

# 中外数据新闻的差异与差距

## ——以全球新冠疫情数据化内容为例



新浪新闻新媒体实验室  
数可视教育公益基金

联合出品

## 研究背景

以新冠疫情为例分析中外数据新闻的差异

# 2020年，新冠疫情席卷全球

截至8月12日，全球220多个国家和地区累计报告：

确诊病例

超过2000万

死亡病例

73.9万

面对这场愈演愈烈的全球公共卫生事件，各国媒体纷纷以新闻报道的形式为其写下注脚。

# 研究背景

以新冠疫情为例分析中外数据新闻的差异

数据新闻报道疫情的优势

高效整合多维度信息

直观展现疫情动态

生动解释信息

研究取样数量

中外机构媒体样本量

16

数据新闻作品样本量

45

挖掘中外媒体关于疫情报道的不同表达方式

# 目录

## CONTENTS

### 一、中外数据新闻作品差异明显

- PC端、移动端，中外偏好各不同
- 可视化设计：国内好静，国外好动
- 国外数据引用更严格、数据源更开放
- 国外媒体对数据的挖掘优于国内

### 二、中外差异从何而来？

- 国内数据新闻起步晚于西方
- 新闻生产模式和市场影响内容呈现
- 不同终端限制了数据新闻作品的呈现
- 国内外的数据开放环境不同

### 三、如何提升数据新闻水平？

- 提升可视化设计水平，合理利用可视化
- 丰富数据来源、规范数据引用、促成数据开放
- 提升对数据的应用和挖掘能力

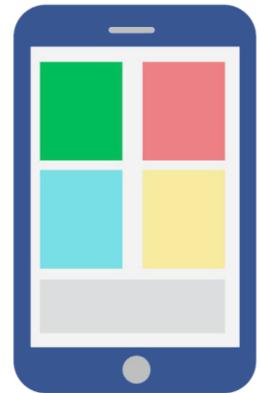
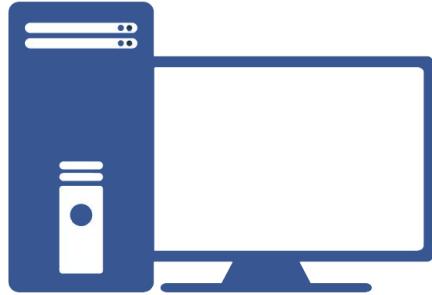


## 一、中外数据新闻作品差异明显

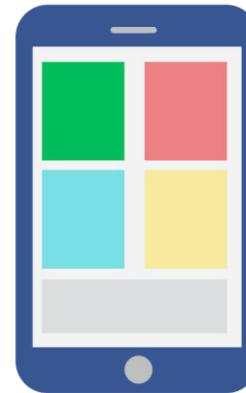
## PART1 中外数据新闻作品差异明显

# PC端、移动端，中外偏好各不同

## 国外作品



## 国内作品



研究发现在呈现终端的选择上，国外的作品偏向PC端和移动端并重，在开发时一般会实现PC端和移动端的兼容。

中国的数据新闻作品则主要基于移动端，部分作品会出现移动端与PC端不兼容的情况。

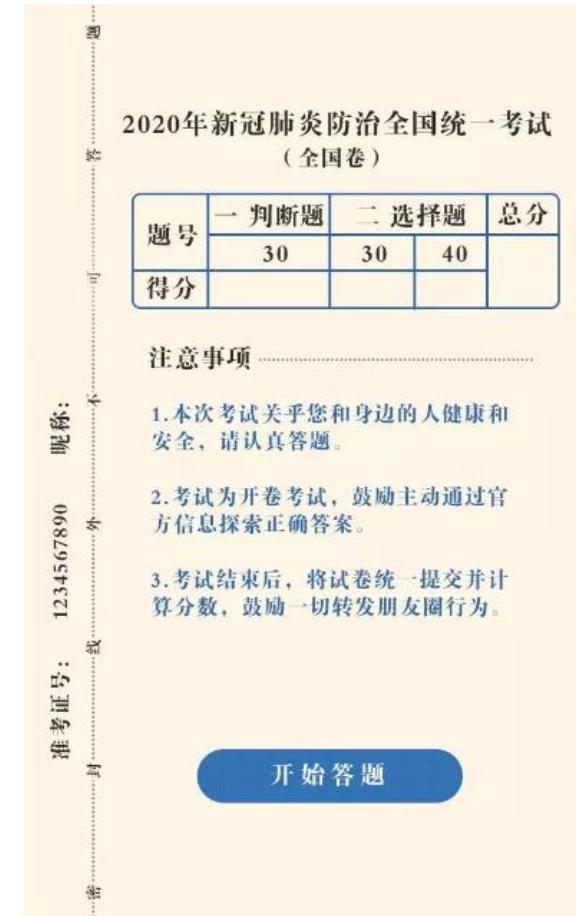
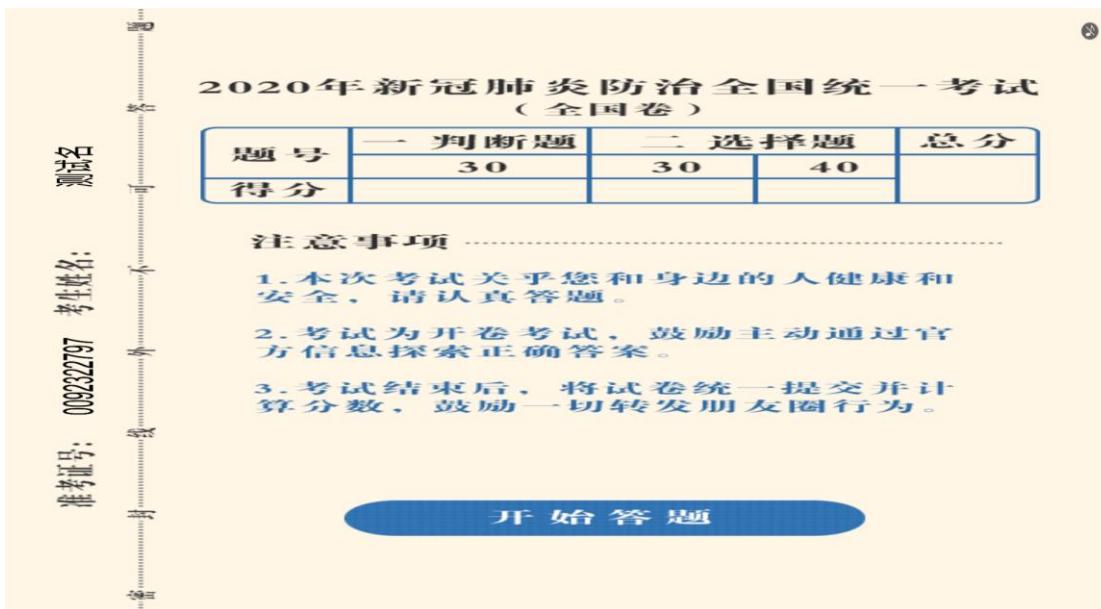
## PART1 中外数据新闻作品差异明显

# PC端、移动端，中外偏好各不同

移动端：正常显示

## 国内作品案例

PC端：比例失调，无法正常阅读



数据新闻作品《2020新冠肺炎防治全国统一考试（全国卷）》，开发时只考虑了移动端界面。

## PART1 中外数据新闻作品差异明显

# 可视化设计：国内好静，国外好动

中国人民大学新闻学院副教授方洁对数据可视化的难易程度进行了划分，从易到难共分为五个层次，分别是：

全文字的静态设计

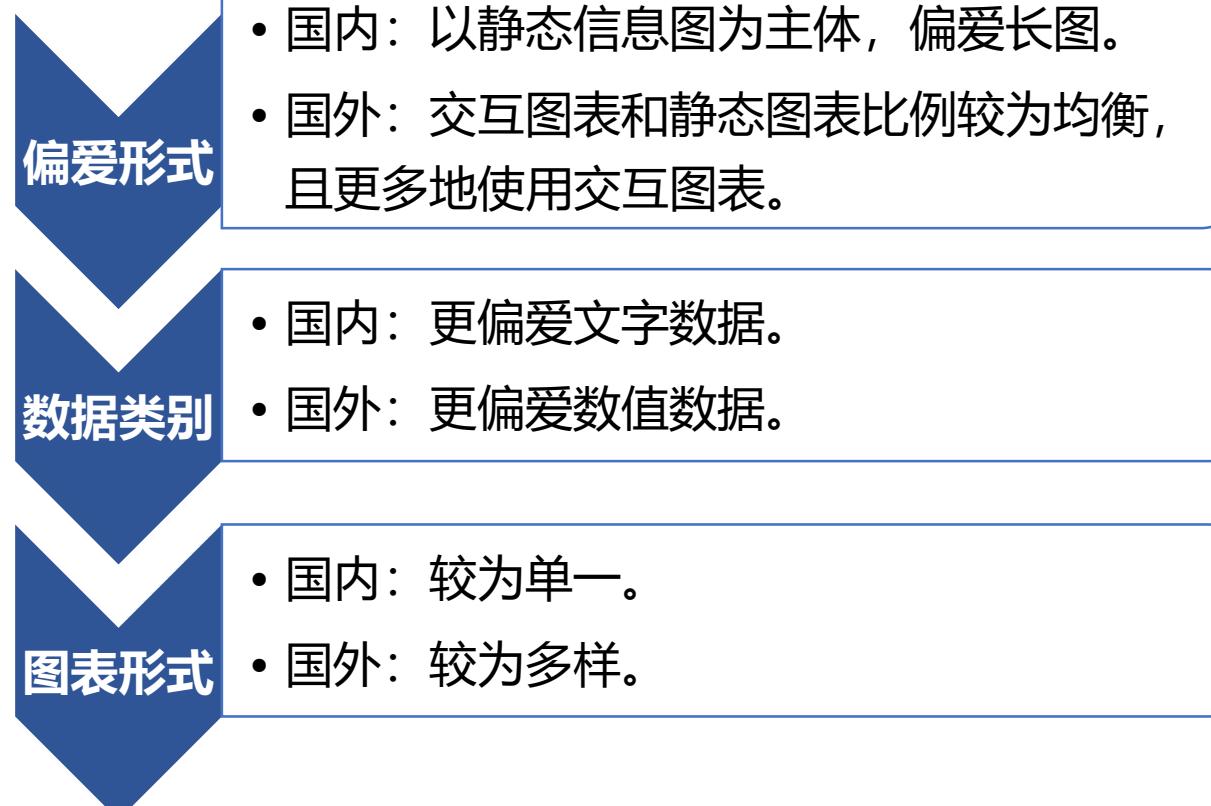
表格或类似表格的图形化处理

结合数据的静态信息图或静态时间轴

具有交互性的时间轴和普通视频

交互地图和3D动画设计以及游戏设计

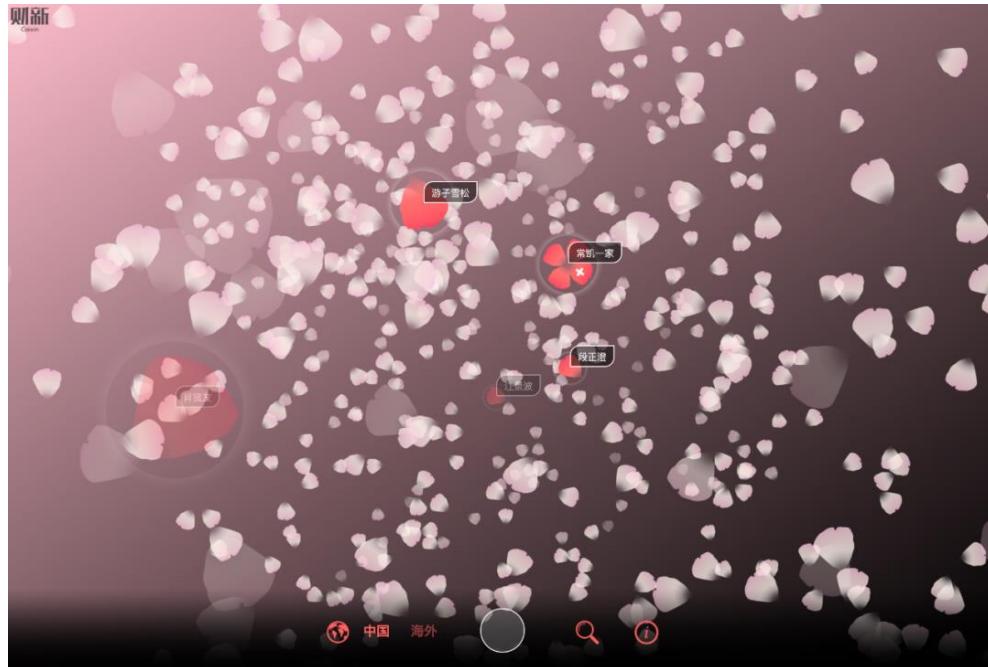
中外数据新闻作品形式与偏好对比：



## PART1 中外数据新闻作品差异明显

# 可视化设计：国内好静，国外好动

## 国内作品案例·对文字数据的应用



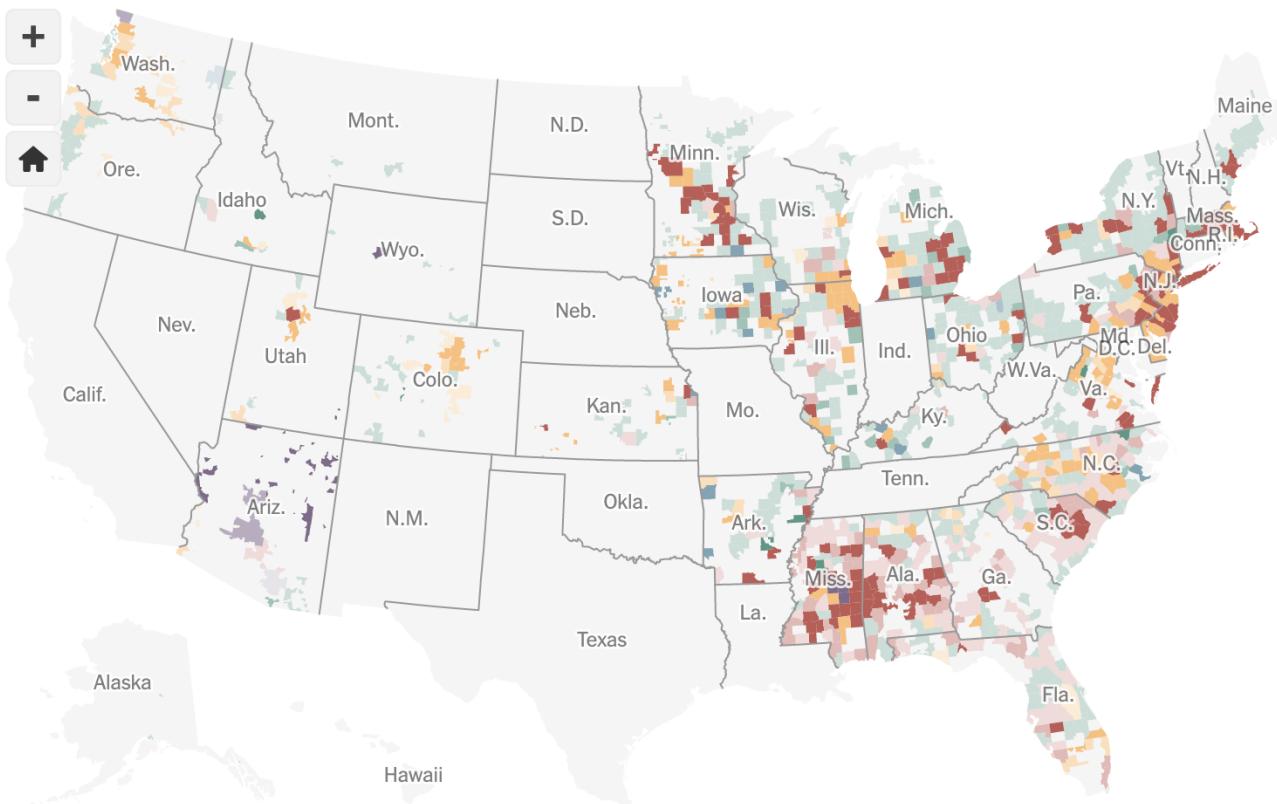
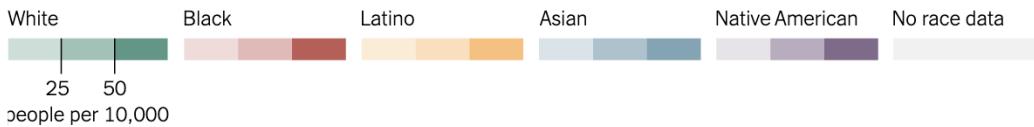
财新《可视化|新冠逝者：献给疫情中离去的生命》，利用交互技术建造了一座献给逝者的纪念馆。未被收录的逝者、已收录的逝者、医护人员、来自同一家庭的逝者，化身为不同的花瓣，点击花瓣上的名字可查看逝者详情。

作品链接：<http://datanews.caixin.com/interactive/2020/THREEJS/blossom/>

## PART1 中外数据新闻作品差异明显

# 可视化设计：国内好静，国外好动

Race or ethnicity with the highest coronavirus rate in each county



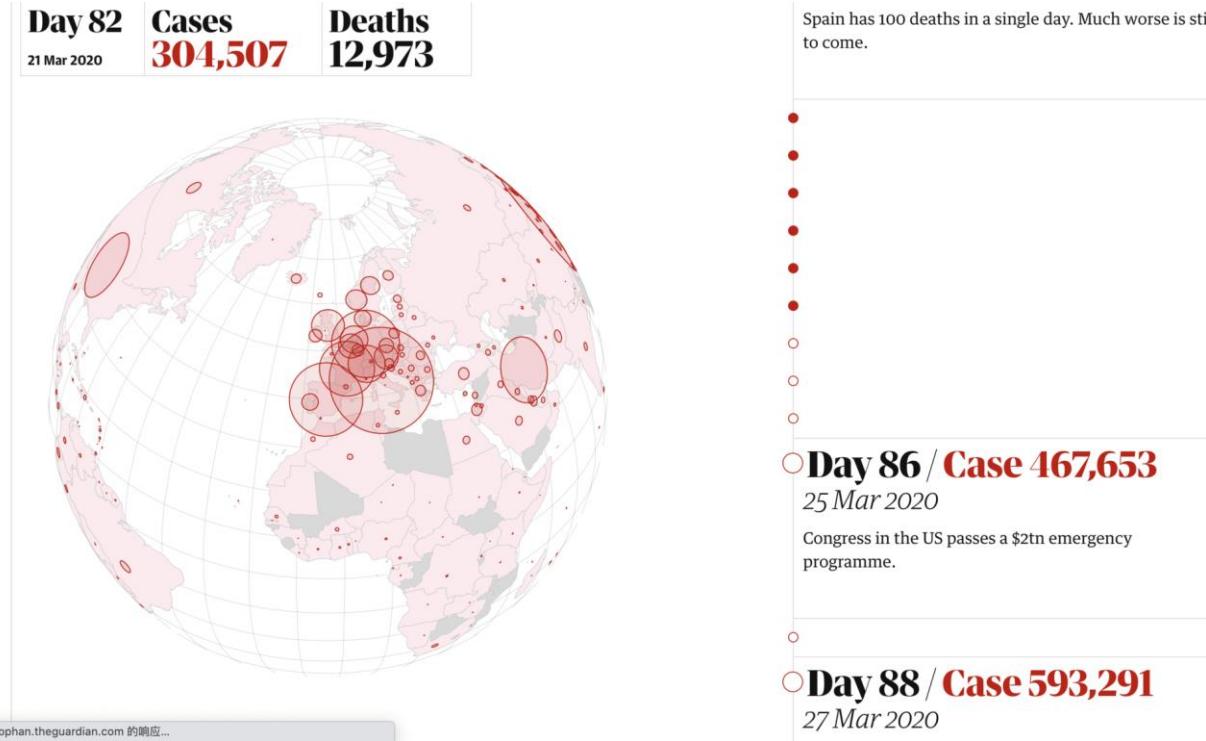
## 国外作品案例·对数值数据的应用

纽约时报的《The Fullest Look Yet at the Racial Inequity of Coronavirus》，以美国疾病控制与预防中心五月公布的最新数据，对拉丁美洲和非洲裔美国居民、白人美国居民的感染率和病死率等数据进行了可视化。

通过该作品可以发现，拉丁美洲和非洲裔美国居民的感染几率是白人邻居的三倍，且拉丁美洲和非洲裔的居民死于该病毒的可能性是白人的两倍。

## PART1 中外数据新闻作品差异明显

### 可视化设计：国内好静，国外好动



### 国外作品案例·对图表的创新应用

卫报作品《How coronavirus spread across the globe - visualised》，对新冠肺炎100天内的全球传播情况进行了可视化再现。

该作品在对疫情传播进行追踪时，采用了球体地图而不是常规的平面地图。

## PART1 中外数据新闻作品差异明显

# 国外数据引用更严格 数据源更开放

## 数据来源

国外数据新闻作品	重视对数据源和参考资料的引用，一般会详细说明数据来源，或是附上数据来源链接以便读者查找。
国内数据新闻作品	较少列出参考资料，也很少对数据进行详细说明，一些作品甚至存在未标注数据来源的情况。

## 数据开放

国外媒体	许多媒体或机构都建立了疫情数据库，并将其开放给公众。如《大西洋月刊》推出的新冠数据追踪计划 (The COVID Tracking Project)，专门从各处收集数据，整理后供公众自行下载。
国内媒体	只发现澎湃美数课在疫情期间对数据进行了开放共享。

# PART1 中外数据新闻作品差异明显 国外数据引用更严格 数据源更开放

## 数据来源·国外作品案例

卫报数据新闻作品《How coronavirus spreads through a population and how we can beat it》结尾处对文中数据模型进行了详细注释。

该作品利用交互技术、模拟动画技术建立了病毒传播模型，内容包括新冠肺炎传播模型；新冠肺炎在有干预条件、无干预条件下的传播模型；麻疹、埃博拉病毒传播模型。

### Notes

The isolation value is not used for most of the case studies as isolation data is not usually available.

1. The Diamond Princess case study is intended to visualise the proportional outcome of Covid-19 infections onboard the Diamond Princess. The susceptible proportion of the population is 18% based on an attack rate of 18% (696 cases / 3711 total passengers) and is not indicative of the true susceptibility figure (which might be closer to 100%). Isolation is not used as there is no data available. The death rate is set at 1.8% based on the current total of deaths (13) over cases (696). The RO used is from [Zhang et al. 2020](#).
2. The no intervention case study is an example that uses figures from various sources to show a hypothetical scenario with no isolation measures. The susceptible proportion of the population is set to 70% based on [comments here](#) and is not indicative of the true susceptibility figure (which might be closer to 100%). The death rate is set at 1% based on various estimates of the case fatality ratio from cases in China ([Verity et al. 2020](#), [Wighton et al. 2020](#)). The RO used is 2.7.
3. The strong intervention case study is an example that uses figures from various sources to show a hypothetical scenario with social distancing and isolation measures covering 40% of the population. The susceptible proportion of the population is set to 70% based on [comments here](#) and is not indicative of the true susceptibility figure (which might be closer to 100%). The death rate is set at 0.66%, lower than the no intervention scenario as it assumes health care would cope better with reduced case load. The RO used is 2.7.
4. The measles case study uses values from [Information is Beautiful](#) and the [CDC](#). Actual RO for measles can [vary widely](#) depending on the country and if studies were done before or after the introduction of the measles vaccine. Case fatality ratios also very widely depending on the country and [outbreak situation](#).
5. The ebola case study uses a RO from [Taylor et al 2016](#), a case fatality rate [from the WHO](#) and sets susceptibility to 30% based on an [attack rate of 30%](#) (actual susceptibility might differ).
6. The Contagion case study uses values mentioned in the movie: RO of 4, fatality rate of 25% to 30%, and "will infect 1 in 12 people on the planet" (8.333%).

# PART1 中外数据新闻作品差异明显 国外数据引用更严格 数据源更开放

## 数据来源·国内作品案例

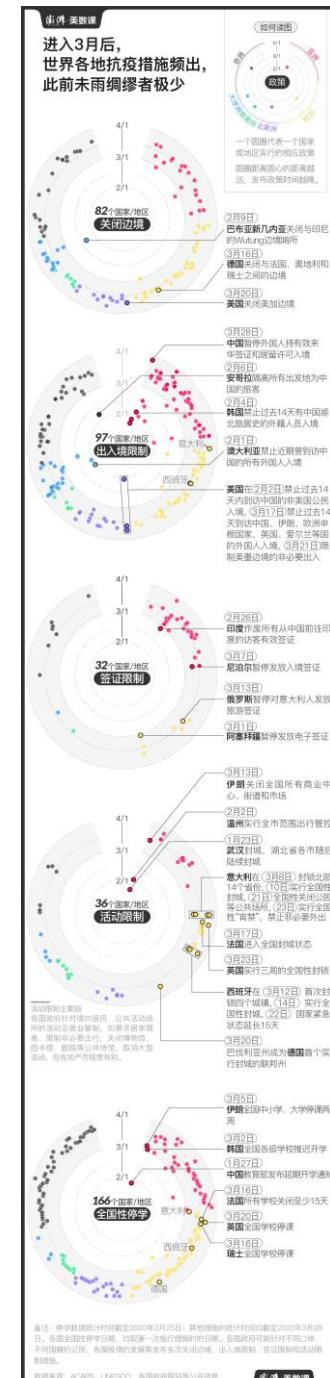
国内的媒体如澎湃新闻在数据引用和注释方面也很规范，例如在作品《多国疫情进入剧烈增长期，全球“封锁”步步加深》中，数据注释写得非常清晰。

但国内数据注释出现的频率总体来说并不高，例如数据新闻作品《72个国家10614例！新冠肺炎疫情全球走势》就未标明数据来源。

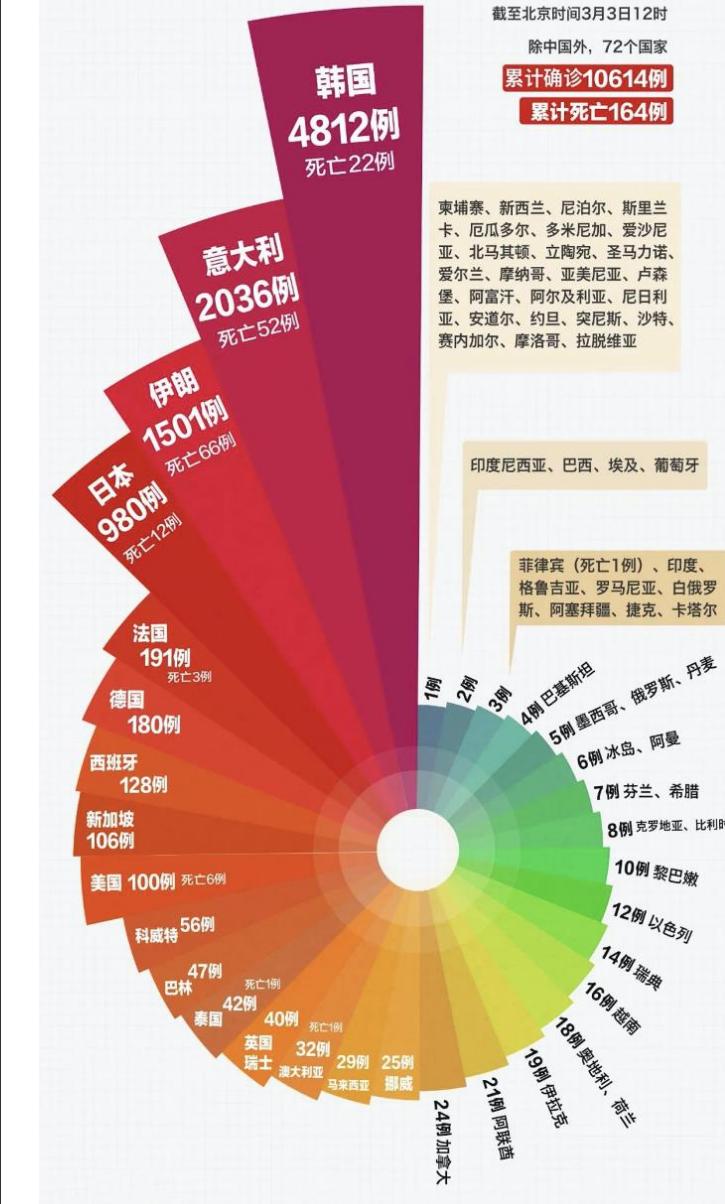
作品链接：

<https://mp.weixin.qq.com/s/03rveYqBjCikxFX3APMaDQ>

<https://mp.weixin.qq.com/s/LmesmSryK8qJPFce8L-NLw>

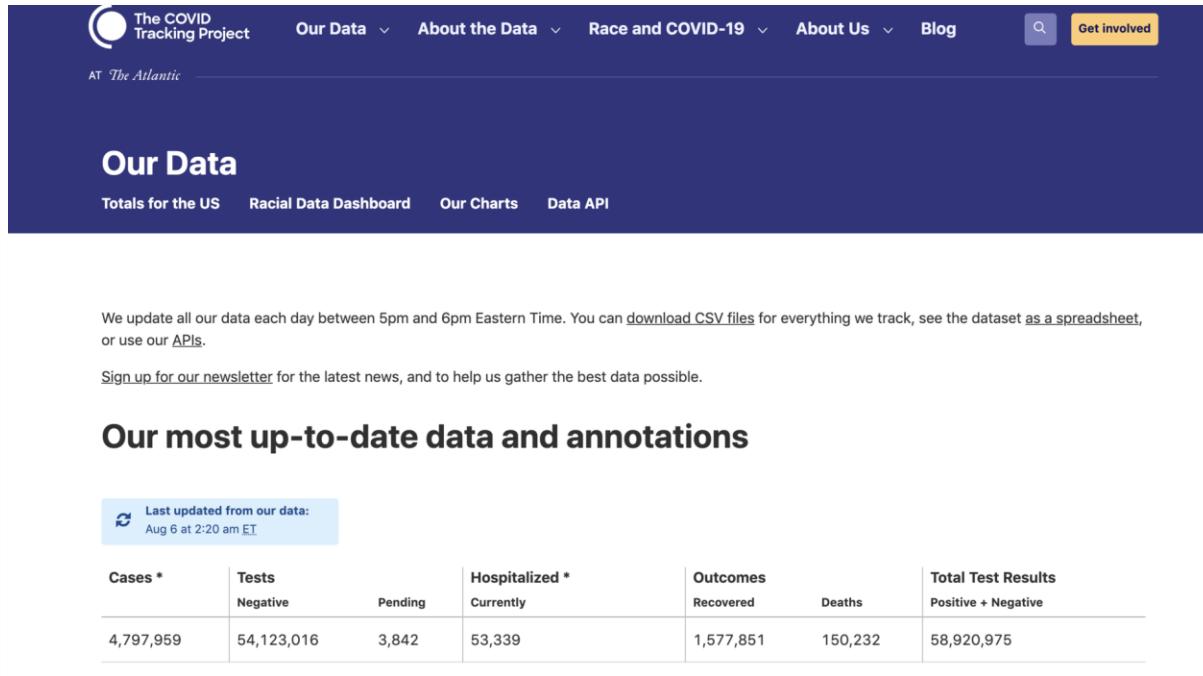


## 新冠肺炎全球疫情形势



# PART1 中外数据新闻作品差异明显 国外数据引用更严格 数据源更开放

## 数据开放·国内外作品案例



The COVID Tracking Project website interface. At the top, there are navigation links: Our Data, About the Data, Race and COVID-19, About Us, and Blog. Below this is a search bar and a 'Get involved' button. The main content area is titled 'Our Data' and includes links to 'Totals for the US', 'Racial Data Dashboard', 'Our Charts', and 'Data API'. A note states: 'We update all our data each day between 5pm and 6pm Eastern Time. You can [download CSV files](#) for everything we track, see the dataset [as a spreadsheet](#), or use our [APIs](#).'. A call to action says: 'Sign up for our newsletter for the latest news, and to help us gather the best data possible.' Below this is a section titled 'Our most up-to-date data and annotations' with a table showing the last update was on Aug 6 at 2:20 am ET. The table details COVID-19 data across various categories like Cases, Tests, Hospitalized, Outcomes, and Total Test Results.

《大西洋月刊》的The COVID Tracking Project, 专门从各地收集、整合、更新数据, 包括每个州获得的阳性和阴性测试结果的数量, 病例的住院和结局数据以及种族和族裔人口统计信息案件和死亡细目。

作品链接: <https://covidtracking.com/data>



A screenshot of a Google Sheets document titled '澎湃新闻新冠疫情数据共享(2020.1.11-2020.7.31)'. The document includes a header with the title, a note about data collection and updates, and a table of COVID-19 data. The table has columns for '报道时间' (Report Date), '省份' (Province), '城市' (City), '新增确诊' (New Confirmed), '新增出院' (New Discharged), '新增死亡' (New Deaths), '消息来源' (Source), '来源链接1' (Link 1), '来源链接2' (Link 2), '来源链接3' (Link 3), and '备注' (Notes). The data shows daily updates from July 31, 2020, across various Chinese provinces and cities, with links to the original sources.

澎湃新闻新冠疫情数据共享, 包括国内外各地的新增确诊、新增出院、新增死亡等指标, 并注明了消息来源和来源链接。

作品链接: <https://shimo.im/sheets/tyWrrrqppYVwQtCW/oURp4>

# PART1 中外数据新闻作品差异明显

## 国外媒体对数据的挖掘优于国内

### 国内媒体

国内媒体中，澎湃新闻和财新两家机构媒体更倾向于对数据进行深度分析、挖掘内在逻辑；网易、新浪等网络媒体在数据信息图和交互数据内容作品等方面表现突出。但在其他部分媒体的数据新闻作品中，仅止步于对数据的可视化呈现。



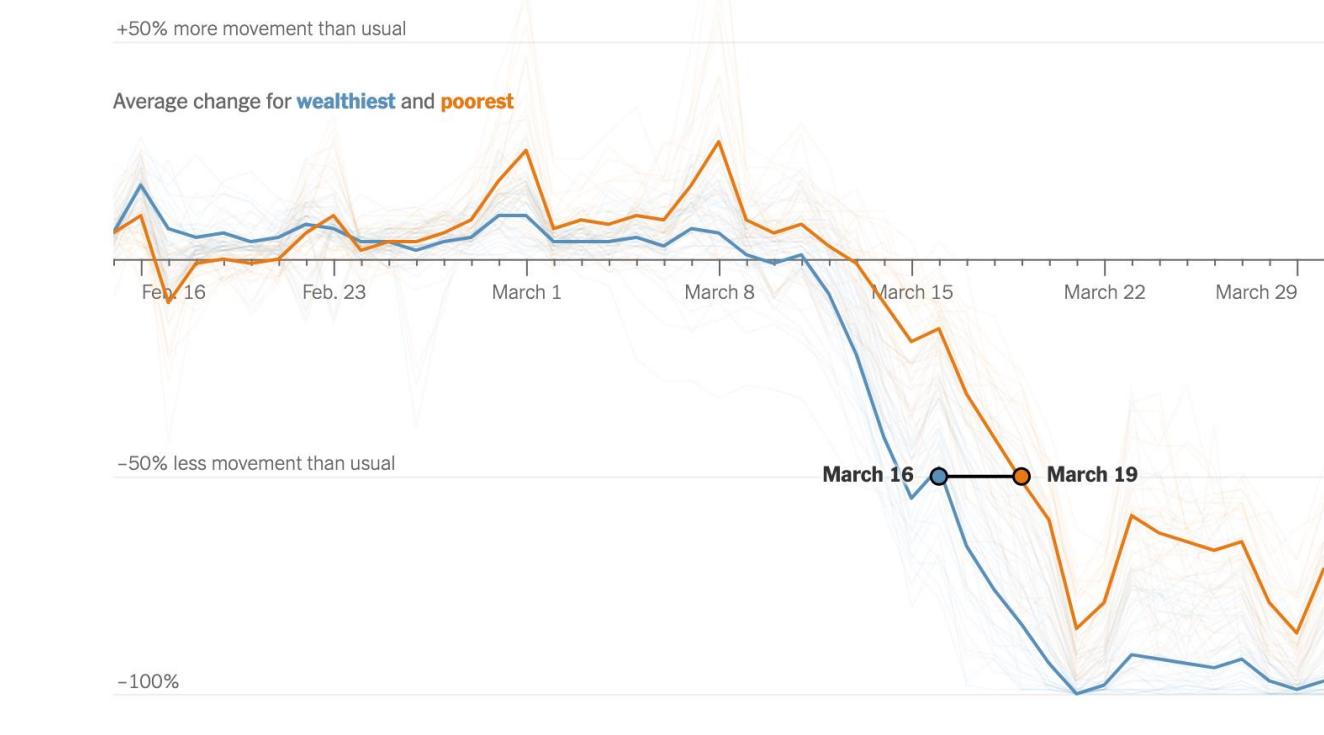
# PART1 中外数据新闻作品差异明显 国外媒体对数据的挖掘优于国内

## 国外机构媒体

### 注重对数据进行多维度展示和分析

如纽约时报的《Location Data Says It All: Staying at Home During Coronavirus Is a Luxury》就基于 Cuebiq 的定位数据，分析了美国不同地区低收入人群和高收入人群的出行数据，揭示了美国高收入人群更早、更多地开始出行限制，低收入人群仍要为了生计出门工作这一社会现象。

But by that date, those in the **highest-income locations** had already cut their movement by nearly half. **Poorer areas** did not see a similar drop until three days later.



作品链接: <https://www.nytimes.com/interactive/2020/04/03/us/coronavirus-stay-home-rich-poor.html>

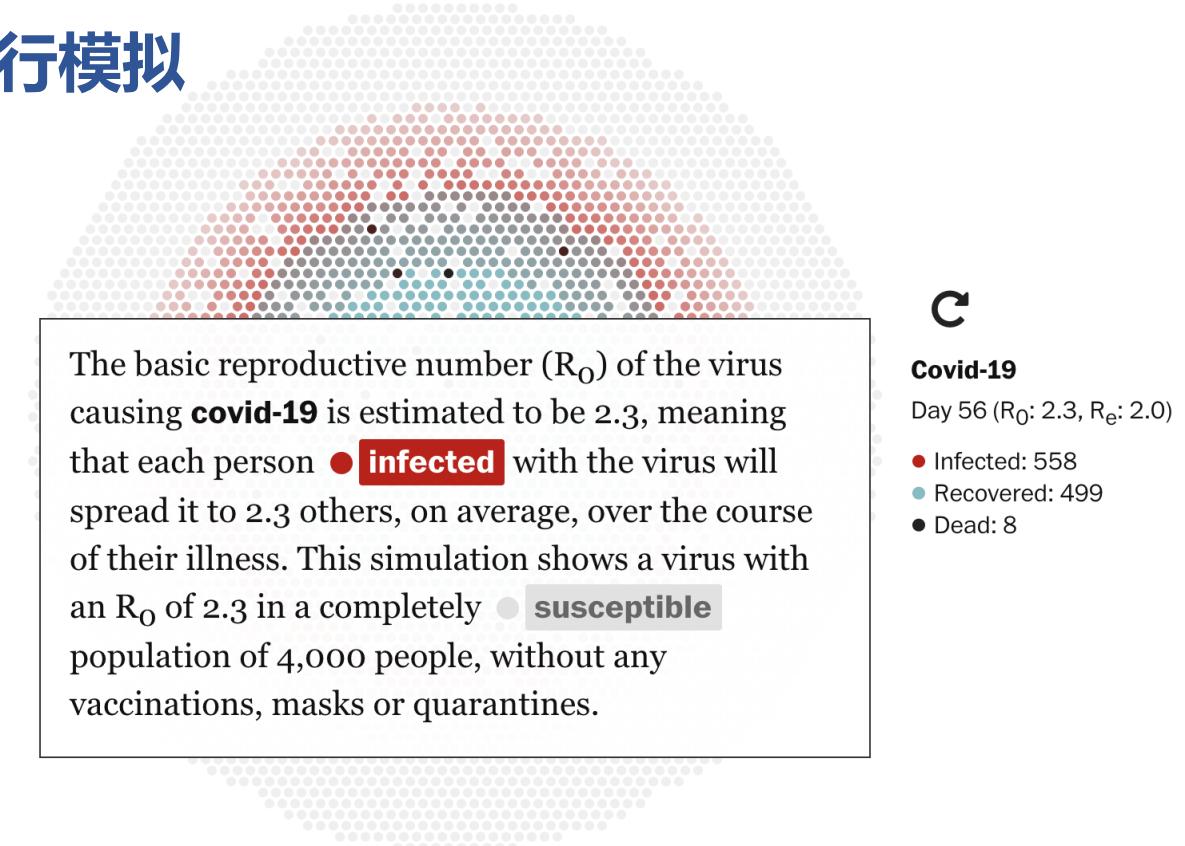
# PART1 中外数据新闻作品差异明显 国外媒体对数据的挖掘优于国内

## 国外机构媒体

通过数学建模**对不同条件下的病毒传播模式进行模拟**

华盛顿邮报《How epidemics like covid-19 end (and how to end them faster)》，利用交互技术、模拟动画技术建立了新冠病毒传播模型。

内容包括根据已有数据建立的新冠肺炎传播模型；对麻疹、埃博拉病毒的流行进行了模型分析；在加入了干预条件，如使用口罩手套等防护品、建立隔离区进行模型分析。



作品链接：[https://www.washingtonpost.com/graphics/2020/health/coronavirus-how-epidemics-spread-and-end/?tid=a\\_inl\\_manual&tidloc=10](https://www.washingtonpost.com/graphics/2020/health/coronavirus-how-epidemics-spread-and-end/?tid=a_inl_manual&tidloc=10)



## 二、中外差异从何而来？

## PART2 中外差异从何而来?

### 国内数据新闻起步晚于西方

**2009年开始**, 《卫报》和《纽约时报》等国外新闻媒体开始对数据新闻进行探索。

**2008年**, “数据新闻”这一概念被Simon Rogers在英国《卫报》上提出。

**2013年**, 随着《纽约时报》的《雪崩》获得普利策奖, 数据新闻引起行业关注, 欧美各大媒体纷纷组建团队, 发力数据新闻。

## PART2 中外差异从何而来?

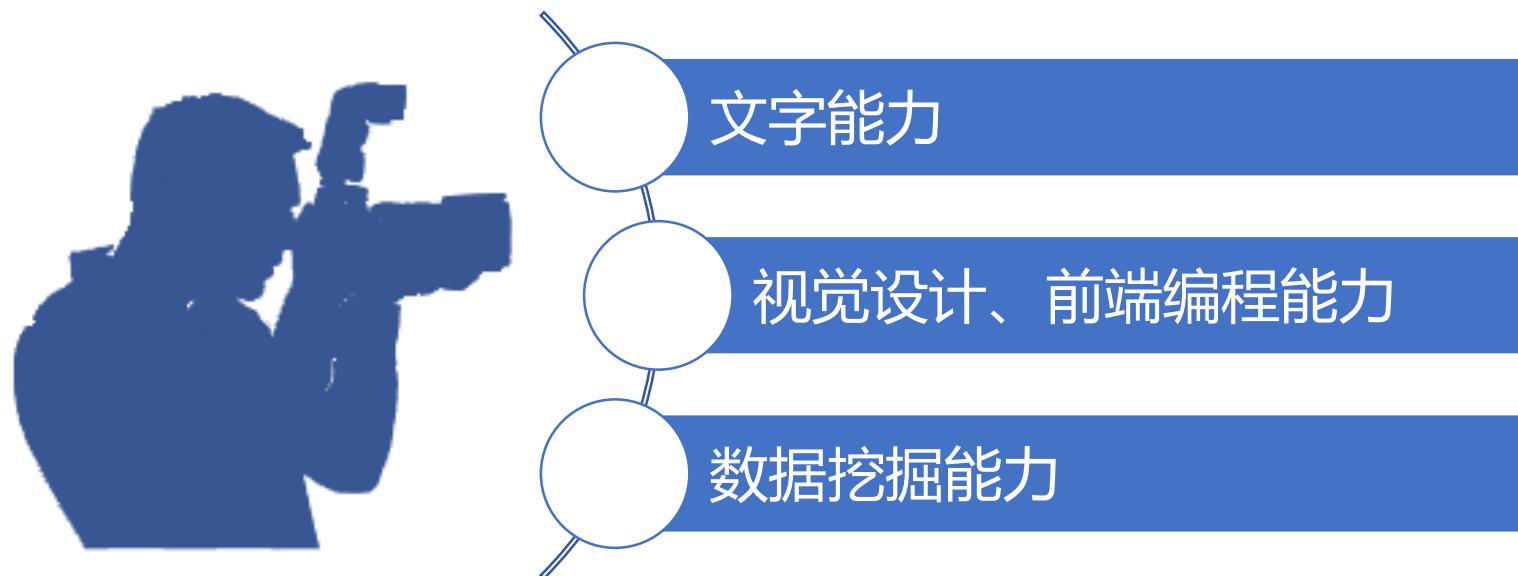
# 国内数据新闻起步晚于西方

早发展数据  
新闻的优势

更丰富的经验积累

更成熟的人才培养机制

截至目前，西方的不少主流媒体都设立了数据新闻部门，团队里对于数据记者的要求往往更加全面：



这使得他们在数据新闻报道的形式  
创新和深度挖掘上游刃有余。

## PART2 中外差异从何而来？

# 国内数据新闻起步晚于西方

## 国内机构媒体



2013年起，财新、澎湃新闻、网易、新浪等中国媒体开始发力数据新闻，和国际媒体相比，起步晚了约四年。

相对而言，中国数据新闻机构的团队建制和规模与西方主流媒体仍有较大差距，数据新闻人才也较为匮乏，这限制了国内数据新闻作品的视觉呈现效果和数据挖掘深度。

## 国内高校



中国高校也于同期开始数据新闻教学，起步较早的有中国传媒大学、中国人民大学、复旦大学等。

## PART2 中外差异从何而来？

# 新闻生产模式和市场影响内容呈现

## 国内外数据新闻实践主体差异

### 国外主体

国外的数据新闻实践主要来自于传统的机构媒体，如英国的《卫报》和美国的《纽约时报》。



### 国内主体

国内的数据新闻生产来源更加广泛，商业性也更加浓厚。

#### 机构媒体



#### 门户网站



#### 自媒体



此外，近年来涌现的一些商业公司也在为网站及新闻机构供应数据新闻。

## PART2 中外差异从何而来？

### 新闻生产模式和市场影响内容呈现

#### 国内外数据新闻成本投入差异

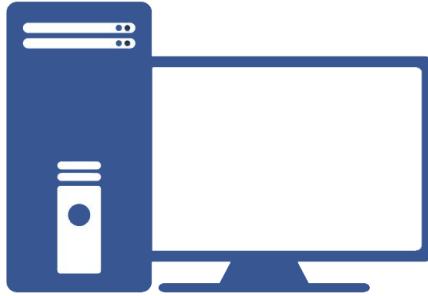
美国机构媒体	较少考虑成本问题，更专注于内容生产。
国内机构媒体	由于和市场关系紧密，编辑部成员需要综合考虑成本预算、以及内容是否能够盈利的问题，因此更加注重阅读量、转发量、点赞量等能被量化的指标。

#### 基于传播效果考虑采取的形式差异

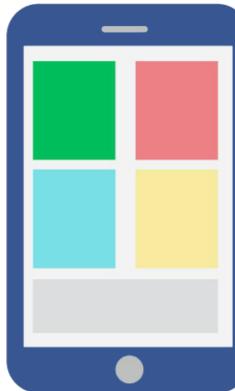
国内媒体	更多采用静态图文、视频等易于传播的呈现方式。
国外媒体	综合考量各种呈现方式，适时采用高成本的交互方式呈现作品。

## PART2 中外差异从何而来？

# 新闻生产模式和市场影响内容呈现

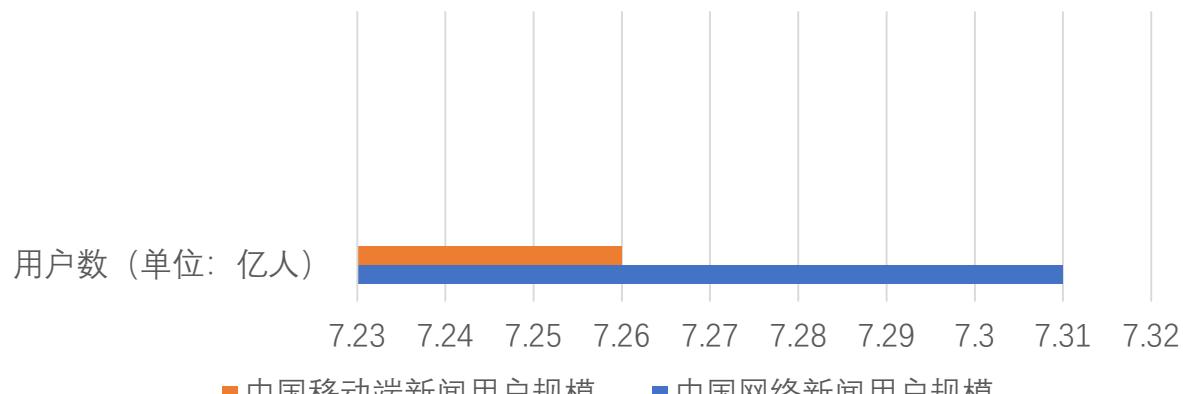


我国互联网时代始于1994年，发展晚于国外，国人的PC端使用粘性相比国外较弱。**步入移动时代后，我国从PC端到移动端的转型也更加彻底。**



中国的数据新闻作品主要依托于移动终端、国外作品多会兼顾PC端和移动端，这样的内容呈现特点也受到市场的影响。

## 2020年3月中国网络新闻用户数据



数据来源：第45次《中国互联网络发展状况统计报告》

## 国内新闻传播的主要途径



移动端用户规模庞大、具有巨大的营销价值和潜力，这使得与市场关系紧密的**中国媒体更倾向于将作品投放在移动终端上**。

## PART2 中外差异从何而来？

# 不同终端限制了数据新闻作品的呈现



## PC端数据新闻作品的三大优势

PC端可以利用鼠标点击实现更复杂、精准的交互和设计

更大的屏幕面积则可以展示承载更大信息量的图表

不同于移动端碎片化的阅读方式，PC端更有利于读者阅读深度作品

因此，对于重视PC端的国外媒体来说，他们的数据新闻作品在交互、信息量、设计等方面都具有优势。

## PART2 中外差异从何而来?

# 不同终端限制了数据新闻作品的呈现

## 国外PC端作品案例



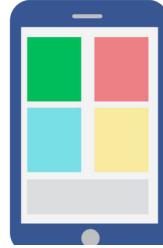
一亩三分地作品《Gobal COVID-19 Tracker & Interactive Charts: Real Time Update& Digestable Information for Everyone》

该作品信息量巨大，读者可以利用鼠标点击查看交互图表详情。内容包括统计数据、疫情地图、美国病例及各州核酸检测机构的联系方式、实时动态、应急物资、趋势、相关视频等。

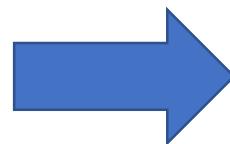
作品链接：<https://coronavirus.1point3acres.com/>

## PART2 中外差异从何而来？

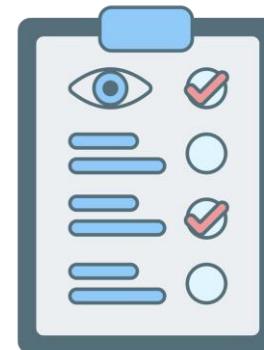
# 不同终端限制了数据新闻作品的呈现



而国内媒体更加偏爱通过移动端呈现作品，适合移动端传播的静态图文成了国内数据新闻常见的呈现方式。



主要传播途径：微信公众号和微博



主要呈现形式：静态图文

如澎湃美数课的《图释两千年传染病史：若瘟疫无法被根除，我们应该如何与之相处？》，就是基于移动端的静态图文。该作品通过梳理人与传染病共生的历史，对新冠疫情之后的境况进行反思。

作品链接：[https://www.thepaper.cn/newsDetail\\_forward\\_6058438](https://www.thepaper.cn/newsDetail_forward_6058438)

人类在20世纪大幅降低了传染病的死亡率



数据来源：澳大利亚健康与福利研究所；Infectious Disease Mortality Trends in the United States, 1980–2014；

## PART2 中外差异从何而来？

# 国内外的数据开放环境不同

## 国外概况

一般而言，欧美国家具有更好的数据开放政策和环境。早在1966年，美国国会就通过了《信息自由法》，规定联邦政府的记录（除某些信息外）原则上向所有公民开放。

除了政府机构，国外民间的数据开放运动也方兴未艾。在这样的环境下，媒体重视数据引用规范，以及向公众开放自建数据库，也成为一种趋势和潮流。

## 国内现状

国内各类数据开放平台正在发展中，在数据检索、用户权限、下载格式等方面有待提升。即使是开放的数据，也往往缺乏统一标准，给数据新闻从业者开发和利用数据带来一定困难。

在这样的环境氛围下，国内的媒体就不太有动力参与到数据开放的运动中来。



### 三、如何提升数据新闻水平？

## PART3 如何提升数据新闻水平?

# 提升可视化设计水平，合理利用可视化

### 问题

我国数据新闻作品设计图表形式单一、静态信息图居多的特点与可视化设计水平不高有关。

### 建议

#### 数据新闻团队

积极引进设计和数据新闻专业的人才

#### 高校

做到新闻素养和技术素养并重，培育良好的数据新闻学习氛围。

### 问题

国内部分媒体存在披着数据新闻的外壳，“为可视化而可视化”的情况。

### 建议

建议媒体在进行数据呈现时，应着重考虑数据可视化是否起到了降低阅读门槛并精准传递信息的作用，而不是单纯用技术炫技，舍本逐末。

## PART3 如何提升数据新闻水平?

### 丰富数据来源

在数据来源方面，**中国数据内容作品的数据来源不够精细、标准化程度不高。**

#### 问题

政府信息存在不够精细、不够标准化等问题。在本次疫情中，各地卫健委公布的疫情数据是媒体最重要的数据来源之一，但却存在数据时段以及颗粒度不匹配的问题。

比如大部分省份在通报疫情情况时都以一天24小时为单位，而某些省份的通报时段则常常为十几个小时的时段，且有些通报时段横跨两天，无法精确得到每日新增病例情况。

#### 建议

因此，政府各部门可以相互协商，对数据通报进行精细化、标准化处理，以便更好地服务公众。

## PART3 如何提升数据新闻水平? 规范数据引用

### 建议

在数据引用方面，国外媒体普遍有较为规范的引用标准，他们或是使用注释的方式在文中直接标注数据来源，或是通过超链接直接链接到数据来源网站。这样的操作有利于读者追溯原始数据内容，国内媒体也应多借鉴参考。

Sources: Bloomberg reporting and [The University of Washington Institute for Health Metrics and Evaluation](#)

Note: Data for Washington State combines data from the Life Care Center in Kirkland, Washington, with other counties in the state.

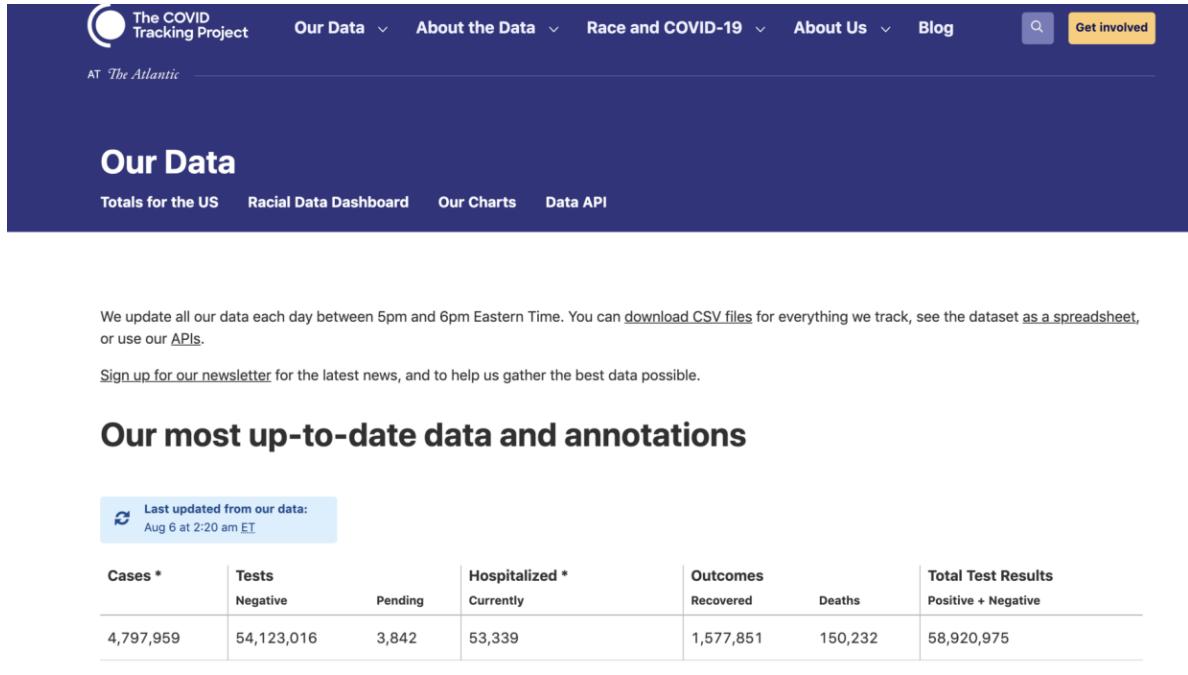
Editors: [Alex Tribou](#), [Mira Rojanasakul](#), [Drew Armstrong](#) and [Heather Smith](#)

彭博《When the Coronavirus Outbreak Could Peak in Each U.S. State》数据来源。

## PART3 如何提升数据新闻水平? 促成数据开放

### 建议

此外，在数据开放方面，国内媒体也可以尝试建立自己的数据库，逐步提高自身的公共服务价值和数据自主性。例如此前提到《大西洋月刊》推出的新冠数据追踪计划 (The COVID Tracking Project) ，就很值得国内媒体学习。



The screenshot shows the 'Our Data' section of the The COVID Tracking Project website. The top navigation bar includes links for 'Our Data', 'About the Data', 'Race and COVID-19', 'About Us', 'Blog', a search icon, and a 'Get involved' button. Below the navigation is a sub-navigation bar with links for 'Totals for the US', 'Racial Data Dashboard', 'Our Charts', and 'Data API'. The main content area features a message about daily updates and download options, followed by a call to sign up for the newsletter. A prominent heading 'Our most up-to-date data and annotations' is displayed above a table. The table has a header row with columns for 'Cases \*', 'Tests', 'Hospitalized \*', 'Outcomes', and 'Total Test Results'. The data row shows values: 4,797,959 cases, 54,123,016 tests (54,123,016 negative, 3,842 pending), 53,339 hospitalized (53,339 currently), 1,577,851 recovered, 150,232 deaths, and 58,920,975 total test results (Positive + Negative). A small note at the bottom of the table states 'Last updated from our data: Aug 6 at 2:20 am ET'.

Cases *	Tests	Hospitalized *	Outcomes	Total Test Results
	Negative	Currently	Recovered	Positive + Negative
4,797,959	54,123,016	3,842	53,339	1,577,851
				150,232
				58,920,975

作品链接: <https://covidtracking.com/data>

## PART3 如何提升数据新闻水平? 促成数据开放

### 国外数据开放案例

《卫报》开办的数据博客栏目，不仅旗下所有数据新闻作品的数据都可供读者下载分析，还为读者提供其他优质的数据资源。



### Other resources

- [CovidVictoria.com](https://CovidVictoria.com) which maps cases and allows zooming
- [The official DHHS dashboard](https://www.hhs.gov/coronavirus/epic-covid-19-data-tracker.html)
- [Covid19Data.com](https://Covid19Data.com) which has a good section on the Victorian outbreak

澳门科技大学学者章戈浩认为，这种与读者之间的“互动性”进一步扩大了数据新闻产品的传播范围和影响力。

## PART3 如何提升数据新闻水平？

# 提升对数据的应用和挖掘能力

## 问题

目前，国内的数据新闻作品往往局限于对数据的简单呈现，较少进行深度挖掘和分析。

## 建议

诚然，简单呈现是数据运用中不可缺少的一部分。但除了对信息的直观展示之外，数据新闻还应在大量的数据中发现关键因素，揭示数据背后的隐含逻辑。

这类数据新闻作品对数据新闻工作者提出了更高的要求：

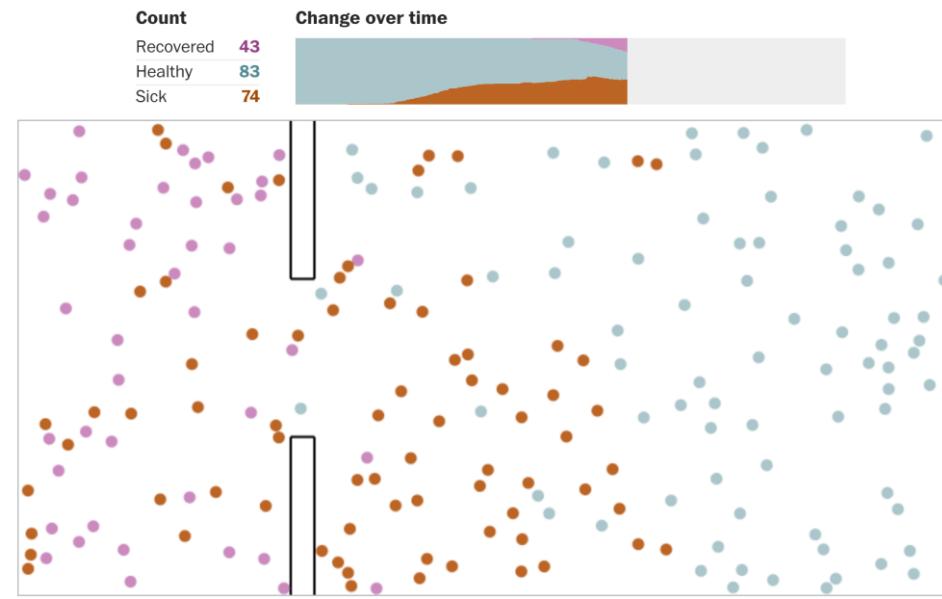
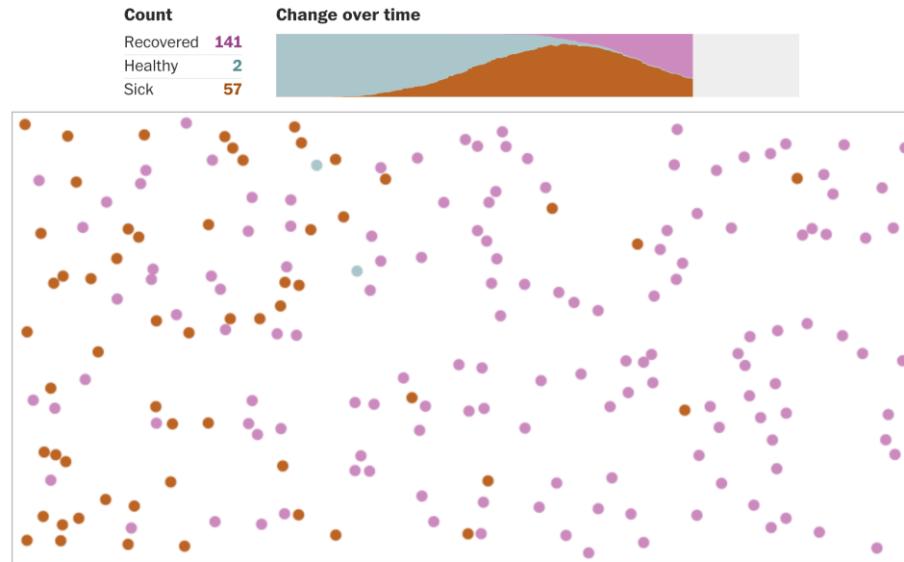
优秀的数据素养，例如收集、分析处理、利用数据能力。

一定的统计学和社会科学知识。

这恰恰是目前国内不少数据新闻从业者所缺乏的。

# PART3 如何提升数据新闻水平? 提升对数据的应用和挖掘能力

## 国外数据挖掘案例



如华盛顿邮报的《Why outbreaks like coronavirus spread exponentially, and how to “flatten the curve”》对病毒传播数据进行建模，模拟了不同社交距离对于病毒传播的影响，并通过可视化将这一复杂过程直观展现。

作品链接：[https://www.washingtonpost.com/graphics/2020/world/corona-simulator/?tid=pm\\_graphics\\_pop\\_b](https://www.washingtonpost.com/graphics/2020/world/corona-simulator/?tid=pm_graphics_pop_b)

# 中外数据新闻的差异与差距

## ——以全球新冠疫情数据化内容为例



新浪新闻新媒体实验室  
数可视教育公益基金

联合出品

撰稿人：黄志敏、姜荟秀、王晓晴、黄桢昕

组稿人：张翔、江华宇