



# ROTEIRO

- Sobrecarga da Informação
- Uniformização de noções e conceitos
- Representação de Dados
- Representação Documentos / Hiperdocumentos
- Aspectos
- Origem de termos
- Documentos
- HiperDocumentos
- Materialização com qualidade

# Sobrecarga da Informação

- **Falta de padronização**
  - informação (diversidade, sem classificação)
  - mídias (diversidades de digitalização e formatação)
  - armazenamento (diferentes SGBDs e algoritmos de armazenamento)
- **Explosão combinatória de ligações de um nó**
- **Interface**

# Sobrecarga da Informação

- **Desorientação**
- **Perda de contexto**
- **Sobrecarga cognitiva**
- **Confiabilidade [Zobel 97]**



# Sistemas Indexados

- **Caracterizados como :**
  - spiders, robots, wanders, crawlers e worms
- **Exemplos :**
  - WebCrawler
  - Lycos
  - AltaVista
  - Infoseek
  - WWWorm e Cadê
- **Melhorado com Sistemas de Meta Busca**

# SRI x SFI

- **SRI interesse a curto prazo e SFI interesse a longo prazo baseado no perfil do usuário**
- **SRI os hiperdocumentos não fluem freqüentemente e SFI tem um fluxo constante de hiperdocumentos**
- **SRI encontram informações relevantes e SFI removem informações irrelevantes**



# SFI baseado na cooperação

- **Utilizar informações encontradas por terceiros.** Ex.GABS[Wittenburg & al. 95],ComMentor[Rösheisen et al. 95], GroupLens[Resnick 94]. Etc.
- **A noção filtragem cooperativa** (Tapestry da Xerox [Goldberg 92])
- **Aval de Especialistas**

# SFI baseado em agentes

- **Confiabilidade dos agentes**
- **Substituição do parecer humano**
- **Como evoluir os agentes**
- **Subscribe “channels”**
- **Tecnologia Push (Webcasting, Netcasting, PointCast, Marimba, BackWeb)**
- **Leitura independente da conexão**





•  
•  
•

# O que é isto ?

11

# INFORMAÇÃO

- **Informação = valor**

11

# INFORMAÇÃO

- Informação = valor + contexto

Diogo tem 11 anos

# INFORMAÇÃO

- Informação = valor + contexto  
+ representação

Diogo tem  $11\frac{1}{2}$  anos

# Informação

- Diogo é meu filho e ele tinha três anos quando eu escrevi esta frase. Por questões didáticas usei o sistema de numeração binária para representar sua idade.
- Poderia usar outras representações :
  - 3
  - Três
  - III

# INFORMAÇÃO

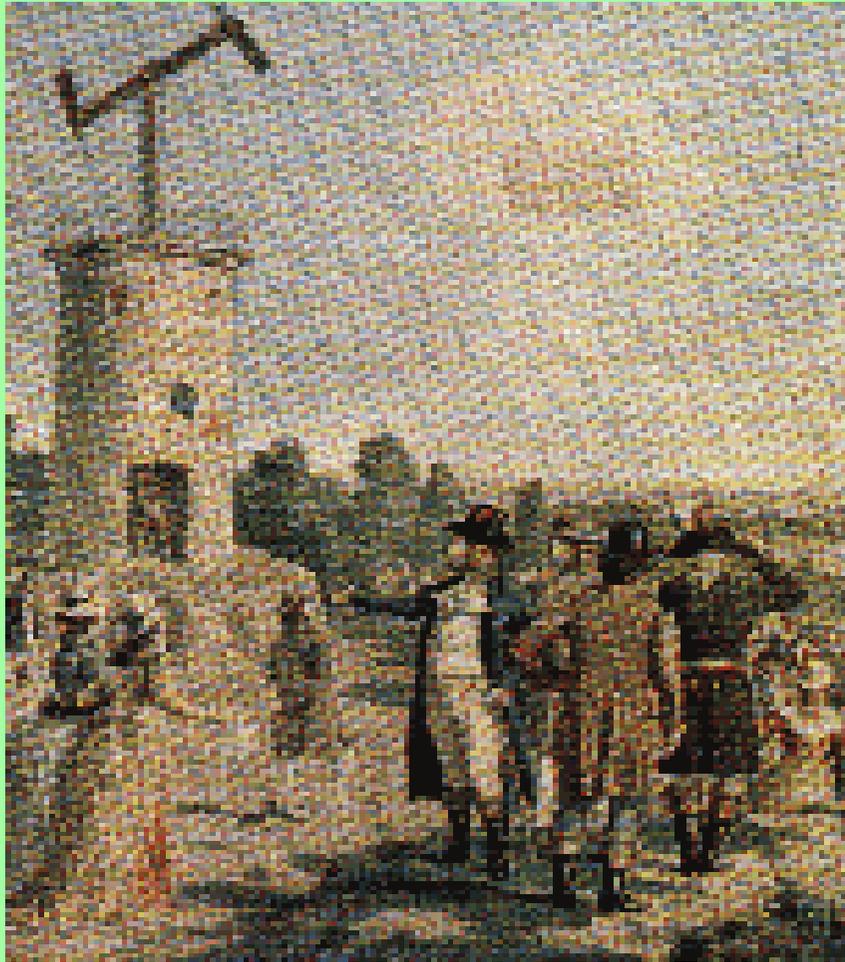
- **Representação**

- Forma : números arábicos, letras, ...

- Sistema : decimal, binário, lingua portuguesa, ....

- **Tempo e Espaço**

# Antena de Chappe



- 1793 até 1830
- 6 horas p/ enviar 25 palavras de Paris a Strasbourg
- +- 500kms
- 52 pessoas em 52 torres

# Treinamento Demorado



- **Palavras e Frases inteiras eram codificadas**

# • • • Telégrafo

- bip longo e bip curto
- Código Morse

— .- = A

- -... = B

--- .-. = C

# Teletipo (TTY)

- **Forma : momento (Com ou Sem corrente)**
- **Sistema : Baudot - Murray**
  - Cinco momentos
  - Exemplos :
    - A = CCSSS ( - ) --> 11000 = 24
    - B = CSSCC ( ? ) --> 10011 = 19
    - C = SCCCS ( : ) --> 01110 = 14

# Computador

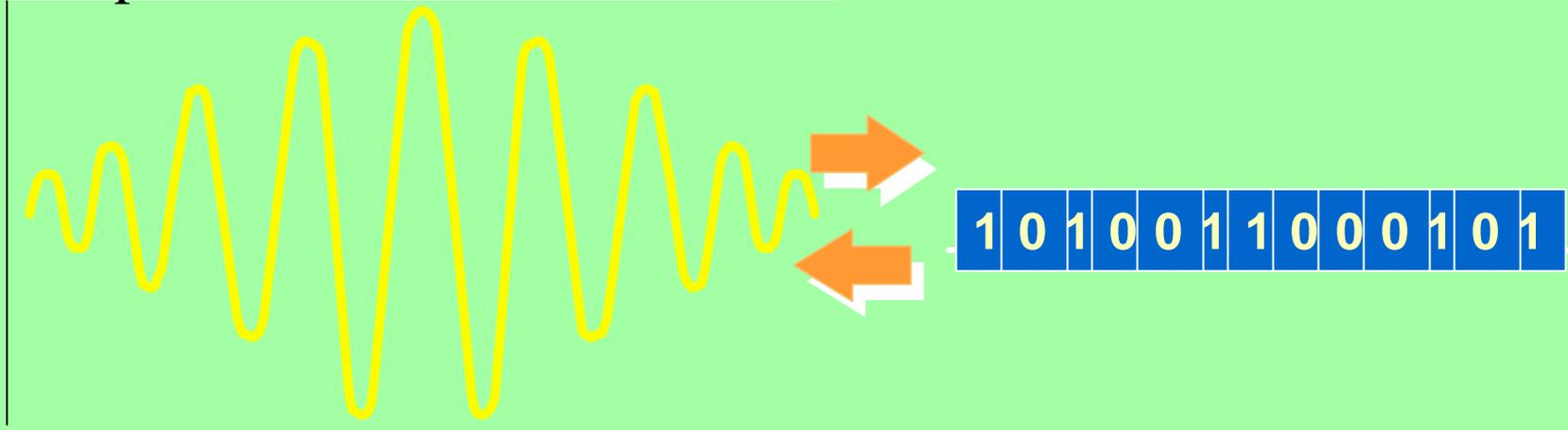
- **Forma : 0 e 1**
- **Sistema : ASCII (7 Bits)**
  - Exemplos :
  - A = 1000001 = 65
  - B = 1000010 = 66
  - C = 1000011 = 67
- **Respeita a ordem alfabética**
- **Não tem falha na seqüência (EBCDIC tinha)**



# Som

- **Conversão analógica / digital**

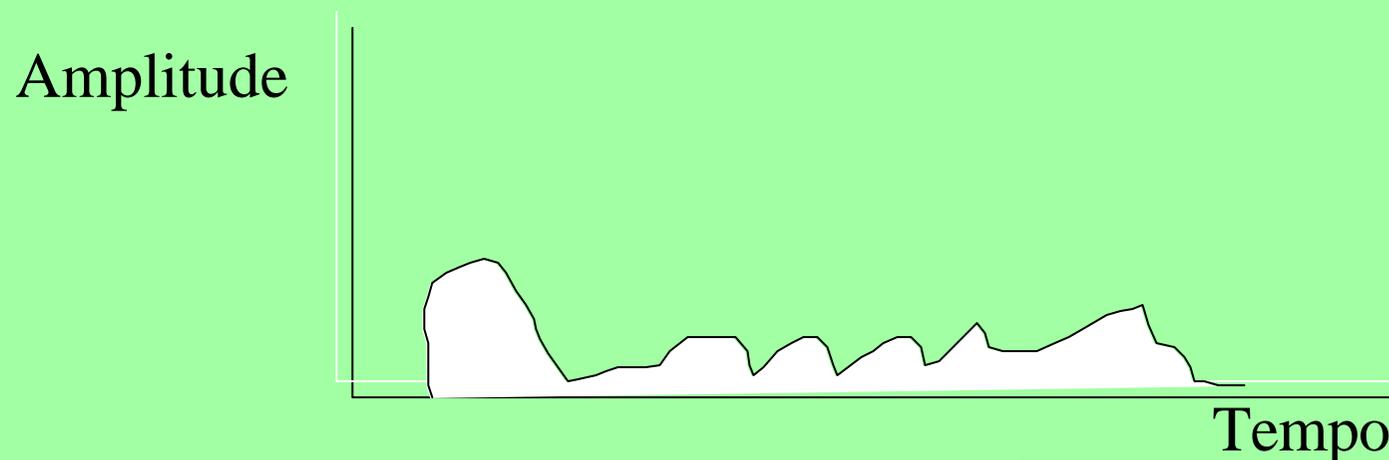
Amplitude



Frequência (vibrações/seg) : 1Hz = 1 vibr./seg  
Ouvido : (20Hz,20KHz)  
Voz : (125Hz,10KHz)

# Som

- **infra som : 0 - 20Hz**
- **ouvido humano : 20 Hz - 20 KHz**
- **Ultra som : 20 KHz - 1 GHz**
- **Hyper som : 1 GHz - 10THz**



# Som - Amostra

- Teorema de Henry Nyquist : para diminuir a perda de qualidade com a digitalização do som se deve ter no mínimo o dobro do máximo da frequência
- CD standard 44,1KHz (2x20KHz)

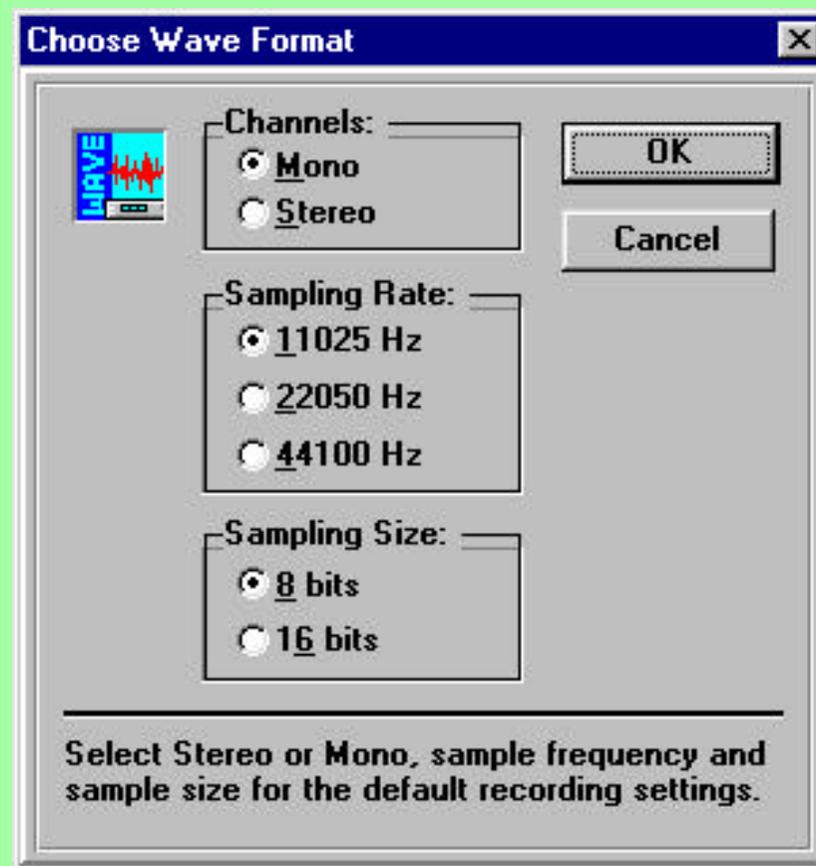
# Som

- **3 variáveis: espaço X qualidade**
  - frequência de captura (=taxa de sampling)
    - 11, 22, 44 Khz
  - tamanho da amostra
    - 8 ou 16 bits
  - mono X estéreo

Pitágoras : Ressonâncias harmônicas (1 corda =  $x$  Hz,  
 $1/2$  corda =  $2x$  Hz) e escala musical de oito notas

Efeito máscara : Os sons fortes mascaram os sons fracos de frequências próximas

# Exemplo de uma interface



# Som

- **1 minuto de áudio**

- 11Khz, 8 bits, mono

 **644 Kbytes**

- 44Khz, 16 bits, estéreoA

 **10 Mbytes**

Digitalização de base  
PCM :

CD-ROM : 44,1Khz 16bits

Rádio FM: 32 Khz 16 bits

Rádio AM : 22 KHz 12 bits

Telefone : 8 KHz 8 bits

# Imagem x Vídeo

- “Uma imagem vale por 1000 palavras”
- Analógico X Digital
- Movimento natural: 24 fps (quadros/s)
- Qualidade X Espaço
  - Foto 20.000.000 pixels (2100x3240)
  - Televisão SECAM 250.000 pixels (625x400)
  - Vídeo VGA 307.000 pixels (640x480)

# Vídeo

- **Variáveis**

- fps (15 a 30)
- tamanho da janela (120 x 160 até 640 x 480)
- cores (256 até 16 milhões)
  - 8, 16 ou 24 bits/pixel



# Vídeo

- **Solução**
  - compressão
  - explorar redundâncias intra-quadro e entre os quadros
  - Quicktime e Video for Windows
  - codecs: INDEO, CINEPAK, MPEG-1







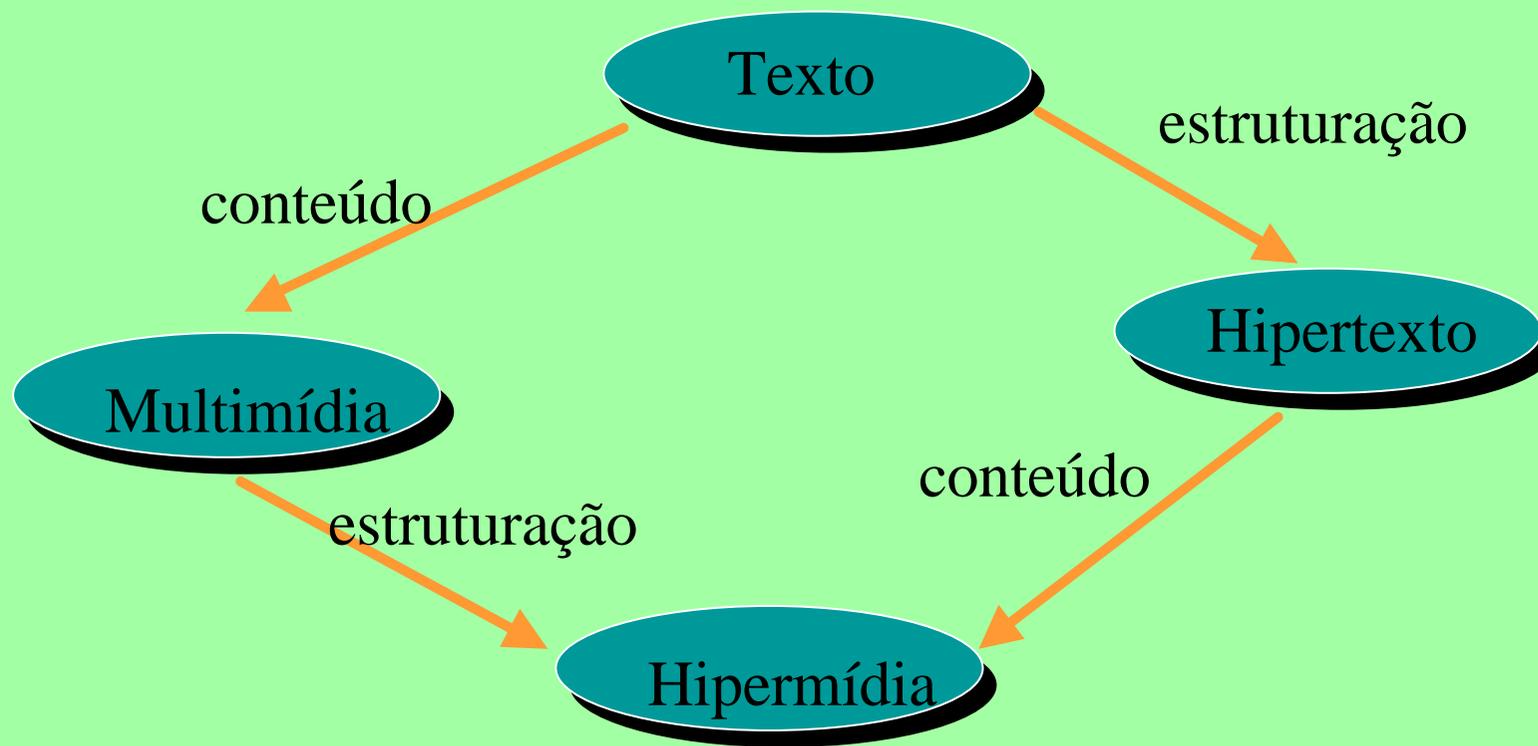
- 
- 
- **Aspectos Importantes em Documentos/Hiperdocumentos**

**Estruturação**

**Conteúdo**

**Acesso**

# Origem dos Termos



# ORIGEM DOS TERMOS

Documentos  
Convencionais

Regras simples de Materialização em vídeo  
Acesso Passivo

Documentos  
Eletrônicos

Regras complexas de materialização  
Acesso interativo/cooperativo  
Multiestruturas  
várias visões

Hiperdocumentos

# Documentos Convencionais

- **Estrutura Lógica**
- **Estrutura de Apresentação**
- **Regras de Materialização / Visualização**
- **Conteúdo Independente de Estruturas**
- **Índices das Referências Internas / Autocontido**





# Hiperdocumentos

- Protocolo de processos sem sobrecarregar o tráfego na rede
- Preenchimento de formulários via WWW distribuindo procedimentos quando possível
- Edição de documentos / hiperdocumentos via WWW distribuindo procedimentos
- Integração com os Bancos de Dados já existentes
- Delimitação da Hiperbase

# Organização da Hiperbase

- **Sistema de Gerência de Hiperdocumentos Ativos**
- **Distribuição dos procedimentos (multiplataformas ou scriptware padrão)**
- **Distribuição das bases de formulários, documentos e hiperdocumentos e conteúdos**
- **Replicação das bases e conteúdos (“mirrors”)**