# 第一届上海市高校开放大数据分析挑战赛暨第一届上海理工大学开放大数据分析大赛

**附录1：题目指南**

**1.全球COVID-19舆情演化分析**

根据新冠肺炎疫情的数值数据、新闻文本数据，挖掘COVID-19疫情爆发前期、中期、后期的舆情走向和公众关注度，寻找相关舆情的重要时间节点，以及随着疫情的发展，舆情的演化趋势，并给出相关分析结果和结论。相关方法：中文文本分词技术、数据标准化、数据挖掘、统计分析、时间序列分析、机器学习和深度学习的方法。

**2.突发公共卫生事件的应对分析**

根据疫情的新闻报道数据，挖掘疫情中政府、媒体、民众的应对策略，以分析在突发公共卫生事件的发展过程中，政府、媒体、民众的应对方法和策略，给出谣言的甄别手段。相关方法：中文文本分词技术、数据标准化、数据挖掘、统计分析、时间序列分析、机器学习和深度学习的方法。

**3.七种可感染人类冠状病毒相关研究的主题分布与演化分析**

利用目前公开的 COVID-19学术文章，寻找可感染人类冠状病毒相关研究的主题分布与演化规律。相关方法：中文文本分词技术、数据标准化、数据挖掘、统计分析、时间序列分析、机器学习和深度学习的方法。

**4. 疫情下商业模式的创新与转型**

以消费购物大数据，反推产品设计、市场营销、渠道运行与建设，生活方式的转变。在疫情大数据的分析基础上，对疫情相关的互联网产品进行创新性设计。

**5.其他参赛团队自行发掘、探索的研究课题**

**附录2**

题 目

团队信息、指导老师信息

1. 作品背景（阐述作品背景、数据分析目的、分析思路等）
2. 数据介绍和分析方法介绍（阐述数据采集与预处理、清洗整理过程）
3. 数据分析环境（介绍数据分析过程所使用的硬件环境，操作系统环境，数据分析软件，编程语言及其版本等等）
4. 数据分析（包括建模过程或分析过程，根据之前分析思路中确定的每项分析内容，鼓励利用传统统计分析、现代机器学习和深度学习等各种数据分析方法，一步步地展开分析。通过图表及文字相结合的方式，形成报告正文。）
5. 数据分析仪表盘的设计（设计一张展示数据分析主要结论的仪表盘）
6. 分析结论与对策建议，或者设计创新性互联网产品（在数据分析结果基础上，以综述性文字给出分析结论；或者在数据分析结果基础上，详细阐述团队设计的疫情相关的互联网产品。）
7. 附录（介绍重要计算方法，专有名词等其他附加内容；若涉及编程，请提供数据分析源代码）

格式要求：

1. 作品原创，严格杜绝抄袭。

2. 提交内容需要涵盖，但不限于上述列表中的内容。

3. 作品中题目小二号宋体居中，团队信息使用仿宋体小四号居中。正文中标题使用四号宋体加粗，正文中段落使用宋体，英文使用Time New Roman，小四号字，20磅间距。

3.提交作品需要同时提交word文档和pdf文档。