信息科技 · 第 6 课

互联网应用中的数据

看新闻、查地图、听音乐、看视频——

这些应用背后，是各种各样的数据，以及把它们送到你眼前的一套传输机制。

课时

1 课时（45 分钟）

学习方式

小组体验 · 合作探究

教学环境

网络教室 · 信息科技学习平台

本课关键词：数据构成 · 数据类型 · Web 服务 · 请求与响应 · 编码 · 流媒体

学完这一课，我们应该能够……

学习目标

通过对互联网中数据的体验和探究：

· 了解互联网中常见的数据构成

· 初步理解互联网应用中的数据传输过程

重点与难点

教学重点：互联网中常见的数据构成

教学难点：互联网应用中的数据传输过程

一句话把握本课

前半节解决“数据有什么”（构成），后半节解决“数据怎么来”（传输）。两条线合起来，学生才能从“会用互联网”走向“懂互联网”。

情境导入

你今天用互联网做了什么？



刷新闻



查路况导航



听歌



聊天 · 看视频

追问三连：

- 这些应用传给你的，到底是什么？
- 它们是同一种东西吗？
- 为什么能在一两秒内就出现在手机上？

认知冲突：我们常说“在网上看视频”，可视频并没有“跑到手机里”。互联网传的是数据；你看到的，是被“还原”出来的内容。

四个典型应用，覆盖最常见的数据形态

同一个应用往往同时包含多种类型的数据——这正是互联网数据构成的特点



新闻网站

图文 + 视频

文字标题正文、配图、现场视频、栏目与时间



地图导览

地图 + 位置

地图瓦片图、经纬度坐标、矢量路网、距离时间



在线音乐

音频 + 歌词

歌曲音频、专辑封面、带时间轴的歌词、歌单



在线词云

文本 → 图形

原始文本、词频统计、字号映射、配色配置

课堂追问：这四个应用中，哪一种数据类型是“你听得见但看不见”的？为什么视频文件总是最大的？

任何一个互联网应用，数据都由三层组成

① 内容数据

你直接看到、听到的：文字、图片、音频、视频。数据量最大，也是用户最关心的部分。

② 属性数据

描述内容“是谁、在哪、何时”：标题、作者、发布时间、经纬度、栏目分类、标签。

③ 交互 / 统计数据

由用户行为产生：点击量、点赞、评分、评论、播放进度、搜索记录。

提问引导（可小组讨论 1 分钟）

· 你在购物 App 里看到的“猜你喜欢”，用到了哪一层数据?

· 一条新闻的“阅读量 10 万 +”，属于内容数据还是统计数据?

六类常见数据类型

数据类型	常见格式	典型应用	特点
文本数据	TXT / HTML / JSON	新闻正文、评论、歌词	体积小、可压缩、易检索
图像数据	JPG / PNG / GIF / SVG	新闻配图、地图瓦片	JPG 有损、 PNG 透明、 SVG 矢量
音频数据	MP3 / AAC / WAV / FLAC	音乐、播客、语音	WAV 无损体积大， MP3 有损体积小
视频数据	MP4 （ H.264 / H.265 ）	短视频、直播、网课	体积最大，含音视频两条流
空间位置数据	经纬度、 GeoJSON	导航、外卖、共享单车	记录“在哪里”，可叠加到地图
结构化数据	数据库表、 CSV 、 JSON	商品列表、成绩、订单	字段固定，便于查询统计

小结：网页或 App 里的“内容”，几乎都能归到这六类之中。

特别看一个：在线词云是怎么“变”出来的

① 收集文本
1000 条新闻标题

② 分词
切分成词语

③ 统计词频
数每个词出现几次

④ 映射视觉
字号 / 颜色

输入：一堆文本

AI 新闻标题、评论区文字、弹幕……
经过分词与词频统计后，每个词都有一个“权重数值”。

输出：一张图形

人工智能 大模型 算法 数据 算力

字号越大 = 出现次数越多

数据可视化思想：把“数值”（词频）映射成“视觉属性”（字号、颜色），让人一眼看出重点。

02

互联网应用中的数据传输

数据存在服务器上，你在屏幕上看到它——中间发生了什么？

以浏览器访问一个网页为例



- **先有地址**
  浏览器不认识域名，只认识 IP 地址，所以要靠 DNS 翻译
- **再有连接**
  HTTP 运行在 TCP 之上，先把“通路”建立起来才谈传数据
- **一次网页 = 几十次请求**
  HTML 、图片、样式、脚本各算一次，这也是图片多的网页更慢的原因

课堂操作： F12 → Network → 刷新，点开任意一条请求，看它的状态码、大小和耗时。

两个角色：服务器与浏览器（B/S 结构）

服务器（Server）

存放网页文件与数据库，被动等待请求；收到请求后，把对应的数据发回去。一台服务器可同时服务成千上万用户。

浏览器 / APP （客户端）

替用户发出请求，并把收到的数据解析、还原成文字、图片、声音与画面。

静态网页 与 动态网页

	静态网页	动态网页
内容来源	事先写好的 HTML 文件	服务器程序现场查询数据库生成
典型例子	学校简介、活动通知	新闻列表、搜索结果、购物车
相同请求	每次返回内容相同	可能因人、因时而不同
地址特征	常见 .html 结尾	常带参数，如 ?id=1024

请求与响应的报文长什么样

请求（浏览器 → 服务器）

```
GET /news/list.html?page=2 HTTP/1.1
Host: www.news.cn
User-Agent: Mozilla/5.0 (iPhone)
Accept: text/html, image/*
```

第一行 = 方法 + 资源路径 + 协议版本；下面是头部，说明我是谁、能接收什么格式。

响应（服务器 → 浏览器）

```
HTTP/1.1 200 OK
Content-Type: text/html; charset=UTF-8
Content-Length: 15380
Last-Modified: Fri, 06 Nov 2026
```

第一行 = 协议版本 + 状态码 + 说明。状态码是服务器的“一句话回复”。

200 成功 | **404 找不到** | **500 服务器出错**

同样的内容，体积可以是几十倍之差

核心公式：文件大小（MB） ≈ 码率（Mbps） × 时长（秒） ÷ 8　　**记忆：码率是比特，除以 8 才是字节**

112 MB

标清 1.5 Mbps · 10 分钟

刷一集网剧的量级

375 MB

高清 5 Mbps · 10 分钟

手机流量要留神

1875 MB

4K 25 Mbps · 10 分钟

接近 1.9 GB

同一首歌（4 分钟）	格式	大小	说明
高保真	WAV	约 40 MB	无损，体积最大
在线音乐	MP3 128 kbps	约 3.8 MB	有损，多数人听不出差别

结论：编码选得好，流量能省一个数量级——这就是“压缩”对互联网的意义。

流媒体技术：为什么视频能“秒开”

传统下载

先把整个文件下载到本地，再开始播放。
缺点：等待久、占存储、无法直播。

流媒体

把文件切成小片，边下载边播放，只保留一小段缓冲。
优点：秒开、自适应清晰度、支持直播。

三个关键词

缓冲 Buffer

先存几秒数据兜住网络抖动；转圈就是在补缓冲

分片传输

视频切成小段按需请求，不用一次拿完

自适应码率

网络好上高清，网络差自动降清晰度

联系上一节：流媒体可行，靠的是“编码压缩体积 + 带宽稳定送达”这两件事。

围绕主题搜集素材，并分类整理

小组任务（12 分钟）：① 选定主题（校园垃圾分类 / 我身边的 AI / 二十四节气）；② 搜集至少 6 个素材，覆盖至少 3 种数据类型；③ 记录类型、格式、大小、来源网址；④ 按类型分类整理，并说明分类理由。

序号	素材名称	数据类型	格式	大小	来源
1					
2					
3					
4					
5					
6					

评价要点：类型覆盖是否够 · 信息是否准确 · 来源是否标注（尊重版权） · 分类是否合理 · 分工是否清楚。

课堂小结

一条主线，三句话

一条主线

数据有哪些
（文本 / 图像 / 音频 / 视频 / 位置 / 结构化）

↓ 怎么传
DNS → TCP → HTTP 请求响应

↓ 为什么快慢不同
数据量 × 编码 × 带宽

三句话记住它

① 一个互联网应用里的数据，往往是多种类型混在一起的。

② 你看到的网页，是“请求—响应”换来的，不是存在手机里。

③ 数据能不能秒开，取决于大小、编码和带宽。

课后作业

① 完成小组素材整理表，下节课带来分享（含来源网址）；② 估算： 40 分钟、码率 8 Mbps 的剧，消耗多少流量?
③ 找一个数据构成复杂的 App ，写出它包含的 3 种数据类型。