

第 1 章 信息可视化启蒙

1.1 可视化启蒙

信息无处不在。小到细胞、分子、原子等传达的信息，大到宇宙、黑洞等给人类带来的未解信息。当今社会，因为信息量的大而广，复杂性清晰可见。科技的飞速发展，也使信息变得更加复杂，信息技术将出现“网罗一切”的特征，人与网络直接错综复杂地连接，相互依存，信息技术覆盖人们全部生活，只要其中一个环节有变动，便会引起一系列的变动。但信息技术又能够被用来预测变动，各个信息的交错、变化形成一个个拉锯战。

数据可视化起源于 20 世纪 50 年代的计算机图形学，当时，人们利用计算机创建出了首批图形图表。那时候人们使用计算机创建图形图表，可视化提取出来的数据可以将数据的各种属性和变量呈现出来。随着计算机硬件的发展，人们欲创建更复杂且规模更大的数字模型，于是发展了数据采集设备和数据保存设备，此时，需要更高级的计算机图形学技术及方法来创建这些规模庞大的数据集。随着数据可视化平台的拓展、应用领域的增加、表现形式的不断变化，以及增加了诸如实时动态效果、用户交互使用等，数据可视化像所有新兴概念一样边界不断扩大。

信息本身可以无限复杂下去，设计师需要把一些复杂的、杂乱无章的信息进行整理归纳，让它们可以清晰地、有条理地以可视化的形式呈现出来，用最快的方式传达出去。信息可视化即把世界上无比复杂多变且难以理解的信息使用文字、图形、图表等各种信息组织方式重组原始信息以提高信息传递的效能，编辑信息内容，使之清晰、明确，能够被简化成人们轻易理解的过程。

1.1.1 关注技术野火

“整个商业领域都因为大数据而重新洗牌，一旦思维转变，数据就能巧妙地被用来激发新产品成型。”

——维克托·迈尔-舍恩伯格

世界已经迈入大数据（big data）时代。随着互联网、物联网、云计算等信息技术的迅猛发展，信息技术与人类世界政治、经济、军事、科研、生活等方方面面不断交叉融合，催生了超越以往任何年代的巨量数据，遍布世界各地的各种智能移动设备、传感器、电子商务网站、社交网络每时每刻都在生成类型各异的数据。截至 2022 年，全球数据产量增长为 81.3ZB，复合年均增长率达 24.6%（<https://wap.seccw.com/Document/detail/id/20277.html>）。大数据具有 4V 特征，即体量巨大（volume）、类型繁多（variety）、时效性高（velocity）及价值（value）高密度低，给人们带来了新的机遇与挑战。

从 SGI（硅图）的首席科学家 John R.Masey 在 1998 年提出大数据概念，到大数据分析技术广泛应用于社会各个领域，如今再也没有人会怀疑大数据分析的力量，而且都在竞相利用大数据来增强企业的竞争力。目前，大数据分析行业仍处于快速发展的初期，每时每刻都有新的变化。从概念到应用、从结构化数据分析到非结构化数据分析，大数据分析技术在不断地进化，出现了“舆情分析—情感分析—预测分析”的变化。

“舆情”通常是指较多群众关于现实社会及社会中各种现象、问题所表达的信念、态度、意见和情绪表现的总和。简而言之，就是社会舆论和民情。一个严格定义是：“舆情”是指在一定的社会空间内，围绕中介性社会事项的发生、发展和变化，作为主体的民众对作为客体的国家管理者产生和持有的社会态度。舆情主要指民众对社会各种具体事物的情绪、意见、价值判断和愿望等。例如，亚马逊上的用户对某商品的评论，商家可以根据用户的评论和反馈为用户提供定制化的服务，甚至可以预测用户的需求，从而达到更加准确的销售目的；再如，新浪微博上粉丝过万的大型零售商等，也可以根据用户发表的微博、微话题、签到地点为用户定制性地推送优惠及新品信息。这些看似庞大无规则的数据，包含着大量的用户标签及潜在的用户肖像，粉丝自身发布的微博含有大量的数据信息，这些信息包含用户的个人爱好、年龄阶段、近期想购买的款式，甚至是自己希望有的款式与功能等。

“情感分析”又可以叫作意见抽取、情感挖掘、主观分析，等等。例如，看完电影对电影好坏的评价、Google Product Search 上面对产品属性的评价，并展示褒贬程度。通过推特（Twitter）上和民意调查得来的数据

进行对比,发现两者对消费者信心的统计结果有很大的相关性,相关度达到80%。同样,利用推特上的公众情绪预测股票,发现CLAM(平静)情绪的指数可以预测三天后道琼斯的指数。

“预测分析”是指使用历史数据、机器学习和人工智能来预测未来会发生什么。世界杯预测、高考预测、电影票房预测、流行病预测……在大数据时代,预测分析已经逐渐在商业社会中得到广泛应用。大数据预测分析技术也成为目前大数据领域最难的一个环节,不仅考验企业的大数据处理技术,更是对数据科学家提出了更高要求。虽然国内仍然在关注舆情分析,但国际上大数据分析的研究已经进入了一个全新的阶段,“预测分析”技术成为最具代表性的未来。

大数据中隐藏着的“广泛需求”是设计的重要来源——大数据带来的新型产品设计机会。例如,日本先进工业技术研究所的坐姿研究。越水重臣教授所做的研究是关于一个人的坐姿。他认为,当一个人坐着的时候,他的身形、姿势和重量分布都可以量化和数据化。他们的团队通过在汽车座椅下部安装360个压力传感器以测量人对椅子的压力方式,并采用0~256这个数值范围对其进行量化,这样就会产生独属于每个乘坐者的精准的数据资料。在实验中,这个系统能够通过人体对座位的压力差异识别出乘坐者的身份,准确率高达98%。当一个人的坐姿转化为数据之后,设计可以介入的点有什么呢?安全座椅设计、疲劳驾驶警示、自动刹车等都可以放到与“坐”相关的所有周边产品……

1.1.2 像哲人那般思考

数据(data)这个词在拉丁文里是“已知”的意思,也可以理解为“事实”。数据可视化是数据产品比较重要的能力。数据可视化已经遍布商业、金融、医学、社会学、政府等各个领域。为什么要将数据可视化?普及数据概念,这是数据产品的需要,海量数据无法直接分析,通过可视化可以更容易、更快速地获得想要的知识。阿里云的DataV针对不同人群落地了不同的数据可视化场景。数据产品所针对的人群是具有一定分析能力、业务理解深度的专业人员,在技术上以通用前端技术为主,强调图表绘制能力,要求数据的准确性,要求可视化展现方式,学习成本低;数据大屏所针对的人群是对数据背景有一定程度了解的受众,在技术上主要运用WebGL、Flash动画软件,性能稳定性要求高,强调视觉冲击和整体信息设计与硬件配合;数据传播所针对的人群是未知背景的普罗大众,技术方面主要是设计主导,信息图及视频实践案例较多,更偏话题性、趣味性。

阿里云发布全球首个一站式大数据平台“数加”,其中包含一款针对

中国县域经济的数据应用产品，能让每个县域的管理者实时了解区域的经济态势、产品特点、内需消费等关键数据的走势，为政府提供支持，堪称师爷——“郡县图治”（见图 1-1）。^①Design Thinking（设计思维）：数据化≠数字化，数据化——把现象变成量化形式的过程；数字化——把模拟数据转化成 0 和 1 表示的二进制码。



图 1-1 阿里云“郡县图治”

大数据带来了“强蝴蝶效应”，该效应是信息可视化设计的重要依据。蝴蝶效应（The Butterfly Effect）是指：20 世纪 70 年代，美国一个叫洛伦兹的气象学家在解释空气系统理论时说，亚马孙雨林一只蝴蝶翅膀偶尔振动，两周后也许就会引起美国得克萨斯州的一场龙卷风。它是指，初始条件十分微小的变化经过不断放大，对未来状态会造成巨大的影响。“强蝴蝶效应”作用在同一只蝴蝶上，也许两周后这只蝴蝶还会引起美国得克萨斯州的一场龙卷风，但是在这之前，人们了解大数据的预测结果后会全部迁移，而且所有的防灾措施都已经准备就绪了。大数据研究专家吴波先生提到，如果大数据成功预测并向大众发布社会的未来，这将严重削弱社会的人际博弈特征和内在的矛盾动力，进而酿造一边倒的趋势风暴，并直接摧毁其所涉及的许多价值领域，对人类将造成不可估量的损失。在商业领域方面，大数据会加快大企业形成垄断，因为只有大企业才有巨额资本支持自身的大数据库建设。如果任由大企业通过大数据形成垄断，这将大大减少人类社会的发展多样性。在社会学领域方面，大数据也可能会

^① 周建丁. 数据可视化：基于阿里云 DataV 的“郡县图治”[EB/OL]（2016-01-22）[2024-06-05]. <https://blog.csdn.net/happytofly/article/details/80121668>

造成社会走向某一个极端。如今，大数据会催化人的某一个喜好和个性，比如，某个用户通过软件收看了某个题材的视频，随后大数据就会将更多的该题材的视频推送过来，大大减少该用户接触其他题材的视频的可能性，形成“信息茧房”，从而进一步形成思维固化。长此以往，正如吴波先生所说的，这个趋势风暴可能会摧毁更多的价值领域，打击社会的多样性发展，造成无法估计的损失。

换言之，大数据能帮助人类提前预测社会趋势，但也可能导致社会趋势改变自然节奏并演变成一场灾难。

可视化推动大数据对设计的赋能：探索数据的审美属性，寻找视觉的社会属性。机制、宏观、围观，人们可以对大数据展开无数创想……但很重要的一点是，可视化不是目的，可视化是手段，设计师应该考量的是可视化作品背后的社会属性，找到大数据对设计的“赋能”能力。

Design Thinking（设计思维）：为何会有爆款和网红？为何有“自动流行”？“友谊的小船”来自哪里？多样且丰富的设计美学能否做到让每个产品都独树一帜？电商时代的社会化和社区化有没有社会属性加持？《超体》是否总有一天会在人们身边出现？

1.1.3 用设计思维构筑设计风格（Design Style with Design Thinking）

具有实用价值的可视化作品不仅是创新十足的统计分析结果，或计算机算法的应用，还必须让使用者能够理解。这就需要一整套规范，包括颜色、形状、线条、层级和平面构成的视觉语言系统，就如同人类一直使用的字母表和字符一样，能更清晰、准确地传达信息。

——马特·伍尔曼

视觉新语言的构成要素：色彩、文字、图像、大小、形状、对比度、透明度、位置、方向、布局、比例、形态和含义……信息具有审美属性，人们的生活正被数据驱动着。如何用设计语言制作让人惊艳的界面、用直观的方式对复杂信息进行准确表达，对数据进行生动讲述是未来的机遇（Opportunity）。

拍信^①，中国增长速度较快的视觉内容品牌，是国内领先的正版图片交易平台，足量高品质正版视频可满足各类企业和个人创作的需求，完美适应自媒体、广告、互联网等企业需求，是追求品质和实惠的不二选择。拍信抓住了国人越来越重视图片版权这个机会点，给重视版权的创作者创

^① 拍信官网 <https://www.paixin.com/>

建了机会，为和他们有品质图片需求的人建立起一个高质量正版图片交易平台（见图 1-2）。



图 1-2 拍信——国内优质的数字创意内容平台

在新媒体环境中，无论是何身份、在何领域，将个人信息可视化、塑造个人品牌已经成为决胜于这个时代的重要一步。首先，要明确你的个人品牌传播的是什么样的内容，受用的人群又是哪些，这就像是新闻学中的基本问题 3W: Who、What 和 Whom，用西方哲学三大终极问题翻译就是：我是谁？我从哪里来？我要到哪里去？每个人各有所长，选择擅长或热爱的事物将其品牌化，并对它进行准确定位。其次，要擅于利用新媒体，评估它能给自己带来多大的影响力。因为传播渠道就是一种手段，它的作用就是尽可能帮助自己制造的内容扩大声量，是自带传播属性的优质内容的一种辅助介质。根据自己想塑造的形象选择合适的渠道很重要，要依靠合适的新媒体建立自己的标签。

当设计师一直学习知识、输出知识时，就需要不断地进行资源互换，将人与人、资源与资源进行嫁接互换，在拥有自己的个人品牌之后，实现阶级跨越。变现，就是一种质变的结果。2017 年，86 岁的百雀羚脑洞大开，用一张全长 427 厘米、美感十足的长图惊艳了不少人（见图 1-3）。^①但随即，有自媒体质疑其实际带来的销售转化低，称截至 5 月 11 日中午 12 点，根据第三方监测平台数据粗算，微信平台总阅读量近 3 000 万，但“3 000 万 + 微信阅读转化却不到 0.000 08”。但根据天猫发布的“2017 双十一品牌销量排行榜”显示，美妆品牌排行榜中百雀羚位居第一。由此可见，变现的方式和时机不仅重要，变现的过程大都会需要一段时间。^②

① 百雀羚推出一镜到底的神广告：化身一九三一美女特工 [EB/OL] (2017-05) [2024-06-05]. <https://www.digitaling.com/projects/21630.html>

② 梁欣萌. 知识经济时代个人品牌的打造 [J]. 国际公关, 2017, 78 (6): 68-69.



图 1-3 百雀羚——美妆品牌宣传长图

1.2 数据的价值 Datawatch show 来自业界的观点分享

Any Data, Delivered at the Speed of Business.

任何数据，任何速度，业务所至，分析所至。

——Datawatch

谷歌公司首席经济学家、美国加州大学伯克利分校的哈尔·瓦里安（Hal Varian）教授指出：“数据正在变得无处不在、触手可及；而数据创造的真正价值，在于我们能否提供进一步的稀缺的附加服务。这种增值服务就是数据分析。”数据的背后隐藏着信息，而信息之中蕴含着知识和智慧。大数据作为具有潜在价值的原始数据资产，只有通过深入分析才能挖掘出所需的信息、知识及智慧。未来人们的决策将日益依赖大数据分析的

结果，而非单纯的经验和直觉。

“你眼前有一堆数据，但你却只能获得很少的信息。”这是 Datawatch 团队来自业界的观点分享。数据应用陷入“三无”的困境——无法抓住、无人问津和无法支持。无法抓住指数据讲究实效性，数据的价值会随时间的流失而下降；无人问津指 70% 的历史数据仅作留存备份；无法支持指 90% 的海量数据被闲置，技术不支持，无法被充分利用。如何发掘数据的最大价值？这里有一串数字，如何以最快的速度发现这些数字里有多少个“3”（见图 1-4）？是通过大小、形状、颜色、密度，还是通过分组、临近、动作、符号（文化特征的）来发现？

1281736875617897654698450698560498	12817 3 6875617897654698450685604982
2867629809858458224509856458935098	8676298098584582245098564589 3 50984
4509809658590910302099059595957725	509809658590910 3 02099059595957725 3
3646750506789045678845789809821377	6467505067890456788457898098216 3 76
65487266490856093294968641	548726649085609 3 2949686897

图 1-4 如何快速发现这些数字里有多少个“3”

可视化让思想变得更加迅速。面对可视化图表，人们的思维是迅速、自动、轻松、即时的，是一种生物本能和潜意识。而面对复杂的数字化文字，人们的思维是缓慢、有意识、需要有一定认知要求的，需要全神贯注，无法在运行多任务的情况下完成识别。如果把人的五感比作信息传输介质，视觉的速率是 1 250MB/s，触觉的速率是 125MB/s，听觉的速率是 12.5MB/s，即一幅图胜过千言万语。人类从外界获得的信息约有 80% 以上来自视觉系统，当大数据以直观的、可视化的图形形式展示在分析者面前时，分析者往往能够一眼洞悉数据背后隐藏的信息并转化成知识以及智慧。对于一个对大数据一窍不通的人来说，让数据实现可视化无非是对使用者了解大数据最方便快捷的方法了，这样一来，大数据可以更贴近用户的使用习惯和使用需求。

下面来看看历史上极具影响力的可视化案例。

1854 年伦敦暴发霍乱，10 天内有 500 人死去，但比死亡更加让人恐慌的是“未知”，当时人们对于医学知识了解甚少，所以人们不知道霍乱的源头和感染分布。流行病专家 John Snow 参与调查病源，他在绘有霍乱流行地区所有道路、房屋、饮用水机井等内容的 1 : 6 500 比例尺地图上用黑杠标注死亡案例，最终地图“开口说话”，显示大街水龙头是传染源。这让更多的人快速了解了霍乱，还使公众意识到城市下水系统的重要性并采取切切实行动。这是首次以地图为基础的空间分析，也是可视化分析的经典案例。