

【初试】2026 年 上海外国语大学 854 数据科学导论之数据科学导论考研精品资料

说明：本套资料由高分研究生潜心整理编写，高清电子版支持打印，考研推荐资料。

一、考研大纲**1. 上海外国语大学 854 数据科学导论考研大纲**

①2025 年上海外国语大学 854 数据科学导论考研大纲。

说明：考研大纲给出了考试范围及考试内容，是考研出题的重要依据，同时也是分清重难点进行针对性复习的推荐资料，本项为免费提供。

二、2026 年上海外国语大学 854 数据科学导论考研资料**2. 《数据科学导论》考研相关资料****(1) 《数据科学导论》[笔记+提纲]**

①上海外国语大学 854 数据科学导论之《数据科学导论》考研复习笔记。

说明：本书重点复习笔记，条理清晰，重难点突出，提高复习效率，基础强化阶段推荐资料。

②上海外国语大学 854 数据科学导论之《数据科学导论》复习提纲。

说明：该科目复习重难点提纲，提炼出重难点，有的放矢，提高复习针对性。

三、电子版资料全国统一零售价

本套考研资料包含以上部分(不含教材)，全国统一零售价：[¥]

四、2026 年研究生入学考试指定/推荐参考书目(资料不包括教材)

上海外国语大学 854 数据科学导论考研初试参考书

《数据科学导论》，杨旭、汤海京、丁刚毅编著，北京理工大学出版社，2014.

《数据科学概论》，覃雄派，中国人民大学出版社，2018.

五、本套考研资料适用学院

语料库研究院

六、本专业一对一辅导(资料不包含，需另付费)

提供本专业高分学长一对一辅导及答疑服务，需另付费，具体辅导内容计划、课时、辅导方式、收费标准等详情请咨询机构或商家。

七、本专业报录数据分析报告(资料不包含，需另付费)

提供本专业近年报考录取数据及调剂分析报告，需另付费，报录数据包括：

①报录数据-本专业招生计划、院校分数线、录取情况分析 & 详细录取名单；

②调剂去向-报考本专业未被录取的考生调剂去向院校及详细名单。

版权声明

编写组依法对本书享有专有著作权，同时我们尊重知识产权，对本电子书部分内容参考和引用的市面上已出版或发行图书及来自互联网等资料的文字、图片、表格数据等资料，均要求注明作者和来源。但由于各种原因，如资料引用时未能联系上作者或者无法确认内容来源等，因而有部分未注明作者或来源，在此对原作者或权利人表示感谢。若使用过程中对本书有任何异议请直接联系我们，我们会在第一时间与您沟通处理。

因编撰此电子书属于首次，加之作者水平和时间所限，书中错漏之处在所难免，恳切希望广大考生读者批评指正。

目录

封面.....	1
目录.....	3
上海外国语大学 854 数据科学导论考研大纲	4
2025 年上海外国语大学 854 数据科学导论考研大纲.....	4
2026 年上海外国语大学 854 数据科学导论考研核心笔记	5
《数据科学导论》考研核心笔记.....	5
第 1 章 引论	5
考研提纲及考试要求	5
考研核心笔记.....	5
第 2 章 数据科学在商业金融领域的应用	14
考研提纲及考试要求	14
考研核心笔记.....	14
第 3 章 数据科学在生物医学领域的应用	23
考研提纲及考试要求	23
考研核心笔记.....	23
第 4 章 数据科学在智慧城市领域的应用	33
考研提纲及考试要求	33
考研核心笔记.....	33
第 5 章 数据科学在影视娱乐领域的应用	42
考研提纲及考试要求	42
考研核心笔记.....	42
第 6 章 数据科学在其它领域的应用实例	48
考研提纲及考试要求	48
考研核心笔记.....	48
第 7 章 数据科学与工程概论	52
考研提纲及考试要求	52
考研核心笔记.....	52
第 8 章 数据科学的未来展望	63
考研提纲及考试要求	63
考研核心笔记.....	63
2026 年上海外国语大学 854 数据科学导论考研复习提纲	65
《数据科学导论》考研复习提纲	65

上海外国语大学 854 数据科学导论考研大纲

2025 年上海外国语大学 854 数据科学导论考研大纲

关注如何在大数据时代背景下，运用各门与数据相关的技术和理论，更好地利用数据，服务于社会。系统介绍与数据科学相关的各个方面的知识，包括数据概念、数据发展史、数据、信息与知识，数据处理基础设施/平台和工具、大数据的特征。

2026 年上海外国语大学 854 数据科学导论考研核心笔记

《数据科学导论》考研核心笔记

第 1 章 引论

考研提纲及考试要求

考点：数据的概念

考点：数据的发展史

考点：大数据时代的来临

考点：大数据对社会所产生的影响

考点：数据科学

考研核心笔记

【核心笔记】序言

“数据科学”从出现到现在已经有三十多年的历史了，其中涉及了很多方面的内容，涵盖数学、统计学、数据工程、模式识别、机器学习、高性能计算、可视化、数据仓库以及数据建模等多个领域的技术和理论。数据科学的最终目的就是从数据中挖掘出有用的信息，让数据增值。

虽然已有三十年历史了，但数据科学仍然是一门新兴的学科，尽管之前运用较多的是在计算智能或者是商业分析方面，但已经慢慢地深入到了人类社会的各个方面。之所以要开设数据科学这门学科，就是为了培养这门学科的专业人员，使他们运用所有可以得到的数据，寻找其背后的故事，从而找到办法让这些数据所蕴含的意义可以轻易地被人们所理解，即便他们不具备数据科学的相关知识。

【核心笔记】数据

1. 数据的概念

数据科学这门学科研究的核心内容就是数据，那究竟什么是数据呢？一提到数据，我们首先想到的会是数字。但数据并不局限于数字，文本、音频、图像、视频都可以是数据。在本书里，我们对数据给出如下的定义：

数据是指以定性或者定量的方式来描述事物的符号记录，是可定义为意义的实体，它涉及事物的存在形式。

简单说来，数据就是人为创造的一种对事物的表示方式，是通过观察或者实验得来的对现实世界中的地方、事件、对象或概念的描述和反映。

数据可以是连续的值，例如声音，称为模拟数据；也可以是不连续（离散）的值，例如成绩，称为数字数据。

2. 数据的发展史

人类历史上最早的有记录的数据，可以追溯到穴居的原始人时期。当时的人类，会在作为居处的洞穴墙壁上，以石器或者骨器刻画来记录数据。这些被记录的数据，或者是简单的记录日期的刻痕，或者是形象化地记载一些日常发生事件的壁画。



图 1-1 结绳记事

图 1-1 所示的是古代印加人采用的一种结绳记事的方法，用来计数或者记录历史。事大，大结其绳；事小，小结其绳。不过，这种记事的方法已经失传，目前还没有人能够了解其全部含义。

随着数字和文字的出现，古人开始以更加明确的形式来记录数据。古埃及人创造了莎草纸，用来进行记录。埃及博物馆中陈列的各种莎草纸文书、图画表明，莎草纸是人类历史上最早、最便利的书写材料之一，是记录古埃及历史的主要载体。

我们的祖先在汉代发明了造纸术（图 1-3），这种数据记录的方式一直延续到现在。这里一定要强调，本书中所提到的数据，不光是涉及数字形式的数据，以文本、图像、语音等方式所记录的数据都是数据科学所研究的内容。造纸术的发明和改进，让文本形式的数据记录方式开始盛行起来。



图 1-3 造纸流程

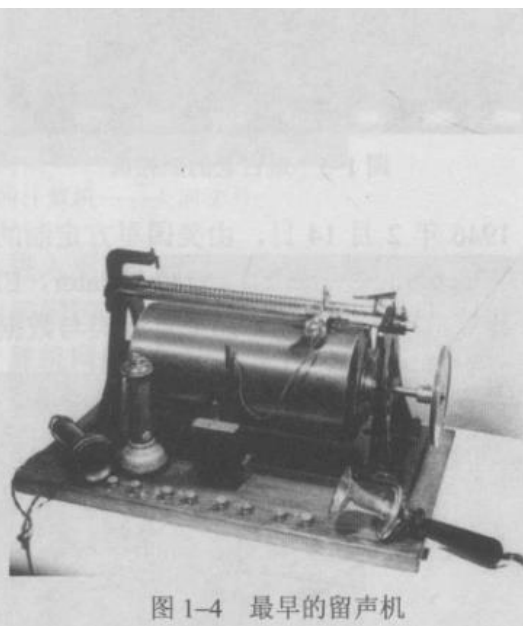


图 1-4 最早的留声机

最早的录音机，也叫留声机，诞生于 1877 年，是发明大王——爱迪生所制造的。爱迪生发现了电话传声器里的模板随着说话声而震动的现象，于是他拿短针做了试验，从中得到了很大的启发。说话的快慢高低能使短针产生相应的颤动。那么，反过来，这种颤动也一定能发出原先的说话声音。于是，他开始研究声音重发的问题。

在 1877 年 8 月 15 日，爱迪生让助手克瑞西按图样制造出一台由大圆筒、曲柄、受话机和模板组成的怪机器。“会说话的机器”诞生的消息，轰动了全世界。1877 年 12 月，爱迪生公开表演了留声机（图 1-4），外界舆论马上把他誉为“科学界之拿破仑”，留声机是 19 世纪最让人振奋的三大发明之一。即将开幕的巴黎世界博览会立即把它作为时新展品展出。