

# Aplicações e tecnologias multimédia

**Maria Teresa Andrade** <sup>(1), (2)</sup>

**(1) INESC Porto**

**(2) FEUP**

**email: [mandrade@fe.up.pt](mailto:mandrade@fe.up.pt), [maria.andrade@inescporto.pt](mailto:maria.andrade@inescporto.pt)**

# Estrutura da apresentação

- Panorama actual
- Análise de sinais audiovisuais
  - Princípios de análise de sinais de áudio
  - Research trends in content processing of music audio signals
- Acesso universal, adaptatividade
- A importância do contexto
- Normas relevantes
- Adaptação sensível ao contexto
  - Experiências, arquitecturas

# Panorama actual

- progresso contínuo da tecnologia
  - grande variedade de dispositivos clientes com distintas capacidades de processamento e apresentação (**heterogeneidade**)
  - **convergência** dos media
    - produtores de conteúdos expandem a sua oferta a múltiplas plataformas
    - mais tipos de receptores e plataformas a suportar o consumo de dados multimédia

# Panorama actual

Diferentes tipos de redes de acesso e rede "core"

Grande diversidade de dispositivos cliente

Grande diversidade de fontes de conteúdo, profissionais e domésticas



# Expectativas do utilizador

- **Mobilidade**
- **Personalização**
- **Interoperabilidade e transparência**
- **Mais qualidade**
- **Mais escolha**
- **Melhores resultados de pesquisas**
- **Consumir qualquer conteúdo, em qualquer terminal, a qualquer instante**
  - de uma forma transparente

# Dificuldades actuais

- **Largura de banda Wireless**
  - necessária maior banda e menor latência e/ou melhores esquemas de compressão
- **Ligações não-fiáveis**
- **Falta de interoperabilidade entre formatos, aplicações e sistemas**
- **Soluções não-escaláveis**
- **Protocolos em constante evolução**
- **Experiência de utilização pobre (dispositivos “handheld” difíceis de usar, interfaces pouco “user-friendly”)**
- **Falta de uma infraestrutura sólida fornecendo todos os (meta)dados necessários**

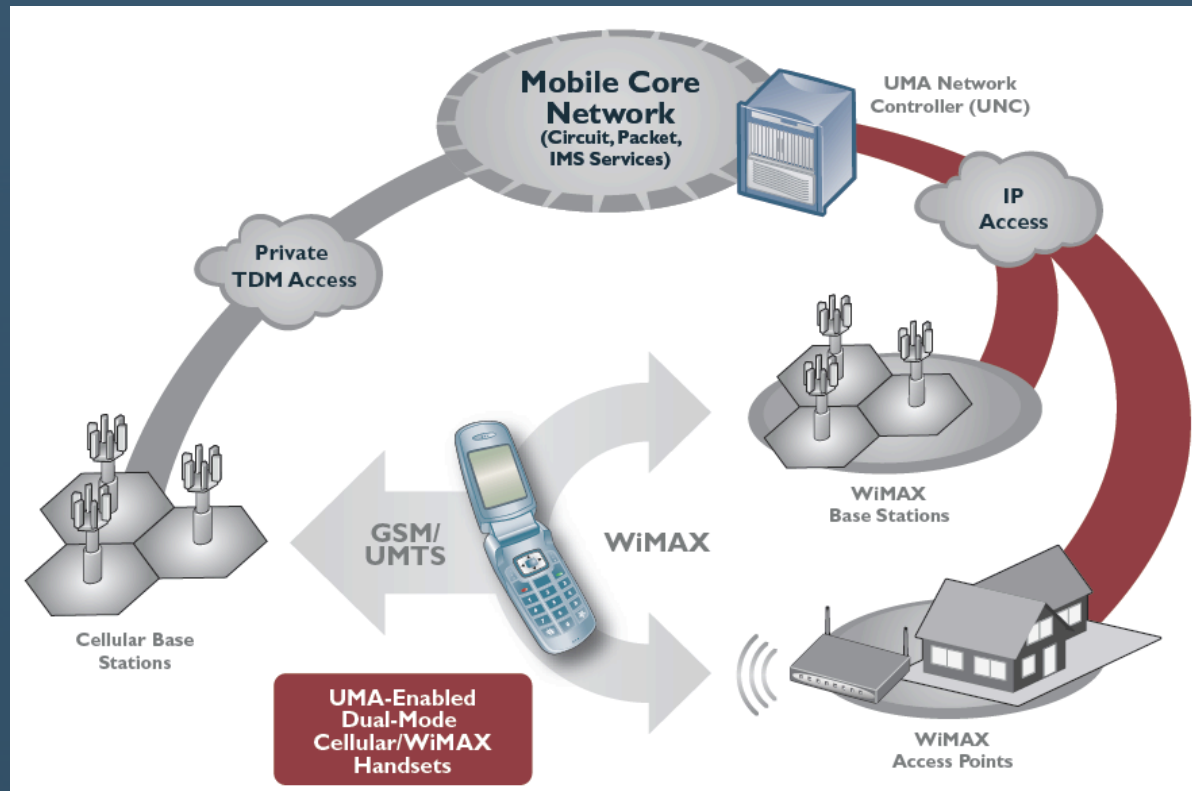
# UMA, o que significa?

- **Universal Multimedia Access**

- Acesso sem restrições a conteúdo multimédia
- A partir de qualquer dispositivo
- Através de qualquer tipo de redes
- Independentemente do formato dos dados originais
- Com garantias de qualidade e eficiência
- Satisfazendo preferências do utilizador

# Não é o 3GPP UMA ...

## Norma 3GPP: Unlicensed Mobile Access (UMA)



- permite aos operadores de telecomunicações, estenderem os seus serviços móveis através de redes de acesso IP (Wi-Fi e banda larga fixa)
- os clientes dos operadores de telecomunicações podem aceder a serviços móveis usando redes de acesso Wi-Fi e banda larga IP (por exemplo utilizando o modem ADSL instalado nas suas casas ou um hot-spot Wi-Fi num local público).



# UMA – de que necessita?

- Conteúdo acessível em rede
- Descrições úteis sobre os conteúdos para permitir pesquisa e localização da informação de interesse
- Múltiplas formas de codificação e representação da informação para adaptação do conteúdo
- Conhecimento do contexto de utilização
- Capacidade de decisão para adaptação

# Metadados - Descrições sobre o conteúdo

- Tipos de descrições sobre o conteúdo
  - semânticas ou de alto nível
  - de baixo nível
- Utilização
  - classificar conteúdo
  - pesquisar conteúdos indo ao encontro do pedido do utilizador (pesquisas úteis)
  - determinar se o conteúdo tem o formato adequado para ser apresentado num determinado dispositivo
  - auxiliar operações de transformação ou adaptação do conteúdo para satisfazer restrições do contexto de consumo
- Uma área de investigação relevante
  - Análise audiovisual

# Metadados de Contexto

- Descrições sobre o contexto de consumo
  - capacidades de terminais
  - características e estado da rede
  - condições/restrições do ambiente de utilização
  - preferências do utilizador
- Utilização
  - determinar se o conteúdo tem formato adequado para ser apresentado em contextos específicos de consumo
  - auxiliar os mecanismos de tomada de decisão de adaptação dos conteúdos

# Abordagens para UMA

- Adaptação de conteúdos baseadas no contexto de utilização
  - criação de perfis e geração a-priori de versões do conteúdo adaptadas a cada perfil; selecção “on-the-fly” da variação adequada
  - geração “on-the-fly” da versão adequada

# Adaptação de conteúdos

- operação efectuada sobre o conteúdo para lhe modificar uma ou mais características, tornando-o adequado ao consumo.
- solução eficiente para permitir o consumo de uma variação do conteúdo, sob condições de rede e do ambiente de utilização que impediam o consumo do conteúdo original.

# Aplicações "Context-aware"

- aplicações com a capacidade de detectar, interpretar e reagir a aspectos relacionados com o contexto de utilização
  - preferências de utilizador, características do ambiente de utilização, capacidades dos terminais ou condições de rede,
- modificando ou adaptando dinamicamente o seu comportamento com base nesses aspectos.
- **requerem motores de decisão de adaptação e motores de adaptação de conteúdos**

# Requisitos para adaptação

- disponibilidade de descrições sobre o conteúdo e sobre o contexto de utilização.
- mecanismos que usem essas descrições para tomar decisões – “**motores de decisão** de adaptação” - e mecanismos que transformem o conteúdo – “**motores de adaptação**”.
- apoio arquitectural para permitir a troca das descrições e a interoperabilidade entre sistemas e modelos de dados distintos.

# Apoio arquitectural

- Infraestrutura baseada em normas e formatos abertos que permita a inter-operação entre plataformas distintas
  - dispositivos receptores, repositórios de conteúdos ou fontes de informação, serviços, motores de pesquisa, motores de decisão e motores de adaptação
- Utilização de um formato comum que permita a re-utilização de metadados ao longo de toda a cadeia de distribuição (ex., MPEG-21)



# Motores de adaptação (1)

- baseados no tipo de operação realizada sobre o conteúdo (**adaptação por transformação**):
  - transcodificação entre formatos distintos
    - ex., mp2 vídeo para h.264
  - escalonamento espacial, temporal e de débito
    - dentro do mesmo formato mas alterando as dimensões espaciais, o nº de imagens por segundo ou o débito
  - "transmoding"
    - transformação de um tipo de media num outro
      - ex., conversão texto-voz
  - filtragem
    - removendo componentes, partes ou tipos de media

# Motores de adaptação (2)

- classificados de acordo com o instante de tempo em que a operam sobre o conteúdo:
  - **adaptação estática**
    - por selecção de uma entre múltiplas versões do conteúdo, previamente criadas
  - **adaptação dinâmica**
    - adaptação por transformação na qual uma nova versão é gerada "on-the-fly" a partir da versão única original

# Tipos de adaptação (3)

- De acordo com o tipo de media em que operam:
  - **“single media”**
    - conteúdo é adaptado em termos de qualidade ou formato, tal como por exemplo, áudio de mp3 para wav
  - **“cross-media”** ou “transmoding”
    - um recurso de um dado tipo ou media é transformado num outro tipo (por ex., audio para texto, vídeo para imagem)

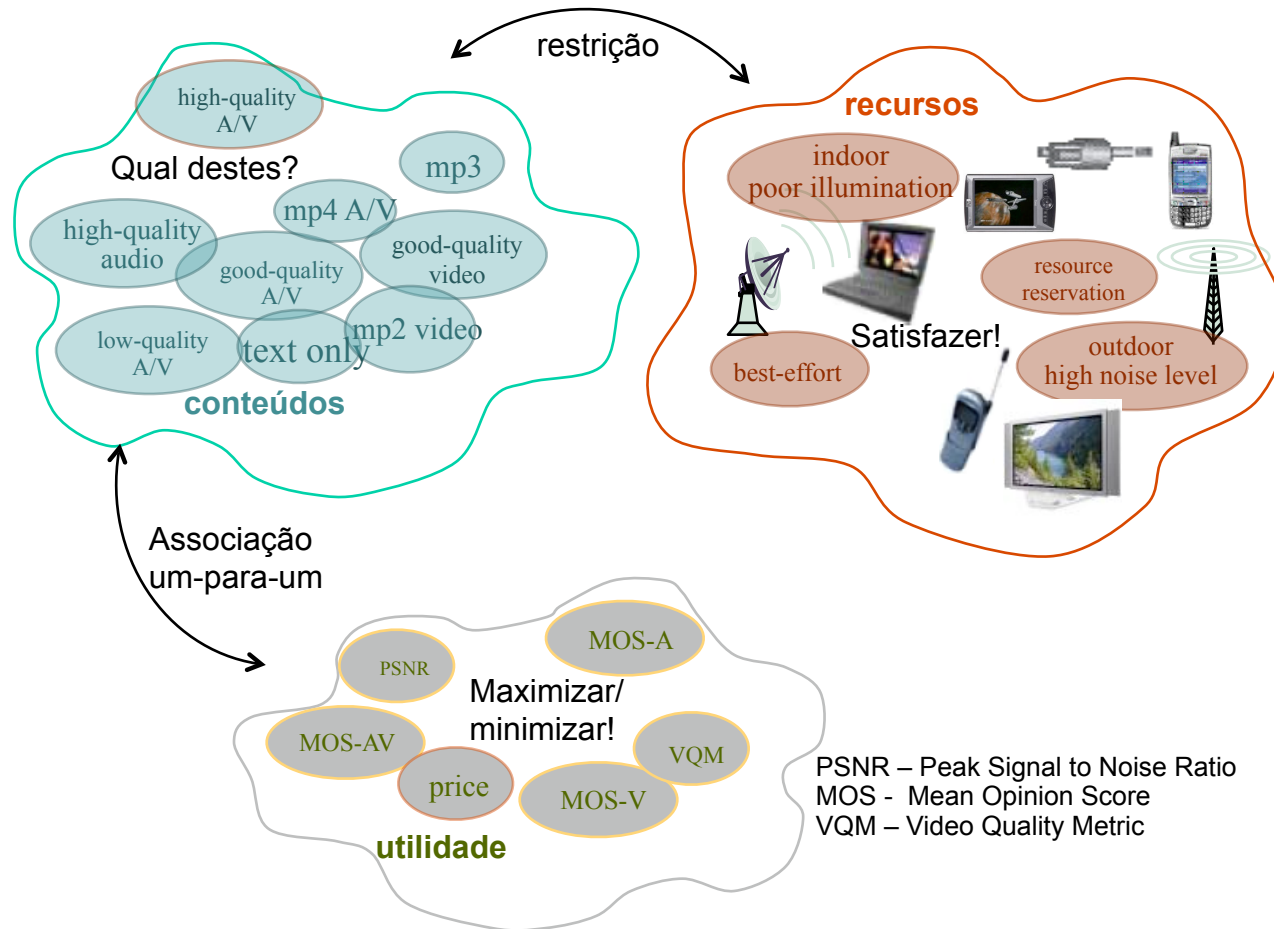
# Tipos da adaptação (4)

- de acordo com a localização
  - “Server-based”
    - servidor tem toda a responsabilidade de reunir informação e tomar decisões
  - “Proxy-Based”
    - servidor passivo, sendo o proxy a tomar decisões e a executar a adaptação
  - “Client-based”
    - a inteligência e capacidades de adaptação estão localizadas no cliente. Não há necessidade de passar descrições de contexto.
  - Adaptação distribuída
    - combinação das anteriores usando também nós da rede

# Motores de decisão

- **Activados por eventos correspondentes a variações das condições do contexto de consumo**
  - Ex., congestão na rede, perda de qualidade, aumento do ruído ambiente, diminuição da intensidade luminosa, etc.
- **Problema de atribuição de recursos otimizando um dado aspecto**
  - Por ex., qualidade, preço, latência, etc
- **Operam num espaço tri-dimensional**
  - Conteúdo, recursos (de rede, computacionais) e utilidade (factor a otimizar)

# Espaço tri-dimensional da decisão de adaptação



# Importância do contexto

- **Contexto de baixo nível**

- Adquirido automaticamente através de sensores físicos ou lógicos
- Ex., probe software a medir a qualidade de um vídeo, módulo software a medir a largura de banda na interface de rede, uma camera a medir a intensidade luminosa, um microfone para detectar volume do som, etc.

- **Contexto de alto nível**

- Descrição de situações reais de consumo de conteúdo usando conceitos semelhantes aos utilizados pelos homens

# MPEG-21

- norma aberta baseada em XML
- conjunto de especificações de formatos e de ferramentas para
  - facilitar a transacção e consumo de conteúdo multimédia em ambientes heterogéneos, satisfazendo características do contexto de utilização, preferências do utilizador, estado das redes e direitos de autor
- Permite representar o contexto de baixo nível e ainda possíveis adaptações
- em fase final de desenvolvimento; constituída por 17 partes, algumas das quais já finalizadas, outras ainda na fase inicial de desenvolvimento.



# MPEG-21 conceitos fundamentais

- **Item Digital (DI – Digital Item)**
  - conjunto de recursos multimédia relacionados com um certo tópico (ou fazendo parte da mesma peça, programa ou serviço) e descrições associadas
  - Inclui um documento XML designado de **DID** (Digital Item Declaration – parte 2 da norma MPEG-21) que permite declarar e descrever as partes constituintes do DI.
- utilizador – qualquer actor (sistema ou pessoa) que interage com o DI

# MPEG-21 DI

- **Item Digital** – objecto complexo

- do ponto de vista conceptual pode ser visto como um pacote de conteúdos multimédia, incluindo metadados, relacionados com um dado tópico

```
<?xml version="1.0" encoding="UTF-8"?>
<DIDL xmlns="urn:mpeg:mpeg21:2002:01-DIDL-NS">
  <item>
    <Descriptor id="final_European_Championship_2004">
      <Component>
        <Resource mimeType="video/mpeg2" ref="videos/final.mpg">
        </Component>
      <Descriptor id="foto_team_Portugal">
        <Component>
          <Resource mimeType="image/jpeg" ref="photos/Portugal.jpg">
          </Component>
        <Descriptor id="interview_team_coach">
          <Component>
            <Resource mimeType="audio/mp3" ref="audio/coach_Portugal.mp3/">
            </Component>
          </item>
        </DIDL>
```

MPEG-21 DID



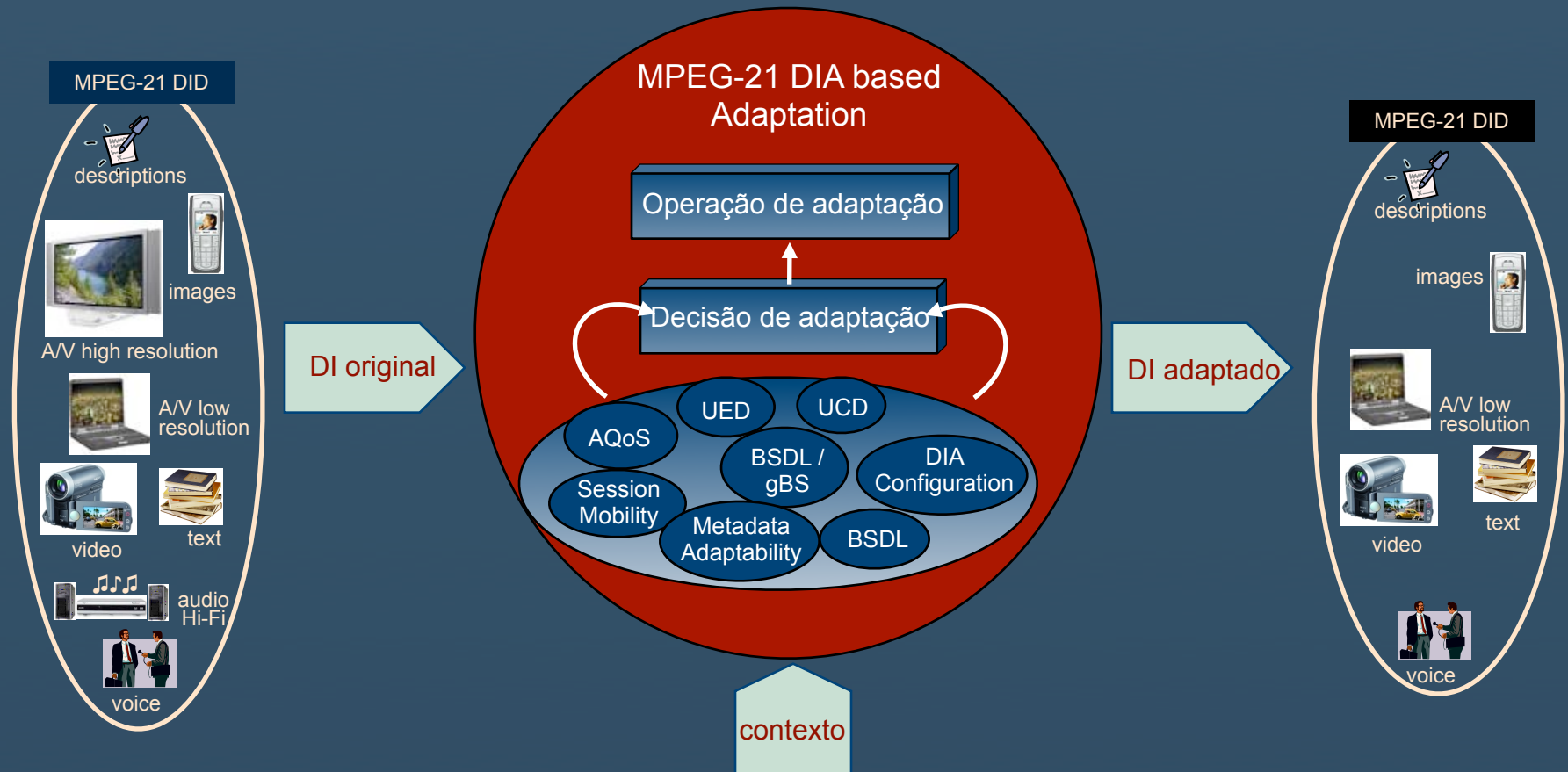
# MPEG-21 - adaptação

- parte 7 (**DIA – Digital Item Adaptation**) - especifica um conj<sup>o</sup> de ferramentas de descrição para auxiliar a adaptação de DIs:
  - descrições do contexto de utilização e da estrutura dos bitstreams;
  - relações entre restrições/limitações de utilização e possíveis operações de adaptação
  - relações entre operações de adaptação e resultado obtido em termos de qualidade

# MPEG-21 - DIA

- Usage Environment Description – **UED**
  - inclui características do utilizador, capacidades dos terminais, características de rede e do ambiente natural de utilização)
- Universal Constraints Description – **UCD**
  - permite descrever restrições de utilização e a optimização das operações de adaptação que satisfazem essas restrições
- Terminal and Network QoS – **AdaptationQoS, AQoS**
  - descreve as relações entre restrições/limitações de qualidade e as operações de adaptação que satisfazem essas restrições, assim como o resultado esperado

# Adaptação com MPEG-21



# Outras normas: IETF, OMA

- **IETF, Internet Content Adaptation Protocol (iCAP)**
  - Baseado em HTTP do tipo Remote Procedure Call (RPC)
  - Os clientes podem enviar mensagens HTTP com instruções para servidores iCAP realizarem certas operações sobre o conteúdo
  - Ex., “virus scanning”, inserção de publicidade, algumas formas de filtragem, etc.
- **OMA, User Agent Profile (UAProf)**
  - Vocabulário usado por clientes WAP (Wireless Access Protocol) para comunicarem as suas capacidades aos servidores
  - Ex., características do hardware (tamanho do écran ou tipo de teclado), características de software (tipo de browser) e preferências (ex, som off/on, escolha de côr, etc)

# Outras normas: W3C

- **Resource Description Framework (RDF)**
  - Especificação genérica para representar metadados
  - Tem 3 formas distintas de representação
    - 1) representação "Triplet", consistindo em conj<sup>o</sup> de valores (Sujeito, Predicado, Objecto)
    - 2) sintaxe XML, normalmente usado para troca de dados RDF entre máquinas
    - 3) representação em grafo

"The creator of <http://www.visnet.org/index.html> is Peter Eisert"

i)	<i>Subject</i>	<i>Predicate</i>	<i>Object</i>
	<a href="http://www.visnet.org/index.html">http://www.visnet.org/index.html</a>	<a href="http://purl.org/dc/elements/1.1/creator">http://purl.org/dc/elements/1.1/creator</a>	<a href="http://www.visnet.org/members/0011">http://www.visnet.org/members/0011</a>

<?xml version="1.0"?>

ii) <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:dc="http://purl.org/dc/elements/1.1/">  
<rdf:Description rdf:about="http://www.visnet.org/index.html">  
<dc:creator>Peter Eisert</dc:creator>  
</rdf:Description>  
</rdf:RDF>

iii) 

```
graph LR; S("http://www.visnet.org/index.html") -- "http://purl.org/dc/elements/1.1/creator" --> O("http://www.visnet.org/members/0011");
```

# Outras normas: W3C

- **Composite Capability/ Preference Profiles (CC/PP)**

- Baseado em RDF para descrever capacidades de clientes e preferências de utilizadores através da informação "user agent" do pedido HTTP
- As preferências do utilizador são enviadas no cabeçalho do pedido HTTP

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:prf="http://www.visnet-noe.org/test-profile-vocabulary#">
  <rdf:Description about="HardwarePlatform">
    <prf:Defaults
      Vendor="VISNET"
      Model="2007"
      Type="PDA"
      Screensize="640x480x24"
      Keyboard="Yes"
      Speaker="Yes"
      memory="32MB"
    </prf:Defaults>
  </rdf:Description>
  <rdf:Description about="SoftwarePlatform">
    <prf:Defaults
      OS="UltraVIS1.0"
      HTMLVersion="4.0"
      JavaScriptVersion="4.0"
      WAPVersion="1.0"
    </prf:Defaults>
  </rdf:Description>
  <rdf:Description about="UserPreferences">
    <prf:Defaults
      Language="English"
    </prf:Defaults>
  </rdf:Description>
</rdf:RDF>
```



# Outras normas: W3C

- **Web Ontology Language (OWL)**

- É a especificação do W3C para desenvolver ontologias
- Ontologias fornecem os meios para definir termos e conceitos representando situações reais e para descrever as relações entre esses termos de uma forma “machine-readable”
- No âmbito do paradigma “Semantic Web”, as ontologias são vistas como o veículo para conseguir a utilização transparente de recursos, para atingir a interoperabilidade

# Utilização de ontologias em sistemas "context-aware"

- Uma ontologia é usada para definir conhecimento sobre um domínio, permitindo uma descrição formal de situações reais específicas nesse domínio
  - os conceitos, qualquer entidade que participa no serviço multimédia, tal como terminal, as redes ou o utilizador, são descritos numa ontologia pelas suas **propriedades**
    - por exemplo, o conceito terminal pode ter as propriedades: "3'" (dimensão do écran), "GPS", "UMTS", etc.
  - e também pelas **relações** entre as propriedades e conceitos

# Utilização de ontologias em sistemas “context-aware”

- Motores de decisão de adaptação podem tirar partido de ontologias
  - Adquirem contexto de baixo nível a partir de sensores que accionam o processo de adaptação
  - Juntamente com os conceitos e regras definidos na ontologia, esses dados são usados para “raciocinar” e inferir contextos de alto nível, ou seja, situações reais.
  - Permitem assim uma análise mais perto da realidade e da forma como o ser humano pensa
- **A adaptação poderá ir mais ao encontro das expectativas do utilizador**

# case study: Adaptação MPEG-21

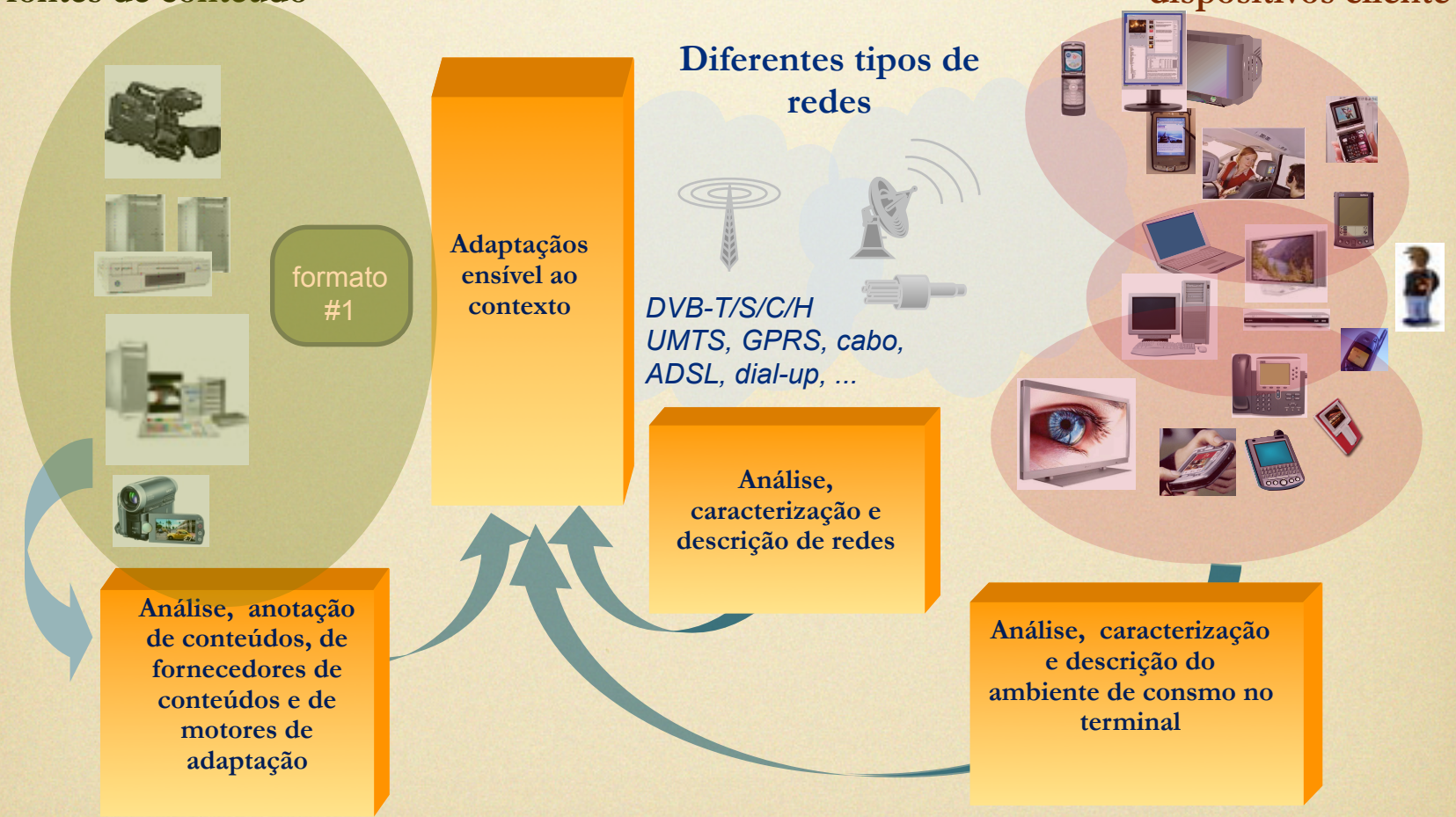
- **conteúdo único na fonte - objecto digital complexo (Digital Item)**
  - para além dos componentes media, inclui documento XML (DID, Digital Item Declaration), descrevendo a estrutura do objecto e fornecendo descrições para auxiliar a sua adaptação
- **recolha de descrições de contexto da rede, terminal e dispositivos de adaptação disponíveis**
- **tomada de decisão de adaptação, tentando maximizar a satisfação do cliente dadas as restrições**
- **adaptação on-the-fly**



# Adaptação de conteúdos

fontes de conteúdo

dispositivos cliente



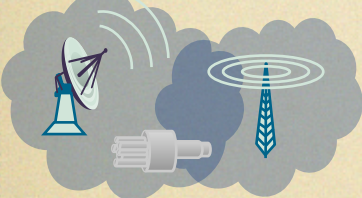


---

Fonte de conteúdo



Recursos de rede



disponibilidade de conteúdos

decisão de adaptação

disponibilidade de recursos



Consumidor



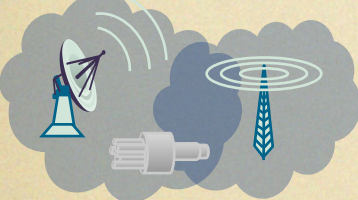


---

Fonte de conteúdo



Recursos de rede



disponibilidade de  
conteúdos

decisão de  
adaptação

disponibilidade  
de recursos



**Consumidor**

pede conteúdo usando palavra-chave  
“Gladiador”



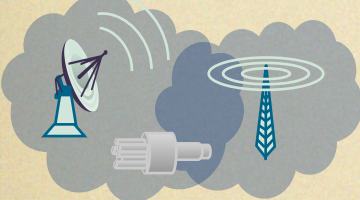


---

Fonte de conteúdo



Recursos de rede



disponibilidade de  
conteúdos

decisão de  
adaptação

disponibilidade  
de recursos



**Consumidor**

pede conteúdo usando palavra-chave  
“Gladiador”



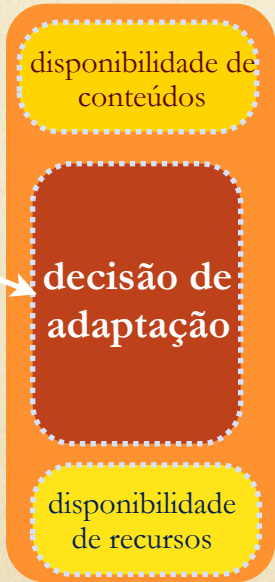


```
<MediaFormat>
  <VisualCoding>
    <Format ...
      MPEG-2 Video Main Profile @ Main Level
    </Format>
    <Frame height="720" width="480" rate="30"/>
    <BitRate>5000000</BitRate>
  </VisualCoding>
</MediaFormat>
```

Fonte de conteúdo



Recursos de rede



**Consumidor**

pede conteúdo usando palavra-chave  
“Gladiador”



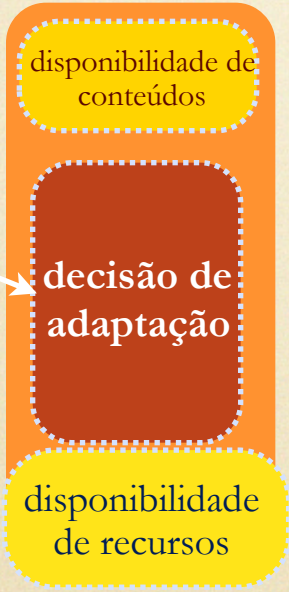
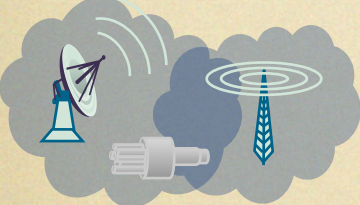


```
<MediaFormat>
  <VisualCoding>
    <Format ...
      MPEG-2 Video Main Profile @ Main Level
    </Format>
    <Frame height="720" width="480" rate="30"/>
    <BitRate>5000000</BitRate>
  </VisualCoding>
</MediaFormat>
```

Fonte de conteúdo



Recursos de rede



**Consumidor**

pede conteúdo usando palavra-chave  
“Gladiador”



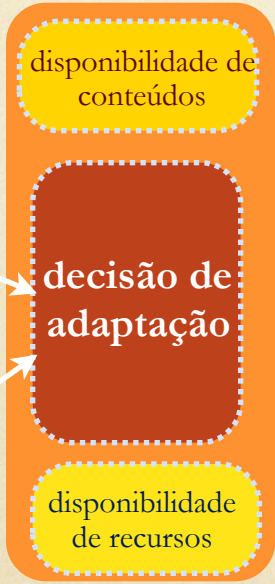


```
<MediaFormat>
  <VisualCoding>
    <Format ...
      MPEG-2 Video Main Profile @ Main Level
    </Format>
    <Frame height="720" width="480" rate="30"/>
    <BitRate>5000000</BitRate>
  </VisualCoding>
</MediaFormat>
```

Fonte de conteúdo



Recursos de rede



```
<DIA>
  ...
  <... "NetworkCapabilityType"
    maxCapacity="640000" minGuaranteed="96000"/>
  <... "NetworkConditionType"
    AvailableBandwidth maximum="460000"/>
</DIA>
```



Consumidor

pede conteúdo usando palavra-chave  
"Gladiador"





```

<MediaFormat>
  <VisualCoding>
    <Format ...
      MPEG-2 Video Main Profile @ Main Level
    </Format>
    <Frame height="720" width="480" rate="30"/>
    <BitRate>5000000</BitRate>
  </VisualCoding>
</MediaFormat>

```

```

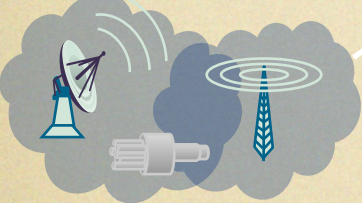
<TerminalCapability xsi:type="CodecCapabilitiesType">
  ...
  <Decoding ...
    MPEG-4 Visual Simple Profile @ Level 1
  </Decoding>
  ...
</TerminalCapability>

```

Fonte de conteúdo



Recursos de rede



Consumidor

pede conteúdo usando palavra-chave  
"Gladiador"



```

<DIA>
  ...
  <... "NetworkCapabilityType"
    maxCapacity="640000" minGuaranteed="96000"/>
  <... "NetworkConditionType"
    AvailableBandwidth maximum="460000"/>
</DIA>

```



```

<MediaFormat>
  <VisualCoding>
    <Format ...
      MPEG-2 Video Main Profile @ Main Level
    </Format>
    <Frame height="720" width="480" rate="30"/>
    <BitRate>5000000</BitRate>
  </VisualCoding>
</MediaFormat>

```

```

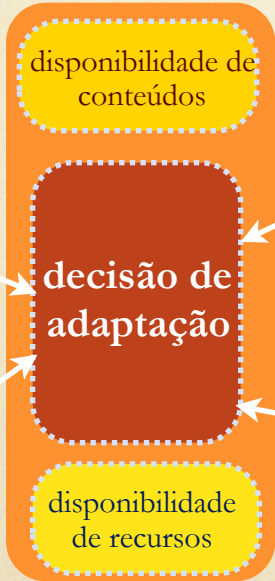
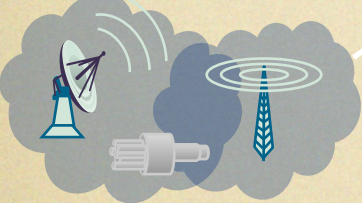
<TerminalCapability xsi:type="CodecCapabilitiesType">
  ...
  <Decoding ...
    MPEG-4 Visual Simple Profile @ Level 1
  </Decoding>
  ...
</TerminalCapability>

```

Fonte de conteúdo



Recursos de rede



```

<DIA>
  ...
  ... "NetworkCapabilityType"
  maxCapacity="640000" minGuaranteed="96000"/>
  ... "NetworkConditionType"
  AvailableBandwidth maximum="460000"/>
</DIA>

```

```

<TerminalCapability xsi:type="CodecCapabilitiesType">
  <Decoding ...
    JPEG
  </Decoding>
  ...
</TerminalCapability>

```

Consumidor  
pede conteúdo usando palavra-chave  
"Gladiador"





```

<MediaFormat>
  <VisualCoding>
    <Format ...
      MPEG-2 Video Main Profile @ Main Level
    </Format>
    <Frame height="720" width="480" rate="30"/>
    <BitRate>5000000</BitRate>
  </VisualCoding>
</MediaFormat>

```

```

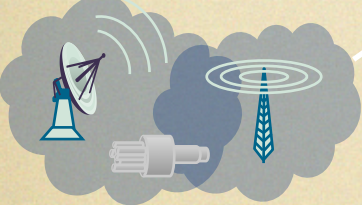
<TerminalCapability xsi:type="CodecCapabilitiesType">
  ...
  <Decoding ...
    MPEG-4 Visual Simple Profile @ Level 1
  </Decoding>
  ...
</TerminalCapability>

```

Fonte de conteúdo



Recursos de rede



```

<DIA>
  ...
  ... "NetworkCapabilityType"
  maxCapacity="640000" minGuaranteed="96000"/>
  ... "NetworkConditionType"
  AvailableBandwidth maximum="460000"/>
</DIA>

```

```

<TerminalCapability xsi:type="CodecCapabilitiesType">
  <Decoding ...
    JPEG
  </Decoding>
  ...
</TerminalCapability>

```



```

<MediaFormat>
  <VisualCoding>
    <Format ...
      MPEG-2 Video Main Profile @ Main Level
    </Format>
    <Frame height="720" width="480" rate="30"/>
    <BitRate>5000000</BitRate>
  </VisualCoding>
</MediaFormat>

```

```

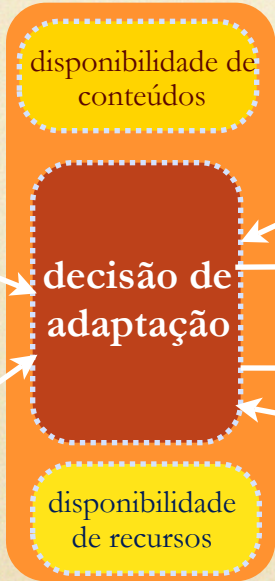
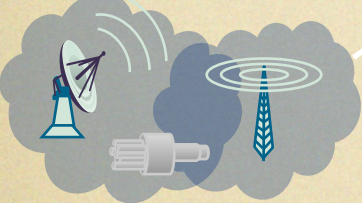
<TerminalCapability xsi:type="CodecCapabilitiesType">
  ...
  <Decoding ...
    MPEG-4 Visual Simple Profile @ Level 1
  </Decoding>
  ...
</TerminalCapability>

```

Fonte de conteúdo



Recursos de rede



```

<TerminalCapability xsi:type="CodecCapabilitiesType">
  ...
  <Decoding ...
    JPEG
  </Decoding>
  ...
</TerminalCapability>

```



```

<DIA>
  ...
  ... "NetworkCapabilityType"
  maxCapacity="640000" minGuaranteed="96000"/>
  ... "NetworkConditionType"
  AvailableBandwidth maximum="460000"/>
</DIA>

```

```

<TerminalCapability xsi:type="CodecCapabilitiesType">
  <Decoding ...
    JPEG
  </Decoding>
  ...
</TerminalCapability>

```