

互联网发展大趋势前瞻之三“明暗相间”： 明网+（深网&暗网）

3. 明暗相间：明网+（深网&暗网）

美剧《纸牌屋》有一段对话，“96%的互联网数据无法通过标准搜索引擎访问，其中大部分属于无用信息，但那上面有一切的东西，儿童贩卖、比特币洗钱、致幻剂、赏金黑客……”剧中所描述的这个“黑暗世界”是真实存在的，常被犯罪分子利用从事非法活动，现已成为各国执法部门联合打击的重点领域。如，2016年11月，北京警方公布了一起大学生用“暗网”传播儿童色情信息的案件。

2016年12月，美国 CBS 电视台播出自制的八集纪录片《暗网》，每集一般由三个小故事构成，通过普通人所不知的网络空间隐秘世界，讲述互联网对人和人性的影响。互联网包罗了海量的信息量，但我们日常能看到的只是冰山一角。那些通过百度、谷歌、雅虎等搜索引擎等可以检索、处于各国政府监管之下的互联网络，是可见互联网，也称作表层网络或明网，就是我们常用的公共互联网。那些通过搜索引擎无法检索的互联网络，是不可见互联网路，也称作深层网络或深网。如，图书书目、病历等网站很多。据统计，明网与深网的数据规模之比为 4%：96%。需要注意的是，深网

之中还藏着暗网，即使知道网址，普通浏览器也无法浏览，各国政府难以追踪痕迹。

“暗网”被喻为“隐藏的服务器”。最初源于美国军事机构，1996年美国海军研究试验所为了保护船只间通讯网络安全，避免被敌军跟踪信号，启动一项旨在通过代理服务器加密传输数据的技术开发，任何使用者都会实时处于匿名状态，而不会泄露身份，其保护密码像洋葱一样层层叠叠，被命名为“洋葱路由”（Onion Routing，简称 Tor）。2004年，美国海军研究实验室财政紧缺，停止对 Tor 的资金支持，并将之开源。



明网信息可以互联共享，暗网信息经过加密，不能通过超链接访问，不能通过百度、谷歌等标准搜索引擎抓取，很难定位和追踪，需要通过动态网页技术才能访问。用户只有通过匿名软件 Tor，才能浏览访问暗网。用户的访问请求会被三重加密，即会在分布全球的成千上万台主机之间随机跳转三次，最终才到达服务器，访问特定网

址，使溯源用户极其困难。所有访问者完全没有身份区别，可以保持完全的匿名性，不会留下任何访问痕迹。浏览和搜索暗网需使用特殊“暗网”专用搜索引擎，如，Onion.City、Onion.to、Not Evil、Memex Deep Web Search Engine 等。2016 年，研究人员开发出新的匿名网络系统 Riffle，修补了 Tor 的一些漏洞，对抗黑客监听和政府追踪的能力更强，而且能使用较少的带宽在匿名用户之间传输大文件。

暗网的域名数量可达可见互联网的 1%，主要由各种贩卖毒品、武器甚至提供杀手雇佣等网站组成。“暗网”上几乎可以找到任何东西，是毒品、军火、色情、伪造证件、假钞甚至杀手等非法交易的温床。如，“丝绸之路”、Agora 都曾是“暗网”最大毒品交易市场，每年交易额不低于 1 亿美元。专项研究报告显示，“暗网”的 57% 由儿童色情、非法金融、毒品市场、贩卖武器、假币等非法内容组成。IS 极端组织网站多数建在“暗网”，以躲避各国政府打击，传播极端思想，持续招募成员。一些国家的异议人士、记者、研究人员经常使用 Tor 浏览器掩藏其网络浏览偏好和地理位置，匿名发布和传播信息。在“阿拉伯之春”中，埃及抗议者通过暗网串联，成功躲避政府部门的技术追踪。2015 年 2 月，美国国防部高级研究计划局（DARPA）研发新搜索引擎 Memex，以更好掌握暗网及分析网络数据关系，帮助执法部门追踪非法活动。

如今，伴随大模型快速发展，其训练数据和算法等的黑箱机制，加之可以千人千面回答问题，使其内容输出进一步黑箱化，成为新的暗网和深网。