



# 数字媒体技术基础

MengYang

[www.smartllv.com](http://www.smartllv.com)

SUNYAT-SENUniversity



机器智能与先进计算教  
育部重点实验室



智能视觉语言  
学习研究组

- 3. 3人工智能+新媒体 7
- 3. 3. 1人工智能+视觉新媒体 7
- 3. 3. 2人工智能+音频新媒体 7
- 3. 3. 3人工智能+文本新媒体 7
- 3. 3. 4人工智能+多新媒体 7

## 第一部分

### 1. 3人工智能+新媒体

## □ 概述

- Web2.0之后媒体变革的起点是移动互联网，但它更大的趋向是媒体智能化，或智媒化。
- 智媒化的特征包括三个方面：
- 万物皆媒：过去的媒体是以人为主导的媒体，而未来，机器及各种智能物体都有媒体化可能。
- 人机合一：智能化机器、智能物体将与人的智能融合，共同作用，构建新的媒体业务模式。
- 自我进化：人机合一的媒介具有自我进化的能力，机器洞察人心的能力、人对机器的驾驭能力互为推进。

- 智媒化将带来以下几方面的可能：用户分析与匹配的场景化、智能化与精准化：智能化的媒体将更好地洞察每个个体用户在特定场景下的行为与需求，并智能推荐其所需要的信息与服务。新闻生产的机器化、智能化与分布式：智能化机器进入到新闻信息的采集、分析、写作等环节，改变现有的生产模式。另一方面，由多元主体在去中心化的模式下完成的协作式报道，在未来将更为普遍。新闻传播的泛在化、智能化与新闻体验的临场化：各种智能物体将成为新闻接收的终端从而为用户提供无所不在的信息获取，而VR / AR等技术，将为人们塑造全新的新闻临场感。互动反馈的传感化与智能化：用户在信息消费过程中的生理反应，将通过传感器直接呈现，用户反馈将进入到生理层面。支持这些新可能的新技术会导致全新的力量对新闻业的“入侵”，这也意味着，智媒化时代将是一个传统传媒业边界消失、格局重塑的时代。

## □ 关键技术

- 基础设施的物联网
- 构成生产要素的大数据
- 优化资源匹配的移动计算
- 加速智能进化的机器学习

# 人工智能+视觉新媒体



## □ 概述



# 人工智能+视觉新媒体



## 国内首个媒体人工智能平台MAGIC介绍





# 人工智能+视觉新媒体



## 国内首个媒体人工智能平台MAGIC影响



## 国内首个媒体人工智能平台MAGIC的视图效果



- 国内首个媒体人工智能平台MAGIC的技术实现概述1（用4页给出一个较新的经典方法的通俗性介绍）



- 国内首个媒体人工智能平台MAGIC的技术实现概述1（用4页给出一个较新的经典方法的通俗性介绍）



- 国内首个媒体人工智能平台MAGIC的技术实现概述1（用4页给出一个较新的经典方法的通俗性介绍）



- 国内首个媒体人工智能平台MAGIC的技术实现概述1（用4页给出一个较新的经典方法的通俗性介绍）



# 人工智能+音频新媒体



## □ 概述



- 加入AI的语音合成（个人音色、情绪等）介绍





- 加入AI的语音合成（个人音色、情绪等）影响



- 加入AI的语音合成（个人音色、情绪等） 音频案例



- 加入AI的语音合成（个人音色、情绪等）技术概述1



- 加入AI的语音合成（个人音色、情绪等）技术概述1



- 加入AI的语音合成（个人音色、情绪等）技术概述1



- 加入AI的语音合成（个人音色、情绪等）技术概述1



# 人工智能+音频新媒体



## □ 概述



- 加入AI的语音合成（个人音色、情绪等）介绍





- 加入AI的语音合成（个人音色、情绪等）影响



- 加入AI的语音合成（个人音色、情绪等） 音频案例



- 加入AI的语音合成（个人音色、情绪等）技术概述1



- 加入AI的语音合成（个人音色、情绪等）技术概述1



# 数字媒体的划分

- 按按来源属性分，则可分成自然媒体（Natural media）和合成媒体

（Synthetic media）。其中自然媒体是指客观世界存在的景物，声音等，经过专门的设备进行数字化和编码处理之后得到的数字媒体，比如数码相机拍的照片，数字摄像机拍的影像，MP3数字音乐、数字电影电视等。合成媒体则是指的是以计算机为工具，采用特定符号，语言或算法表示的，由计算机生成（合成）的文本，音乐，语音，图象和动画等，比如用3D制作软件制作出来的动画角色。

# 传统数字媒体

- ❑ 数字图形（矢量图等）
- ❑ 图像（图像搜索引擎）
- ❑ 视频（优酷、youtube）



## □ 数字音频



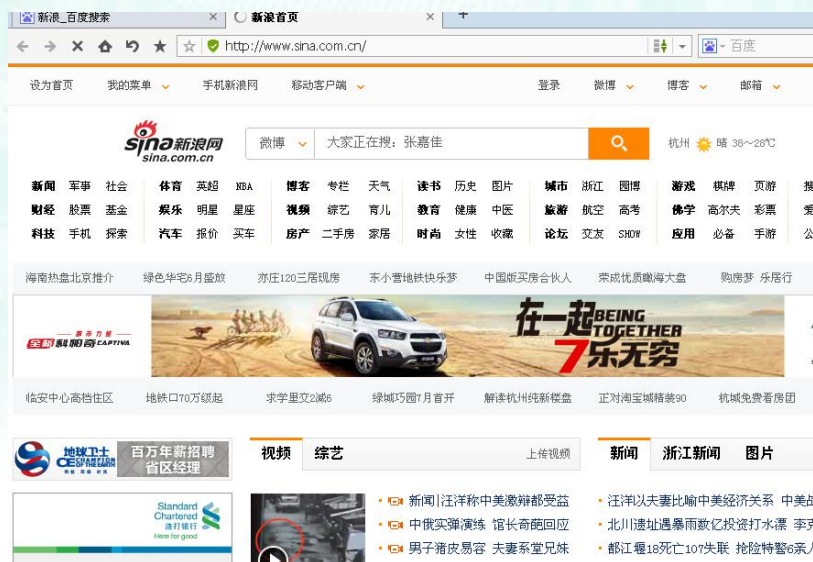
## ❑ 数字杂志新闻小说





## 网页

## 网页的超文本传输协议，超媒体传输协议





# 使用HTML设计网页

- ❑ 超文本标记语言（HyperTextMarkupLanguage, 简称HTML）是Web页面的标记性语言。通过使用一组约定的标记符号，对Web上的各种信息进行标记，浏览器会解释这些标记符号并以它们指定的格式把相应的内容显示在屏幕上，而标记符号本身不会在屏幕上显示出来。



# HTML文档的基本结构



```
MyFirstHTMLPage.txt - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
<html>
  <head>
    <title>使用记事本编写HTML文件</title>
  </head>
  <body>
    <p>
      我的第一个HTML网页
    </p>
  </body>
</html>
```

The screenshot shows a Notepad window titled "MyFirstHTMLPage.txt - 记事本". The menu bar includes "文件(F)", "编辑(E)", "格式(O)", "查看(V)", and "帮助(H)". The text content is an HTML document structure. Red boxes are drawn around the code to illustrate nesting: the outermost box is around the entire document from `<html>` to `</html>`; a second box is around the `<head>` and `</head>` section; a third box is around the `<body>` and `</body>` section; and a fourth box is around the `<p>` and `</p>` paragraph. This visualizes the "nested" structure of HTML documents.

- HTML文档拥有一种“嵌套”结构

# HTML标记说明

- ❑ 标记往往成对出现，开始标记与结束标记之间的部分为标记的内容，如

`<html>...</html>;`

- ❑ 开始标记和结束标记一般配对使用，但有些标记可以省略结束标记，如

`<br><img><input>;`

- ❑ 标记可以嵌套，但不能交叉嵌套；
- ❑ 在开始标记中可以包含对属性值的设置。

# 使用XML表达数据

- ❑ XML (eXtensible Markup Language) 可扩展标记语言，将网络上的文档规范化并赋予标记一定的含义，同时XML又不仅仅只是标记语言，它还提供了一个标准，用户可以利用这个标准定义新的标记语言。

- 可扩展标记语言XML (eXtensible Markup Language) 是为了克服HTML缺乏灵活性和伸缩性的缺点以及SGML (Standard Generalized Markup Language) 过于复杂、不利于软件应用的缺点而发展起来的一种元标记语言，是未来网络科技的第二代HTML语言。
  - HTML的标记是给定的，用户不能增加新标记是不易扩展的；
  - XML标记是可以扩展的，用户可以根据需要定义新的标记。
- XML的标记是区分大小写的。一个XML文档只能有一个根标记，其他标记分层嵌套从而形成一颗标记树。XML元素由一个开始标记和一个结束标记组成，它可以包含其他子元素，HTML中的某些元素不一定需要一个结束标记（比如，LI和BR），并且某些属性不需要用引号引起来。

**【例】**使用XML在定义员工的姓名，出生日期与电话号码等信息。

编辑XML源文件。在记事本或其他文本编辑器中输入以下代码，以XMLFile.xml存盘。

```
<?xml version="1.0" encoding="utf-8" ?>
<?xml-stylesheet type="text/xsl" href="XSLTFile.xsl"?>
<personlist>
  <person>
    <name>李明</name>
    <birth>06/10/1982</birth>
    <telephone>66666666</telephone>
  </person>
  <person>
    <name>王华</name>
    <birth>12/12/1973</birth>
    <telephone>88888888</telephone>
  </person>
</personlist>
```

# 传统数字媒体的特点

- ❑ 一点到多点；
- ❑ 大众广播式；
- ❑ 效果非智能化；
- ❑ 传播内容非海量；







# 大数据时代数字媒体

- ❑ 社交媒体类
- ❑ 短视频类
- ❑ 大众媒体类
- ❑ 虚拟场景类
- ❑ 语言交互类
- ❑ ...



# 社交媒体类

- ❑ 微信、
- ❑ 微博、
- ❑ facebook、
- ❑ 朋友圈



# 短视频类

- 抖音、
- 快手、
- 哔哩哔哩



# 大众媒体类

- ❑ 电视台客户端
- ❑ 公众号
- ❑ 服务号
- ❑ 官方微信、微博



# 虚拟场景类

- ❑ 虚拟音乐会、
- ❑ 虚拟博物馆



# 语言交互类

- ❑ 微软小冰、
- ❑ 智能客服、
- ❑ 电话机器人、
- ❑ 智能音箱、
- ❑ IBMwatson虚拟助手





# 大数据时代数字媒体特点

- ❑ 多点到多点；
- ❑ 精准传播；
- ❑ 效果智能化；
- ❑ 传播内容海量化；



## 第二部分

### 1. 4数字媒体的软件介绍





# 视觉媒体类软件 (Graphics and Image)

- ❑ **Adobe Illustrator**: a powerful publishing tool from Adobe. Uses vector graphics; graphics can be exported to Web.
- ❑ **Adobe Photoshop**: the standard in a graphics, image processing and manipulation tool.
  - Allows layers of images, graphics, and text that can be separately manipulated for maximum flexibility.
  - **Filterfactory** permits creation of sophisticated lighting-effects filters.
- ❑ **Macromedia Fireworks**: software for making graphics specifically for the web.
- ❑ **Macromedia Freehand**: a text and web graphics editing tool that supports many bitmap formats such as GIF, PNG and JPEG.



# 视觉媒体类软件 (Video)

- ❑ **Adobe Premiere:** an intuitive, simple video editing tool for non-linear editing, i.e., putting video clips into any order:
  - Video and audio are arranged in "tracks".
  - Provides a large number of video and audio tracks, superimpositions and virtual clips.
  - A large library of built-in transitions, filters and motions for clips → effective multimedia productions with little effort.
- ❑ **Adobe After Effects:** a powerful video editing tool that enables users to add and change existing movies. Can add many effects: lighting, shadows, motion blurring; layers.
- ❑ **Final Cut Pro:** a video editing tool by Apple; Macintosh only.

# 视觉媒体类软件

- ❑ AI换脸软件：ZAO，Deepfake，
- ❑ 智能视频剪辑：阿里云AI智能集锦—首席剪辑官“快影”上任CCTV5客户

(

[https://v.youku.com/v\\_show/id\\_XMzgZ0DMxNDMwMA==.html](https://v.youku.com/v_show/id_XMzgZ0DMxNDMwMA==.html)），爱奇艺AI剪辑师

(

<https://www.iqiyi.com/articles/2019-04-18-13>)



# 音频媒体类软件 (Music Sequencing and Notation)



## ❑ **Cakewalk**: now called ProAudio.

- The term sequencer comes from older devices that stored sequences of notes ("events", in MIDI).
- It is also possible to insert WAV files and Windows MCI commands (for animation and video) into music tracks (MCI is a ubiquitous component of the Windows API.)

(MCI: Media Control Interface 媒体控制接口)

## ❑ **Cubase**: another sequencing/editing program, with capabilities similar to those of Cakewalk. It includes some digital audio editing tools.

## ❑ **Macromedia Soundedit**: mature program for creating audio for multimedia projects and the web that integrates well with other Macromedia products such as Flash and Director.



# 音频媒体类软件 (Digital Audio)



- **Digital Audio** tools deal with accessing and editing the actual sampled sounds that make up audio:
  - **CoolEdit**: a very powerful and popular digital audio toolkit; emulates a professional audio studio — — multitrack productions and sound file editing including digital signal processing effects.
  - **Sound Forge**: a sophisticated PC-based program for editing audio WAV files.
  - **Pro Tools**: a high-end integrated audio production and editing environment — — MIDI creation and manipulation; powerful audio mixing, recording, and editing software.





# 音频媒体类软件

- ❑ AI换声软件：<https://www.modulate.ai/>,  
(  
[https://www.sohu.com/a/299042233\\_120044877](https://www.sohu.com/a/299042233_120044877))
- ❑ AI音频：网易网易AI音频的《醒来》？  
(<https://zhuanlan.zhihu.com/p/336741136>)

# 其他媒体类软件 (Multimedia Authoring)



- ❑ **Macromedia Flash**: allows users to create interactive movies by using the score metaphor, i.e., a timeline arranged in parallel event sequences.
- ❑ **Macromedia Director**: uses a movie metaphor to create interactive presentations—very powerful and includes a built-in scripting language, **Lingo**, that allows creation of complex interactive movies.
- ❑ **Authorware**: a mature, well-supported authoring product based on the **Iconic/Flow-control** metaphor.
- ❑ **Quest**: similar to Authorware in many ways, uses a type of flowcharting metaphor. However, the flowchart nodes can encapsulate information in a more abstract way (called **frames**) than simply subroutine levels.

# 其他媒体类软件

- ❑ 聊天机器人平台：

- ❑ Chatopera

- ❑ 网址：<https://bot.chatopera.com>

- ❑ 简介：Chatopera机器人平台包括知识库、多轮对话、意图识别和语音识别等组件，标准化聊天机器人开发，支持企业OA智能问答、HR智能问答、智能客服和网络营销等场景







# Announcement for all

- ❑ 期末考试（60%）+平时成绩（40%）
- ❑ 平时成绩=平时作业+考勤
- ❑ 杜绝抄袭，按时提交
- ❑ 认定为抄袭者，此次作业0分
- ❑ 不按时提交者，此次作业0分
- ❑ 考勤不到者，按照学校规定处理
- ❑ 期末考试为大作业