* era, pois, uma via para que estes grupos se ligassem a outros grupos, identicamente apetrechados e espalhados pelo mundo, para troca de mensagens e para partilha de documentos digitais de interesse comum. Assim se compreende porque é que, de 1984 para 1989, passou de 1.000 para 100.000 o número de computadores na *Internet*. Destes fervilhantes 5 anos, resultaram também dois fenómenos determinantes: as *comunidades virtuais* e as *comunidades hacker*.

Comunidades virtuais: Grupos de investigação e desenvolvimento em domínios da engenharia e da computação, espalhados pelo mundo, viram na *Internet* a oportunidade de aumentarem a qualidade e a quantidade da sua produção científica e tecnológica, ao criarem repositórios comuns de informação digital, partilháveis quase em tempo real entre os seus membros. Para além do enorme benefício do *correio electrónico* e do *protocolo FTP* para a simples troca de mensagens e de documentos, podiam submeter relatórios, artigos ou resultados aos seus pares remotos e, assim, recolher pareceres e opiniões com mais rapidez e eficiência do

que através do correio tradicional ou através de encontros pessoais, muito mais caros e muito mais espaçados no tempo.

Muitos grupos que não aderiram a estas autodesignadas *comunidades virtuais* ou que tardaram em aderir ficaram de tal forma desfasados e marginalizados, que várias linhas de investigação em curso, simplesmente foram interrompidas por terem sido ultrapassadas. O centro da produção científica estavam a deslocar-se de cada grupo de investigação e desenvolvimento para a respectiva *comunidade virtual*, em que os próprios grupos passavam a constituir parte integrante.

Comunidades Hacker: Jovens estudantes das universidades, particularmente competentes em programação de computadores e especialmente fascinados pela computação em rede, dedicavam longas horas extracurriculares a tentar obter, através da *Internet*, documentos ou código de aplicações informáticas a cujo acesso não estavam autorizados. Para o conseguir, produziam aplicações computacionais ímpares capazes de lhes proporcionar os meios, ilegítimos, naturalmente, para tomar o controlo remoto de um qualquer computador na *Internet* e recolher cópias do material aí armazenado. Foram designados por *Hackers (purgadores)*. Numa variante de maior gravidade criminal, designada por *Crackers (destruidores)*, a intenção era a de puro vandalismo informático; as aplicações produzidas, que vieram a resultar nos actuais *vírus informáticos*, destinavam-se a eliminar, modificar ou arruinar conteúdos ou sistemas. *Hackers* e *Crackers*, tendo aparecido no início dos anos 80, contudo, é a partir de 1984 que assumem um preocupante cariz de comunidade, ao agirem, de forma muitas vezes concertada, através das redes ligadas à *Internet*. Por essa época, estas actividades ainda não tinham a carga criminal que os actuais ordenamentos jurídicos já lhes conferem. Todavia, a carga imoral era (obviamente) indiscutível, embora eventualmente não fosse assimilada de imediato por esses jovens de espírito rebelde e curiosidade avassaladora. Cobiçados pelas organizações informáticas em pleno crescimento, grande parte desses *Hackers*, brilhantes programadores, passada a fase da “*pirataria*”, viriam a aplicar a experiência e a competência resultantes nos futuros desenvolvimentos da computação em geral e da *Internet* em particular.

A massa crítica: Por tudo isto, chegamos a 1989 com uma *Internet* em crescimento acelerado, com centenas de comunidades virtuais a produzirem abundantes mananciais de documentação científica e uma massa enorme de jovens ansiosos por aceder a esse precioso património, de forma eficiente, rápida e, de preferência, gratuita. É de referir que o vertiginoso aumento do tráfego de dados na *Internet* motivou, neste mesmo ano, o alargamento da banda máxima no seu “eixo principal” de 56.000 bits/segundo para 1.544.000 bits/segundo.

4. Documentos na Internet (1989)

Em 1989, os milhões de documentos armazenados nos computadores da *Internet* ainda não eram de fácil acesso. Entre diversos factores de dificuldade, destacam-se três: o primeiro

consistia na diversidade de plataformas computacionais e sistemas operativos; o segundo consistia no desconhecimento geral do conteúdo em documentação relevante de cada computador; o segundo consistia na volatilidade de conteúdos, isto é, um documento podia existir num dia e deixar de existir no dia seguinte, podia até ser modificado a ponto de o seu sentido ficar subvertido. Isto, sem referir factores que tinham a ver com aspectos de privacidade ou de exclusividade da informação armazenada, só acessível a membros conhecedores de uma chave de acesso.

Inicialmente, os conteúdos eram trocados por *correio electrónico*. Depois, o *correio electrónico* passou a ser utilizado para divulgar listas descritivas de documentos armazenados e das instruções técnicas de como os obter através da rede. As comunidades virtuais mantinham documentos especiais com índices actualizados e detalhados incluindo descrições e instruções. Os membros dessas comunidades sabiam como aceder a esses índices e, quando encontravam um documento que lhes interessasse, executavam comandos do *protocolo FTP* que permitiam a sua obtenção. Um dos factores de crescimento das comunidades virtuais foi a menor dificuldade de acesso “comunitário” a documentos.

Os membros das comunidades virtuais, incluindo alguns *ex-Hackers*, sonhavam com soluções que permitissem manter, automaticamente, índices gerais, globais e universais de todos os documentos ou outros conteúdos relevantes da Internet e que fosse possível pesquisá-los por múltiplas combinações de palavras-chave. Embora a enorme diversidade computacional (incluindo a diversidade em plataformas, em sistemas operativos e em hábitos de organização de documentos digitais) diluísse expectativas em tais soluções, é exactamente neste ano de 1989 que se iniciam os esforços para atingir aquele objectivo.

5. O Archie (1989-1990)

Um primeiro esforço para produzir índices automáticos para a Internet foi realizado em 1989 por Peter Deutsch e pelos seus alunos da Faculdade de Computação da Universidade McGill em Montreal, tendo como base de referência todos os documentos armazenados em *servidores FTP anónimos* (computadores que disponibilizam cópias de todo o seu conteúdo sem exigência de chave de acesso) sobre sistema operativo *UNIX*.

Considerando que uma parte importante dos computadores da Internet eram *servidores FTP anónimos* sobre *UNIX*, Peter Deutsch e os seus alunos desenvolveram um sistema baseado em *UNIX*, a que chamaram *Archie*, e que foi instalado num computador da faculdade, também com sistema *UNIX*. O *Archie* procurava permanentemente esses servidores e actualizava um índice de todos os nomes de ficheiros neles encontrados. Para realizar uma pesquisa, o sistema *Archie* era então operado, através de comandos *UNIX* e palavras-chave. Estas palavras-chave eram então comparadas com o índice do *Archie* e deste se extraía (caso existisse) a lista de ficheiros e respectivos servidores que satisfazia a pesquisa realizada. Embora limitado à

pesquisa de nomes de ficheiros e exigindo conhecimentos de operação *UNIX*, o seu sucesso foi imediato e crescente, pelo que em 1990 foi disponibilizado na *Internet*, por operação remota *Telnet*, primeiro, no tal computador da faculdade, depois, em vários computadores.

6. O WAIS (1989-1991)

Em 1989, Brewster Kahle, da *Thinking Machines Corp.*, desenvolveu o *WAIS (Wide Area Information Server – Servidor de Informação de Área Ampla)*, capaz de indexar a totalidade das palavras de todos os documentos de texto armazenados num determinado computador e permitir a pesquisa do ficheiro de índices resultante através de palavras-chave simples ou de expressões booleanas com essas palavras-chave. Utilizando um *sistema pericial*, o *WAIS* atribuía uma pontuação de relevância, graduada de 0 a 1000, a cada resultado da comparação da expressão de pesquisa com os índices de ocorrências de palavras nos documentos e apresentava os resultados da pesquisa como uma lista dos ficheiros de documentos por ordem decrescente de relevância para com a pesquisa realizada. A partir das informações desses resultados, o utilizador podia escolher e receber o texto que entendesse.

Brewster Kahle era especialista em inteligência artificial, tendo sido aluno de Marvin Minsky e de Danny Hillis. O *sistema pericial* com que equipou o *WAIS* utilizava uma *máquina de inferência* destinada a extrair regras de relevância lexical a partir da análise permanente do universo de palavras, das respectivas frequências de ocorrência, das suas relações de proximidade e das palavras-chave utilizadas nas pesquisas efectuadas pelos utilizadores.

Em princípios de 1990, o sistema começou a expandir-se a vários computadores, mas o processo em nada mudava: o utilizador apenas pesquisava num computador *WAIS* de cada vez. Por essa razão, foi achado conveniente que cada computador *WAIS* fosse temático: por exemplo, havia um computador *WAIS* para toda a documentação do Dow Jones (Bolsa de Nova Iorque), outro para os arquivos do New York Times, outro para a Enciclopédia Britânica, etc.

Nos finais de 1990, Brewster Kahle disponibilizou na *Internet* uma aplicação gratuita (*cliente WAIS*) para acesso igualmente gratuito a alguns desses computadores *WAIS* sobre temas computacionais, como manuais de programação, especificações e normas sobre sistemas, artigos de conferências, relatórios da *Internet*, etc. Dado o sucesso imediato, em 1991, decide comercializar uma aplicação e uma licença de acesso para pesquisar documentos, simultaneamente, em todos os *servidores WAIS*. No seu máximo, o sistema *WAIS* manteve índices em cerca de 600 computadores. Assim, para acelerar as pesquisas, os índices de cada computador passaram a ser centralizados num único *mega-índice* replicado em vários computadores (os *servidores WAIS*).

Embora de aprendizagem demorada e complicado de utilizar, o *WAIS* cumpria as normas de pesquisa bibliográfica expressas na directiva ANSI Z39.50, recebendo, assim, o reconhecimento de uma verdadeira biblioteca digital.

7. O gopher (1989-1991)

Em 1991, na Universidade de Minnesota, um grupo liderado por Mark MacCahill, um programador avançado, apresentou na Internet o primeiro interface simples e eficiente para pesquisa de documentos de texto ou outros ficheiros em geral armazenados nos computadores da rede.

Tudo começou nos finais de 1989. A rede informática do campus da Universidade dispunha já de um enorme conjunto de documentos armazenados em ficheiros de texto espalhados pelos diversos computadores. Querendo organizar esse valioso espólio, e generalizar o seu acesso a todo o campus, a Universidade propôs ao seu departamento de computação que se desenvolvesse um sistema que centralizasse todos estes documentos e que disponibilizasse uma estrutura de directórios classificativos, permanentemente actualizada, que permitisse a qualquer utilizador da rede do campus pesquisar e aceder aos documentos com eficiência e comodidade. Esta proposta motivou um aceso debate entre um grupo dos mais conservadores, cultores das *arquitecturas mainframe* e do *centralismo da informação*, e os mais revolucionários, que incluíam o grupo de Mark MacCahill, cultores dos *pequenos sistemas* e da *informação distribuída*, e das *ligações generalizadas em rede*, em consonância com as emergentes *arquitecturas cliente-servidor*.

Os primeiros, defendendo a instalação imediata de um grande computador central na rede do campus e apoiados nesta iniciativa pelos dirigentes universitários, ganharam o debate inicialmente. Todavia, os segundos desafiaram-se a si próprios a desenvolver, num prazo curto, um protótipo funcional e demonstrativo das vantagens dum *sistema de informação distribuída* sobre um *sistema de informação centralizada*. Eles tinham consciência de que a utilização de pequenos sistemas diminuía os custos e os riscos inerentes à salvaguarda da segurança, integridade e consistência de uma grande massa de informação num grande sistema único e, além disso, menos eficiente; a informação poderia ser automaticamente replicada em sistemas redundantes, que, por serem de pequena dimensão, eram muito mais baratos.

O protótipo de demonstração, apresentado em 1990, foi designado por *gopher*, numa alusão à mascote da Universidade de Minnesota, uma variedade americana de esquilo com a particularidade de escavar extensas galerias por debaixo do solo. A grande facilidade de utilização dos interfaces e os mais baixos custos envolvidos na manutenção da informação em pequenos sistemas diluíram a oposição dos conservadores. Num sistema *gopher*, o utilizador não necessitava de possuir especiais competências em matéria de computação, bastando-lhe apenas escolher itens de menus sucessivos até à obtenção, automática, da cópia do documento

pretendido. Em relação aos documentos, apenas uma pequena modificação: para que um documento fosse compatível com o sistema *gopher*, apenas devia ser-lhe acrescentado um cabeçalho especial com os títulos internos e com as palavras-chave, em estrutura, qualidade e quantidade suficientes para determinar a sua natureza.

Tendo estabelecido o conceito de *espaço gopher* (*gopherspace*), como um conjunto de computadores com o *servidor gopher*, cada um destes computadores, uma vez acedido apresentava ao utilizador um interface hierárquico de menus e submenus temáticos que conduzia o utilizador até à obtenção automática do documento pretendido, estivesse nesse mesmo computador ou noutra computador do mesmo *espaço gopher*. O utilizador apenas necessitava de ter no seu computador pessoal um *cliente gopher* para realizar o acesso ao menu e a leitura dos documentos seleccionados.

Em 1991 o *gopher* é disponibilizado na Internet e, no período de pouco mais de 3 anos, os *espaços gopher* da Internet contavam já com mais de 10.000 *gophers*, dando acesso a vários milhões de diferentes documentos. Grande parte do sucesso dos *gophers* deve-se à criação, na Universidade do Nevada, em Reno, de um sistema de pesquisa automática, o *Veronica*. Quando um computador *Veronica* era acedido através de um *cliente gopher*, disponibilizava ao utilizador uma hierarquia de menus para a totalidade dos *espaços gopher* da Internet, como se se tratasse de um único *espaço gopher*.

A rápida divulgação de interfaces gráficas com janelas (*Mac-OS*, *NeXTStep*, *X-Windows* e *MS-Windows*) transforma os respectivos *clientes gopher* em percursos da *Internet hipermediática* que se avizinha. Os menus ganham o aspecto de hipertextos com conectores seleccionáveis e *clicáveis* por rato, organizados em topologia de árvore conectiva. Os alvos das conexões são, na sua grande maioria, ficheiros de texto simples. Todavia, embora menos frequentemente, acediam também a ficheiros de outras naturezas mediáticas, como imagens e som, ou mesmo código executável de aplicações informáticas.

Bibliografia:

Não é apresentada qualquer bibliografia já que um dos trabalhos propostos aos alunos consiste na pesquisa de referências bibliográficas, em formato impresso ou digital, que tratem do tema exposto no presente texto.