

目 录

一、中国研究数据服务平台（CNRDS）简介	1
二、CNRDS 公司特色数据库	3
上市公司经营研究	4
（一）中国创新专利研究数据库-CIRD	4
（二）中国审计研究数据库-CARD	8
（三）中国券商与分析师数据库-CBAS	10
（四）环境、社会和治理数据库-CESG	12
（五）中国家族企业研究数据库-CFFD	14
（六）公司并购重组数据库-CMAD	16
（七）监管问询数据库-CRID	17
（八）中国上市公司内部控制研究数据库-ICRD	18
（九）上市公司参控股公司数据库-CPCD	20
（十）关键审计事项数据库-KAMD	22
（十一）中国上市公司供应链研究数据库-SCRD	23
（十二）中国上市公司异常收支数据库-ARED	25
（十三）投资者关系管理数据库-IRMD	26
（十四）海外上市研究数据库-COLD	27
（十五）中国法学研究数据库-CNLAWS	28
（十六）上市公司盈余与信息质量数据库-EIQD	30
（十七）上市公司专利引用数据库-CITE	31
（十八）董监高责任保险数据库-CDOD	33
（十九）员工持股计划数据库-ESOP	34
（二十）上市公司精准扶贫研究数据库-TPAD	35
（二十一）上市公司海外经营数据库-COOD	36
（二十二）上市公司土地市场信息数据库-LMID	37
（二十三）上市公司年度信息数据库-CAID	38
（二十四）上市公司区块链投资数据库-BLOCK	39
（二十五）绿色专利研究数据库-GPRD	40
（二十六）上市公司承诺事项数据库-CMIT	43

（二十七）上市公司委托贷款数据库-CELD	44
（二十八）上市公司环境治理数据库-CEGD	45
（二十九）上市公司金融化数据库-CFID	46
（三十）上市公司银行借款数据库-LOAN	47
（三十一）对赌协议数据库-VAMD	48
（三十二）一致行动人数据库-PACD	49
（三十三）上市公司风投持股数据库-VCSD	50
（三十四）上市公司数字化转型数据库-CTTD	51
（三十五）上市公司公告统计数据库-CCAS	52
（三十六）上市公司 ESG 评级数据库-ESGR	53
（三十七）公司上市前专利数据库-PCID	54
（三十八）上市公司股权质押数据库-EPCD	56
上市公司新闻舆情	58
（一）中国上市公司财经新闻数据库-CFND	58
（二）中国上市公司股吧评论数据库-GUBA	60
（三）网络搜索指数数据库-WSVI	61
（四）中国上市公司社交媒体数据库-CSMD	62
（五）上市公司并购新闻数据库-MAND	63
（六）上市公司财务舞弊新闻数据库-FFND	64
（七）上市公司高管新闻数据库-CEND	65
上市公司文本信息	66
（一）年报文本语气数据库-ARTD	66
（二）管理层讨论与分析数据库-CMDA	67
（三）上市公司业绩说明会数据库-ECCD	68
（四）中国上市公司网上路演数据库-CNRD	69
（五）中国上市公司澄清公告数据库-CCAD	70
（六）年报关键词研究数据库-AKRD	71
（七）MD&A 关键词研究数据库-MKRD	72
（八）上市公司监管意见数据库-CROD	73
（九）CSR 报告文本数据库-CSRT	74
（十）分析师报告文本数据库-TDAR	75
（十一）IPO 申报稿文本数据库-IPOT	76

资本市场人物特征	78
(一) 董事长与总经理研究数据库-CCEO	78
(二) 上市公司其他高管数据库-OECD	80
(三) 独立董事研究数据库-IDRD	81
(四) 董秘信息数据库-CSBD	83
(五) 基金经理研究数据库-FMRD	84
(六) 金融机构高管数据库-FIED	85
(七) 董事会研究数据库-BOARD	86
银行及金融研究	87
(一) 银行研究数据库-CBRD	87
(二) 中国商业银行分支机构数据库-CCBD	89
(三) 中国风险投资和私募股权数据库-VCPE	91
(四) 中国金融理财研究数据库-CFSD	93
(五) 基金公司研究数据库-FCRD	95
(六) 券商投行数据库-CIBD	97
(七) 保险机构研究数据库-IIRD	98
(八) 信托机构研究数据库-CTID	99
(九) 金融中介处罚数据库-FIPD	100
(十) 商业银行数字化数据库-CBDD	101
公司债券信息研究	102
(一) 发债企业财务报表数据库-BCFS	102
(二) 发债企业财务报表附注数据库-BNFS	103
(三) 上市公司发债研究数据库-BIRD	105
社会经济组织研究	107
(一) 政府审计数据库-CGAD	107
(二) 中国高校科技统计数据库-USTS	108
(三) 中国非营利组织数据库-CNGO	110
(四) 中国商会研究数据库-CCCD	112
(五) 高新技术企业研究数据库-HTED	113
(六) 全国高校专利研究数据库-CUID	114
(七) 战略联盟研究数据库-SARD	115
(八) 排污许可研究数据库-DPRD	116

(九) 小巨人企业数据库-SGED	117
三、CNRDS 经济特色数据库	118
宏观经济研究系列	119
(一) 全球夜间灯光数据库-GNLD	119
(二) 中国高铁航线数据库-CRAD	121
(三) 中国环境数据库-CEDS	122
(四) 中国商品交易市场统计数据库-CEMD	124
(五) 中国人口和就业统计数据库-PESD	125
(六) 脱贫数据库-LPOP	126
(七) 中国固定资产投资统计数据库-CIFA	127
(八) 最低工资标准数据库-WAGE	128
(九) 中国住户调查数据库-HSLD	129
(十) 碳中和研究数据库-CNLD	130
(十一) 影子银行研究数据库-SBRD	132
(十二) 绿色金融研究数据库-GFRD	133
(十三) 国家统计标准数据库-CNSD	134
(十四) 环境处罚研究数据库-EPRD	135
(十五) 共同富裕研究数据库-CPRD	136
(十六) 中国人口普查分县数据库-CDPC	138
(十七) 数字经济政策数据库-DEPD	140
(十八) 环境注意力数据库-GEAD	141
(十九) 金融科技研究数据库-FTCH	143
(二十) 工业机器人研究数据库-IRRD	144
区域经济研究系列	146
(一) 城市统计数据库-CCSD	146
(二) 中国县域统计数据库-COTY	148
(三) 中国区域经济统计数据库-REGN	149
(四) 中国城乡建设统计数据库-URCS	151
(五) 中国农村统计数据库-CRSD	153
(六) 区域碳排放数据库-RCED	155
产业经济研究系列	157

（一）产业政策数据库-IPRD	157
（二）中国房地产统计数据库-RESA	158
（三）中国能源统计数据库-CESD	159
（四）中国工业统计数据库-CISD	160
（五）中国高技术产业统计数据库-HISD	161
（六）中国第三产业统计数据库-TISD	162
（七）数字经济研究数据库-DERD	163
（八）工业企业专利数据库-IIED	165
（九）电子商务研究数据库-ECRD	167
对外经济研究系列	170
（一）中国对外直接投资数据库-OFDI	170
（二）一带一路研究数据库-BRRD	172
（三）海外并购数据库-COMA	173
（四）中国贸易外经统计数据库-TEES	174
（五）中国海关贸易数据库-CCTD	175
（六）外国在华专利研究数据库-FCID	177
（七）专利国际化研究数据库-PIRD	181
人文社科研究系列	183
（一）中国文化研究数据库-CCRD	183
（二）儒家文化数据库-CFCN	185
（三）中国旅游统计数据库-CTSD	186
（四）中国科技统计数据库-CSTD	187
（五）全国景点数据库-ATRA	188
（六）中国民政统计数据库-CASD	189
（七）中国卫生健康统计数据库-CHSD	190
（八）中国环境统计数据库-ENVS	191
（九）宗族文化数据库-CLAN	193
（十）中国火炬统计数据库-TORCH	194
（十一）教育经费统计数据库-EFSD	195
（十二）中国价格统计数据库-CPSD	196
（十三）中国知识产权研究数据库-IPRR	197
（十四）中国劳动经济数据库-CLSD	198

(十五) 人口老龄化数据库-PARD	200
财政金融研究系列	201
(一) 中国税收研究数据库-CTRD	201
(二) 中国保险统计数据库-INSD	203
(三) 中国行政审批数据库-CAAD	204
(四) 政府和社会资本合作数据库-PPPD	205
(五) 中国地方债数据库-CLGD	206
(六) 财政预决算数据库-FBFA	208
四、CNRDS 基础数据库	209
上市公司股票基础数据	210
(一) 中国上市公司股价数据库-CNSP	210
(二) 中国股票异常交易数据库-CAST	210
(三) 中国特殊处理股票数据库-CSTS	210
(四) 中国融资融券数据库-CMTD	210
(五) 中国公司 IPO 综合数据库-CIPO	211
(六) 中国上市公司增发与配股数据库-CSEO	211
上市公司财务基础数据	212
(一) 上市公司业绩预告数据库-CEPD	212
(二) 财务报告披露时间数据库-FRDT	212
(三) 中国公司财务报表数据库-CNFS	212
(四) 上市公司财务报表附注数据库-NFSD	212
(五) 上市公司财务指标数据库-CNFI	212
(六) 中国上市公司盈利预测数据库-CEFD	213
上市公司治理基础数据	214
(一) 上市公司基本信息数据库-CBID	214
(二) 中国公司治理数据库-CCGD	214
(三) 审计意见与费用数据库-AUDIT	214
(四) 中国上市公司管理层变更数据库-MTDB	214
(五) 公司与高管违规处罚数据库-VPCE	214
(六) 高管薪酬与激励数据库-ECEI	215
(七) 中国上市公司关联交易数据库-CRTD	215
(八) 中国公司股权研究数据库-CERD	215

（九）机构投资者持股研究数据库-IORD	215
（十）中国公司股利分红数据库-CCDD	215
（十一）中国上市公司诉讼仲裁数据库-CLAD	216
（十二）中国上市公司内部人交易数据库-CITD	216
经济研究基础数据	217
（一）中国宏观经济研究数据库（年度）-MACRO	217
（二）中国宏观经济研究数据库（季度）-MACROQ	217
（三）中国宏观经济研究数据库（月度）-MACROM	217
（四）中国区域经济研究数据库-CRED	217
（五）中国债券研究数据库-BOND	217
五、最新学术动态	218
六、数据定制	219
七、数据合作	221

一、中国研究数据服务平台（CNRDS）简介

1. 平台概况

中国研究数据服务平台（*Chinese Research Data Services Platform, CNRDS*）是一个高质量、开放式、平台化的中国经济、金融与商学研究的综合数据平台。**CNRDS** 的宗旨在于借鉴 **WRDS** 等国外一流商学院打造的数据平台，构建中国特色的研究数据资源，促进中国研究数据的规范化、平台化、国际化，继而推动中国研究向纵深化、精细化方向发展。

CNRDS 平台提供的数据库主要来源于三个途径：1) 平台自有数据，从采集、整理到呈现均有该领域国内外知名专家和学者深度参与；2) 由相关领域的专家和学者提供，并由 **CNRDS** 提供技术支持和形式审核以及后续维护；3) 各专业领域的数据库公司提供原始数据，**CNRDS** 再邀请相关领域的专家和学者对数据质量进行严格审核。

我们相信每一位学者在支持和推动中国研究数据和资源构建上迈出的小小一步，都将成为中国学术研究发展的一大步。

2. 平台特色

CNRDS 特色库紧跟学术热点和学术前沿，提供市场尚无或者获取难度较大的特色研究数据；基础库参阅各领域重点文献，整合了学术研究过程中可能用到的大部分基础数据。

CNRDS 数据库的特点可归纳为：

- 国内外知名高校教授合作开发
- 特色数据库满足各类研究需求
- 界面与模块设计契合学术研究

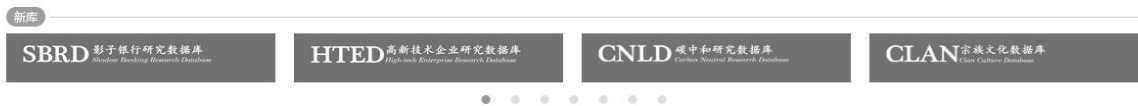
3. 平台网址

www.cnrds.com

4. 适用对象

中国经济、金融与商学研究方向的教授、学者；利用资本市场数据构建投资决策的金融投资机构；撰写报告的企业。

5. 平台界面



经济与金融 会计

期刊简称	论文标题	发表日期	相关库
JFE	Issuance overpricing of China's corporate debt securities	2022年04月	
JFQA	Housing Wealth as Precautionary Saving: Evidence from...	2022年03月	RESD
JEF	The diversification benefits and policy risks of accessing...	2022年03月	
JCF	The party school education and corporate innovation: Ev...	2022年02月	CIRD,CCEO
JBF	The role of a long-term investor-underwriter relationship...	2022年02月	CIPO
JCF	Serving multiple 'masters': Evidence from the loan decis...	2022年02月	CBRD
JCF	The effect of geographic proximity on corporate tax avoi...	2022年02月	CAID

第1页, 共51页 < >

二、CNRDS 公司特色数据库

为了满足不同学科的研究，方便不同领域学者能够迅速查找、浏览和下载所需数据，**CNRDS** 参照国际标准并结合中国资本市场数据结构，基于 **CNRDS** 平台已有数据库及强大的研发能力，设立了公司特色库，主要提供中国资本市场参与机构（包括上市公司、金融机构以及其他社会经济组织等）的特色研究数据。**CNRDS** 公司特色数据库涉及的主题丰富多元，紧跟国内外中国资本市场问题研究的热点和前沿，目前已设立六个研究系列，具体如下：

上市公司经营研究 主要包括上市公司经营过程中涉及的有关数据，例如内部控制、创新专利、投资者关系管理、企业社会责任等数据。

上市公司新闻舆情 主要包括与上市公司相关的新闻报道、股吧评论和网络搜索等信息，并采用计算机技术对这类数据进行了深度清洗、加工和判断。

上市公司文本信息 主要采用已有文献的主流文本分析方法，对与上市公司相关的公告、年报等各类文本信息进行全面地分析与展示。

资本市场人物特征 主要囊括了资本市场发展中涉及到的诸多参与人员，例如上市公司高管、基金经理等多维度的个人特征数据。

银行及金融研究 主要包括资本市场的重要金融机构及金融行为数据，例如银行机构、基金公司、券商投行、风险投资等数据。

社会经济组织研究 主要列示了资本市场主流参与者之外的另一类重要组织机构的数据，例如，慈善基金会、商会、高校等机构的数据。

为保证数据质量及契合学术研究需要，**CNRDS** 公司特色数据库从数据采集、数据整理和数据呈现均有该领域知名的国内外教授和学者深度参与，并结合 **CNRDS** 平台严格的数据管理和建库流程，可较大程度为研究人员提供高水准、高质量、专业性的学术特色数据。

CNRDS 公司特色数据库具有以下三个特点：

- 国内外知名教授深度参与，保证研究数据质量
- 涉及不同研究领域数据，满足多种研究需求
- 各特色库自成体系，充分契合学术研究习惯

上市公司经营研究

（一）中国创新专利研究数据库-CIRD

1.数据库简介

中国创新专利研究数据库（*Chinese Innovation Research Database, CIRD*）主要基于对不同上市公司、非上市公司和不同地区的专利申请和授权情况而开发的专业数据库。*CIRD* 不仅考虑了不同主体申请或授权的专利数据，其中区分了发明、实用新型和外观设计三类类别的专利，还包括了诸如独立申请还是联合申请、专利的分类号汇总和统计等重要信息。

CIRD 包括七个模块的数据：

（1）上市公司专利申请与获得模块提供中国所有上市公司专利申请与授权情况，包含上市公司专利申请情况、上市公司专利获得情况、上市公司专利申请后获得情况、上市公司专利获得年限情况四张数据表。该模块包括自 1990 年以来每年申请和获得授权的专利数量；区分为每年独立和联合申请和获得授权的专利数量；区分上市公司子公司本身、联营公司、合营公司和集团公司合计的专利申请和获得授权的数量；专利申请后 1 至 4 年授权情况及授权专利从申请到授权的平均年限等数据。

（2）上市公司专利分类号信息模块提供中国所有上市公司专利分类号的汇总和统计情况，该模块区分为申请和获得授权专利的分类号情况；包括了发明和实用新型两类重要专利的分类号；区分上市公司本身，子公司、联营公司和合营公司等不同类型公司的专利分类号；包括分类号汇总和分类号数量两个重要字段。

（3）上市公司专利分类号信息（细分）模块提供中国所有上市公司专利分类号的汇总和根据 IPC 大类细分的统计情况，该模块区分为申请和获得授权专利的分类号情况；包括了发明和实用新型两类重要专利的分类号；区分上市公司本身，子公司、联营公司和合营公司等不同类型公司的专利分类号；包括分类号汇总和分类号数量两个重要字段。

（4）上市公司研发费用模块提供中国所有上市公司自 2006 年以来每年的研发支出情况以及 2017 年以来每年研发费用明细情况。

（5）非上市公司专利申请与获得模块提供所有中国非上市公司的专利信息，该模块包括自 1985 年以来非上市公司每年申请和获得授权的专利数量；区分为每年独立和联合申请和获得授权的专利数量。

（6）科研机构专利申请与获得模块提供中国所有科研机构（包括高校）的专利信息，该模块包括自 1986 年以来科研机构每年申请和获得授权的专利数量；区分为每年独立和联合申请和获得授权的专利数量。

（7）各省市创新专利情况提供中国各省、地级市的专利信息，该模块包括了中国各省份、地级市每年的专利申请数量，并区分为发明、实用新型和外观设计的专利数量。

2.数据库特色

- 全面的公司专利信息

CIRD 对上市公司及其子公司、联营企业和合营企业各自的专利数量进行统计，精确反映公司集团的专利情况，是目前唯一能够做到如此详尽的数据库。

- 提供公司专利分类号信息

针对发明专利和实用新型专利，**CIRD** 包含了其对应的专利分类号及其统计结果，可为专利研究的深度扩展提供数据支持。

- 提供地域创新力信息

CIRD 提供了不同省市的专利申请与授权情况，可反映一个地区的创新力情况。

3. 典型数据举例

(1) 上市公司专利申请数量

股票代码	会计年度	公司类型	当年独立申请的发明数量	当年独立申请的实用新型数量	当年独立申请的外观设计数量	当年联合申请的发明数量	当年联合申请的实用新型数量	当年联合申请的外观设计数量
000030	2010	上市公司本身	2	8	6	1	13	5
000030	2011	上市公司本身	2	6	2	1	3	13
000030	2012	上市公司本身	1	2	6	1	13	6
000030	2013	上市公司本身	6	13	6	2	5	1
000032	2012	子公司	3	6	7	1	2	1
000063	2005	上市公司本身	1006	137	26	4	0	2

(2) 各省市专利申请情况

省份名称	地市名称	会计年度	当年申请的发明数量	当年申请的实用新型数量	当年申请的外观设计数量
安徽省	安庆市	2011	165	421	115
安徽省	安庆市	2012	239	799	221
安徽省	安庆市	2013	1105	1471	234
安徽省	安庆市	2014	3294	2476	305

4. 数据库内容

数据模块	数据表	数据字段
上市公司专利申请与获得	上市公司专利申请情况	股票代码、会计年度、公司类型、申请时间、当年独立申请的发明数量、当年独立申请的实用新型数量、当年独立申请的外观设计数量、当年联合申请的发明数量、当年联合申请的实用新型数量、当年联合申请的外观设计数量
	上市公司专利获得情况	股票代码、会计年度、公司类型、获得时间、当年独立获得的发明数量、当年独立获得的实用新型数量、当年独立获得的外观设计数量、当年联合获得的发明数量、当年联合获得的实用新型数量、当年联合获得的外观设计数量
	上市公司专利申请后获得情况	股票代码、会计年度、更新日期、专利申请个数、专利申请_获得个数、专利申请_当年获得个数、专利申请_第二年获得个数、专利申请_第三年获得个数、专利申请_第四年获得个数、发明申请个数、发明申请_获得个数、发明申请_当年获得个数、发明申请_第二年获得个数、发明

数据模块	数据表	数据字段
		申请_第三年获得个数、发明专利_第四年获得个数、实用新型申请个数、实用新型申请_获得个数、实用新型申请_当年获得个数、实用新型申请_第二年获得个数、实用新型申请_第三年获得个数、实用新型申请_第四年获得个数、外观设计申请个数、外观设计申请_获得个数、外观设计申请_当年获得个数、外观设计申请_第二年获得个数、外观设计申请_第三年获得个数、外观设计申请_第四年获得个数
	上市公司专利获得年限情况	股票代码、会计年度、专利获得、专利获得_平均年限、发明专利获得、发明专利获得_平均年限、发明专利获得_公开平均年限、实用新型获得、实用新型获得_平均年限、外观设计获得、外观设计获得_平均年限
上市公司专利分类号信息	上市公司发明专利申请专利分类号	股票代码、会计年度、公司类型、申请时间、发明专利 A 类、发明专利 A 类数量、发明专利 B 类、发明专利 B 类数量、发明专利 C 类、发明专利 C 类数量、发明专利 D 类、发明专利 D 类数量、发明专利 E 类、发明专利 E 类数量、发明专利 F 类、发明专利 F 类数量、发明专利 G 类、发明专利 G 类数量、发明专利 H 类、发明专利 H 类数量
	上市公司实用新型申请专利分类号	股票代码、会计年度、公司类型、申请时间、实用新型申请 A 类、实用新型申请 A 类数量、实用新型申请 B 类、实用新型申请 B 类数量、实用新型申请 C 类、实用新型申请 C 类数量、实用新型申请 D 类、实用新型申请 D 类数量、实用新型申请 E 类、实用新型申请 E 类数量、实用新型申请 F 类、实用新型申请 F 类数量、实用新型申请 G 类、实用新型申请 G 类数量、实用新型申请 H 类、实用新型申请 H 类数量
	上市公司发明专利获得专利分类号	股票代码、会计年度、公司类型、获得时间、发明获得 A 类、发明获得 A 类数量、发明获得 B 类、发明获得 B 类数量、发明获得 C 类、发明获得 C 类数量、发明获得 D 类、发明获得 D 类数量、发明获得 E 类、发明获得 E 类数量、发明获得 F 类、发明获得 F 类数量、发明获得 G 类、发明获得 G 类数量、发明获得 H 类、发明获得 H 类数量
	上市公司实用新型获得专利分类号	股票代码、会计年度、公司类型、获得时间、实用新型获得 A 类、实用新型获得 A 类数量、实用新型获得 B 类、实用新型获得 B 类数量、实用新型获得 C 类、实用新型获得 C 类数量、实用新型获得 D 类、实用新型获得 D 类数量、实用新型获得 E 类、实用新型获得 E 类数量、实用新型获得 F 类、实用新型获得 F 类数量、实用新型获得 G 类、实用新型获得 G 类数量、实用新型获得 H 类、实用新型获得 H 类数量
上市公司专利分类号信息(细分)	上市公司发明专利申请专利细分分类号	股票代码、会计年度、公司类型、申请时间、发明专利 A01、发明专利 A01 数量、发明专利 A21、发明专利 A21 数量、发明专利 A22、发明专利 A22 数量、发明专利 A23、发明专利 A23 数量……
	上市公司实用新型申请专利细分分类号	股票代码、会计年度、公司类型、申请时间、实用新型申请 A01、实用新型申请 A01 数量、实用新型申请 A21、实用新型申请 A21 数量、实用新型申请 A22、实用新型申请 A22 数量……
	上市公司发明专利获得专利细分分类号	股票代码、会计年度、公司类型、获得时间、发明获得 A01、发明获得 A01 数量、发明获得 A21、发明获得 A21 数量、发明获得 A22、发明获得 A22 数量、发明获得 A23、发明获得 A23 数量……
	上市公司实用新型获得专利细分分类号	股票代码、会计年度、公司类型、获得时间、实用新型获得 A01、实用新型获得 A01 数量、实用新型获得 A21、实用新型获得 A21 数量、实用新型获得 A22、实用新型获得 A22 数量……

数据模块	数据表	数据字段
上市公司研发费用	上市公司研发支出	股票代码、会计年度、研发人员数量、研发人员数量占比、研发支出占营业收入比例、研发支出、资本化研发支出、资本化研发支出占研发支出的比例、研发支出占当期净资产比例、费用化研发支出、资本化研发支出占当期净利润的比例
	上市公司研发费用明细	股票代码、会计年度、研发费用明细科目、研发费用投入额
非上市公司专利申请与获得	非上市公司专利申请情况	公司名称、会计年度、当年独立申请的发明数量、当年独立申请的实用新型数量、当年独立申请的外观设计数量、当年联合申请的发明数量、当年联合申请的实用新型数量、当年联合申请的外观设计数量
	非上市公司专利获得情况	公司名称、会计年度、当年独立获得的发明数量、当年独立获得的实用新型数量、当年独立获得的外观设计数量、当年联合获得的发明数量、当年联合获得的实用新型数量、当年联合获得的外观设计数量
科研机构专利申请与获得	科研机构专利申请情况	单位名称、会计年度、当年独立申请的发明数量、当年独立申请的实用新型数量、当年独立申请的外观设计数量、当年联合申请的发明数量、当年联合申请的实用新型数量、当年联合申请的外观设计数量
	科研机构专利获得情况	单位名称、会计年度、当年独立获得的发明数量、当年独立获得的实用新型数量、当年独立获得的外观设计数量、当年联合获得的发明数量、当年联合获得的实用新型数量、当年联合获得的外观设计数量
各省市创新专利情况	各省市专利申请情况	省份名称、地市名称、会计年度、当年申请的发明数量、当年申请的实用新型数量、当年申请的外观设计数量
	各省市专利获得情况	省份名称、地市名称、会计年度、当年获得的发明数量、当年获得的实用新型数量、当年获得的外观设计数量

5.数据范围

数据模块	数据起始时间
上市公司专利申请与获得	1990 年
上市公司专利分类号信息	1990 年
上市公司专利分类号信息（细分）	1990 年
上市公司研发费用	2006 年
非上市公司专利申请与获得	1985 年
科研机构专利申请与获得	1985 年
各省市创新专利情况	1985 年

(二) 中国审计研究数据库-CARD

1. 数据库简介

中国审计研究数据库 (*Chinese Audit Research Database, CARD*) 基于中注协历年发布的会计师事务所综合评价前百家信息、会计师事务所和审计师相关信息以及上市公司审计单位、审计意见和审计费用等信息而建立的专业数据库。*CARD* 共有 6 张表, 其中, 会计师事务所信息表涵盖了 2003 年以来上市公司审计事务所的业务收入、处罚情况、注册会计师人数等情况; 注册会计师信息表涵盖了 2014 年以来, 上市公司签字注册会计师的个人特征信息, 在该表制作过程中, 我们严格对审计师的同名同姓情况进行了处理, 确保审计师个人信息的准确性; 审计师能力与专长模块包含审计师能力和审计师专长两张表, 审计师能力表主要包含会计师事务所和注册会计师的地区与行业审计能力, 审计师专长表则包含会计师事务所和注册会计师的地区与行业专长。其中, *CARD* 对行业专长的计算参考了国内外权威文献, 分别采用行业市场份额法和行业组合份额法, 从客户资产总额、营业收入和审计费用三个方面对会计师事务所和注册会计师的行业专长进行了展现。而对地区专长的计算也参考了行业专长的计算方法。最后, *CARD* 还包括上市公司审计意见和审计费用两张表格。

2. 数据库特色

- 数据来源权威

CARD 的所有数据均来自官方披露, 包括了诸多详实信息。

- 计算了审计能力与审计专长

CARD 从多个维度计算了会计师事务所和注册会计师的地区专长和行业专长等重要指标。

3. 典型数据举例

注册会计师信息

审计师姓名	审计公司股票代码	会计年度	所在会计师事务所	全科合格时间	学校类型	性别	学位
丁亚轩	002504	2018	中兴财光华会计师事务所(特殊普通合伙)	1995	211	男	本科
丁亚轩	002504	2019	中兴财光华会计师事务所(特殊普通合伙)	1995	211	男	本科
丁兆栋	300183	2018	中兴华会计师事务所(特殊普通合伙)	1997	其他	男	大专
丁兆栋	002073	2019	中兴华会计师事务所(特殊普通合伙)	1997	其他	男	大专

4. 数据库内容

数据模块	数据表	数据字段
会计师事务所信息	会计师事务所信息	会计师事务所名称、股票代码、会计年度、事务所排名、是否国际四大、是否前十大、事务所业务收入、是否受到处罚、注册会计师人数、人均业务收入、师均业务收入、当年审计的上市公司数量
注册会计师信息	注册会计师信息	注册会计师姓名、股票代码、会计年度、会计师事务所名称、共同签字注册会计师、签字排名、批准注册时间、全科合格时间、学校类型、是否会计相关专业、性别、学位、当年审计的公司数量、当年审计的行业类型、当年审计的地区、是否上市公司独董、任职独董的公司代码

数据模块	数据表	数据字段
审计师能力与专长	行业专长_客户总资产维度	会计师事务所名称、股票代码、会计年度、注册会计师姓名、共同签字注册会计师、签字排名、事务所行业专长 1、事务所行业专长 1SQRT、事务所行业专长 2、事务所行业专长 2SQRT、注册会计师行业专长 1、注册会计师行业专长 1SQRT、注册会计师行业专长 2、注册会计师行业专长 2SQRT
	行业专长_客户营业收入维度	会计师事务所名称、股票代码、会计年度、注册会计师姓名、共同签字注册会计师、签字排名、事务所行业专长 3、事务所行业专长 3SQRT、事务所行业专长 4、事务所行业专长 4SQRT、注册会计师行业专长 3、注册会计师行业专长 3SQRT、注册会计师行业专长 4、注册会计师行业专长 4SQRT
	行业专长_客户审计费用维度	会计师事务所名称、股票代码、会计年度、注册会计师姓名、共同签字注册会计师、签字排名、事务所行业专长 5、事务所行业专长 5SQRT、事务所行业专长 6、事务所行业专长 6SQRT、注册会计师行业专长 5、注册会计师行业专长 5SQRT、注册会计师行业专长 6、注册会计师行业专长 6SQRT
	行业专长_客户数量维度	会计师事务所名称、股票代码、会计年度、注册会计师姓名、共同签字注册会计师、签字排名、上市公司所在行业上市公司数、事务所当年审计的该行业上市公司数、注册会计师当年审计的该行业上市公司数、事务所行业审计能力、注册会计师行业审计能力
	地区专长_客户总资产维度	会计师事务所名称、股票代码、会计年度、注册会计师姓名、共同签字注册会计师、签字排名、事务所地区专长 1（基于总资产）、事务所地区专长 1SQRT、事务所地区专长 2、事务所地区专长 2SQRT、注册会计师地区专长 1、注册会计师地区专长 1SQRT、注册会计师地区专长 2、注册会计师地区专长 2SQRT
	地区专长_客户营业收入维度	会计师事务所名称、股票代码、会计年度、注册会计师姓名、共同签字注册会计师、签字排名、事务所地区专长 3、事务所地区专长 3SQRT、事务所地区专长 4、事务所地区专长 4SQRT、注册会计师地区专长 3、注册会计师地区专长 3SQRT、注册会计师地区专长 4、注册会计师地区专长 4SQRT
	地区专长_客户审计费用维度	会计师事务所名称、股票代码、会计年度、注册会计师姓名、共同签字注册会计师、签字排名、事务所地区专长 5、事务所地区专长 5SQRT、事务所地区专长 6、事务所地区专长 6SQRT、注册会计师地区专长 5、注册会计师地区专长 5SQRT、注册会计师地区专长 6、注册会计师地区专长 6SQRT
	地区专长_客户数量维度	会计师事务所名称、股票代码、会计年度、注册会计师姓名、共同签字注册会计师、签字排名、上市公司所在地区上市公司数、事务所当年审计的该地区上市公司数、注册会计师当年审计的该地区上市公司数、事务所地区审计能力、注册会计师地区审计能力
审计意见与费用	审计单位与审计意见	股票代码、会计年度、会计师事务所、审计师、审计意见、审计意见说明
	审计单位与审计费用	股票代码、会计年度、会计师事务所、审计费用

5.数据范围

数据模块	数据起始时间
会计师事务所信息	2003 年
注册会计师信息	2014 年
审计师能力与专长	2014 年
审计意见与费用	1990 年

（三）中国券商与分析师数据库-CBAS

1.数据库简介

中国券商与分析师数据库 (*Chinese Broker and Financial Analysts System, CBAS*) 是上海经禾信息技术有限公司与海外知名高校的金融与会计学学者合作, 参照 *IBES* 等国际分析师数据库联合开发的有关中国券商与分析师研究的专业学术数据库。*CBAS* 包括盈利预测、荐股评级、调研数据、券商数据、分析师报告文本分析 5 大模块。

盈利预测模块和荐股评级模块 此两大模块从 40 万份分析师报告中提取了盈利预测、荐股评级、目标股价等信息, 是目前为止较全面详实的分析师预测和荐股数据库。

调研数据模块 此模块提供调研日期、参与人员、参与单位、接待人员等基本信息。

券商数据模块 此模块从券商年报中提取了券商财务信息、大股东及其持股比例信息、大股东性质信息、董事、监事、高管个人信息等独特数据。

分析师报告文本分析模块 此模块是大数据技术在学术数据库上的应用, 我们对每份分析师报告提供了用 *LDA* 主题模型提取的概率最高的 5 个主题, 以及文本句子数量和财务信息句子数量, 是目前为止唯一用大数据技术对分析师报告信息进行提炼的数据库。

2.数据库特色

- 分析师及券商唯一编码

CBAS 给数据库中 100 多家券商以及 6000 多名分析师分配了独特编码, 是目前为止唯一给券商和分析师独特编码的数据库, 有利于研究者进行数据匹配和深度挖掘。

- 覆盖全面的调研信息

CBAS 包括 2007 年以来的分析师调研信息, 既包括来自年报的分析师调研情况, 也包括来自调研报告的调研情况。

- 独有的券商相关信息

CBAS 包括了国内主要券商的大股东、高管、财务数据信息等。

3.典型数据举例

券商股东构成

券商名称	券商代码	会计报告日期	大股东名称	大股东持股比例	大股东性质	最终控制人
爱建证券有限责任公司	B00001	2014-12-31	上海陆家嘴金融发展有限公司	0.51136	地方国有	地方国资局
安信证券股份有限公司	B00002	2014-12-31	中国证券投资者保护基金有限责任公司	0.2588	中央国有	国资委
渤海证券股份有限公司	B00006	2014-12-31	天津市泰达国际控股(集团)有限公司	0.536695	地方国有	地方国资局

4.数据库内容

数据模块	数据表	数据字段
盈利预测	预测情况	股票代码、报告日期、盈利预测结束日期、分析师姓名、分析师代码、是否明星分析师、券商名称、券商代码、EPS 预测、P/E 比率预测、净利润预测、主营业务收入预测、实际盈余公告日
	实际情况	股票代码、盈利预测结束日期、真实 EPS、真实净利润、真实营业利润
荐股评级	评级	股票代码、报告日期、分析师姓名、分析师代码、券商名称、券商代码、原始报告中的荐股评级、标准化荐股意见、标准化荐股意见代码、是否是明星分析师
调研数据	访客	股票代码、公司简称、公开时间、调研时间、参与单位、参与单位类别、参与人员
	接待人员	股票代码、公司简称、调研时间、公开时间、调研接待人员 n 姓名、调研接待人员 n 职务
	调研内容	股票代码、公司简称、调研时间、调研报告内容
券商数据	股东构成	券商名称、券商代码、会计报告日期、大股东名称、大股东占股比例、大股东性质、最终控制人
	董监高	券商名称、券商代码、董监高姓名、职位类型、性别、出生年份、国籍、学历/学位、现任或者曾任、工作经历起止日期、工作单位、工作职位、是否全职、会计报告日期、个人简介
	IPO 承销排名	券商名称、券商代码、会计年度、当年承销 IPO 总募集资金金额、当年承销 IPO 总募集资金排名、当年承销 IPO 总募集资金占当年 IPO 市场总募集资金比例、当年 IPO 承销费用、当年 IPO 承销费用排名、当年 IPO 承销数量、当年联席保荐 IPO 数量、当年 IPO 承销数量排名
	财务报表	券商名称、券商代码、合并会计报表或母公司会计报表、会计报告日期、营业收入、营业利润、营业外收入、营业外支出、利润总额、所得税、净利润、少数股东损益、归属于母公司所有者的净利润、基本每股收益 EPS、稀释每股收益 EPS、货币资金、客户资金存款、交易性金融资产……
	财务比率	券商名称、券商代码、合并会计报表或母公司会计报表、会计报告日期、证券公司净资本、营业收入、营业收入增长率、手续费及佣金净收入、手续费及佣金净收入在营业收入中的占比、利息净收入……
分析师报告文本分析	分析师报告文本分析	股票代码、报告日期、分析师姓名、分析师代码、券商名称、券商代码、分析师报告主题、句子数目、财务信息句子数目

5.数据范围

数据模块	数据起始时间
盈利预测	2005 年
荐股评级	2005 年
调研数据	2007 年
券商数据	2006 年
分析师报告文本分析	2006 年

(四) 环境、社会和治理数据库-CESG

1.数据库简介

环境、社会和治理(ESG)近年来受到各方面的关注, ESG 要求企业必须要超越把利润作为企业唯一目标的传统理念, 强调在生产过程中对环境、对社会的贡献。企业落实社会责任, 实现企业经济责任和社会责任的平衡, 会提高企业的竞争力, 为企业树立良好的声誉和形象, 增强投资者信心。环境、社会和治理数据库 (Environmental、Social and Governance Database of Listed Company, **CESG**) 旨在为科研人员提供中国上市公司环境、社会和治理领域相关数据。**CESG** 的设计参考了国内外知名 ESG 数据库的设计思路, 从六个方面用 58 个细分指标对上市公司的环境、社会和治理情况进行衡量。

2.数据库特色

- 设计科学

CESG 参考了国外知名 ESG 数据库的设计标准, 并且结合了我国企业的实际情况, 设计了 ESG 的六个方面: “慈善、志愿者活动以及社会争议”、“公司治理”、“多样化”、“雇员关系”、“环境”以及“产品”, 共有 58 个字段。每个方面都从“优势”和“关注”两个角度设计字段。

- 直观

CESG 的大部分字段为哑变量, 非常直观地表达了上市公司 ESG 不同维度的特征, 使用者可以很快地构建全面或者某一方面的 ESG 指数。

3.典型数据举例

(1) 慈善、志愿者活动以及社会争议优势

股票代码	公司简称	会计年度	捐赠总额	支持教育	支持慈善	志愿者活动	带动就业	促进当地经济	其他优势
000001	平安银行	2012	92 万元	1	1	0	0	0	0
000002	万科 A	2012	1720 万元	1	1	1	1	1	1
000009	中国宝安	2012	110.256 万元	1	1	0	0	0	1

(2) 雇员关系关注

股票代码	公司简称	会计年度	雇员安全纠纷	裁员	其他关注
000527	美的电器	2012	0	0	1
000528	柳工	2012	0	0	1
000800	一汽轿车	2012	1	0	0

4.数据库内容

数据模块	数据表	数据字段
慈善、志愿者活动以及社会争议	优势	股票代码、公司简称、会计年度、捐赠总额、支持教育、支持慈善、志愿者活动、国际援助、带动就业、促进当地经济、其他优势
	关注	股票代码、公司简称、会计年度、融资纠纷
公司治理	优势	股票代码、公司简称、会计年度、CSR 报告全面性、CSR 报告页数、CSR 专栏、CSR 领导机构、CSR 愿景、CSR 培训、每股社会贡献值、可靠性保证、其他优势
	关注	股票代码、公司简称、会计年度、会计违规

多样化	优势	股票代码、公司简称、会计年度、党员、女性高管、女性董事席位、创新人力资源项目、其他优势
	关注	股票代码、公司简称、会计年度、没有女性高管
雇员关系	优势	股票代码、公司简称、会计年度、员工参股、员工福利、安全管理体系、安全生产培训、职业安全认证、职业培训、员工沟通渠道、其他优势
	关注	股票代码、公司简称、会计年度、雇员安全纠纷、裁员
环境	优势	股票代码、公司简称、会计年度、环境有益的产品、减少三废的措施、循环经济、节约能源、绿色办公、环境认证、环境表彰、其他优势
	关注	股票代码、公司简称、会计年度、环境处罚、污染物排放
产品	优势	股票代码、公司简称、会计年度、质量体系、售后服务、客户满意度调查、质量荣誉、专利数目、研发支出、研发人员比例、技术人员比例、反腐败措施、战略共享、诚信经营理念、其他优势
	关注	股票代码、公司简称、会计年度、产品纠纷

5.数据范围

数据模块	数据起始时间
全部模块	2006年

(五) 中国家族企业研究数据库-CFFD

1. 数据库简介

中国家族企业研究数据库 (*Chinese Family Firm Research Database, CFFD*)，是对中国上市公司家族企业进行调查和研究而建立的专业数据库。不同于以往将实际控制人可以追溯到个人的民营企业确定为家族企业，*CFFD* 所确定的家族企业包括三个标准：（1）实际控制人是一个自然人或某个以血缘、姻缘为连结的家族；（2）该自然人或家族直接或间接是上市公司第一大股东；（3）至少两个或以上家族成员在该上市公司或者关联公司持股或任职。*CFFD* 包括四张数据表：家族企业基本信息、*IPO* 时家族企业持股情况、家族企业控制链以及家族企业继承人信息。其中，家族企业基本信息包括实际控制人、成立和上市时间在内的一些企业基本特征信息；*IPO* 时家族企业持股情况包括企业在上市前后实际控制人的持股变化，家族其他相关人员的名称、持股和职务信息等；家族企业控制链则包括一个完整时间序列内家族企业实际控制人和相关人员的持股层级和持股比例情况；最后，家族企业继承人信息包括家族企业的创始人和继承人信息。

2. 数据库特色

- 精确的家族企业定义

相比以往的家族企业定义，*CFFD* 对家族企业的定义更为准确，真正可以追溯到哪个家族对企业形成实际控制。

- 详细的家族企业成员持股信息

对 *IPO* 前后家族企业实际控制人、家族企业成员的持股和亲缘关系信息进行了详细列示，有助于发现家族企业的内部关系和持股关系。

- 时间序列上的控制链信息

根据年报和公告信息，构建了家族企业实际控制人在时间序列上的持股变化，动态跟踪家族企业所有权的变化情况。

- 家族继承人信息

丰富的家族企业继承人个人特征信息，为研究家族企业传承问题提供重要数据支持。

3. 典型数据举例

IPO 时家族企业持股情况

股票代码	股东姓名	招股前持股比例 (%)	招股后持股比例 (%)	在公司的职位	关系说明
002010	徐冠巨	32.78	24.58	董事长	徐传化系徐冠巨、徐观宝之父
002010	徐观宝	18.89	14.17	董事	徐传化系徐冠巨、徐观宝之父
002010	徐传化	18.33	13.75	创始人	徐传化系徐冠巨、徐观宝之父
002194	孟庆南	39.83	29.8	董事长	孟庆南与王丽丽系夫妻关系
002194	王丽丽	39.83	29.8	副董事长、总经理	孟庆南与王丽丽系夫妻关系
002194	孟凡博	12	8.98	董事	孟凡博系孟庆南和王丽丽之子
002194	王凯	4.88	3.65	董事	王凯系王丽丽之弟
002194	王凌	0.06	0.04	副总经理	王凌系王丽丽之妹
002194	王莉萍	0.05	0.04	中层管理人员	王莉萍系王丽丽之妹

4.数据库内容

数据模块	数据表	数据字段
家族企业基本信息	家族企业基本信息	股票代码、公司中文名称、会计年度、实际控制人、实际控制人类别、是否家族企业、法定代表人、成立日期、首发上市日期、同公司 B 股代码、同公司 H 股代码、注册地址、办公地址、所在省份、所在地级市、所在县级市、经度、纬度、董事会秘书、证券事务代表、行业名称 A、行业代码 A、行业名称 B、行业代码 B
IPO 时家族企业持股情况	IPO 时家族企业持股情况	股票代码、公司名称、序号、实际控制人、股东名称、招股前持股比例 (%)、招股后持股比例 (%)、法人股东的大股东、法人股东的大股东持股比例 (%)、法人股东的大股东职位、在上市公司的职位、关系说明
家族企业控制链	家族企业控制链	股票代码、公司名称、统计日期、实际控制人、实际控制人类别、亲属关系备注、持股层级数、序号、一级股东、一级股东持股比例 (%)、二级股东、二级股东持股比例 (%)、三级股东、三级股东持股比例 (%)、四级股东、四级级股东持股比例 (%)、五级股东、五级股东持股比例 (%)、六级股东、六级股东持股比例 (%)、七级股东、七级股东持股比例 (%)、八级股东、八级股东持股比例 (%)、九级股东、九级股东持股比例 (%)、十级股东、十级股东持股比例 (%)、十一级股东、十一级股东持股比例 (%)、十二级股东、十二级股东持股比例 (%)、十三级股东、十三级股东持股比例 (%)
家族企业继承人信息	家族企业继承人信息	股票代码、公司名称、控股年份、创始人姓名、继承人姓名、与创始人关系、职位、任职开始日期、任职结束日期、性别、出生日期、国籍、籍贯/出生地、学历、职称与资格证书、个人简介

5.数据范围

数据模块	数据起始时间
家族企业基本信息	2000 年
IPO 时家族企业持股情况	截面数据
家族企业控制链	1999 年
家族企业继承人信息	截面数据

(六) 公司并购重组数据库-CMAD

1. 数据库简介

公司并购重组数据库 (*Mergers and Acquisitions Database, CMAD*) 是基于上市公司及其子公司等相关方的兼并、收购和重组事件整理而成的专业数据库。*CMAD* 收录了公司并购事件的基本信息, 并购相关方 (包括买方、买方和标的方) 的基本信息、公司股东和管理层等, 详细和完整地展示了并购过程中的重要信息。该数据库总共分为两个大模块, 并购事件信息模块主要包括并购时间、并购方式、并购是否成功、并购金额等字段; 并购涉及方信息主要包括涉及方基本信息、股东信息和高管等信息。

2. 数据库特色

- 数据量巨大、字段丰富

CMAD 囊括了近几十年来的十万多例并购事件, 采用简约风格设计了大量重要字段。

- 极为详细的并购涉及方信息

CMAD 对并购涉及方的信息进行了深度挖掘, 为学者进行此类研究提供了重要数据支持。

3. 典型数据举例

并购事件基本信息

股票代码	买方名称	标的方名称	卖方名称	首次公告日	重组类型
000001	美国新桥投资集团公司	平安银行股份有限公司	深圳国际信托投资有限公司	2004-05-31	协议转让
000002	上海万科浦东置业有限公司	上海南都置地有限公司	上海中桥基建(集团)股份有限公司	2005-03-05	协议转让
000005	深圳市恒裕(实业)集团有限公司	深圳市中环星苑房地产开发有限公司	深圳世纪星源物业发展有限公司	2017-12-28	协议转让

4. 数据库内容

数据模块	数据表	数据字段
并购事件信息	并购事件基本信息	事件 ID、股票代码、公司名称、买方名称、买方代码、买方与披露方关系、标的方名称、标的方代码、标的方与披露方关系、卖方名称、卖方代码、卖方与披露方关系、最新公告日、首次公告日、实施进度、重组类型、支付方式、交易金额、币种、标的类型、股权转让比例 (%)、控制权是否变更
并购涉及方信息	涉及方基本信息	事件 ID、股票代码、相关方、相关方身份、注册资本、经营状态、成立日期、公司类型、所属行业、曾用名、注册地址、所在省份、所在市
	涉及方股东信息	事件 ID、股票代码、上市公司交易地位、相关方、相关方身份、序号、股东、持股比例 (%)、股东类型
	涉及方高管信息	事件 ID、股票代码、上市公司交易地位、相关方、相关方身份、序号、姓名、职务

5. 数据范围

数据模块	数据起始时间
全部模块	1994 年

（七）监管问询数据库-CRID

1.数据库简介

监管问询数据库（*Regulatory Inquiries Database, CRID*）是基于交易所等监管机构向上市公司问询有关其经营或治理相关问题的文件整理而成的专业数据库。*CRID* 的数据主要来源于上交所和深交所，并对问询文件内容和公司回复内容进行梳理和展示。此外，根据已有研究，本数据库还提取了与监管问询以及公司回复相关的日期时间等重要字段。

2.数据库特色

- 字段丰富

CRID 除了包括直接来源于交易所的字段数据，还对 *PDF* 文本内容以及相关日期等进行了提取和展示。

3.典型数据举例

监管问询

股票代码	函件类别	问询标题	发函日期	限期回复日期	公司回复日期
300043	非许可类重组问询函	关于对浙江互动娱乐股份有限公司的重组问询函	2014-12-9	2014-12-11	2014-12-11
002324	非许可类重组问询函	关于对浙江普利特股份有限公司的重组问询函	2015-1-9	2015-1-13	2015-1-13
002354	非许可类重组问询函	关于对浙江科冕木业股份有限公司的重组问询函	2015-1-12	2015-1-15	2015-1-19
002072	非许可类重组问询函	关于对浙江凯瑞德股份有限公司的重组问询函	2015-2-5	2015-2-10	2015-2-10
000035	非许可类重组问询函	关于对浙江中国天楹股份有限公司的重组问询函	2015-2-3	2015-2-5	2015-2-10
002164	年报问询函	关于对浙江宁波东力股份有限公司的重组问询函	2015-3-9	2015-3-11	2015-3-11
002160	年报问询函	关于对浙江常铝股份股份有限公司的重组问询函	2015-3-13	2015-3-20	2015-3-19

4.数据库内容

数据模块	数据表	数据字段
监管问询	监管问询	股票代码、公司简称、函件类别、问询标题、问询内容、公司回复内容、发函日期、限期回复日期、公司回复日期

5.数据范围

数据模块	数据起始时间
全部模块	2014 年

（八）中国上市公司内部控制研究数据库-ICRD

1.数据库简介

中国上市公司内部控制研究数据库 (*Internal Control Research Database of Chinese Companies, ICRD*) 是以研究为目的而设计的专业学术研究数据库, 旨在为国内外从事上市公司内部控制研究的学者和机构从业人员提供精确、完整的学术研究数据。本数据库涵盖中国上市公司有关内部控制方面的重要数据信息, 根据 *COSO* 内部控制框架的基本要素, 不仅包括上市公司内部控制报告和年报中有关内部控制的内容, 而且综合了其他与内部控制相关的内容, 完整地体现了企业内部控制的总体状况。

2.数据库特色

- 基于 *COSO* 内控

ICRD 的基础模块设计基于 *COSO* 内控对公司内部控制的整体要求和评价, 较为准确地反映了企业内部控制各个方面, 有助于从不同角度研究企业内部控制情况。

- 全面完善

ICRD 还考虑了当前我国企业内部控制的各个方面, 全面、完善地考察了企业内部控制机制。

3.重要数据举例

内部风险披露

股票代码	会计年度	是否披露经营管理风险	是否披露技术创新风险	是否披露财务风险	是否披露人力资源风险	是否披露安全和环境风险
300301	2015	1	0	0	1	0
300351	2015	1	0	1	0	0
002572	2015	1	1	0	0	0
002168	2015	1	0	0	1	0
300298	2015	1	1	0	1	0
002349	2015	1	0	0	1	1
002522	2015	1	0	1	0	0

4.数据库内容

数据模块	数据表	数据字段
内部控制环境	董监高持股	股票代码、会计年度、管理层持股比例、董事会持股比例、监事会持股比例
	独立董事比例	股票代码、会计年度、董事总人数、独立董事人数、独立董事比例
	监事执行情况	股票代码、会计年度、监事会人数、职工监事人数
	两职兼任情况	股票代码、会计年度、董事长与总经理兼任情况
	企业社会责任	股票代码、会计年度、是否发布社会责任报告、是否有慈善捐赠、公司是否关注环境、安全和产品质量
	负面事件	股票代码、会计年度、是否卷入诉讼纠纷、公司是否受到监管部门惩罚、高管是否受到监管部门处罚、公司是否被媒体披露重大负面消息

数据模块	数据表	数据字段
风险管理	内部风险披露	股票代码、会计年度、是否披露经营管理风险、是否披露技术创新风险、是否披露财务风险、是否披露人力资源风险、是否披露安全和环境风险、是否披露其他内部风险、前三年是否有显著的损失
	外部风险披露	股票代码、会计年度、是否披露市场竞争风险、是否披露产品价格风险、是否披露原材料风险、是否披露行业风险、是否披露任何经济环境风险、是否披露政策风险、是否披露其他外部风险
	风险应对	股票代码、会计年度、风险数量、披露的风险类型、披露的风险内容、风险应对策略
控制活动	经营与业绩控制	股票代码、会计年度、存货是否有损失、资产减值损失是否增加、是否有业绩分析报告
	事故安全控制	股票代码、会计年度、是否有严重的生产事故、是否有严重的污染事故、是否有严重的资产安全事故
信息与沟通	信息交流	股票代码、会计年度、董事会秘书是否“金牌秘书”、是否有机制管理与投资者的关系、公司网站上是否有联系投资者的链接、是否有其他与投资者交流的平台
	信息及时性	股票代码、会计年度、定期报告是否公布于预定日期
	反舞弊情况	股票代码、会计年度、公司是否有重述、修正或补充的财务报表、公司是否有任何信息泄漏
监督活动	内部控制披露	股票代码、会计年度、是否有内部控制评价报告、内部控制是否有效、是否有提高内部控制的计划
	内部控制缺陷	股票代码、会计年度、是否有缺陷识别的标准、内部控制是否存在缺陷、是否有改正内部缺陷的计划
	内部审计情况	股票代码、会计年度、是否披露内控审计报告、审计意见类型、审计意见、非标准意见说明、会计师事务所、注册会计师名称

5.数据范围

数据模块	数据起始时间
全部模块	2010年

（九）上市公司参控股公司数据库-CPCD

1.数据库简介

上市公司参控股公司数据库（*Controlled & Participated Company Database of Listed Company, CPCD*）基于上市公司年报、公告等来源，采集和整理上市公司的子公司、合营企业、联营企业等数据而开发建立的专业数据库。目前，上市公司通常在其财务报告中披露参控股公司信息，但是所披露的信息较为简单，仅可以获得列表信息，为了丰富该领域的研究数据，我司还从各个渠道对参控股公司的其他重要信息进行补充和完善，例如从国家企业信用信息公示系统获取了相关工商信息等。该数据库主要分为六个模块：参控股公司列表、参控股公司基本信息、参控股公司股东信息、参控股公司管理人员、参控股公司法律诉讼、参控股公司对外投资。其中参控股公司列表包含了时间序列上所有上市公司的参控股公司名称、持股比例等重要字段。需要说明的是，由于有关参控股公司的其他信息来自于其他渠道，因此诸如参控股公司基本信息等表中的参控股公司可能无法完全覆盖参控股公司列表中的全部公司。

2.数据库特色

- 唯一提供上市公司子公司、合营公司和联营公司的数据库

CPCD 主要基于年报数据，提供了时间序列上完整和丰富的上市公司参控股公司信息，数据量近70万条，是目前市场上唯一提供此类数据批量下载的数据库。

- 提供了上市公司参控股公司的多维度信息

CPCD 除了提供上市公司参控股列表信息以及持股比例等基本信息，还提供了各个参控股公司的重要信息，例如参控股公司的工商信息、股东、高管和法律裁决信息等。

3.典型数据举例

参控股公司列表

股票代码	会计年度	参控股公司名称	参控关系	参控比例	被参控股公司注册资本
000001	1996	深圳市元盛实业有限公司	子公司	100.00%	2101 万元
000001	1997	深圳市元盛实业有限公司	子公司	100.00%	2010 万元
000001	1998	上海银涛高尔夫俱乐部有限公司	子公司	75.00%	2000 万元
000001	1998	深圳市元盛实业有限公司	子公司	100.00%	2010 万元
000001	1998	上海元盛房地产有限公司	子公司	100.00%	1500 万元
000001	1998	大连元盛房屋开发有限公司	子公司	100.00%	1000 万元

4.数据库内容

数据模块	数据表	数据字段
参控股公司列表	参控股公司列表	股票代码、会计年度、序号、参控股公司名称、参控关系、参控比例、被参控股公司注册资本、货币单位
参控股公司基本信息	参控股公司基本信息	股票代码、参控股公司名称、法定代表人、注册资本、货币单位、经营状态、成立日期、登记机关、所属地区、所属行业、企业地址、经营范围
参控股公司股东信息	参控股公司股东信息	股票代码、参控股公司名称、参控股公司股东、持股比例、股东类型

数据模块	数据表	数据字段
参控股公司 管理人员	参控股公司 管理人员	股票代码、参控股公司名称、高管姓名、高管职务
参控股公司 法律诉讼	参控股公司 法律诉讼	股票代码、参控股公司名称、序号、案件名称、发布时间、案件编号、案件身份、执行法院
参控股公司 对外投资	参控股公司 对外投资	股票代码、参控股公司名称、被投资企业名称、被投资企业注册资本、货币单位、出资比例、被投资企业成立日期、被投资企业状态

5.数据范围

数据模块	数据起始时间
参控股公司列表	1990 年
参控股公司基本信息	截面数据
参控股公司股东信息	截面数据
参控股公司管理人员	截面数据
参控股公司法律诉讼	2007 年
参控股公司对外投资	截面数据

(十) 关键审计事项数据库-KAMD

1.数据库简介

关键审计事项数据库 (Key Audit Matters Database, **KAMD**) 是基于上市公司审计报告中披露的关键审计事项段整理而成的专业数据库。关键审计事项,是指注册会计师根据职业判断认为对本期财务报表审计最为重要的事项,从注册会计师与治理层沟通过的事项中选取。为了扩充报告使用者对公司价值判断的信息,同时增强报告使用者的决策相关性和有用性,2016年12月23日中国财政部发布了《中国注册会计师审计准则第1504号——在审计报告中沟通关键审计事项》,要求审计师披露在审计过程中遇到重要的、可能存在重大错误、需要特别关注的事项。**KAMD**根据上市公司年报披露关键审计事项的情况,及时搜集和整理了相关数据,便于学者获得该领域的详细研究数据。

2.数据库特色

- 数据的前沿性

目前,国际上已有多篇论文研究关键审计事项,国内对于该领域尚处于起步阶段,**KAMD**的推出紧跟了研究前沿和热点。

3.典型数据举例

关键审计事项

股票代码	会计年度	关键审计事项类型	关键审计事项内容	应对方法
000002	2016	存货的可变现净值的评估	2016年12月31日,贵公司及其子公司(以下合称“贵集团”)已完工开发产品、在建开发产品及拟开发产品(以下统称“存货”)的账面价值合计金额重大。该等存货按照成本与可变……	与评价存货的可变现净值相关的审计程序中包括以下程序:……
000002	2016	土地增值税的计提	贵集团应缴纳的主要税项之一为土地增值税。贵集团销售开发的房地产需要就土地增值额按照超率累进税率30% - 60%缴纳土地增值税。……	与评价土地增值税的计提相关的审计程序中包括以下程序:评价管理层与计量预计的土地增值税相关的内部控制的设计和运行有效性;……

4.数据库内容

数据模块	数据表	数据字段
关键审计事项	关键审计事项	股票代码、会计年度、关键审计事项类型、关键审计事项内容、应对方法

5.数据范围

数据模块	数据起始时间
全部模块	2016年

（十一）中国上市公司供应链研究数据库-SCRD

1.数据库简介

中国上市公司供应链研究数据库（*Supply Chain Research Database of Chinese Companies, SCRD*）基于上市公司财务报告的供应商和客户信息而建立的专业数据库。*SCRD* 包含中国上市公司供应商信息、客户信息、供应商及客户公司信息三个模块，共 8 张表。其中，供应商信息包括前五大主要供应商名称及采购额、供应商总采购信息；客户信息包括前五大主要客户名称及收入额、客户总收入信息；供应商及客户公司信息包括公司基本信息、股东信息、对外投资、主要人员。*SCRD* 包括 2001 年至今上市公司的供应商和客户信息。

2.数据库特色

- 数据完整、全面

相比于其他类似数据库，无论从数据质量还是数量来看，*SCRD* 所包括的上市公司客户与供应商信息更为准确和完整。

3.典型数据举例

（1）供应商名称及采购额

股票代码	公司简称	序号	会计年度	供应商名称	供应商采购额 (元)	供应商采购额占比
000038	深大通	1	2012	山东菏建建筑集团有限公司	10520967.18	16.4%
000038	深大通	2	2012	兖州市建筑安装工程总公司	9883378.17	15.4%
000038	深大通	3	2012	青岛德泰建设工程有限公司	6400911	9.98%
000038	深大通	4	2012	山东港基建设集团有限公司	5378335.7	8.38%
000038	深大通	5	2012	天元建设集团有限公司	4023947.87	6.27%

（2）公司基本信息

公司名称	公司编号	是否上市	公司股票代码	注册资本 (万元)	所属地区	成立日期	所属行业
北京迪信通商贸股份有限公司	AZ0000807	1	06188.HK	66666	北京市	2001-05-31	批发和零售业
北京点众科技股份有限公司	AZ0000808	0		4104	北京市	2011-09-15	科学研究和技术服务业
北京电力设备总厂有限公司	AZ0000809	0		34226	北京市	1981-06-09	制造业
北京鼎汉技术股份有限公司	AZ0000812	1	300011.SZ	55865	北京市	2002-06-10	科学研究和技术服务业
北京东方富林液化气体有限公司	AZ0000816	0		100	北京市	2009-06-18	批发和零售业
北京东方石油化工有限公司	AZ0000818	0		562369	北京市	2002-12-25	制造业

4.数据库内容

数据模块	数据表	数据字段
主要供应商信息	供应商名称及采购额	股票代码、公司简称、序号、会计年度、供应商名称、供应商采购额、供应商采购额占比 (%)
	供应商总采购信息	股票代码、公司简称、会计年度、前五大供应商采购总额、前五大供应商采购额占比 (%)
主要客户信息	客户名称及收入	股票代码、公司简称、序号、会计年度、客户名称、客户收入、客户收入占比 (%)
	客户总收入信息	股票代码、公司简称、会计年度、前五大客户收入、前五大客户收入占比 (%)
供应商及客户公司信息	公司基本信息	公司名称、公司编号、公司分类、是否上市、公司股票代码、注册资本、经营状态、公司类型、核准日期、所属地区、曾用名、人员规模、实缴资本、成立日期、所属行业、登记机关、企业地址、经营范围
	公司股东信息	公司名称、公司编号、公司分类、是否上市、公司股票代码、序号、股东名称、持股比例 (%)、认缴出资额、认缴出资日期、实缴出资额、实缴出资日期
	公司对外投资	公司名称、公司编号、公司分类、是否上市、公司股票代码、序号、被投资企业名称、被投资法定代表人、注册资本、出资比例 (%)、成立日期、经营状态
	公司主要人员	公司名称、公司编号、公司分类、是否上市、公司股票代码、序号、管理人员名称、职务

5.数据范围

数据模块	数据起始时间
全部模块	2001 年

（十二）中国上市公司异常收支数据库-ARED

1. 数据库简介

中国上市公司异常收支数据库（*Abnormal Revenue and Expenditure Database, ARED*）涵盖我国上市公司高管在职消费、营业外收支明细等非正常收支数据。*ARED* 包括高管在职消费和营业外收支 2 个模块，共 4 张表。高管在职消费模块依据前人研究经验，区分为直接法下和间接法下的高管在职消费数据，这些数据主要包括业务招待费、差旅费、办公费、会议费、交际应酬费等多类高管在职消费支出，其中，根据这些科目的对象又将其归类为管理费用和销售费用下的数据；营业外收支明细模块包括公司的营业外收入明细和营业外支出明细模块。

2. 数据库特色

- 紧跟学术研究

ARED 紧跟该领域最新研究成果，归纳整理了研究高管在职消费和营业外收支用到的研究数据。

3. 典型数据举例

高管在职消费_直接法

股票代码	会计年度	差旅费-销售费用（元）	差旅费-管理费用（元）	办公费-销售费用（元）	办公费-管理费用（元）	董事会会费（元）
000713	2013	10305939.67	1116423.6	1820774.03	2972753.02	487224
000713	2014	10093385.54	1006174.5	1537123.12	2324831.78	652417.37
000713	2015	9606328.04	887482.41	1708404.77	3024887.01	560960

4. 数据库内容

数据模块	数据表	数据字段
高管在职消费	直接法	股票代码、会计年度、业务招待费-管理费用、差旅费-销售费用、差旅费-管理费用、办公费-销售费用、办公费-管理费用、交际应酬费-销售费用、交际应酬费-管理费用、通讯费-销售费用、通讯费-管理费用、出国培训费、董事会会费、小车费-销售费用、小车费-管理费用、会议费、高管总人数、主营业务收入
	间接法	股票代码、会计年度、管理费用、董事/监事/高管成员薪酬、坏账准备、存货跌价准备、无形资产摊销、总资产、存货、固定资产、营业收入
营业外收支明细	营业外收入明细	股票代码、公司简称、会计年度、非流动资产处置利得、债务重组利得、非货币性资产交换利得、接受捐赠、政府补助、罚没收入、赔偿款收入、特殊项目、其他
	营业外支出明细	股票代码、公司简称、会计年度、非流动资产处置损失、债务重组损失、非货币性资产交换损失、罚没支出、对外捐赠、特殊项目、其他

5. 数据范围

数据模块	数据起始时间
高管在职消费	1991 年
营业外收支明细	2006 年

(十三) 投资者关系管理数据库-IRMD

1. 数据库简介

投资者关系管理是上市公司通过与投资者进行有效沟通，促进投资者或潜在投资者对公司的了解和认同，让投资者发现公司价值，并保持对公司长期认可，提升投资者忠诚度，进而实现公司价值最大化的有效管理手段。投资者关系管理数据库 (*Investor Relationship Management Database, IRMD*) 包含公司基本信息、投资者网络互动、投资者实地调研、投资者关系等四个模块，共 8 张表。

2. 数据库特色

- 多维和详尽

IRMD 从多个维度整理了投资者关系的相关信息，尤其是投资者网络互动为研究提供了新思路。

3. 典型数据举例

公司与投资者互动情况统计

股票代码	公司简称	年度	被提问数	回复数
600004	白云机场	2012	23	17
600017	日照港	2012	99	37
600019	宝钢股份	2012	805	578

4. 数据库内容

数据模块	数据表	数据字段
公司基本信息	公司地址信息	股票代码、公司简称、公司中文名称、公司简介、所属省份、所属地级市、所属区县、注册地址、办公地址、邮编
投资者网络互动	用户与公司问答	股票代码、公司简称、浏览用户/注册用户、用户名、提问时间、提问内容、上市公司是否回复、回复时间、回复内容
	公司互动情况统计	股票代码、公司简称、年度、被提问数、回复数
投资者实地调研	访客	股票代码、公司简称、公开时间、调研时间、参与单位、参与单位类别、参与人员
	接待人员	股票代码、公司简称、公开时间、调研时间、调研接待人员 n 姓名、调研接待人员 n 职务
	调研内容	股票代码、公司简称、调研时间、调研报告内容
投资者关系	是否有其他与投资者交流的平台	股票代码、公司简称、公司主页是否有投资者关系管理、公司主页是否有投资者咨询、公司主页是否有微信微博
	公司董秘是否金牌董秘	股票代码、公司简称、统计日期、董秘姓名、届数

5. 数据范围

数据模块	数据起始时间
公司基本信息	1991 年
投资者网络互动	2009 年
投资者实地调研	2011 年
投资者关系	2006 年

(十四) 海外上市研究数据库-COLD

1. 数据库简介

海外上市研究数据库 (*Overseas Listing Research Database, COLD*) 是基于公司公告、招股说明书、年度报告等数据整理而成的研究中国公司海外上市的专业学术数据库。*COLD* 涵盖了中国公司在海外股票市场, 例如纽约证券交易所、纳斯达克证券交易所、伦敦证券交易所的上市情况。*COLD* 包括 *IPO* 概况、公司股东、公司高管、财务报表状况和审计单位等共 7 张表。*IPO* 概况表总结了与公司海外上市相关的上市日期、所在股票交易所、承销商等信息; 公司股东表包括股东名称、持股数和持股比例等字段; 公司高管包括高管姓名、职位、出生年份等字段; 审计单位表包括会计师事务所名称、审计意见等字段。此外, *COLD* 还包含上市后的三大财务报表信息。

2. 数据库特色

- 全面的数据和字段

COLD 不仅提供了市场上已有的海外上市公司的财务报表数据等信息, 还提供较难获得但对于研究又不可或缺的公司高管、股东和审计单位等重要数据。

3. 典型数据举例

海外 IPO 概况

股票代码	公司中文名称	上市地点	上市日期	新股发行数量	首发募集资金
ATHM	汽车之家	纽约证券交易所	2013-12-11	7820000	132940000
ATIF	亚洲时代	纳斯达克	2019-05-03	2074672	10373360
ATV	橡果国际	纽约证券交易所	2007-05-03	6700000	103850000

4. 数据库内容

数据模块	数据表	数据字段
IPO 概况	IPO 概况	股票代码、公司简称、公司中文名称、公司英文名称、成立年份、注册资本、经营范围、注册地址、办公地址、上市地点、上市板、交易币种、首发上市日期、退市日期、新股发行数量、首发募集资金、首发价格
公司股东	公司股东	股票代码、公司名称、统计日期、股东排名、股东名称、持有股份、持股比例(%)
公司高管	公司高管	股票代码、公司名称、高管姓名、职位、任职日期、离职日期、出生年份、性别、国籍、个人简介
资产负债表	资产负债表	股票代码; 截止日期; 货币类型; 现金及现金等价物……
利润表	利润表	股票代码; 截止日期; 货币类型; 营业收入; 其他业务收入……
现金流量表	现金流量表	股票代码; 截止日期; 货币类型; 税前利润……
审计单位	审计单位	股票代码、会计年度、会计师事务所名称、审计意见

5. 数据范围

数据模块	数据起始时间
海外上市研究	1990 年

(十五) 中国法学研究数据库-CNLA

1. 数据库简介

中国法学研究数据库 (*Chinese Law Research Database, CNLAW*) 是面向广大高校和学者以及实务界人士, 旨在提供完整、精确的法学研究数据。此数据库全面涵盖法学研究所需要的各种信息, 不仅收录了海量的公司法律裁决信息, 还对上市公司诉讼仲裁事件进行汇总。*CNLAW* 包含上市公司法律裁决、知识产权法律裁决、上市公司诉讼仲裁 3 个数据模块。

2. 数据库特色

- 数据维度多样

CNLAW 包括海量上市公司法律裁判信息、上市公司诉讼仲裁信息的多维度信息。

- 内容详细完善

CNLAW 对各个维度的信息进行了详细地描述, 比如, 在上市公司诉讼仲裁模块中, 不仅对一审进行详细描述, 更针对是否上诉, 对二审也进行了细致描述, 甚至还追踪了后续的执行情况。

3. 典型数据举例

上市公司诉讼仲裁

公司简称	公告日期	案件名称	原告方	被告方	诉讼类型	涉案金额	起诉日期
益佰制药	2004-12-24	借款合同纠纷	交通银行上海浦东分行	上海源通医药有限公司	民事诉讼	1000万元	2004-09-20
中原环保	2004-12-15	赔偿其担保损失纠纷	国泰君安证券股份有限公司	白鸽(集团)股份有限公司	民事诉讼	3998万元	2001-12-31

4. 数据库内容

数据模块	数据表	数据字段
上市公司法律裁决	民事一审(原被告)	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、原告、被告、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据
上市公司法律裁决	民事一审(申请与被申请)	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、申请人、被申请人、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据
	民事二审(上诉与被上诉)	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、上诉人、被上诉人、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据
	民事再审(再审申请与被申请)	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、再审申请人、被申请人、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据
	民事再审(申诉与被申诉)	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、申诉人、被申诉人、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据

数据模块	数据表	数据字段
上市公司法律裁决	执行裁决	股票代码、案件年份、案件名称、案号、案件类型、审理法院、法院层级、裁判日期、申请执行人、被执行人、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据
	行政一审	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、原告、被告、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据
	行政二审	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、上诉人、被上诉人、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据
	行政审查	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、申请执行人、被执行人、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据
知识产权法律裁决	一审程序	股票代码、案件年份、案件名称、案号、审理法院、法院层级、裁判日期、案由、审理程序、原告、被告、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据
	二审程序	股票代码、案件年份、案件名称、案号、审理法院、法院层级、裁判日期、案由、审理程序、上诉人、被上诉人、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据
	再审程序	股票代码、案件年份、案件名称、案号、审理法院、法院层级、裁判日期、案由、审理程序、再审申请人、被申请人、判决结果、判决结果分类、处罚措施、涉案金额、涉案最大金额、法律依据
上市公司诉讼仲裁	上市公司诉讼仲裁统计	股票代码、公司简称、当事人类型、公告日期、案件名称、案件简介、原告方、被告方、诉讼类型、涉案金额、币种、起诉日期、一审受理法院、一审判决日期、一审判决内容、是否上诉、上诉方、二审受理法院、二审判决日期、二审判决内容、执行情况

5.数据范围

数据模块	数据起始时间
全部模块	1995 年

(十六) 上市公司盈余与信息质量数据库-EIQD

1.数据库简介

上市公司盈余与信息质量数据库 (*Earnings and Information Quality Database of Listed Companies, EIQD*) 基于对上市公司盈余管理与信息披露质量相关文献的考察和研究, 从上市公司财务报告数据中整理出与盈余质量研究相关的指标数据, 并从监管机构获得上市公司信息披露质量考评结果, 从而设置应计盈余管理、真实盈余管理、线下项目管理、盈余重述、微利企业、信息披露质量 6 个模块, 为该领域的研究提供极为准确和便捷的数据支持。

2.数据库特色

- 盈余质量衡量的多样性

EIQD 列示了企业盈余质量的五种衡量方法, 列出了计算各盈余质量指标的相关财务数据。

3.典型数据举例

线下项目管理

股票代码	会计年度	投资收益 (元)	营业外收入 (元)	政府补助 (元)	营业外支出 (元)
000792	2014	12615400.02	643703543.38	640519739.76	59551191.54
000792	2015	21518806.84	607009071.18	557972806.18	68243210.62
000792	2016	305263629.97	264192814.27	60993079.96	9866209.77
000792	2017	43898973.2	10300931.84	3334200.00	30814675.67

4.数据库内容

数据模块	数据表	数据字段
应计盈余管理	应计盈余管理	股票代码、会计年度、货币资金、应收账款、一年内到期的非流动资产、流动资产合计、固定资产、固定资产清理、无形资产……
真实盈余管理	真实盈余管理	股票代码、会计年度、经营活动现金流入小计、经营活动现金流出小计、经营活动产生的现金流量净额、营业总收入、主营业务收入……
线下项目管理	线下项目管理	股票代码、会计年度、投资收益、营业外收入、政府补助、营业外支出、2012 行业名称、2012 行业代码
盈余重述	盈余重述	股票代码、会计年度、会计差错发生年度 1、会计差错发生年度 2、盈余变更方向、因税调整、2012 行业名称、2012 行业代码
微利企业	微利企业	股票代码、会计年度、资产收益率 A (%)、资产收益率 B (%)、资产收益率 C (%)、净资产收益率 A (%)、净资产收益率 B (%)、净资产收益率 C (%)、2012 行业名称、2012 行业代码
信息披露质量	信息披露质量	股票代码、考评年度、考评结果

5.数据范围

数据模块	数据起始时间
全部模块	1991 年

(十七) 上市公司专利引用数据库-CITE

1.数据库简介

上市公司专利引用数据库 (*Patent Citation Database of Listed Companies, CITE*) 是基于对中国上市公司发明专利的被引用情况整理而成的专业数据库。本数据库的发明专利引用原始数据主要来自 *Google Patent*, 在处理过程中考虑了上市公司及其参控股公司的名称匹配、专利自引用、公司更名等多种情况, 由此形成了发明专利引用情况和实用新型专利引用情况两个模块, 包含发明专利引用基本信息、发明申请专利的被引用信息、发明申请专利的被引用数量、发明授权专利的被引用信息、发明授权专利的被引用数量、实用新型专利引用基本信息、实用新型授权专利的被引用信息和实用新型授权专利的被引用数量八张表, 涵盖了专利引用研究中的诸多重要字段。

2.数据库特色

- 详细的专利引用方和被引用方详情

CITE 既展现了上市公司专利作为被引用方的详细信息, 又包括引用方的信息, 为研究人员提供丰富的专利各方信息。

- 多个维度定义专利被引用数量

CITE 使用剔除自引用后的专利被引用数量、分年度累计被引用数量等多个维度衡量了专利被引用数量。

3.典型数据举例

(1) 发明专利引用基本信息

股票代码	被引用专利号	被引用专利分类号	被引用专利申请日	专利所有权人	专利引用数量	非专利引用数量	被引用次数
000002	CN102877661A	E04H1/02(2006.01)I	2012-09-15	西安万科企业有限公司	2	1	1
000002	CN102877660A	E04H1/02(2006.01)I	2012-09-15	西安万科企业有限公司	4	1	
000009	CN103427688A	H02M7/48(2007.01)I	2012-05-24	武汉永力电源技术有限公司	3		
000009	CN102709553B	H01M4/58(2010.01)I	2012-05-28	深圳市贝特瑞新能源材料股份有限公司	6		15

(2) 申请专利的被引用数量

股票代码	被引用专利号	被引用专利申请日	引用年度	公司类型	发明类型	各年被引用次数	剔除自引用的各年被引用次数	各年累计被引用次数
000002	CN101565984	2008-04-21	2012	上市公司本身	独立发明	1	1	1
000002	CN101565988	2008-04-21	2012	上市公司本身	独立发明	3	3	4

000002	CN101818 527A	2010-06-04	2012	子公司	独立 发明	8	8	12
000002	CN101818 539A	2010-06-04	2012	子公司	独立 发明	5	5	7
000002	CN101818 541A	2010-06-04	2012	子公司	独立 发明	1	1	1
000002	CN101831 985A	2010-06-04	2012	子公司	独立 发明	1	1	2

4.数据库内容

数据模块	数据表	数据字段
发明专利引用情况	发明专利引用基本信息	股票代码、被引用专利号、被引用专利分类号、被引用专利申请日、专利所有人、专利引用数量、非专利引用数量、被引用次数
	发明专利的被引用信息	股票代码、被引用专利号、被引用专利分类号、被引用专利申请日、公司类型、被引用专利申请人、发明类型、引用专利号、引用专利申请日期、引用专利授权日期、引用专利所有人
	发明专利的被引用数量	股票代码、被引用专利号、被引用专利申请日、引用年度、公司类型、发明类型、各年被引用次数、剔除自引用的各年被引用次数、各年累计被引用次数、剔除自引用的各年累计被引用次数
	发明专利的被引用信息	股票代码、被引用专利号、被引用专利分类号、被引用专利申请日、公司类型、被引用专利权人、发明类型、被引用专利授权公告日、引用专利号、引用专利申请日期、引用专利授权日期、引用专利所有人
	发明专利的被引用数量	股票代码、被引用专利号、被引用专利申请日、引用年度、公司类型、发明类型、各年被引用次数、剔除自引用的各年被引用次数、各年累计被引用次数、剔除自引用的各年累计被引用次数
实用新型专利引用情况	实用新型专利引用基本信息	股票代码、被引用专利号、被引用专利分类号、被引用专利申请日、专利所有人、专利引用数量、非专利引用数量、被引用次数
	实用新型授权专利的被引用信息	股票代码、被引用专利号、被引用专利分类号、被引用专利申请日、公司类型、被引用专利权人、发明类型、被引用专利授权公告日、引用专利号、引用专利申请日期、引用专利授权日期、引用专利所有人
	实用新型授权专利的被引用数量	股票代码、被引用专利号、被引用专利申请日、引用年度、公司类型、发明类型、各年被引用次数、剔除自引用的各年被引用次数、各年累计被引用次数、剔除自引用的各年累计被引用次数

5.数据范围

数据模块	数据起始时间
发明专利引用情况	1991 年
实用新型专利引用情况	1992 年

(十八) 董监高责任保险数据库-CDOD

1.数据库简介

董监高责任保险数据库 (*Directors and Officers Liability Insurance Database, CDOD*) 是基于沪深 A 股上市公司董事会、股东大会、年报等公告中公布的购买董监高责任保险信息而建立的专业数据库。*CDOD* 记录了 2002 年以来上市公司是否投保董监高责任保险、是否公布相关议案、议案公告日期、会议或文件标题、相关表述、保费金额、保金金额、保险年限、承保公司等重要字段, 弥补了该领域研究的数据空白。

2.数据库特色

- 数据全面, 样本量多

CDOD 遍历了上市公司所有董事会、股东大会和年报等公告中可能出现的董监高责任保险购买信息, 尽量收集完整的样本数据, 不遗漏相关信息。

- 字段丰富

CDOD 列示了上市公司是否公布相关议案、议案公告日期、会议或文件标题、相关表述、保费金额、承保公司等重要字段, 为研究提供重要支持。

3.典型数据举例

董监高责任保险

股票代码	议案公告日	公告标题	保费金额(万元)	保金金额(万元)	承保公司
600690	2009/6/11	2008 年度股东大会会议资料	24.7	13000	丘博保险(中国)有限公司
600779	2018/6/8	2017 年年度股东大会决议公告	35	10000	华泰财产保险有限公司
601588	2012/12/22	2012 年第二次临时股东大会决议公告	19	10000	美亚财产保险有限公司
601872	2010/5/19	2009 年度股东大会会议资料	15	3000	美国美亚保险公司深圳分公司

4.数据库内容

数据模块	数据表	数据字段
董监高责任保险	董监高责任保险	股票代码、股票简称、会计年度、是否投保董监高责任保险、是否公布相关议案、议案公告日期、公告标题、相关表述、保费金额、保金金额、保险年限、承保公司

5.数据范围

数据模块	数据起始时间
全部模块	2002 年

(十九) 员工持股计划数据库-ESOP

1.数据库简介

员工持股计划数据库 (*Employee Stock Ownership Plans Database, ESOP*) 基于上市公司发布的员工持股计划公告, 从中提取关键信息而建立的专业数据库。由于上市公司一般先发布员工持股计划草案, 而后实施截止时公布员工持股计划的完成情况, 因此, 在设计 *ESOP* 时既考虑了计划草案中的信息, 也包括了实施完成的信息。*ESOP* 主要包括两张表, 基本信息表和持股详情表。其中, 基本信息表包括了员工持股计划公告日、预计持股数量、实际持股数量、持有人数、高管认购比例以及董监高是否参与等重要字段; 持股详情表则包括了每个员工持股计划中的参与者姓名、职务和认购股份等信息。

2.数据库特色

- 详细的员工持股计划数据

ESOP 既涵盖了员工持股计划的基本信息, 根据研究需要, 还提取和整理了一些有价值的字段, 例如, 高管认购比例、董监高是否参与以及详细列示了具体参与人的持股情况。

3.典型数据举例

员工持股计划基本信息表

股票代码	首次公告日	高管认购比例 (%)	董监高是否参与	股份锁定日	锁定期	存续期
000035	2016/3/25	49.31	1	2016/12/3	36 个月	60 个月
000038	2017/9/30	71.8	1	2018/4/14	12 个月	24 个月
000040	2017/8/8	7.79	1	2017/11/15	12 个月	24 个月
000042	2016/4/16	41.31	1	2016/7/14	12 个月	24 个月
000056	2015/8/26	32.7	1	2015/12/21	12 个月	18 个月
000078	2017/9/29	36.66	1	2017/12/5	12 个月	48 个月

4.数据库内容

数据模块	数据表	数据字段
基本信息表	基本信息表	股票代码、股票简称、首次公告日、股东大会公告日、实施进度、实施完成日、资金总额、预计持股数量、拟发行价格、实际资金规模、实际持股数量、占总股本的比例 (%)、成交均价、持有人数、员工认购比例 (%)、高管认购比例 (%)、董监高是否参与、股份锁定日、锁定期、存续期
持股详情表	持股详情表	股票代码、首次公告日、持有人、职务、认购份额数量、认购份额占计划比例 (%)、认购股票数量、认购股票占计划比例 (%)

5.数据范围

数据模块	数据起始时间
全部模块	2014 年

（二十）上市公司精准扶贫研究数据库-TPAD

1.数据库简介

上市公司精准扶贫研究数据库（*Targeted Poverty Alleviation Database of Listed Company, TPAD*）是基于上市公司历年财务报告与企业社会责任报告中上市公司精准扶贫相关信息而建立的专业数据库，该数据库涵盖了上市公司从概括性说明到具体实施，再到后续跟踪等一系列精准扶贫措施，系统地反映了2016年以来沪深两市的公司每年在履行精准扶贫社会责任方面的相关工作。

2.数据库特色

- 较为全面的数据来源

TPAD 综合了上市公司财务报告与企业社会责任报告两个来源的企业精准扶贫责任履行情况。

- 科学合理的设计

TPAD 在既定的信息中进行筛选、整理，结合国家级贫困县相关内容，在一定程度上反映了企业在履行精准扶贫社会责任时对相关地区的关注度。

3.典型数据举例

上市公司精准扶贫分项投入

股票代码	会计年度	项目	指标	计量单位	金额
000001	2016	产业发展脱贫	投入金额	万元	60
000001	2016	转移就业脱贫	职业技能培训投入金额	万元	2
000001	2016	健康扶贫	贫困地区医疗卫生资源投入金额	万元	55
000001	2016	生态保护扶贫	投入金额	万元	4
000001	2016	兜底保障	“三留守”人员投入金额	万元	6
000001	2016	社会扶贫	定点扶贫工作投入金额	万元	95

4.数据库内容

数据模块	数据表	数据字段
上市公司精准扶贫基本信息	上市公司精准扶贫基本信息	股票代码、公司简称、会计年度、扶贫项目名称、扶贫地点、是否涉及国家级贫困县、项目需求金额（万元）、公司投资额（万元）、是否有后续扶贫计划、后续计划投入金额（万元）
上市公司精准扶贫总体情况	上市公司精准扶贫总体情况	股票代码、公司简称、会计年度、总金额（万元）、资金（万元）、物资折款（万元）、帮助建档立卡贫困人口脱贫数（人）、是否有后续扶贫计划
上市公司精准扶贫分项投入	上市公司精准扶贫分项投入	股票代码、公司简称、会计年度、项目、指标、计量单位、数量/开展情况
上市公司精准扶贫获奖情况	上市公司精准扶贫获奖情况	股票代码、公司简称、会计年度、序号、获奖情况、获奖级别

5.数据范围

数据模块	数据起始时间
全部模块	2016年

(二十一) 上市公司海外经营数据库-COOD

1. 数据库简介

上市公司海外经营数据库 (*Overseas Operation Database of Listed Companies, COOD*) 是基于中国上市公司年度报告中披露的海外关联参控股公司以及海外分部收入等信息整理而成的专业数据库。*COOD* 包括两张表, 海外参控股公司表囊括了年报中披露的海外子公司、联合营企业等公司数据, 提供诸如公司名称、注册资本、持股比例、总资产、净利润等重要数据; 海外经营收入表主要来源于年报中的地区分部报告, 整理了所有上市公司海外经营分部的收入和成本情况。

2. 数据库特色

- 多维度展现上市公司海外经营

COOD 包括微观层面的海外参控股公司经营情况, 还从分部报告情况综合展示了海外经营情况。

- 时间序列长, 数据量丰富

海外参控股公司数据从 2010 年开始, 海外经营收入从 2001 年开始, 提供了多个年度数据。

3. 典型数据举例

(1) 海外参控股公司

股票代码	统计日期	海外参控股公司名称	注册地	经营范围	持股比例 (%)
000002	2018-12-31	Vanke Holdings USA LLC	美国	投资	100
000005	2013-12-31	首冠国际有限公司	香港	投资控股	100
000007	2014-12-31	香港港众投资有限公司	香港	投资	100

(2) 海外经营收入

股票代码	公司简称	会计年度	地区	主营业务收入 (万元)	主营业务成本 (万元)
000008	神州高铁	2015	海外	510.943369	315.861014
000012	南玻 A	2001	美国	10319.8677	8086.6072

4. 数据库内容

数据模块	数据表	数据字段
海外参控股公司	海外参控股公司	股票代码、公司简称、统计日期、海外参控股公司名称、公司重要程度、与上市公司关系、注册地、国家区域名称、注册资本、经营范围、直接持股 (%)、间接持股 (%)、持股比例合计 (%)、表决权比例 (%)、取得方式、币种、总资产、负债总额、净资产总额、营业收入、营业利润、净利润、期末总投资额、企业所得税率
海外经营收入	海外经营收入	股票代码、公司简称、会计年度、地区、主营业务收入 (万元)、主营业务成本 (万元)

5. 数据范围

数据模块	数据起始时间
海外参控股公司	2010 年
海外经营收入	2001 年

（二十二）上市公司土地市场信息数据库-LMID

1.数据库简介

上市公司土地市场信息数据库 (*Land Market Information Database of List Company, LMID*) 是基于全国各市县自然资源和规划局公布的国有土地出让信息, 并匹配上市公司及其参控股公司信息而开发的专业数据库。*LMID* 囊括了自 1999 年以来上市公司及其参控股公司国有土地出让相关信息, 字段涵盖土地所在省份、地级市、县级市、土地位置、土地面积、成交价格等重要字段。截至 2018 年底, *LMID* 共收录了上市公司国有土地出让近 3 万条数据, 为该领域研究提供了极大的数据支持。

2.数据库特色

- 来源权威可靠

LMID 数据主要来源于各省、市、县自然资源和规划局所公布的土地交易信息, 数据源权威可靠。

- 字段丰富, 匹配精确

LMID 几乎囊括研究上市公司国有土地出让所需要的数据, 对原有土地出让数据字段进行了最大限度保留, 并精确匹配了与上市公司及其参控股公司相关的国有土地出让情况。

3.典型数据举例

上市公司国有土地出让

股票代码	土地使用权人	参控关系	省份	土地面积 (公顷)	成交价格 (万元)	合同签订 日期
603993	洛阳栾川钼业集团股份有限公司	上市公司本身	河南省	9.6392	2966.94	2012/9/13
603996	中新科技集团股份有限公司	上市公司本身	浙江省	3.2031	3359	2015/2/4
603997	沈阳继峰汽车零部件有限公司	子公司	辽宁省	2.79448	1344.14	2015/6/15
603997	天津继峰汽车零部件有限公司	子公司	天津市	2.85541	1808	2017/1/6
603997	重庆继涵科技有限公司	子公司	浙江省	4.73629	1620	2018/7/4

4.数据库内容

数据模块	数据表	数据字段
上市公司国有 土地出让	上市公司国有 土地出让	股票代码、土地使用权人、参控关系、省份、地级市、县级市、土地面积、土地来源、土地用途、供地方式、土地使用年限、行业分类、土地级别、成交价格、约定容积率_下限、约定容积率_上限、合同签订日期、约定开工时间、约定竣工时间

5.数据范围

数据模块	数据起始时间
全部模块	1999 年

(二十三) 上市公司年度信息数据库-CAID

1. 数据库简介

上市公司年度信息数据库 (*Annual Information Database of Listed Company, CAID*) 基于上市公司各年度财务报告, 提供了 2000 年至今中国 A 股和 B 股市场所有上市公司时间序列上的基本信息。*CAID* 共包括上市公司年度信息表与上市公司人员结构表两张表, 反映了不同年度上市公司基本信息的变化与人员结构的变化, 可为相关研究提供较为完善的数据支持。

2. 数据库特色

- 字段设置更为合理

CAID 囊括了与上市公司基本信息相关的几乎所有字段; 在人员结构方面, 对不同标准的上市公司人员结构划分进行了统一, 例如: 某些上市公司在进行学历划分时, 将博士与硕士学位的人员分开进行统计, 而有些公司则统一为研究生以上学历, *CAID* 对此类数据进行了统一。

3. 典型数据举例

上市公司年度信息

股票代码	会计年度	公司简称	所在省份	所在地级市	董事会秘书
000001	2012	平安银行	广东省	深圳市	李南青
000001	2013	平安银行	广东省	深圳市	李南青
000002	2012	万科 A	广东省	深圳市	谭华杰
000002	2013	万科 A	广东省	深圳市	谭华杰
000004	2012	国农科技	广东省	深圳市	杨斌
000004	2013	国农科技	广东省	深圳市	杨斌

4. 数据库内容

数据模块	数据表	数据字段
上市公司年度信息	上市公司年度信息	股票代码、会计年度、公司简称、股票上市证券交易所、公司中文名称、公司英文名称、法定代表人、成立日期、首发上市日期、同公司 B 股代码、同公司 H 股代码、注册地址、办公地址、所在省份、所在地级市、所在县级市、经度、纬度、董事会秘书、证券事务代表、行业名称 A、行业代码 A、行业名称 B、行业代码 B
上市公司人员结构	上市公司人员结构	股票代码、会计年度、公司简称、类别名称、人数 (人)

5. 数据范围

数据模块	数据起始时间
上市公司年度信息	2000 年
上市公司人员结构	2010 年

（二十四）上市公司区块链投资数据库-BLOCK

1.数据库简介

上市公司区块链投资数据库（*Blockchain Investment Database of Listed Company, BLOCK*）是基于对中国上市公司公告、投资者问答和新闻报道等与上市公司区块链投资相关内容的整理和分析，从而建立的专业数据库。目前，很多上市公司股票成为区块链概念股，这些概念股中不乏有真正实质性涉足区块链技术的公司，但也有很多蹭热度和炒概念的公司。为此，**BLOCK**从上市公司公告内容、投资者问答与新闻报道内容等，深度挖掘沪深A股上市公司进行区块链投资的情况，以期为学术研究和实务投资提供数据支持。

2.数据库特色

- 数据准确，样本量多

BLOCK通过上市公司公告内容判断的“区块链概念”公司远多于其他数据库。

- 数据可溯源，权威性高

BLOCK数据来源于公告、投资者问答平台和新闻报道，具有较高的权威性。

3.典型数据举例

涉及区块链的上市公司

股票代码	公告日期	是否涉及区块链	是否有区块链收入	相关表述
600093	2019-10-12	涉及	是	公司在供应链管理中引入区块链技术，使参与方的底层资产信息更加真实
600146	2019-10-09	不涉及	否	经公司自查，目前公司不涉及区块链业务
601916	2019-10-15	涉及	是	本行充分运用互联网、大数据、区块链等新技术，创新推出“池化”和“线上化”金融服务
603022	2019-10-30	不涉及	否	公司本次收购的标的公司不涉及区块链业务

4.数据库内容

数据模块	数据表	数据字段
涉及区块链的上市公司	涉及区块链的上市公司	股票代码、股票简称、公告日期、公告标题、是否涉及区块链、是否有区块链收入、相关表述 1、相关表述 2、相关表述 3
涉及区块链的投资者问答	涉及区块链的投资者问答	股票代码、公司简称、用户名、浏览用户/注册用户、提问时间、提问内容、上市公司是否回复、回复时间、回复内容、是否说明不涉及区块链、数据来源
涉及区块链的新闻报道	涉及区块链的新闻报道	股票代码、股票简称、新闻类型、报道日期、新闻标题、新闻内容、报道媒体

5.数据范围

数据模块	数据起始时间
全部模块	2016 年

(二十五) 绿色专利研究数据库-GPRD

1.数据库简介

绿色专利研究数据库 (*Green Patent Research Database, GPRD*) 是结合中国专利数据以及世界知识产权局公布的绿色专利分类号标准而开发的专业数据库。*GPRD* 主要分为: 上市公司绿色专利申请与获得、上市公司上市前绿色专利申请与获得、上市公司绿色专利分类号信息、上市公司绿色发明专利引用、上市公司绿色实用新型专利引用、各省市绿色专利申请与获得、非上市公司专利申请与获得、科研机构绿色专利申请与获得这八个大的模块, 在这些模块下又按照不同的研究内容进行细分, 能够满足不同研究人员多维度的研究需求。

2.数据库特色

- 数据来源权威、内容丰富

GPRD 对于绿色专利的划分标准完全遵循了世界知识产权局的绿色专利标准, 对来源于国家知识产权局和 *Google Patent* 的专利进行了全面地梳理与筛查, 最后形成了包括上市公司、非上市公司、科研机构和地域绿色专利, 涵盖绿色专利的申请与授权情况以及被引用情况。

- 地域绿色专利创新情况

GPRD 不仅提供了每一年度各省市的绿色专利申请或授权数量, 还结合当年度该地区创新专利总数计算出了绿色专利所占百分比, 较为客观地反映出了各地区对绿色创新的重视情况。

3.典型数据举例

上市公司绿色专利申请情况

股票代码	会计年度	公司类型	当年独立申请的绿色发明数量	当年独立申请的绿色实用新型数量	当年联合申请的绿色发明数量	当年联合申请的绿色实用新型数量
000009	2012	子公司	5	5	0	0
000009	2012	集团公司合计	5	5	0	0
000012	2012	上市公司本身	1	1	0	1
000012	2012	子公司	0	2	0	2

4.数据库内容

数据模块	数据表	数据字段
上市公司绿色专利申请与获得	上市公司绿色专利申请情况	股票代码、会计年度、公司类型、申请时间、当年独立申请的绿色发明数量、当年独立申请的绿色实用新型数量、当年联合申请的绿色发明数量、当年联合申请的绿色实用新型数量
	上市公司绿色专利获得情况	股票代码、会计年度、公司类型、获得时间、当年独立获得的绿色发明数量、当年独立获得的绿色实用新型数量、当年联合获得的绿色发明数量、当年联合获得的绿色实用新型数量
	上市公司绿色专利申请后获得情况	股票代码、会计年度、更新日期、专利申请个数、专利申请_获得个数、专利申请_当年获得个数、专利申请_第二年获得个数、专利申请_第三年获得个数、专利申请_第四年获得个数、发明申请个数、

数据模块	数据表	数据字段
		发明申请_获得个数、发明申请_当年获得个数、发明申请_第二年获得个数、发明申请_第三年获得个数、发明申请_第四年获得个数、实用新型申请个数、实用新型申请_获得个数、实用新型申请_当年获得个数、实用新型申请_第二年获得个数、实用新型申请_第三年获得个数、实用新型申请_第四年获得个数
	上市公司绿色专利获得年限情况	股票代码、会计年度、专利获得、专利获得_平均年限、发明专利获得、发明专利获得_平均年限、发明专利获得_公开平均年限、实用新型获得、实用新型获得_平均年限
上市公司绿色专利分类号信息	上市公司绿色发明专利分类号	股票代码、会计年度、公司类型、申请时间、发明申请 A 类、发明申请 A 类数量、发明申请 B 类、发明申请 B 类数量……
	上市公司绿色实用新型申请专利分类号	股票代码、会计年度、公司类型、申请时间、实用新型申请 A 类、实用新型申请 A 类数量、实用新型申请 B 类、实用新型申请 B 类数量……
	上市公司绿色发明获得专利分类号	股票代码、会计年度、公司类型、获得时间、发明获得 A 类、发明获得 A 类数量、发明获得 B 类、发明获得 B 类数量……
	上市公司绿色实用新型获得专利分类号	股票代码、会计年度、公司类型、获得时间、实用新型获得 A 类、实用新型获得 A 类数量、实用新型获得 B 类、实用新型获得 B 类数量……
上市公司绿色发明专利引用	绿色发明专利引用基本信息	股票代码、被引用专利号、被引用专利分类号、被引用专利申请日、专利所有权人、专利引用数量、非专利引用数量、被引用次数
	申请的绿色发明专利的被引用信息	股票代码、被引用专利号、被引用专利分类号、被引用专利申请日、公司类型、被引用专利申请人、发明类型、引用专利号、引用专利申请日期、引用专利授权日期、引用专利所有权人
	申请的绿色发明专利的被引用数量	股票代码、被引用专利号、被引用专利申请日、引用年度、公司类型、发明类型、各年被引用次数、剔除自引用的各年被引用次数、各年累计被引用次数、剔除自引用的各年累计被引用次数
	授权的绿色发明专利的被引用信息	股票代码、被引用专利号、被引用专利分类号、被引用专利申请日、公司类型、被引用专利权人、发明类型、授权公告日、引用专利号、引用专利申请日期、引用专利授权日期、引用专利所有权人
	授权的绿色发明专利的被引用数量	股票代码、被引用专利号、被引用专利申请日、引用年度、公司类型、发明类型、各年被引用次数、剔除自引用的各年被引用次数、各年累计被引用次数、剔除自引用的各年累计被引用次数
上市公司绿色实用新型专利引用	绿色实用新型专利引用基本信息	股票代码、被引用专利号、被引用专利分类号、被引用专利申请日、专利所有权人、专利引用数量、非专利引用数量、被引用次数
	授权的绿色实用新型专利的被引用信息	股票代码、被引用专利号、被引用专利分类号、被引用专利申请日、公司类型、被引用专利权人、发明类型、授权公告日、引用专利号、引用专利申请日期、引用专利授权日期、引用专利所有权人
	授权的绿色实用新型专利的被引用数量	股票代码、被引用专利号、被引用专利申请日、引用年度、公司类型、发明类型、各年被引用次数、剔除自引用的各年被引用次数、各年累计被引用次数、剔除自引用的各年累计被引用次数
各省市绿色专利申请与获得	各省市绿色专利申请情况	省份名称、地市名称、会计年度、当年申请的绿色发明数量、绿色发明占地区年度申请的发明总数百分比、当年申请的绿色实用新型数量、绿色实用新型占地区年度申请的实用新型总数百分比

数据模块	数据表	数据字段
	各省市绿色专利获得情况	省份名称、地市名称、会计年度、当年获得的绿色发明数量、绿色发明占地区年度获得的发明总数百分比、当年获得的绿色实用新型数量、绿色实用新型占地区年度获得的实用新型总数百分比
非上市公司绿色专利申请与获得	非上市公司绿色专利申请情况	公司名称、会计年度、当年独立申请的绿色发明数量、当年独立申请的绿色实用新型数量、当年联合申请的绿色发明数量、当年联合申请的绿色实用新型数量
	非上市公司绿色专利获得情况	公司名称、会计年度、当年独立获得的绿色发明数量、当年独立获得的绿色实用新型数量、当年联合获得的绿色发明数量、当年联合获得的绿色实用新型数量
科研机构绿色专利申请与获得	科研机构绿色专利申请情况	单位名称、会计年度、当年独立申请的绿色发明数量、当年独立申请的绿色实用新型数量、当年联合申请的绿色发明数量、当年联合申请的绿色实用新型数量
	科研机构绿色专利获得情况	单位名称、会计年度、当年独立获得的绿色发明数量、当年独立获得的绿色实用新型数量、当年联合获得的绿色发明数量、当年联合获得的绿色实用新型数量

5.数据范围

数据模块	数据起始时间
上市公司绿色专利申请与获得	1990 年
上市公司上市前绿色专利申请与获得	1993 年
上市公司绿色专利分类号信息	1990 年
上市公司绿色发明专利引用	1993 年
上市公司绿色实用新型专利引用	1994 年
各省市绿色专利申请与获得	1986 年
非上市公司绿色专利申请与获得	1985 年
科研机构绿色专利申请与获得	1985 年

（二十六）上市公司承诺事项数据库-CMIT

1.数据库简介

上市公司承诺事项数据库（*Commitments Database of Listed Companies, CMIT*）是基于上海证券交易所和深圳证券交易所公布的上市公司承诺事项与履行相关情况整理而成的专业数据库。*CMIT* 记录了上市公司经营过程中做出的各类承诺事项、是否履行以及履行时间等信息。

2.数据库特色

- 数据来源权威、字段全面

CMIT 的数据来源于官方监管机构，来源权威，字段全面。

3.典型数据举例

上市公司承诺事项

股票代码	承诺类型	做出承诺时间	承诺主体名称	承诺事项类别	预计履行完成时间
600000	再融资	2017-09-04	上海国际集团有限公司	股份限售、股份锁定	2020-09-03
600005	再融资	2010-12-23	武汉钢铁（集团）公司	资产注入	2016-04-07
600006	权益变动和收购	2014-03-17	东风汽车公司	股份减持相关承诺	2015-03-16
600008	股权分置改革	2006-04-14	北京首都创业集团有限公司	其他	2014-06-25
600008	其他	2012-04-20	北京首创股份有限公司	解决同业竞争	2017-04-19

4.数据库内容

数据模块	数据表	数据字段
上市公司承诺事项	上市公司承诺事项	股票代码、公司简称、承诺类型、做出承诺时间、承诺主体名称、承诺主体类型、承诺事项类别、目前履行状态、预计履行完成时间、实际履行完成时间、备注

5.数据范围

数据模块	数据起始时间
全部模块	1993 年

(二十七) 上市公司委托贷款数据库-CELD

1.数据库简介

上市公司委托贷款数据库 (*Entrust Loans Database of Listed Companies, CELD*) 是根据上市公司委托贷款公告整理而成的专业数据库。上市公司委托贷款在一定程度上体现了公司层面的影子银行行为, 因此, 通过对委托贷款公告的深度挖掘, *CELD* 详细展示了上市公司委托贷款过程中的贷款人、借款人、借贷金额、借款利率等重要字段, 为相关领域研究提供了重要数据支持。

2.数据库特色

- 数据字段丰富

CELD 的数据来源于上市公司公告, 并对公告内容进行了全方位挖掘, 字段全面。

3.典型数据举例

上市公司委托贷款

股票代码	公告日期	贷款人	借款人	借贷金额	是否构成关联交易	是否含有抵押担保条款
000011	2009/10/20	深圳市投资控股有限公司	深圳市皇城地产有限公司	1.5 亿元	是	否
000037	2010/12/28	上市公司	深南电(中山)电力有限公司	1 亿元	否	是
000042	2014/9/25	国元证券股份有限公司	上海深长城地产有限公司	4.8 亿元	否	是
000043	2014/7/18	深圳格兰云天酒店管理有限公司	上市公司	5000 万元	是	否

4.数据库内容

数据模块	数据表	数据字段
上市公司委托贷款	上市公司委托贷款	股票代码、公司简称、公告标题、公告日期、委托贷款方向、贷款人、借款人、借贷金额、委托贷款期限、借款利率、发放银行/机构、借款人与上市公司关系、贷款人与上市公司关系、是否构成关联交易、是否含有抵押担保条款

5.数据范围

数据模块	数据起始时间
全部模块	2006 年

(二十八) 上市公司环境治理数据库-CEGD

1. 数据库简介

上市公司环境治理数据库 (*Environmental Governance Database of Listed Company, CEGD*) 是根据上市公司环境治理相关披露信息整理而成专业数据库。**CEGD** 重点排污企业名单根据年报中的“上市公司及其子公司是否属于环境保护部门公布的重点排污单位/规定的重污染行业”确定, 如果上市公司及其子公司属于重点排污单位, 并且公布了具体的污染物排放情况, 则据此整理而成“上市公司环境排放明细”表; 上市公司排污环保费用和上市公司环境治理补助则来自于财务报表附注中的管理费用明细和政府补助明细; 上市公司环保处罚来自上市公司发布的环境处罚公告。

2. 数据库特色

- 丰富的上市公司环境数据

CEGD 从是否重点排污企业、排污费、环保支出、排放明细、环境治理补助以及环保处罚等多个维度展示了上市公司环境治理情况。

3. 典型数据举例

上市公司环境治理补助

股票代码	统计日期	环境治理项目	政府补助金额
000007	2015-12-31	节约能源和循环经济发展专项基金	100000
000009	2014-12-31	尘烟污染综合整治专项补助资金	720000
000009	2014-12-31	有机废弃物综合利用技术孵化及应用研发补助	400000
000009	2014-12-31	高效节能型化成分选设备研究补助	100000
000009	2014-12-31	尾矿渣综合利用生产蒸压加气混凝土砌块技改项目补助	105000
000009	2014-12-31	红柱石尾矿渣综合利用项目经费补助	130000

4. 数据库内容

数据模块	数据表	数据字段
重点排污企业名单	重点排污企业名单	股票代码、会计年度、是否重点排污
上市公司排污环保费用	上市公司排污环保费用	股票代码、统计日期、项目名称、发生额
上市公司环境排放明细	上市公司环境排放明细	股票代码、会计年度、公司或子公司名称、主要污染物的名称、排放方式、排放口数量、排放口分布情况、排放浓度、执行的污染物排放标准、排放总量、核定的排放总量、超标排放情况
上市公司环境治理补助	上市公司环境治理补助	股票代码、统计日期、环境治理项目、政府补助金额
上市公司环保处罚	上市公司环保处罚	股票代码、公告日期、处罚对象、关系、违规行为、处分类型、处分措施、处理人、处罚金额

5. 数据范围

数据模块	数据起始时间
全部模块	2008 年

(二十九) 上市公司金融化数据库-CFID

1. 数据库简介

上市公司金融化数据库 (*Financialization Database of Listed Company, CFID*) 是根据上市公司在经营过程中将闲置资金用于委托理财、证券投资等事项整理而成的专业数据库。*CFID* 主要包括三个大模块: 并购基金、委托理财和证券投资, 其中, 并购基金表根据上市公司发布的公告整理而成, 包括上市公司发起或参与设立并购基金的出资额、持股比例、投资领域等内容; 委托理财表和证券投资表则根据上市公司年报披露的委托理财信息和证券投资信息整理而成。

2. 数据库特色

- 独特的并购基金数据

CFID 对并购基金公告进行了深度提取和整理, 满足特色学术研究需求。

- 全面的字段内容

CFID 遵照公告和年报原文, 保留了几乎所有重要字段。

3. 典型数据举例

上市公司投资理财

股票代码	会计年度	受托人名称	委托理财金额	起始日期	终止日期	实际收益
000009	2012	建行南京分行	10000000	2012/11/9	2013/4/1	135200
000009	2012	建行南京分行	15000000	2012/11/9	2013/4/10	278200
000036	2012	中行深圳分行	60000000	2012/1/17	2012/2/17	270000
000036	2012	中行深圳分行	80000000	2012/2/17	2012/3/19	320000
000036	2012	中行深圳滨河支行	99000000	2012/7/10	2012/7/31	150000
000036	2012	中行深圳滨河支行	100000000	2012/8/1	2012/8/31	240000

4. 数据库内容

数据模块	数据表	数据字段
并购基金	并购基金	股票代码、公司简称、公告日期、股东名称、基金名称、出资额、基金规模、货币类型、持股比例(%)、股东构成、投资期限、退出期、投资领域、基金管理公司
委托理财	委托理财	股票代码、会计年度、受托人名称、委托理财金额、货币类型、起始日期、终止日期、理财期限、报酬确定方式、实际收回本金、实际获得收益
证券投资	证券投资	股票代码、会计年度、被投资证券代码、被投资证券名称、证券类型、投资类型、投资金额、期末持有数量、期末账面价值

5. 数据范围

数据模块	数据起始时间
全部模块	2007 年

（三十）上市公司银行借款数据库-LOAN

1.数据库简介

上市公司银行借款数据库 (*Bank Loan Database of Listed Company, LOAN*) 是基于上市公司公布的各类借贷融资公告整理而成的专业数据。上市公司在经营过程中经常涉及三类借贷融资行为：银行授信、借款担保和银行贷款，*LOAN* 通过深度挖掘这三类公告内容，整理了与研究相关的关键字段，期待为该领域研究提供重要数据支持。其中，银行授信模块主要包括上市公司获得银行授信及授信额度的数据；借款担保模块包括上市公司为关联方借款提供担保的数据；银行贷款模块则包括上市公司向银行获得贷款金额及贷款期限的数据。

2.数据库特色

- 模块清晰、字段丰富

LOAN 非常清晰地将公司借贷行为分为了授信、担保和贷款行为，为相关研究提供了不同的数据维度，并且根据每个模块的特点设置了丰富的数据字段。

3.典型数据举例

授信

股票代码	公告日期	申请授信公司名称	授信银行名称	授信金额(万元)	申请授信公司与上市公司关系	授信期限(月)	计价货币
000002	2007-06-20	万科企业股份有限公司	中国建设银行股份有限公司	2000000	本公司		人民币元
000004	2020-03-24	深圳爱加密科技有限公司	招商银行股份有限公司	1000	孙公司	12	人民币元
000005	2016-11-15	深圳世纪星源股份有限公司	广发银行股份有限公司	7000	本公司	12	人民币元

4.数据库内容

数据模块	数据表	数据字段
授信	授信	股票代码、公告日期、申请授信公司名称、授信银行名称、授信金额、申请授信公司与上市公司关系、授信期限、计价货币
担保	担保	股票代码、公司简称、统计日期、交易日期、担保方公司名称、被担保方公司名称、授担保方式、担保金额、计价货币、担保起始日期、担保终止日期、是否关联交易
贷款	贷款	股票代码、公告日期、借款公司名称、贷款银行名称、贷款金额、贷款期限、贷款公司与上市公司关系、计价货币

5.数据范围

数据模块	数据起始时间
授信	2003 年
担保	2003 年
贷款	2001 年

(三十一) 对赌协议数据库-VAMD

1.数据库简介

对赌协议数据库 (*Valuation Adjustment Mechanism Database, VAMD*) 基于上市公司并购重组公告中的估值调整机制整理而成的专业数据库。*VAMD* 整理了出让方、投资方和标的方名称, 及其在各个报告期所承诺的净利润和实际净利润, 较为直观体现了并购重组过程中的对赌情况, 保留了对赌协议的主要关键信息, 为该领域研究提供了重要数据支持。

2.数据库特色

- 展示了对赌协议的关键信息

VAMD 从 14 个字段较为清晰、简约地展现了对赌协议的主要关键信息。

3.典型数据举例

对赌协议

证券代码	首次公告日	交易出让方	交易投资方	事件角色	交易标的物	承诺报告期	承诺净利润	承诺补偿方式
000975	2012-05-04	侯仁峰, 王水, 李红磊	银泰黄金股份有限公司	投资方	内蒙古玉龙矿业股份有限公司 69.4685% 股权	2012-12-31	280000000	现金补偿
000975	2012-05-04	侯仁峰, 王水, 李红磊	银泰黄金股份有限公司	投资方	内蒙古玉龙矿业股份有限公司 69.4685% 股权	2013-12-31	310000000	现金补偿
000975	2012-05-04	侯仁峰, 王水, 李红磊	银泰黄金股份有限公司	投资方	内蒙古玉龙矿业股份有限公司 69.4685% 股权	2014-12-31	340000000	现金补偿

4.数据库内容

数据模块	数据表	数据字段
对赌协议	对赌协议	证券代码、公司简称、首次公告日、并购事件、方案进度、交易出让方、交易投资方、事件角色、交易标的物、承诺报告期、承诺净利润、币种、承诺补偿方式、实际净利润

5.数据范围

数据模块	数据起始时间
对赌协议	2010 年

（三十二）一致行动人数据库-PACD

1.数据库简介

一致行动人数据库 (*Persons Acting in Concert Database, PACD*) 是基于上市公司主要股东的一致行动人信息制作的专业数据库。通过仔细阅读和整理上市公司财务报告和公告等文件, *PACD* 展示了上市公司主要股东中是否存在一致行动人、一致行动人名称和持股比例等重要信息, 为该领域研究提供了重要数据支持。

2.数据库特色

- 数据详细和全面

PACD 的数据来源于上市公司财务报告, 详细展示了研究中的常用字段。

3.典型数据举例

上市公司一致行动人

股票代码	会计年度	一致行动人名称	持股位次	一致行动标识	持股数量	持股比例	合计持股比例
000001	2014	中国平安保险(集团)股份有限公司	第1大股东	A	5734892419	50.2	58.95
000001	2014	中国平安人寿保险股份有限公司	第2大股东	A	728793600	6.38	58.95
000001	2014	中国平安人寿保险股份有限公司-传统-普通保险产品	第4大股东	A	270649974	2.37	58.95
000002	2014	全国社保基金一零三组合	第8大股东	A	79999794	0.72	1.37
000002	2014	中国建设银行-博时主题行业股票证券投资基金	第9大股东	A	72000000	0.65	1.37
000006	2014	深圳市人民政府国有资产监督管理委员会	第1大股东	A	296031373	21.93	34
000006	2014	深圳市资本运营集团有限公司	第2大股东	A	162966811	12.07	34

4.数据库内容

数据模块	数据表	数据字段
一致行动人	上市公司一致行动人	股票代码、公司简称、会计年度、是否存在一致行动人、一致行动人名称、持股位次、一致行动标识、持股数量、持股比例、合计持股比例

5.数据范围

数据模块	数据起始时间
一致行动人	2012年

(三十三) 上市公司风投持股数据库-VCSD

1. 数据库简介

上市公司风投持股数据库 (*VC Shareholders Database of Listed Company, VCSD*) 根据公司上市前后的股东持股情况, 采用权威方法匹配出上市公司风投持股情况的专业数据库。*VCSD* 包括 *IPO* 前风投持股情况、上市后风投持股情况、风险投资机构基本信息和历年风险投资机构列表四个模块。

2. 数据库特色

- 数据权威可靠

VCSD 的数据来源于上市公司历年财务报告和权威研究报告, 对于风险投资机构的判断参考了已有关权威期刊论文的研究成果。

3. 典型数据举例

上市后风投持股情况

股票代码	公司名称	统计日期	风险投资股东	持股数量 (股)	持股比例 (%)	股份性质
300001	青岛特锐德电气股份有限公司	2012-06-30	青岛德锐投资有限公司	99750000	49.78	民营企业
300001	青岛特锐德电气股份有限公司	2012-12-31	新疆古月杨股权投资合伙企业 (有限合伙)	2505000	1.25	民营企业
300001	青岛特锐德电气股份有限公司	2012-12-31	青岛德锐投资有限公司	99750000	49.78	民营企业

4. 数据库内容

数据模块	数据表	数据字段
IPO 前风投持股情况	IPO 前风投持股情况	股票代码、公司名称、公告日期、IPO 阶段、股东排名、风险投资机构编号、风险投资股东、持股数量 (股)、持股比例 (%)、股份性质
上市后风投持股情况	上市后风投持股情况	股票代码、公司名称、统计日期、股东排名、风险投资机构编号、风险投资股东、持股数量 (股)、持股比例 (%)、股份性质
风险投资机构基本信息	风险投资机构基本信息	企业名称、企业编号、登记状态、成立日期、注册资本 (万)、实缴资本 (万)、币种、企业类型、营业期限、纳税人资质、所属行业、所属地区、登记机关、人员规模、参保人数、企业地址、经营范围
历年风险投资机构列表	历年风险投资机构列表	年份、风险投资机构名称、成立时间、地区

5. 数据范围

数据模块	数据起始时间
IPO 前风投持股情况	2009 年
上市后风投持股情况	2009 年
风险投资机构基本信息	截面数据
历年风险投资机构列表	2002 年

（三十四）上市公司数字化转型数据库-CDTD

1.数据库简介

上市公司数字化转型数据库（*Digital Transformation Database of Listed Company, CDTD*）是综合已有研究成果以及数字经济领域的关键词汇，整理了 100 多个与数字化相关的关键词，从数字化转型的衡量方式、关联方数字化和外部关注数字化三个维度对上市公司数字化进行了全面展示。

2.数据库特色

- 提供全面的上市公司数字化转型数据

CDTD 不仅从文本角度衡量了上市公司的数字化转型，还从其他多个维度对数字化转型进行展示。

3.典型数据举例

无形资产数字化

股票代码	会计年度	数字化技术无形资产	无形资产	数字化技术无形资产占比
000001	2012	552707000.00	5878379000.00	9.4024%
000004	2012	3717.00	10620297.79	0.0350%

4.数据库内容

数据模块	数据表	数据字段
上市公司数字化转型衡量	年报文本数字化	股票代码、会计年度、数字化转型词汇、词频统计、出现该词汇的句子数
	MD&A 文本数字化	股票代码、会计年度、数字化转型词汇、词频统计、出现该词汇的句子数
	无形资产数字化	股票代码、会计年度、数字化技术无形资产、无形资产、数字化技术无形资产占比
上市公司关联方数字化	集团公司数字化	股票代码、会计年度、参控股公司名称、参控股公司总数、数字化相关参控股公司数量、数字化占比
	供应商数字化	股票代码、公司简称、会计年度、供应商名称、公司编号、公司股票代码、数字化转型词汇、词频统计、出现该词汇的句子数
	客户数字化	股票代码、公司简称、会计年度、客户名称、公司编号、公司股票代码、数字化转型词汇、词频统计、出现该词汇的句子数
	投资者问答数字化	股票代码、公司简称、会计年度、提问回复总数、数字化相关提问数量、数字化相关回复数量、数字化提问比例、数字化回复比例
上市公司外部关注数字化	分析师报告数字化	股票代码、会计年度、数字化转型词汇、词频统计、出现该词汇的句子数、相关报告数、报告总数
	新闻报道数字化	股票代码、会计年度、数字化转型词汇、词频统计、相关报道次数

5.数据范围

数据模块	数据起始时间
全部模块	2007 年

(三十五) 上市公司公告统计数据库-CCAS

1.数据库简介

上市公司公告统计数据库 (*Announcement Statistics Database of Chinese Listed Company, CCAS*) 是根据上市公司在证券交易所披露的公告整理统计而成的专业数据库。**CCAS** 共有全部公告、一般公告、定期报告、股权股本、交易提示、增发配股、重大事项等共七张表,统计了自 2001 年以来上市公司发布各类公告的次数,并将其归类统计。希望 **CCAS** 的推出能为学者研究提供更多数据支持。

2.数据库特色

- 数据全面

CCAS 对上市公司自 2001 年以来发布的公告进行了详细而全面的归类并统计其次数,有助于减少学者的数据搜集时间和成本。

3.典型数据举例

重大事项

股票代码	会计年度	利润分配	收购兼并	关联交易	抵押借贷	担保事项	投资理财	高管变动
000008	2012		5	19		2	2	1
000009	2012	1	1	2		5	2	1
000010	2012	1			1	5	1	1

4.数据库内容

数据模块	数据表	数据字段
全部公告	全部公告	股票代码、会计年度、全部公告
一般公告	一般公告	股票代码、会计年度、股东大会、董事会公告、监事会公告、权证公告、独立董事意见、H 股公告、工作述职报告、股东回报规划、法律纠纷、问询函及回复、公司制度文件、基本信息变更、中介机构报告、投资者关系活动记录……
定期报告	定期报告	股票代码、会计年度、年度报告、半年度报告、一季度报告、三季度报告、业绩预告、业绩快报、补充更正、财务报告
股权股本	股权股本	股票代码、会计年度、股本变动、股权变动、股份增持、股份减持、股权回购、质押冻结、股权分置改革
交易提示	交易提示	股票代码、会计年度、停牌公告、复牌公告、交易异动、澄清公告、风险提示、特别处理、暂停上市、终止上市、恢复上市
增发配股	增发配股	股票代码、会计年度、增发预案、增发说明书、增发获准、增发发行、增发结果、增发上市、其他增发公告、配股预案、配股说明书、配股获准、配股发行、配股结果、配股上市、其他配股公告
重大事项	重大事项	股票代码、会计年度、利润分配、股权激励、资产重组、收购兼并、资产减值、合同协议、资金投向、关联交易、抵押借贷、担保事项、投资理财、高管变动、政策影响、违规违纪、重大损失、月度经营情况、其他重大事项

5.数据范围

数据模块	数据起始时间
全部模块	2001 年

（三十六）上市公司 ESG 评级数据库-ESGR

1.数据库简介

上市公司 ESG 评级数据库 (*ESG Rating Database of Listed Company, ESG-R*) 基于 *ISO 26000*、*GRI Standards*、*SASB Standards* 等国际 ESG 披露标准和国内外知名 ESG 数据库的设计思路，并结合中国 ESG 信息披露的相关政策，构建了独特的中国企业的 ESG 评分体系。*ESG-R* 覆盖全部 A 股上市公司，通过对海量数据的分析、整合、与处理，呈现了 2007 年以来上市公司年度的 ESG 评分与排名情况，为学者的相关研究提供数据支持。

2.数据库特色

- 长时间跨度

ESG-R 涵盖 2007 年以来的上市公司年度数据

- 全范围覆盖

ESG-R 覆盖了 A 股所有上市（不包括已退市公司）

- 海量数据支持

基于 CNRDS 平台相关数据，以及海量底层原始数据

- 专业数据处理

利用专业手段对原始数据及计算流程进行处理

- 中国特色的指标

指标构建结合我国国情

3.典型数据举例

上市公司 ESG 评级

股票代码	公司简称	会计年度	ESG 得分	ESG 排名	E 得分	E 排名	S 得分	S 排名	G 得分	G 排名
000001	平安银行	2012	28.68331	499	5.2807	1054	29.968	508	44.620	58
000002	万科 A	2012	47.12574	31	5.3619	1028	65.449	1	40.796	123
000006	深振业 A	2012	26.29568	641	1.6187	1761	35.340	256	38.797	157

4.数据库内容

数据模块	数据表	数据字段
上市公司 ESG 评级	上市公司 ESG 评级	股票代码、公司简称、会计年度、ESG 得分、ESG 排名、E 得分、E 排名、S 得分、S 排名、G 得分、G 排名

5.数据范围

数据模块	数据起始时间
上市公司 ESG 评级	2007 年

(三十七) 公司上市前专利数据库-PCID

1. 数据库简介

公司上市前专利数据库 (*Innovation Database of Pre-IPO Companies, PCID*) 是结合国家知识产权局专利数据, 基于研究上市公司上市前的专利申请和授权情况的目的而开发的专业数据库。*PCID* 包括公司上市前专利申请与获得、公司上市前绿色专利申请与获得以及公司上市前数字经济专利申请与获得三个模块。其中, 公司上市前专利申请与获得模块统计了公司上市前专利申请与获得情况、专利申请后获得情况, 以及专利获得年限情况; 公司上市前绿色专利申请与获得和公司上市前数字经济专利申请与获得模块分别统计了专利分类号符合《国际专利分类表》中绿色专利定义、以及符合国家统计局 2021 年发布的《数字经济及其核心产业统计分类》中数字经济专利定义的上市公司上市前专利情况。

2. 数据库特色

- 多维度展示上市公司上市前的专利情况

PCID 统计了上市公司上市前历年专利申请与获得的情况、专利申请后的授权情况、各类专利的授权年限, 以及公司上市前的绿色专利及数字经济专利申请与授权情况数据。

3. 典型数据举例

(1) 公司上市前专利申请情况

股票代码	会计年度	申请时间	当年独立申请的发明数量	当年独立申请的实用新型数量	当年独立申请的外观设计数量	当年联合申请的发明数量	当年联合申请的实用新型数量	当年联合申请的外观设计数量
603183	2007	上市前	0	0	0	1	1	0
603183	2012	上市前	3	10	0	0	0	0
603183	2013	上市前	3	0	0	0	1	0
603183	2014	上市前	1	3	0	0	0	0
603183	2015	上市前	1	1	0	0	0	0
603183	2016	上市前	2	1	0	0	0	0

(2) 公司上市前专利获得年限情况

股票代码	会计年度	专利获得	专利获得_平均年限	发明专利获得	发明专利获得_平均年限	发明专利获得_公开平均年限	实用新型获得	实用新型获得_平均年限	外观设计获得	外观设计获得_平均年限
300651	2011	4	0.597	0			4	0.597	0	
300651	2012	14	4.393	3	3.130	0	4	0.625	7	0.638
300651	2013	6	2.618	4	2.020	0	2	0.597	0	
300651	2014	37	1.235	0			2	0.647	35	0.587
300651	2015	8	2.641	3	2.167	0	5	0.474	0	
300651	2016	6	2.546	1	1.700	1.419	2	0.418	3	0.426

4.数据库内容

数据模块	数据表	数据字段
公司上市前专利申请与获得	公司上市前专利申请情况	股票代码、会计年度、申请时间、当年独立申请的发明数量、当年独立申请的实用新型数量、当年独立申请的外观设计数量、当年联合申请的发明数量、当年联合申请的实用新型数量、当年联合申请的外观设计数量
	公司上市前专利获得情况	股票代码、会计年度、获得时间、当年独立获得的发明数量、当年独立获得的实用新型数量、当年独立获得的外观设计数量、当年联合获得的发明数量、当年联合获得的实用新型数量、当年联合获得的外观设计数量
	公司上市前专利申请后获得情况	股票代码、会计年度、更新日期、专利申请个数、专利申请_获得个数、专利申请_当年获得个数、专利申请_第二年获得个数、专利申请_第三年获得个数、专利申请_第四年获得个数、发明申请个数、发明申请_获得个数、发明申请_当年获得个数、发明申请_第二年获得个数、发明申请_第三年获得个数、发明申请_第四年获得个数、实用新型申请个数、实用新型申请_获得个数、实用新型申请_当年获得个数、实用新型申请_第二年获得个数、实用新型申请_第三年获得个数、实用新型申请_第四年获得个数、外观设计申请个数、外观设计申请_获得个数、外观设计申请_当年获得个数、外观设计申请_第二年获得个数、外观设计申请_第三年获得个数、外观设计申请_第四年获得个数
	公司上市前专利获得年限情况	股票代码、会计年度、专利获得、专利获得_平均年限、发明专利获得、发明专利获得_平均年限、发明专利获得_公开平均年限、实用新型获得、实用新型获得_平均年限、外观设计获得、外观设计获得_平均年限
公司上市前绿色专利申请与获得	公司上市前绿色专利申请情况	股票代码、会计年度、申请时间、当年独立申请的绿色发明数量、当年独立申请的绿色实用新型数量、当年联合申请的绿色发明数量、当年联合申请的绿色实用新型数量
	公司上市前绿色专利获得情况	股票代码、会计年度、获得时间、当年独立获得的绿色发明数量、当年独立获得的绿色实用新型数量、当年联合获得的绿色发明数量、当年联合获得的绿色实用新型数量
公司上市前数字经济专利申请与获得	公司上市前数字经济专利申请情况	股票代码、会计年度、申请时间、当年独立申请的数字经济发明数量、当年独立申请的数字经济实用新型数量、当年联合申请的数字经济发明数量、当年联合申请的数字经济实用新型数量
	公司上市前数字经济专利获得情况	股票代码、会计年度、获得时间、当年独立获得的数字经济发明数量、当年独立获得的数字经济实用新型数量、当年联合获得的数字经济发明数量、当年联合获得的数字经济实用新型数量

5.数据范围

数据模块	数据起始时间
公司上市前专利申请与获得	1992 年
公司上市前绿色专利申请与获得	1993 年
公司上市前数字经济专利申请与获得	1992 年

(三十八) 上市公司股权质押数据库-EPCD

1. 数据库简介

上市公司股权质押数据库 (*Equity Pledge of Listed Companies Database, EPCD*) 提供了上市公司股权质押信息以及质押股东的基本信息。*EPCD* 包括股权质押情况和质押股东基本信息两个模块, 其中, 股权质押情况模块提供了上市公司股权质押的相关数据, 包括股东排名、性质、是否为上市公司实际控制人等股东信息, 质押方、质押股份、质押比例等质押详情信息, 以及股东累计质押股份、累计质押股比例等累计质押信息。质押股东基本信息模块提供了进行质押操作的机构股东的工商信息数据。

2. 数据库特色

- 全面的股权质押数据

EPCD 不仅提供了不同质押事由下的具体质押信息, 还提供了截至该次股东累计质押股权的股数、占股东所持总股数比例等数据。

- 详细的股东信息

EPCD 提供了股东排名、股东性质和是否为实际控制人等质押股东信息, 此外, 还提供了机构质押股东详细的工商信息, 对于研究股权质押行为与股东性质、股东持股比等因素间的关系具有重要意义。

3. 典型数据举例

股权质押情况

股票代码	年份	公告日期	股东名称	股东类型	股东性质	股东排名	持股数量	持股比例	是否实际控制人	是否控股股东	是否质押回购
002414	2019	2019-07-27	黄立	个人	民营	2	267187500	28.53	是	否	否
002607	2022	2022-12-06	王振东	个人	民营	3	760775418	12.34	否	否	否
300252	2017	2017-05-17	黄昌华	个人	民营	1	151656675	34.12	是	否	否
400175	2016	2016-01-21	郝镇熙	个人	民营					否	否
832056	2022	2022-11-01	周建彬	个人	民营					否	
002607	2022	2022-04-29	鲁忠芳	个人	民营	2	921148498	14.94	是	是	否

4.数据库内容

数据模块	数据表	数据字段
股权质押情况	股权质押情况	股票代码、年份、公告日期、股东名称、股东类型、股东性质、股东排名、持股数量、持股比例、是否实际控制人、是否控股股东、质押方、质押方类型、是否质押回购、质押股份性质、解押状态、质押详情_质押股数、质押详情_占所持股比例、质押详情_占总股本比例、质押起始日期、质押截止日期、解押日期、累计质押_质押股数、累计质押_占所持股比例、累计质押_占总股本比例、质押事由、证监会行业
股东基本信息	股东基本信息	企业名称、登记状态、法定代表人、注册资本、成立日期、核准日期、所属省份、所属城市、所属区县、参保人数、企业类型、所属行业、曾用名、企业地址、经营范围

5.数据范围

	数据起始时间
股权质押情况	2001 年
股东基本信息	截面数据

上市公司新闻舆情

（一）中国上市公司财经新闻数据库-CFND

1. 数据库简介

中国上市公司财经新闻数据库 (*Financial News Database of Chinese Listed Companies, CFND*) 是国内首个采用人工智能算法来采集、整理和分析上市公司财经新闻的大数据数据库。本数据的数据来源涵盖 400 多家网络媒体和 500 多家报纸刊物。海量的上市公司财经新闻数据成为本数据库的核心竞争力，以海量数据为基础，通过数据清洗、专家判断等手段，结合人工智能算法为学者提供上市公司财经新闻基本信息、财经新闻量化统计信息以及新闻相似度等。

CFND 包括网络财经新闻和报刊财经新闻 2 个模块，7 张表。网络财经新闻包括了来自 400 多家重要网络媒体的新闻报道数据，其中，最主要的是包括二十家主流网络财经媒体的新闻报道，如：和讯网、新浪财经、东方财富网、腾讯财经、网易财经、凤凰财经、中国经济网、搜狐财经、金融界、华讯财经、FT 中文网、全景网、中金在线、中国证券网、证券之星、财新网、澎湃新闻、第一财经、21CN 财经频道、财经网。这二十家网络财经媒体不仅在财经新闻报道数据和质量上属于国内前列，而且也是投资者经常浏览和关注的财经网站，因此其报道的新闻具有重要价值。此外，为了丰富数据来源，提供更多有用的新闻信息，该模块还包括了其他 400 余家大型重要网站、行业网站或地方性网站。该模块的新闻数据量为从 2002 年开始至今全部财经新闻，囊括了个股新闻、宏观经济报道、行业报道等，截止至 2018 年底，累计达 4000 多万条原始新闻。

报刊财经新闻包括了来自 500 多家重要报纸媒体的新闻数据，其中，最主要的是包含了研究常用的八大主流财经报纸：中国证券报、上海证券报、第一财经日报、21 世纪经济报道、中国经营报、经济观察报、证券日报和证券时报。这八大财经报纸不仅具有财经新闻报道及时、质量高、影响力大等特点，还是国内经济、管理和商学研究媒体问题时经常用到的数据来源。同时，为了体现新闻报道的全面性，该模块还包括其他 500 余家重要报纸刊物，主要包括重要性中央报纸，地方性晨报、日报、晚报以及其他财经报刊。该模块的原始新闻数量为从 1994 年开始至今约 4000 多万条新闻报道。

无论是网络财经新闻还是报刊财经新闻，为了匹配得到上市公司的相关新闻，我们在几千万条原始新闻的基础上，经过数据清洗与匹配、新闻情感判断、相似新闻计算等方法，得到了最终较为科学的结果。

2. 数据库特色

- 新闻报道数量巨大

CFND 不仅涵盖网络上的主流财经网站、国家级网站、地方网站、行业网站，还包括主流财经报刊、国家级报刊等重要报刊新闻，总共 400 多家网络媒体和 500 多家报纸刊物，几千万条新闻数据，可以充分满足研究需要。

- 新闻情感识别准确

CFND 使用计算机领域成熟的文本情感判断方法，文章正、中、负面语气识别率较高，样本内的准确率高达 85%，可以较好地捕捉上市公司好、坏消息。

3. 典型数据举例

网络新闻量化统计（按自然日）

股票代码	统计日期	全部正面新闻数	全部中性新闻数	全部负面新闻数	原创正面新闻数	原创中性新闻数	原创负面新闻数
000002	2016-05-25	17	28	6	3	3	3
000002	2016-05-26	14	33	16	2	4	6
000002	2016-05-27	15	34	32	4	3	2
000002	2016-05-28	2	6	10	1	0	3
000002	2016-05-29	9	1	9	0	0	1

4. 数据库内容

数据模块	数据表	数据字段
网络财经新闻	网络财经新闻基本信息	股票代码、公司简称、行业名称、行业代码、新闻 ID、报道时间、报道媒体、报道媒体所在地区、来源媒体、新闻情感、是否原创、采集网址、正文句子数、标题是否提及、正文提及句子数、正文提及次数、提及全部公司数、提及全部公司句子数、全部公司提及次数
	网络财经新闻量化统计（按自然日）	股票代码、公司简称、统计日期、标题出现该公司的新闻总数、内容出现该公司的新闻总数、正面新闻数（全部）、中性新闻数（全部）、负面新闻数（全部）、正面新闻数（原创）、中性新闻数（原创）、负面新闻数（原创）
	网络财经新闻量化统计（按交易时间）	股票代码、公司简称、统计日期、时间段、星期时间、是否交易时间段、标题出现该公司的新闻总数、内容出现该公司的新闻总数、正面新闻数（全部）、中性新闻数（全部）、负面新闻数（全部）、正面新闻数（原创）、中性新闻数（原创）、负面新闻数（原创）
	网络财经新闻相似度	新闻 1_ID、新闻 1_标题、新闻 1_所属公司代码、新闻 2_ID、新闻 2_标题、新闻 2_所属公司代码、新闻相似度
报刊财经新闻	报刊财经新闻基本信息	股票代码、公司简称、行业名称、行业代码、新闻 ID、报道时间、报道媒体、报道媒体所在地区、新闻情感、是否原创、正文句子数、标题是否提及、正文提及句子数、正文提及次数、提及全部公司数、提及全部公司句子数、全部公司提及次数
	报刊财经新闻量化统计	股票代码、公司简称、统计日期、标题出现该公司的新闻总数、内容出现该公司的新闻总数、正面新闻数（全部）、中性新闻数（全部）、负面新闻数（全部）、正面新闻数（原创）、中性新闻数（原创）、负面新闻数（原创）
	报刊财经新闻相似度	新闻 1_ID、新闻 1_标题、新闻 1_所属公司代码、新闻 2_ID、新闻 2_标题、新闻 2_所属公司代码、新闻相似度

5. 数据范围

数据模块	数据起始时间
网络财经新闻	2002 年
报刊财经新闻	1994 年

(二) 中国上市公司股吧评论数据库-GUBA

1.数据库简介

股吧评论数据库 (*Stock Comments Database, GUBA*) 是基于我国网络股吧论坛中不同类型的帖子评论进行统计研究而建立的专业性财经文本数据库。该数据库主要根据中国最大的股吧论坛网民对上市公司及其他股吧帖子进行文本分析、数量统计等, 从中得到维度众多、数量丰富的研究数据。

GUBA 分为上市公司股吧帖子统计和其他股吧帖子统计 2 个模块, 囊括上市公司股吧帖子统计(自然日)、上市公司股吧帖子统计(交易时间)、其他股吧帖子统计(自然日)和其他股吧帖子统计(交易时间) 4 张表。上市公司股吧帖子统计(自然日)和其他股吧帖子统计(自然日)根据自然日时间(0:00-24:00)对一天中的发帖总量、阅读总量和评论总量进行统计; 上市公司股吧帖子统计(交易时间)和其他股吧帖子统计(交易时间)则根据交易时间统计了不同时间段内的发帖情况, 这种计算方法可以直接反映交易时间内和非交易时间内投资者情绪对股市的影响。*GUBA* 采用机器学习方法对每个帖子的正面、负面和中性情感进行判断, 并统计当天该上市公司所在股吧的正面、负面和中性帖子总量, 这大大丰富了学者的研究维度。*GUBA* 数据库包括 2008 年以来至今所有股吧帖子的统计。

2.数据库特色

- 唯一且全面

GUBA 是目前唯一一个对股吧论坛帖子进行分析统计的数据库, *GUBA* 不仅统计了大量的帖子数据, 还从不同时间维度进行分析, 这可为拓展研究广度和深度提供数据支持。

3.典型数据举例

股吧文章统计 (自然日)

股票代码	统计日期	帖子总量	正面帖子量	负面帖子量	中性帖子量	阅读数	评论数
000665	2016-07-14	44	12	12	20	24117	42
000665	2016-07-15	26	5	7	14	18997	23
000665	2016-07-16	14	6	2	6	7341	1

4.数据库内容

数据模块	数据表	数据字段
上市公司股吧帖子统计	上市公司股吧帖子统计(自然日)	股票代码、公司简称、统计日期、帖子总量、正面帖子量、负面帖子量、中性帖子量、阅读数、评论数
	上市公司股吧帖子统计(交易时间)	股票代码、公司简称、统计日期、时间段、星期时间、是否交易时间段、帖子总量、正面帖子量、负面帖子量、中性帖子量、阅读数、评论数
其他股吧帖子统计	其他股吧帖子统计(自然日)	股吧名称、统计日期、帖子总量、正面帖子量、负面帖子量、中性帖子量、阅读数、评论数
	其他股吧帖子统计(交易时间)	股吧名称、统计日期、时间段、星期时间、是否交易时间段、帖子总量、正面帖子量、负面帖子量、中性帖子量、阅读数、评论数

5.数据范围

数据模块	数据起始时间
全部模块	2008 年

（三）网络搜索指数数据库-WSVI

1.数据库简介

网络搜索指数数据库 (*Web Search Volume Index of Chinese Listed Companies, WSVI*) 以各种网络搜索指数为基础, 综合新闻舆情等信息, 可以反映网民情绪、公司搜索热度等行为, 是衡量上市公司关注度及其变化情况的关键指标。*WSVI* 数据库收集和整理了 2011 年以来, 以我国上市公司的股票代码、公司简称、公司全称为关键字的网络搜索指数数据。

2.数据库特色

- 唯一

目前市场上尚无提供网络搜索指数的数据库, *WSVI* 数据库是市场上唯一比较完善的该领域研究数据库。

- 全面

WSVI 数据库不仅包括全部公司关键字的网络搜索指数加总值, 为了研究需要, 该库还包括以股票代码为关键字的搜索值。

3.典型数据举例

上市公司网络搜索指数 (全国)

股票代码	行业代码	指数值 (股票代码)	指数值 (全部)	关键字个数	统计日期
000001	J66	6702	18273	2	2011-01-04
000001	J66	6371	17234	2	2011-01-05
000001	J66	6236	17135	2	2011-01-06
000001	J66	6658	16447	2	2011-01-07

4.数据库内容

数据模块	数据表	数据字段
上市公司网络搜索指数 (全国)	上市公司网络搜索指数 (全国)	股票代码、行业代码、指数值 (股票代码)、指数值 (全部)、关键字个数、统计日期

5.数据范围

数据模块	数据起始时间
上市公司网络搜索指数 (全国)	2011 年

（四）中国上市公司社交媒体数据库-CSMD

1.数据库简介

中国上市公司社交媒体数据库 (*Social Media Database of Chinese Listed Companies, CSMD*) 是基于对上市公司的社交媒体信息调查研究而建立的专业数据库。该库涵盖了券商微信公众号文章信息和财经类微信公众号文章信息。前者主要是指中国主要券商在其公众号上发布的对上市公司的推荐文章信息，后者主要指民间财经类公众号对上市公司的报道和分析等。

2.数据库特色

- 前沿

微信信息近期成为会计金融学术界研究热点，但目前还没有相应的数据库。**CSMD** 紧跟学术研究前沿，采用数据挖掘方法对这类信息进行文本分析、数量统计和分类规整等，可以为学术界提供全面、准确的财经文本数据。

3.典型数据举例

券商微信公众号

文章标题	发布时间	获取时间	来源公众号名称	文章字数	标题包含公司	正文包含公司
【重点推荐】恒逸石化 (000703) 化纤估值洼地，“供应链金融”增色	2015-06-01	2016-08-24	国泰君安证券研究	675	000703	000703
【重点推荐】中兴通讯 (000063) 中兴 2014: 从困顿反转到再次腾飞	2014-04-14	2016-08-24	国泰君安证券研究	698	000063	000063
【地产行业】韩思怡: 万科 A (000002) B 转 H 获证监会核准, 比价效应对 A 股利好明显	2014-03-04	2016-08-29	申万宏源研究	985	000002	000002 600340

4.数据库内容

数据模块	数据表	数据字段
微信公众号 财经新闻	券商微信公众号	文章标题、文章内容、发布时间、获取时间、来源公众号名称、文章字数、标题包含公司、正文包含公司、文章包含的所有公司
	财经微信公众号	文章标题、文章内容、发布时间、获取时间、来源公众号名称、文章字数、标题包含公司、正文包含公司、文章包含的所有公司

5.数据范围

数据模块	数据起始时间
全部模块	2012 年

（五）上市公司并购新闻数据库-MAND

1.数据库简介

上市公司并购新闻数据库 (*Mergers and Acquisition News Database, MAND*) 是基于上市公司新闻库, 对新闻事件类型进行判断识别后提取的数据库。市场上出现的并购重组新闻或者类似的小道消息、传闻等, 都会影响上市公司股票的走势, 因此对上市公司并购重组新闻的识别以及真实性判断具有重大意义。本数据库针对国内上市公司的新闻数据, 获取并购重组事件发生前后的新闻, 对并购新闻数据进度深度挖掘, 抽取出并购各方名称、并购金额和转让股权比例等多个重要字段, 同时判断该新闻是否属于小道消息或相关传闻。

2.数据库特色

- 精确的并购新闻提取

MAND 从海量上市公司新闻中准确提取出与并购重组相关的新闻, 同时使用计算机处理和人工交叉核对流程, 识别精度较高。

- 丰富的并购新闻字段

为了满足学者对并购领域的相关研究需求, *MAND* 展示了较为丰富的并购相关字段。

3.典型数据举例

上市公司并购新闻

股票代码	新闻日期	新闻标题	新闻来源	转让比例	并购价格
600533	2015/12/2	栖霞建设拟 13.84 亿收购河北银行 8.78%股份	中国财经时报网	8.78%	13.84 亿元
002464	2016/9/21	转型再升级 金利科技拟 18 亿收购微屏软件完善游戏产业链	证券日报网	93%	18.41 亿元
300282	2014/1/15	汇冠股份拟 10.7 亿元收购旺鑫精密	中国经济网	100%	10.7 亿元
600869	2014/6/25	远东电缆拟 1.8 亿并购艾能电力	中国经济网	51%	1.8 亿元
600378	2014/7/23	增强天科股份控制权 中化资产要约收购 1742 万股	中国经济网	5.86%	2.24 亿元

4.数据库内容

数据模块	数据表	数据字段
上市公司 并购新闻	上市公司 并购新闻	股票代码、公司简称、新闻日期、新闻 ID、新闻来源、买方名称、卖方名称、标的方名称、转让比例、并购价格、货币单位

5.数据范围

数据模块	数据起始时间
全部模块	2007 年

（六）上市公司财务舞弊新闻数据库-FFND

1.数据库简介

上市公司财务舞弊新闻数据库（*Financial Fraud News Database of Listed Companies, FFND*）是基于上市公司财务舞弊的相关媒体报道，从中分离出三类重要媒体报道，包括财务舞弊相关正式公告发布之前的媒体报道、因财务舞弊被立案调查期间的媒体报道以及财务舞弊相关正式公告发布后的媒体报道。具体来说，这三类财务舞弊新闻分别是：（1）媒体质疑类：该新闻是媒体对上市公司存在财务舞弊的质疑，上市公司并没有发布公告也没有被监管机构立案调查，只是媒体等民间机构对上市公司的一种纯粹质疑。（2）立案调查类：上市公司发生财务丑闻，监管机构介入调查或者做出处罚。（3）发布公告类：上市公司发布公告承认存在财务舞弊，或因收到处罚通知而发布公告，或者对财务舞弊的传言进行澄清。

2.数据库特色

- 精确的财务舞弊新闻识别和分类

FFND 采用机器学习和人工识别的方法，从海量新闻数据中准备识别出了上市公司的三类财务舞弊新闻，数据质量较高。

3.典型数据举例

上市公司财务舞弊新闻

股票代码	新闻类别	文章标题	发布时间	新闻来源
000040	立案调查类	宝安地产深陷前任“虚假陈述”索赔案	2014-09-04	网易
000048	立案调查类	康达尔涉嫌虚假陈述 遭证监会立案调查	2012-12-03	网易
002131	发布公告类	业绩变脸解释站不住脚 利欧股份或涉嫌虚假陈述	2012-07-18	和讯网
002200	发布公告类	绿大地独董质疑公司财务造假	2011-10-28	金融界
601311	媒体质疑类	骆驼集团涉嫌财务造假 可能虚构6个亿以上收入	2011-04-17	网易
601388	媒体质疑类	怡球资源财务数据存蹊跷 涉嫌虚增利润数百万	2012-11-09	网易
601566	媒体质疑类	九牧王存货跌价计提过低 涉嫌虚增利润	2013-07-05	财经网

4.数据库内容

数据模块	数据表	数据字段
上市公司财务舞弊新闻	上市公司财务舞弊新闻	股票代码、公司简称、新闻类别、发布时间、新闻来源

5.数据范围

数据模块	数据起始时间
全部模块	2001年

（七）上市公司高管新闻数据库-CEND

1.数据库简介

上市公司高管新闻数据库 (*Executives News Database of Listed Company, CEND*) 是基于中国财经大数据研究系统 (*CFBD*) 记录的来自 400 多家网络新闻媒体和 500 多家报纸新闻媒体所获得的 9000 多万条原始新闻数据, 通过确定公司高管姓名、上市公司关键词和相关规则等, 使用机器学习和人工检查分析新闻标题、新闻内容, 与相关领域专家讨论, 从而识别出与上市公司高管相关的新闻报道。*CEND* 包括高管网络新闻和高管报纸新闻两张表, 前者主要提取自上市公司网络新闻报道, 后者提取自上市公司报纸新闻报道。这些来源网站与报纸均包括了财经类较重要的来源。

2.数据库特色

- 精确的高管新闻匹配

CEND 除从几千万条新闻数据中较精确的匹配出了与上市公司高管相关的新闻, 并从多个维度列出该高管新闻的相关字段, 例如新闻情感、相似新闻数量、新闻内容等。

3.典型数据举例

上市公司高管报刊新闻

股票代码	高管姓名	性别	新闻来源	新闻日期	新闻标题	相似新闻数量	新闻情感
000670	史浩樑	男	证券时报	2014-05-22	深耕智能产品芯片设计 掘金互联网时代大潮	4	1
002405	孙玉国	男	证券时报	2014-05-23	腾讯携四维图新布局智能 地图	2	0
002405	吴劲风	男	证券时报	2014-05-23	腾讯携四维图新布局智能 地图	2	0
600645	夏亮	男	证券时报	2014-05-23	中源协和重组方案确定 8 亿收购执诚生物	0	1

4.数据库内容

数据模块	数据字段
上市公司高管网络新闻	股票代码、公司简称、高管编码、高管姓名、职务、性别、报道媒体、报道时间、新闻 ID、新闻情感、相似新闻数量
上市公司高管报刊新闻	股票代码、公司简称、高管编码、高管姓名、职务、性别、报道媒体、报道时间、新闻 ID、新闻情感、相似新闻数量

5.数据范围

数据模块	数据起始时间
上市公司高管网络新闻	2000 年
上市公司高管报刊新闻	2004 年

上市公司文本信息

(一) 年报文本语气数据库-ARTD

1. 数据库简介

年报文本语气数据库 (*Annual Report's Tone Database, ARTD*) 基于上市公司 2007 年以来的年报文本信息, 根据已有中文语调研究成果, 采用多种方法计算上市公司年报文本的语气, 并建立消极词汇数、积极词汇数、年报语调等多个字段, *ARTD* 采用包括 LM 词典和台湾大学情感词典在内的多种前沿词典和方法计算年报文本语气, 为该领域的深度研究提供了数据支撑。

2. 数据库特色

- 准确和权威的情感判断词典

ARTD 制作过程中参考了发表在权威期刊上的多篇文献, 并借鉴这些已有研究成果使用的词典, 对年报语调进行了多字段展示。

- 数据唯一、全面

随着财经类文本语气研究成为热点, 为了减少学者搜集和整理数据的成本, 提供了多年度完整的年报文本信息, 为该领域研究提供了数据支持。

3. 典型数据举例

年报文本语气

股票代码	会计年度	LM 词典积极词汇数	LM 词典消极词汇数	台大词典积极词汇数	台大词典消极词汇数	基于 LM 词典的 TONE1	基于台大词典的 TONE
000001	2008	3828	3529	4021	1543	0.00479	0.44536
000001	2009	4111	3897	4467	1607	0.00321	0.47086
000001	2010	4535	4064	4896	1584	0.00683	0.51111
000001	2011	5772	4817	6334	1944	0.01108	0.53032
000001	2012	5216	4963	6116	1999	0.00307	0.50733
000001	2013	4069	3365	4752	1311	0.01230	0.56754

4. 数据库内容

数据模块	数据表	数据字段
年报文本语气	年报文本语气	股票代码、会计年度、年报字数、年报词汇数、年报句子数、LM 词典积极词汇数、LM 词典消极词汇数、台大词典积极词汇数、台大词典消极词汇数、基于 LM 词典的 TONE1、基于 LM 词典的 TONE2、基于台大词典的 TONE

5. 数据范围

数据模块	数据起始时间
年报文本语气	2007 年

(二) 管理层讨论与分析数据库-CMDA

1.数据库简介

管理层讨论与分析数据库(*Management's Discussion and Analysis Database, CMDA*)是针对上市公司财务报告中管理层讨论与分析模块而建立的专业数据库, 数据库涵盖了从 1999 年至今所有上市公司的年报、半年报和季报中与经营讨论与分析相关的数据资料, 并采用人工智能算法对文本内容涉及的正面、负面词汇进行判断识别。其中, 正面词汇和负面词汇的判断基础主要是根据 *Tim Loughran and Bill McDonald (2011)* 的英文正负面词典翻译为中文词典, 再结合中文语境进行扩充和完善得到。*CMDA* 可以有效帮助使用者分析上市公司所披露和预期的经营文本数据。

2.数据库特色

- 全面的管理层讨论与分析内容

CMDA 包括上市公司发布的季报、半年报和年报中的管理层讨论与分析内容, 可以满足研究的需求。

- 精确的正负面词汇统计

CMDA 基于 *Tim Loughran and Bill McDonald* 词典, 结合中文语境设定的正负面词汇字典, 既保证了词性判断的权威性, 又符合了中文的特殊语法结构。

3.典型数据举例

管理层经营讨论和分析

股票代码	会计年度	经营分析时间	经营讨论与分析内容	正面词汇数量	负面词汇数量	句子数量
002160	2016	2016-12-31	公司近 3 年(包括本报告期)的普通股股利分配方案(预案)……	18	3	6
300401	2016	2016-12-31	2016 年度,对公司来说是很不平凡的一年……	27	9	6
300399	2016	2016-03-31	报告期内,行业移动信息服务及保险产品……	38	3	7
603718	2015	2015-06-30	报告期内总体经营情况公司主营业务为动物疫苗的研发……	98	31	29

4.数据库内容

数据模块	数据表	数据字段
管理层经营讨论和分析	管理层经营讨论和分析	股票代码、公司简称、会计年度、经营分析时间、经营讨论与分析内容、正面词汇数量、负面词汇数量、句子数量、文字数量、词汇数量

5.数据范围

数据模块	数据起始时间
管理层经营讨论和分析	1999 年

（三）上市公司业绩说明会数据库-ECCD

1.数据库简介

上市公司业绩说明会数据库 (*Earnings Communication Conferences Database of Chinese Listed Companies, ECCD*) 是基于上市公司每年在交易所发布的业绩说明会上的提问与回答内容的文本信息进行整理和分析生成的专业文本分析数据库。**ECCD** 涵盖了从 2008 年至今的业绩说明会情况, 采用计算机技术对提问和回复内容涉及的正负面词汇以及语调进行识别、判断; 其中, 文本内容的正面词汇和负面词汇的判断基础主要是根据 *Tim Loughran and Bill McDonald* (2011) 的英文正负面词典翻译的中文词典结合中文语境进行扩充和完善得到。

2.数据库特色

- 精确的正负面词汇统计

ECCD 的词性判断基于 *Tim Loughran and Bill McDonald* 词典, 并结合中文语境设定的正负面词汇字典, 既保证了词性判断的权威性, 又符合了中文的特殊语法结构。

3.典型数据举例

业绩说明会问答文本分析

股票代码	会计年度	提问内容	提问时间	回答正面词汇数比例	回答负面词汇数比例	回答净正面语调
000002	2013	公司 13 年乃至今后几年最大的挑战是什么?	2013-08-09	0.0417	0.2917	-0.7500
000002	2013	贵公司相对其它同行业的公司优越在那里?	2013-08-09	0.1271	0.0508	0.4286
000002	2013	在管理经营中, 董事长最关注的问题是什么?	2013-08-09	0.0345	0.1724	-0.6667
000002	2013	请问 5 月份和半年业绩预报可能在 6 月 13 日前出吗?	2014-05-29	0.1765	0.0588	0.5000

4.数据库内容

数据模块	数据表	数据字段
业绩说明会概况	业绩说明会概况	股票代码、公司简称、会计年度、标题、举办时间、嘉宾姓名、嘉宾职务、嘉宾简介
业绩说明会问答文本分析	业绩说明会问答文本分析	股票代码、会计年度、问题序号、提问内容、提问时间、回答人、回答时间、回答内容、正面词汇数量、负面词汇数量、正面词汇数比例、负面词汇数比例、净正面语调

5.数据范围

数据模块	数据起始时间
业绩说明会概况	2016 年
业绩说明会问答文本分析	2008 年

（四）中国上市公司网上路演数据库-CNRD

1.数据库简介

中国上市公司网上路演数据库（*Net Roadshow Database of Chinese Listed Companies, CNRD*）提供了我国上市公司在 *IPO* 过程中的网上路演信息。这些信息主要来自于在主板、中小板和创业板上市的公司，在证券交易所指定的网站披露的路演信息。**CNRD** 包含以下三个数据模块，分别是公司基本信息、参与嘉宾信息以及路演问答信息。其中，公司基本信息包括股票代码、公司简称、所在板块、路演时间、公司简介、投资要点和风险提示等字段；嘉宾信息包括嘉宾姓名、嘉宾所在单位及任职以及其个人简介等字段；问答信息则包括投资者对上市公司的提问、回答的嘉宾以及回答内容等字段。

2.数据库特色

- 数据全面

涵盖上市公司网上路演过程中详尽的嘉宾信息和问答信息，数据字段全面、来源权威。

3.典型数据举例

参与嘉宾信息

股票代码	公司简称	所在板块	嘉宾姓名	嘉宾所在单位名称	嘉宾所在单位任职
002001	新和成	中小板	胡柏藩	浙江新和成股份有限公司	董事长
002001	新和成	中小板	张方治	浙江新和成股份有限公司	总经理
300100	双林股份	创业板	邬建斌	双林股份	董事长
300100	双林股份	创业板	曹靖	双林股份	财务总监
600908	无锡银行	沪市主板	任晓平	无锡农村商业银行股份有限公司	董事长
600908	无锡银行	沪市主板	邵辉	无锡农村商业银行股份有限公司	行长

4.数据库内容

数据模块	数据表	数据字段
公司基本信息	公司基本信息	股票代码、公司简称、所在板块、路演时间、公司简介、投资要点、风险提示
参与嘉宾信息	参与嘉宾信息	股票代码、公司简称、所在板块、嘉宾姓名、嘉宾所在单位、嘉宾所在单位任职、嘉宾个人简介
路演问答信息	路演问答信息	股票代码、公司简称、所在板块、问题、回答、回答人

5.数据范围

数据模块	数据起始时间
全部模块	2004 年

(五) 中国上市公司澄清公告数据库-CCAD

1. 数据库简介

中国上市公司澄清公告数据库 (*Clarification Announcement Database of Chinese Listed Company, CCAD*) 考察了上市公司针对市场传闻等信息发布的澄清公告, 并将其转化为文本格式, 提取其中的关键字段, 从而建立的专业数据库。**CCAD** 包含了诸多重要字段, 其中, 传言日期除了提取自公告内容外, 我们还根据传闻内容, 通过网络搜索等找到相应的传言开始日期; 传言事项则根据传言内容所述的事项, 对这些事项进行科学分类; 传言类型则根据是否负面传言进行分类; 传言媒体和传言标题均提取自公告内容中。

2. 数据库特色

- 准确的文本信息, 丰富的字段内容

CCAD 精确地将 *PDF* 格式的公告内容转为文本信息, 学者可从中挖掘研究信息。此外, 对澄清公告的重要字段进行了提取和整合, 尤其是传言日期均由手工收集整理, 可以体现传言的开始日期。

3. 典型数据举例

上市公司澄清公告

股票代码	公司简称	发布日期	传言日期	传言类型	传播媒体
600535	天士力	2017/3/8	2017/3/17	1	中国证券报
600810	神马股份	2017/2/17	2017/2/17	-1	大众证券报、证券时报
600129	太极集团	2017/2/8	2016/12/17	-1	中国经营报
600275	武昌鱼	2017/1/25	2017/1/24	-1	东方财富网股吧
002560	通达股份	2017/1/25	2017/1/23	1	中国证券网

4. 数据库内容

数据模块	数据表	数据字段
上市公司澄清公告	上市公司澄清公告	股票代码、公司简称、发布日期、公告内容、传言日期、传言类型、传播媒体、传言标题

5. 数据范围

数据模块	数据起始时间
上市公司澄清公告	1996 年

（六）年报关键词研究数据库-AKRD

1.数据库简介

年报关键词研究数据库（*Keyword Research Database of Annual Report, AKRD*）是基于上市公司年报文本而开发的专业数据库。参考权威期刊论文研究成果，*AKRD*总结了八类常用年报关键词出现频次，分别为：企业数字化转型、文本确定程度、风险信息披露程度、企业合作文化、管理者短视主义、文本情感程度、未来事项披露以及研发文本信息。

2.数据库特色

- 精确的数据处理过程

AKRD 严格按照已有学术文献的处理方法，从关键词的选取到计算，均严格确保其准确性。

3.典型数据举例

企业数字化转型

股票代码	会计年度	文本类型	关键词	关键词数量（全部）	关键词数量（去表）
000670	2014	年报	云计算	12	12
300248	2014	年报	移动互联网	68	58
600656	2014	年报	互联网金融	12	6

4.数据库内容

数据模块	数据表	数据字段
企业数字化转型	企业数字化转型	股票代码、会计年度、文本类型、关键词、关键词数量（全部）、关键词数量（去表）
文本确定程度	文本确定程度	股票代码、会计年度、文本类型、关键词、关键词数量（全部）、关键词数量（去表）
风险信息披露程度	风险信息披露程度	股票代码、会计年度、文本类型、关键词、关键词数量（全部）、关键词数量（去表）
企业合作文化	企业合作文化	股票代码、会计年度、文本类型、关键词、关键词数量（全部）、关键词数量（去表）
管理者短视主义	管理者短视主义	股票代码、会计年度、文本类型、关键词、关键词数量（全部）、关键词数量（去表）
文本情感程度	文本情感程度	股票代码、会计年度、文本类型、关键词、关键词数量（全部）、关键词数量（去表）
未来事项披露	未来事项披露	股票代码、会计年度、文本类型、关键词、关键词数量（全部）、关键词数量（去表）
研发文本信息	研发文本信息	股票代码、会计年度、文本类型、关键词、关键词数量（全部）、关键词数量（去表）

5.数据范围

数据模块	数据起始时间
全部模块	2007年

(七) MD&A 关键词研究数据库-MKRD

1.数据库简介

MD&A 关键词研究数据库 (*Keyword Research Database of MD&A, MKRD*) 是基于上市公司管理层讨论与分析文本而开发的专业数据库。参考权威期刊论文研究成果, *MKRD* 总结了八类常用管理层讨论与分析文本关键词出现频次。在对 MD&A 文本进行深度整理和清洗的基础上, *MKRD* 计算这些关键词出现频次时, 以期为学术研究提供多元化的数据支持。

2.数据库特色

- 精确的数据处理过程和全面的数据字段

MKRD 严格按照已有学术文献的处理方法, 从关键词的选取到计算, 均严格确保其准确性。*MKRD* 不仅提供了全文文本的计算结果, 还计算了去除表格后的关键词数量, 便于学者进行多维度研究。

3.典型数据举例

企业数字化转型

股票代码	会计年度	文本类型	关键词	关键词数量 (全部)	关键词数量 (去表)
000016	2020	管理层讨论与分析	物联网	6	5
000166	2020	管理层讨论与分析	金融科技	5	3
000977	2020	管理层讨论与分析	人工智能	7	7

4.数据库内容

数据模块	数据表	数据字段
企业数字化转型	企业数字化转型	股票代码、会计年度、文本类型、关键词、关键词数量 (全部/去表)
文本确定程度	文本确定程度	股票代码、会计年度、文本类型、关键词、关键词数量 (全部/去表)
风险信息披露程度	风险信息披露程度	股票代码、会计年度、文本类型、关键词、关键词数量 (全部/去表)
企业合作文化	企业合作文化	股票代码、会计年度、文本类型、关键词、关键词数量 (全部/去表)
管理者短视主义	管理者短视主义	股票代码、会计年度、文本类型、关键词、关键词数量 (全部/去表)
文本情感程度	文本情感程度	股票代码、会计年度、文本类型、关键词、关键词数量 (全部/去表)
未来事项披露	未来事项披露	股票代码、会计年度、文本类型、关键词、关键词数量 (全部/去表)
研发文本信息	研发文本信息	股票代码、会计年度、文本类型、关键词、关键词数量 (全部/去表)

5.数据范围

数据模块	数据起始时间
全部模块	2007 年

(八) 上市公司监管意见数据库-CROD

1. 数据库简介

上市公司监管意见数据库 (*Regulatory Opinions Database of Listed Company*, **CROD**) 是基于中国证券监督管理委员会发布的 IPO 审核反馈意见整理而成的专业数据。参考已有研究文献, 提供了股票代码、意见内容、大类问题个数、小类问题个数和文章总字数等关键字段。

2. 数据库特色

- 数据来源权威

CROD 直接来源于证监会官网, 对数据最大限度保留, 并匹配了相关的股票代码。

- 多维度反馈意见信息

CROD 从大类问题数、小类问题数和总字数等几个字段反映了 IPO 监管意见的情况。

3. 典型数据举例

IPO 审核意见

股票代码	IPO 意见标题	公告时间	大类问题数	小类问题数	问题总字数
000630	铜陵有色金属集团铜冠矿山建设股份有限公司首次公开发行股票申请文件反馈意见	2020-11-27	3	34	12924
0002809	广东红墙新材料股份有限公司首次公开发行股票申请文件的反馈意见	2015-11-27	3	39	6004
605005	合兴汽车电子股份有限公司首次公开发行股票申请文件反馈意见	2020-04-17	4	36	11801

4. 数据库内容

数据模块	数据表	数据字段
IPO 审核意见	IPO 审核意见	股票代码、公司名称、IPO 意见标题、公告时间、IPO 审核意见全文内容、大类问题数、小类问题数、问题总字数

5. 数据范围

数据模块	数据起始时间
IPO 审核意见	2015 年

（九）CSR 报告文本数据库-CSRT

1.数据库简介

CSR 报告文本数据库 (*Corporate Social Responsibility Report Text, CSRT*) 根据上市公司的企业社会责任报告文本制作而成的专业数据库。*CSRT* 分析了上市公司的 CSR 报告, 展示了 CSR 报告基本信息、情感语调、可读性和相似度四个维度的重要信息。其中, 基本信息展示了以 PDF 格式展现的 CSR 报告的大小、页数、字数和词数等重要信息; 情感语调则基于 LM 词典和 NTUST 词典计算了 CSR 报告的 TONE 值; 可读性参考前沿研究, 从多个指标展示了报告的可读性; 最后, 从相同行业内不同公司的相似度和同一公司不同年度的相似度展示了 CSR 报告的相似情况。

2.数据库特色

- 数据分析准确、字段丰富

CSRT 基于上市公司发布的企业社会责任报告文本, 从多个角度精准提取了相关信息。

3.典型数据举例

基本信息

股票代码	公司简称	会计年度	总字数 1	总字数 2	总词数 1	总词数 2	报告大小	总页数	句子数
000001	平安银行	2012	36477	30005	19753	12238	1274KB	70	766
000002	万科 A	2012	47663	38612	25383	15946	5949KB	81	1617
000006	深振业 A	2012	6647	5337	3535	2182	489KB	14	124
000009	中国宝安	2012	11402	9580	6286	3728	442KB	15	213
000012	南玻 A	2012	6927	5966	3780	2296	322KB	8	117

4.数据库内容

数据模块	数据表	数据字段
基本信息	基本信息	股票代码、公司简称、会计年度、总字数 1、总字数 2、总词数 1、总词数 2、报告大小、总页数、句子数
语调分析	语调分析	股票代码、公司简称、统计日期、会计年度、LM 词典积极词汇数、LM 词典消极词汇数、基于 LM 词典的 TONE1、基于 LM 词典的 TONE2、台大词典积极词汇数、台大词典消极词汇数、基于台大词典的 TONE1、基于台大词典的 TONE2
可读性分析	可读性分析	股票代码、公司简称、会计年度、分句中的平均字数、副词和连词的比例、每一千字中常用字字数、每一千字中次常用字字数
同行业相似度	同行业相似度	股票代码、公司简称、股票代码 1、公司简称 1、文本序号 1、文本序号 2、统计日期、会计年度、文本类型、所属行业、行业内公司相似度
同公司相似度	同公司相似度	股票代码、公司简称、统计日期、会计年度、文本类型、所属行业、公司不同年度相似度

5.数据范围

数据模块	数据起始时间
全部模块	2006 年

（十）分析师报告文本数据库-TDAR

1.数据库简介

分析师报告文本数据库 (*Text Database of Analyst Report, TDAR*) 是基于对分析师报告文本内容进行深度挖掘而开发的专业数据库。*TDAR* 包括五个模块, 分别为基本信息、可读性、前瞻性、相似度和情感语调。其中, 基本信息提供了报告标题、文件大小、字数和句子数等基本信息; 可读性包含学术研究中常用的文本可读性衡量指标; 前瞻性则基于机器学习方法, 对分析师报告中的前瞻性句子进行计算; 相似度计算了每篇报告与其发布前三个月的该公司所有分析师报告的相似情况; 情感语调基于词典法和机器学习方法计算了分析师报告文本的情感倾向。希望 *TDAR* 的推出能该领域研究提供文本数据支持。

2.数据库特色

- 提供多维度和准确的分析师文本指标

TDAR 从多个维度展示了分析师报告文本, 并且引入机器学习等计算机方法计算前瞻性和情感语调等, 拓展了研究范畴。

3.典型数据举例

分析师报告前瞻性

编号	股票代码	发布日期	前瞻性句子数	句子总数
1201210291	000001	2012-10-29	3	49
1201210269	000001	2012-10-26	4	83
1201209261	000001	2012-09-26	18	186

4.数据库内容

数据模块	数据表	数据字段
分析师报告 基本信息	分析师报告 基本信息	编号、股票代码、报告标题、发布日期、总字数 1、总字数 2、总词数 1、总词数 2、文件大小、页数、句子数
分析师报告 可读性	分析师报告 可读性	编号、股票代码、发布日期、总字数 1 (个)、总字数 2 (个)、总词数 1 (个)、总词数 2 (个)、句子数 (个)、分句中的平均字数 (个) ……
分析师报告 前瞻性	分析师报告 前瞻性	编号、股票代码、发布日期、前瞻性句子数、句子数
分析师报告 相似度	分析师报告 相似度	编号、股票代码、发布日期、平均相似度
分析师报告 情感语调	分析师报告 情感语调 (词典法)	编号、股票代码、发布日期、LM 词典积极词汇数、LM 词典消极词汇数、台大词典积极词汇数、台大词典消极词汇数、LM 词典积极语句数量、LM 词典消极语句数量、台大词典积极语句数量、台大词典消极语句数量
	分析师报告 情感语调 (机器学习)	编号、股票代码、发布日期、积极性语句数量、中性语句数量、消极性语句数量、句子数

5.数据范围

数据模块	数据起始时间
全部模块	2006 年

(十一) IPO 申报稿文本数据库-IPOT

1. 数据库简介

IPO 申报稿文本数据库 (*Text Database of IPO Application Documents, IPOT*) 是基于公司 **IPO** 申报稿的文本信息进行整理和分析形成的专业文本分析数据库。**IPOT** 整理了从 2007 年至今的公司申报文本数据资料, 通过参考已有文献, 分别从 **IPO** 申报稿文本的基本信息、可读性、情感分析、相似度以及公司 **IPO** 申报情况 5 个方面进行了人工智能分析。其中, 基本信息提供了报告标题、文件大小、字数和句子数等基本信息; 可读性包含学术研究中常用的文本可读性衡量指标; 情感语调基于词典法计算了相关文本语气; 同公司相似度计算了每篇报告与该公司发布的所有 **IPO** 申报稿的相似情况; 公司 **IPO** 申报情况提供了该公司发布申报稿的次数、以及公司是否成功上市等数据。

2. 数据库特色

- 数据可靠、全面

IPOT 整理了从 2007 年至今的 A 股 IPO 申报稿的文本资料, 为该领域的学术研究提供了可靠、全面的数据支持。

- 数据精确、严谨

IPOT 参考已有学术文献的处理方法, 对相关文本进行了精确的清洗、处理和计算, 可直接用于学术研究。

3. 典型数据举例

(1) 同公司相似度

股票代码	公司简称	文本序号 1	文本序号 2	公告日期 1	公告日期 2	相似度
A22193.SZ	竞远安全	536	4496	2023-01-30	2022-06-24	0.99749339
A22193.SZ	竞远安全	4498	536	2023-01-16	2023-01-30	0.99997514
A22193.SZ	竞远安全	4498	4496	2023-01-16	2022-06-24	0.99748689
A22193.SZ	竞远安全	4498	4497	2023-01-16	2022-11-04	0.99876320
301391.SZ	卡莱特	353	2992	2022-03-31	2021-12-29	0.99972713
301391.SZ	卡莱特	353	2993	2022-03-31	2022-02-24	0.99981564
301391.SZ	卡莱特	2992	353	2021-12-29	2022-03-31	0.99972713

(2) 公司 IPO 申报情况

更新日期	股票代码	公司简称	是否上市公司	发布 IPO 申报稿次数
2023.08	001201.SZ	东瑞股份	1	3
2023.08	001202.SZ	炬申股份	1	2
2023.08	001203.SZ	大中矿业	1	4
2023.08	A11046.SZ	招金励福	0	3
2023.08	A11048.SZ	武桥股份	0	1
2023.08	A10013.SH	先临三维	0	2

4.数据库内容

数据模块	数据表	数据字段
IPO 申报稿的基本信息	IPO 申报稿的基本信息	股票代码、公司简称、报告标题、文本序号、公告日期、文本长度 1、文本长度 2、总词数 1、总词数 2、文件大小、页数、句子数
IPO 申报稿的可读性	IPO 申报稿的可读性	股票代码、公司简称、文本序号、公告日期、专业词密度、分句中的平均字数、副词和连词的比例、每一千字中常用字字数(个)、每一千字中次常用字字数(个)
IPO 申报稿的情感分析	IPO 申报稿的情感分析	股票代码、公司简称、文本序号、公告日期、LM 词典积极词汇数、LM 词典消极词汇数、台大词典积极词汇数、台大词典消极词汇数、基于 LM 词典的 TONE1、基于 LM 词典的 TONE2、基于台大词典的 TONE1、基于台大词典的 TONE2
同公司相似度	同公司相似度	股票代码、公司简称、文本序号 1、文本序号 2、公告日期 1、公告日期 2、相似度
公司 IPO 申报情况	公司 IPO 申报情况	更新日期、股票代码、公司简称、是否上市公司、发布 IPO 申报稿次数

5.数据范围

数据模块	数据起始时间
全部模块	2007 年

资本市场人物特征

（一）董事长与总经理研究数据库-CCEO

1.数据库简介

董事长与总经理研究数据库 (*Chairman and CEO Research Database, CCEO*) 是基于对中国上市公司年报、招股说明书、公告等多方来源数据的考察，并通过网络搜索、计算机提取和人工核对等多种方式补充完善个人特征相关信息而建立的专业数据库。*CCEO* 主要包括董事长与总经理基本信息、教育背景、任职信息三张表，详细列示了上市公司董事长与总经理的姓名、年龄、性别、简历、就读高校、所学专业、离任职信息、薪酬情况、职务信息、职称、工作经验等信息。

2.数据库特色

- 多维

CCEO 包括上市公司高管的多维度特征数据，包括年龄、性别、学历、专业、海外经历、当前任职情况等多维度信息。

- 详细

CCEO 对董事长与总经理的毕业院校的分类详细到其所就读的大专院校、本科院校、硕士院校、博士院校、MBA 院校以及博士后院校等；对董事长与总经理专业背景的描述则包括其大学所学专业、是否拥有海外教育背景等。

- 数据丰富

CCEO 包含的董事长与总经理籍贯/出生地、职称等字段的非空值远远多于市场上的已有数据库，可为学者研究提供便利。

3.典型数据举例

董事长与总经理基本信息

股票代码	高管姓名	高管职务	性别	出生日期	国籍	籍贯或出生地
000997	林学杰	总经理	男	1960	中国	福建省福州市
002028	杨小强	总经理	男	1963	中国	北京市海淀区
002123	左强	董事长	男	1970	中国	北京市海淀区
002151	赵耀升	总经理	男	1959	中国	北京市海淀区
002208	张金生	总经理	男	1963	中国	北京市宣武区

4.数据库内容

数据模块	数据表	数据字段
董事长与总经理研究	基本信息	股票代码、公司简称、统计日期、高管编码、高管姓名、高管职务、高管职务编码、性别、出生日期、国籍、籍贯、籍贯所在地区编码、出生地、出生地所在地区编码、籍贯或出生地、籍贯或出生地所在地区编码、籍贯或出生地数据来源、学历、职称与资格证书、董事长与总经理个人简介、是否兼任董事长和 CEO、是否兼任董事长和总经理、海外背景、学术背景

数据模块	数据表	数据字段
董事长与总经理研究	教育背景	股票代码、公司简称、统计日期、董事长与总经理编码、董事长与总经理姓名、所学专业、是否有海外教育背景、最高学历/学位、专科院校、专科院校代码、专科院校入学时间、专科院校毕业时间、本科院校、本科院校代码、本科院校入学时间、本科院校毕业时间、硕士院校、硕士院校代码、硕士院校入学时间、硕士院校毕业时间、博士院校、博士院校代码、博士院校入学时间、博士院校毕业时间、博士后院校、博士后院校代码、博士后院校入学时间、博士后院校毕业时间、MBA 院校名称、MBA 院校代码、MBA 院校入学时间、MBA 院校毕业时间、访学院校名称、访学院校代码、访学院校访问时间、访学院校结束时间、数据来源
	任职信息	股票代码、公司简称、统计日期、董事长与总经理编码、董事长与总经理姓名、董事长与总经理职务、董事长与总经理职务编码、任职开始日期、任职结束日期、报告期薪酬、年末持股数量

5.数据范围

数据模块	数据起始时间
基本信息	1999 年
教育背景	2004 年
任职信息	1999 年

(二) 上市公司其他高管数据库-OECD

1.数据库简介

上市公司其他高管数据库 (*Other Executives Database of Listed Companies, OECD*) 是基于中国上市公司除董事长、总经理、CEO 之外的其他高管简介、招股说明书、公告等来源披露的其他高管信息, 从多个维度进行其特征信息的整合和展示的专业数据库。*OECD* 主要包括上市公司其他高管基本信息、教育背景、任职信息三张表。

2.数据库特色

- 全面

OECD 不仅包括上市公司其他高管个人基本信息, 还包括教育背景等重要信息, 涵盖多个重要字段, 例如所学专业、毕业院校、报告期薪酬、年末持股数量等。

3.典型数据举例

上市公司其他高管教育背景

股票代码	统计日期	姓名	本科院校	院校代码	硕士院校	院校代码	博士院校	院校代码
000608	2012/12/31	李鸣	北京大学	01-008	耶鲁大学	FA-032	北京大学	01-008
000608	2013/12/31	李鸣	北京大学	01-008	耶鲁大学	FA-032	北京大学	01-008
000608	2014/12/31	李鸣	北京大学	01-008	耶鲁大学	FA-032	北京大学	01-008

4.数据库内容

数据模块	数据表	数据字段
上市公司其他高管	基本信息	股票代码、公司简称、统计日期、高管编码、高管姓名、高管职务、高管职务编码、性别、出生日期、国籍、学历、职称与资格证书、学术背景、是否在其他上市公司兼职、任职的上市公司家数
	教育背景	股票代码、公司简称、统计日期、高管编码、高管姓名、所学专业、是否有海外教育背景、最高学历/学位、专科院校、专科院校代码、本科院校、本科院校代码、硕士院校、硕士院校代码、博士院校、博士院校代码、博士后院校、博士后院校代码、MBA 院校名称、MBA 院校代码、访学院校名称、访学院校代码、数据来源
	任职信息	股票代码、公司简称、统计日期、高管编码、高管姓名、高管职务、高管职务编码、是否在职、任职开始日期、任职结束日期、报告期薪酬、年末持股数量

5.数据范围

数据模块	数据起始时间
全部模块	2010 年

（三）独立董事研究数据库-IDRD

1.数据库简介

独立董事研究数据库 (*Independent Directors Research Database, IDRD*) 是基于对独立董事的个人简介、新闻报道、公司公告、监管机构公告等多方数据来源的考察, 从多个维度对独立董事信息进行整合和展示的专业数据库。独立董事研究从本世纪初就成为学术研究前沿与热点, 并逐渐成为公司治理领域极为重要的一个研究方向, 然而, 目前市场上罕有单独成库的独立董事研究相关的数据, 并且所涉及的数据深度和广度也远远不够。因此, *IDRD* 对该领域的研究数据进行了补充和扩展。独立董事研究数据库包括独董背景信息、独董新闻报道、独董意见、独董处罚情况、独董辞任职、独董比例、独董交叉任职及董事网络关系八张表格。

2.数据库特色

- 多维度独立董事动态信息

IDRD 从多个权威渠道获得了多维度独董动态信息, 例如, 新闻报道、处罚、辞任职信息等, 这为独立董事研究的扩展提供了重要数据支持。

- 提供董事网络关系等前沿数据

采用 *Pajek* 软件对董事网络关系进行精确计算, 提供了研究常用的董事网络关系指标。

3.典型数据举例

(1) 独董任辞职

股票代码	公司简称	统计日期	独董编码	独董姓名	任职开始日期	任职结束日期
000001	平安银行	2012/12/31	1200088	马林	2011/5/25	2016/5/19
000001	平安银行	2012/12/31	1200089	刘雪樵	2011/5/25	2014/1/20
000002	万科 A	2012/12/31	1200150	齐大庆	2008/4/23	2014/3/27
000002	万科 A	2012/12/31	1200159	陈茂波	2011/3/31	2012/8/30
000004	国农科技	2012/12/31	1200247	李良智	2010/3/31	2013/5/9
000004	国农科技	2012/12/31	1200248	潘秀玲	2010/3/31	2013/5/9

(2) 董事网络关系

股票代码	会计年度	董事编码	董事名称	程度中心度	接近中心度	中介中心度	结构洞
000001	2012	1200056	李敬和	34	0.158619	0.001139	0.906828
000001	2012	1200068	王利平	41	0.161909	0.001049	0.907248
000001	2012	1200070	姚波	41	0.161909	0.001049	0.907248
000001	2012	1200071	顾敏	41	0.161909	0.001049	0.907248
000001	2012	1200072	卢迈	28	0.158786	0.001199	0.893468

4.数据库内容

数据模块	数据表	数据字段
独立董事	独董背景信息	股票代码、公司简称、统计日期、独董编码、独董姓名、是否异地独董、是否有法律背景、是否兼任其他上市公司独董、任职独董的上市公司数量、是否其他上市公司高管(不包括独董)、是否有行业专长

独立董事	独董新闻报道	股票代码、公司简称、独董编码、独董姓名、职务、性别、报道媒体、报道时间、新闻标题、相似新闻数量、新闻情感、新闻类型
	独董异议	股票代码、公司简称、会计年度、独董姓名、意见标题、异议类别、发表异议原因、异议事项
	独董处罚情况	股票代码、公司简称、公告日期、独董编码、独董姓名、违规行为、违规类型、处理人、处分类型
	独董任辞职	股票代码、公司简称、统计日期、独董编码、独董姓名、任职开始日期、任职结束日期
	独董比例	股票代码、公司简称、会计年度、董事总人数、独立董事人数、独立董事比例(%)
	独董交叉任职	股票代码、公司简称、统计日期、独董编码、独董姓名、职务、任职开始日期、任职结束日期、是否交叉任职
	董事网络关系	股票代码、会计年度、董事编码、董事名称、是否独立董事、程度中心度、接近中心度、中介中心度、结构洞

5.数据范围

数据模块	数据起始时间
独董背景信息	1993 年
独董新闻报道	2002 年
独董异议	2004 年
独董处罚情况	2002 年
独董任辞职	2010 年
独董比例	1999 年
独董交叉任职	1993 年
董事网络关系	2006 年

（四）董秘信息数据库-CSBD

1.数据库简介

董秘信息数据库 (*Secretary of the Board Database, CSBD*) 是基于公开披露的上市公司董秘简介数据以及新财富年度金牌董秘评选整理而成的专业数据库。*CSBD* 除了包含上市公司董秘的个人基本信息外, 还包含了兼任情况、海外背景、学术背景、是否为金牌董秘等多维度信息。

2.数据库特色

- 多维

CSBD 包括上市公司董秘的个人多维度特征数据, 覆盖面广, 涵盖年龄、性别、学历、职业背景、职业资格证书、任职情况等丰富信息。

3.典型数据举例

上市公司董秘信息

股票代码	统计日期	高管编码	董秘姓名	是否为金牌董秘	性别	出生日期	学历
000002	2006-12-31	1200129	肖莉	1	女	1963	硕士
000002	2015-12-31	1200152	谭华杰	0	男	1973	本科
000002	2016-12-31	1209171	朱旭	0	女	1975	硕士
000003	2000-12-31	1200178	彭卫	0	男	1969	硕士
000004	2010-12-31	1200218	蒋伟诚	0	男	1970	硕士
000004	2010-12-31	1200256	杨斌	0	男	1972	硕士

4.数据库内容

数据模块	数据表	数据字段
上市公司董秘信息	上市公司董秘信息	股票代码、公司简称、统计日期、高管编码、董秘姓名、董秘的其他职务、是否为金牌董秘、性别、出生日期、国籍、学历、职称与资格证书、个人简介、海外背景、学术背景

5.数据范围

数据模块	数据起始时间
全部模块	2010年

(五) 基金经理研究数据库-FMRD

1.数据库简介

基金经理研究数据库 (*Fund Manager Research Database, FMRD*) 是基于对基金经理的简介、负责的基金等多方数据来源的考察,对基金经理个人信息进行整合和展示的专业数据库。主要包括基金经理基本信息、工作经历、管理基金和教育背景四张表,包含基金经理所在任职机构名称、职位、所持基金情况、学历/学位、毕业院校等重要信息。

2.数据库特色

- 完善、全面的基金经理信息

FMRD 不仅包括了基金经理的个人基本信息,还包括了基金经理的教育背景、工作经历、管理基金等重要信息。

3.典型数据举例

基金经理工作经历

基金经理姓名	基金代码	任职机构名称	职位	入职日期	离职日期
费逸	270021	广发基金管理有限公司	研究员	2010-07-06	2017-07-18
林高榜	110012	易方达基金管理有限公司	基金经理助理	2016-12-14	2017-05-10
刘晓龙	270025	广发基金管理有限公司	基金经理	2010-11-23	2017-03-29
谈洁颖	519956	长信基金管理有限责任公司	基金经理	2012-07-23	2017-03-28
徐冶丹	004585	北京鹏扬投资管理有限公司	基金经理	2017-06-02	2017-07-14

4.数据库内容

数据模块	数据表	数据字段
基金经理	基金经理基本信息	基金经理姓名、性别、出生年份、学历/学位、简介
	基金经理工作经历	基金经理姓名、基金代码、基金简称、任职机构名称、职位、入职日期、离职日期
	基金经理管理基金	基金经理姓名、基金代码、基金名称、基金公司名称、投资类型、开始日期、结束日期
	基金经理教育背景	基金经理姓名、基金代码、基金简称、学历/学位、专业、毕业学校

5.数据范围

数据模块	数据起始时间
全部模块	截面数据

（六）金融机构高管数据库-FIED

1.数据库简介

金融机构高管数据库 (*Financial Institution Executives Database, FIED*) 是从银行、券商和信托公司的年报以及官方网站等多方数据来源进行整理和分析后而建立的专业数据库。主要包括券商高管、银行高管、信托公司高管 3 个模块, 共 3 张表, 包含券商、银行和信托公司高管的职务、兼职单位以及简介信息等重要字段。*FIED* 包括几乎所有从 2000 年至今的券商高管、银行高管、信托公司高管个人特征数据, 大大减少了学者搜集和整理数据的时间。

2.数据库特色

- 提供深度的金融机构高管特征数据

FIED 从银行、券商和信托公司的年报以及官方网站等获得了这些金融机构的高管简介信息, 并在此基础上进行深度的整理和分析, 提供研究中常用的重要字段。

3.典型数据举例

银行高管

银行高管姓名	银行代码	银行名称	统计日期	职位	性别	年龄
陈鹏飞	100029	抚顺银行股份有限公司	2015-12-31	监事、副行长	男	33
付蕾	100029	抚顺银行股份有限公司	2015-12-31	副行长	女	46
李向文	100029	抚顺银行股份有限公司	2015-12-31	非执行董事	男	41

4.数据库内容

数据模块	数据表	数据字段
金融机构高管	券商高管	统计日期、券商名称、券商代码、董监高姓名、职位类型、性别、出生年份、国籍、学历/学位、现任或者曾任、工作经历起止日期、工作单位、职位、是否全职、个人简介
	银行高管	银行高管姓名、银行代码、银行名称、统计日期、职位、性别、年龄、出生日期、任职期间、持有股份、是否领取薪酬、兼职单位的名称、兼职单位的职务、任职的股东单位名称、在股东单位的职务、个人简介
	信托公司高管	信托公司代码、信托公司名称、统计日期、信托公司高管姓名、性别、年龄、学历/学位、专业、职位、选任日期、兼职单位名称、兼职单位职务、任职的股东单位名称、任职的股东单位职务、金融从业年限、所推举的股东名称、推举股东的持股比例 (%)、个人简介

5.数据范围

数据模块	数据起始时间
券商高管	2007 年
银行高管	2000 年
信托公司高管	2004 年

(七) 董事会研究数据库-BOARD

1. 数据库简介

董事会研究数据库 (*Board Research Database, BOARD*) 是基于上市公司财务报告中董事会相关披露信息整理而成的专业数据库。*BOARD* 包括董事会基本特征、董事会下设专门委员会、董事出席会议情况和控股股东委派董事四张表格, 其中, 基本特征表主要从董事会多样性维度对董事个人特征以及董事会总体特征进行展示; 董事会下设专门委员会列示了上市公司设立的专门委员会名称以及本年的会议次数情况; 董事出席会议情况则包含了董事当年出席董事会和股东大会的次数情况; 最后, 通过阅读各个董事的简历情况, 判断哪些董事是由控股股东委派的, 从而形成控股股东委派董事表。

2. 数据库特色

- 提供准确的董事会研究数据

BOARD 不仅数据来源于上市公司年报, 而且严格按照学术需求和规范进行制作, 为学术研究提供重要数据支持。

3. 典型数据举例

董事会下设专门委员会

股票代码	公司简称	统计日期	委员会名称	召开会议次数
000001	平安银行	2021-12-31	薪酬与考核委员会	4
000001	平安银行	2021-12-31	提名委员会	4
000001	平安银行	2021-12-31	关联交易控制委员会	7

4. 数据库内容

数据模块	数据表	数据字段
董事会基本特征	董事会基本特征	股票代码、公司简称、统计日期、董事姓名、职务、性别、出生日期、国籍、学历、职称与资格证书、董事总人数、独立董事比例 (%)、CEO 是否兼董事长、CEO 是否兼董事、董事平均年龄、女性董事比例 (%)、董事会持股数量、董事会持股比例 (%)、设置专业委员会的数量、是否设立审计委员会、董事会会议次数
董事会下设专门委员会	董事会下设专门委员会	股票代码、公司简称、统计日期、委员会名称、召开会议次数
董事出席会议情况	董事出席会议情况	股票代码、公司简称、统计日期、董事姓名、是否独立董事、本年应参加董事会次数、亲自出席次数、现场出席次数、以通讯方式参加次数、以其他方式参加次数、委托出席次数、缺席次数、是否连续两次未亲自参加会议、出席股东大会的次数、本年应参加的股东大会次数
控股股东委派董事	控股股东委派董事	股票代码、公司简称、统计日期、是否有控股股东委派董事、控股股东名称、委派董事序号、委派董事姓名、职务、性别、出生日期、国籍、学历、职称与资格证书

5. 数据范围

数据模块	数据起始时间
全部模块	2007 年

银行及金融研究

(一) 银行研究数据库-CBRD

1. 数据库简介

银行研究数据库 (*Chinese Bank Research Database, CBRD*) 是面向广大高校和学者以及实务界人士, 对中国商业银行进行深入考察与了解建立的详细、准确的研究数据库。该数据库涵盖了我国国有银行, 股份制银行、城市商业银行、农村商业银行等银行。*CBRD* 包含银行基本信息、银行财务报表、银行信贷资产分析、银行重要业务指标等 4 个模块。其中, 银行基本信息包括银行概况、股东构成、高管特征、审计信息、员工与股东信息等; 银行财务报表模块包括了三张常用报表; 银行信贷资产分析对信贷资产结构, 十大贷款客户, 贷款行业分布以及贷款担保方式进行了统计; 银行重要业务指标对一些重要指标进行了罗列。

2. 数据库特色

- 全面、准确

CBRD 不仅包括中国目前已经上市的银行机构, 还包括大量未上市的银行机构数据, 进行多方位信息来源的综合, 并详细列示了与研究相关的大量字段。为确保非上市银行数据的准确性, 我们进行了逐项、严格的校对工作, 采用多种方法进行细致查验。

3. 典型数据举例

银行股东构成

银行代码	银行名称	统计日期	序号	股东名称	期末股份数 (股)	期末持股比例 (%)
100136	上海银行	2012-12-31	1	上海联和投资有限公司	814803808	19.24
100136	上海银行	2012-12-31	2	香港上海汇丰银行有限公司	338720000	8
100136	上海银行	2012-12-31	3	中国建银投资有限责任公司	296380000	7
100136	上海银行	2012-12-31	4	上海市黄浦区国有资产总公司	130409244	3.08
100136	上海银行	2012-12-31	5	上海商业银行有限公司	127020000	3

4. 数据库内容

数据模块	数据表	数据字段
银行基本信息	银行概况	银行代码、银行名称、成立日期、银行性质、注册资本、是否上市、省级名称、注册地址、经营范围、法定代表人、总经理
	股东构成	银行代码、银行名称、统计日期、序号、股东名称、持股金额、期末股份数、期末持股比例 (%)、期初股份数、期初持股比例 (%)、股份质押数、股份冻结数
	高管特征	银行高管姓名、银行代码、银行名称、统计日期、职位、性别、年龄、出生日期、任职期间、持有股份、是否领取薪酬、兼职单位的名称、兼职单位的职务、任职的股东单位名称、在股东单位的职务、个人简介
	审计信息	银行代码、银行名称、统计日期、会计师事务所名称、审计日期、审计师姓名、审计费用

数据模块	数据表	数据字段
银行基本信息	员工与股东信息	银行代码、银行名称、统计日期、股东户数、机构股东数、自然人股东数、员工总数、高级职称员工数、高级职称员工占比(%)、中级职称员工数、中级职称员工占比(%)、初级职称员工数、初级职称员工占比(%)、中层以上管理人员数量、银行业务人员数量、博士研究生学历人数、硕士研究生学历人数、研究生学历占比(%)、大学本科学历人数、大学本科学历占比(%)、大专及以上学历人数、大专及以上学历占比(%)、大专及以下学历人数、大专及以下学历占比(%)、退养员工人数
银行财务报表	资产负债表	银行代码、银行名称、统计日期、现金及存放中央银行款项、存放同业款项、拆出资金、交易性金融资产、衍生金融资产、应收利息、发放贷款及垫款、可供出售金融资产、固定资产、资产总计、同业和其他金融机构存放款项、向中央银行借款、拆入资金、吸收存款、应交税费、应付利息、负债合计、实收资本、资本公积金、盈余公积金、未分配利润、一般风险准备、归属于母公司所有者权益合计、所有者权益合计、负债与所有者权益总计
	利润表	银行代码、银行名称、统计日期、营业收入、利息净收入、利息收入、利息支出、手续费及佣金净收入、手续费及佣金收入、手续费及佣金支出、营业支出、营业税金及附加、管理费用、营业利润、营业外收入、营业外支出、利润总额、所得税、净利润、归属于母公司所有者的净利润、少数股东损益、基本每股收益、稀释每股收益
	现金流量表	银行代码、银行名称、统计日期、客户存款和同业存放款项净增加额、向中央银行借款净增加额、向其他金融机构拆入资金净增加额、收取利息和手续费净增加额、收到其他与经营活动有关的现金、经营活动现金流入小计、客户贷款及垫款净增加额、经营活动产生的现金流量净额、投资活动产生的现金流量净额、筹资活动产生的现金流量净额
银行信贷资产分析	信贷资产结构	银行代码、银行名称、统计日期、正常类贷款、关注类贷款、不良类贷款
	十大贷款客户	银行代码、银行名称、统计日期、序号、客户名称、贷款余额、占贷款总额比例(%)、占资本净额比例(%)
	贷款行业分布	银行代码、银行名称、统计日期、序号、行业种类、贷款余额、占贷款总额比例(%)
	贷款担保方式	银行代码、银行名称、统计日期、质押贷款余额、抵押贷款余额、保证贷款余额、信用贷款余额
银行重要业务指标	银行重要业务指标	银行代码、银行名称、统计日期、存款总额、贷款总额、资本充足率(%)、核心资本充足率(%)、拆入资金比例(%)、拆出资金比例(%)、单一最大客户贷款比例(%)、最大十家客户贷款比例(%)、利息回收率(%)、逾期贷款、贷款呆帐准备金、不良贷款余额、不良贷款比例、不良贷款拨备覆盖率、资本净额、核心资本净额、成本收入比(%)

5.数据范围

数据模块	数据起始时间
银行基本信息	1992年
银行财务报表	2004年
银行信贷资产分析	2000年
银行重要业务指标	2003年

（二）中国商业银行分支机构数据库-CCBD

1.数据库简介

中国商业银行分支机构数据库 (*Branch Database of Chinese Commercial Bank, CCBD*) 涵盖了中国商业银行分支机构的成立日期、所在地、分支机构名称等信息。**CCBD** 分为三个数据模块，银行分支机构信息模块的数据主要来自中国银行业监督管理委员会对银行机构颁（换）发许可证的相关信息，包括机构设立情况、机构失控情况和机构退出情况三张表格。银行网点分布信息模块的数据主要来自高德地图以计算机技术获得的银行网点分布情况。由于银行网点分布信息除了包括银行分支机构信息，还包括诸如银行营业厅、ATM 取款机等多种类型的银行服务机构，因此其数量要远多于银行分支机构信息，可为学术研究提供更多支持。行政处罚模块的数据主要来自中国银行保险监督管理委员会所公布的行政处罚信息。

2.数据库特色

- 全面的银行分支机构和网点信息

CCBD 既包括官方批准的银行分支机构信息，又包括来自各个省市县的银行网点分布情况，可以充分反映一个地区的银行竞争、银行服务等情况。

3.典型数据举例

银行分支机构信息_机构设立情况

机构分类	机构名称	机构地址	机构所在地	金融许可证 发证日期	批准成立 日期
商业银行	中国工商银行股份有限公司 天津市分行	天津市河西区 围堤道 123 号	天津市	2021-12-24	1984-11-22
商业银行	中国工商银行股份有限公司 河北省分行	河北省石家庄 市中山西路 188 号	河北省-石家 庄市	2021-12-01	1985-02-15
商业银行	中国工商银行股份有限公司 辽宁省分行	辽宁省沈阳市 和平区南京北 街 88 号	辽宁省-沈阳 市	2021-12-29	1984-01-01
商业银行	中国工商银行股份有限公司 黑龙江省分行	黑龙江省哈尔 滨市道里区中 央大街 218 号	黑龙江省-哈 尔滨市	2021-10-08	1984-01-01
商业银行	中国工商银行股份有限公司 上海市分行	上海市浦东新 区银城路 8 号	上海市	2021-10-15	1984-01-01
商业银行	中国工商银行股份有限公司 江苏省分行	南京市中山南 路 408 号	江苏省-南京 市	2021-09-23	1984-01-01
商业银行	中国工商银行股份有限公司 安徽省分行	合肥市芜湖路 189 号	安徽省-合肥 市	2021-09-15	1988-04-08

4.数据库内容

数据模块	数据表	数据字段
银行分支机构信息	机构设立情况	机构编码、机构分类、机构名称、机构地址、机构所在地、金融许可证发证日期、批准成立日期
	机构失控情况	机构编码、机构分类、机构名称、机构地址、机构所在地、金融许可证发证日期、批准成立日期、失控原因、失控时间
	机构退出情况	机构编码、机构分类、机构名称、机构地址、机构所在地、金融许可证发证日期、批准成立日期、退出日期
银行网点分布信息	银行网点分布	省份名称、市县名称、银行网点名称、银行代码、银行网点地址、获取时间
行政处罚	行政处罚	行政处罚决定书文号、处罚对象类型、被处罚单位名称、被处罚个人所在单位、主要违法违规事实（案由）、行政处罚依据、行政处罚决定、作出行政处罚决定的机关名称、作出行政处罚决定的日期

5.数据范围

数据表	数据起始时间
机构设立情况	2007 年
机构失控情况	2007 年
机构退出情况	2007 年
银行网点分布	截面数据
行政处罚	2003 年

(三) 中国风险投资和私募股权数据库-VCPE

1.数据库简介

中国风险投资和私募股权数据库(*Chinese Venture Capital and Private Equity Database, VCPE*)是以研究为目的而设计的,面向国内外从事风险和私募股权投资研究的学者,旨在提供完整、精确的学术研究数据。该数据库全面覆盖股权投资研究所需要的各种信息,不仅收录了基金信息、投资机构信息、被投企业信息,还涵盖了投资事件及退出事件的各个层面信息。

2.数据库特色

- 内容详细完善

VCPE 对各个维度的信息进行了详细描述,提供了中国股权投资市场情报、投资机构及融资企业等各个层面信息。

- 学术性强

考虑到本数据库的学术导向,我们在设计数据库时综合考虑了学者对风险投资和私募股权研究的需要,并对相关模块的字段进行了调整。

3.典型数据举例

投资事件

投资方名称	融资方简称	投资金额	币种	股权比例%	融资时间	融资轮次
广发信德	云意电气	2240.00	人民币元	4.00	2010-08-27	A
未披露	时光一百	3000.00	人民币元		2012-03	Angel
云月投资	四川智强食品	3000.00	美元	60.00	2012-03-21	A
雷石久隆	佳利电子	1500.00	人民币元	29.00	2010-09-28	A
昆吾九鼎投资	宾西牛业	9000.00	人民币元		2012-03-26	A
中科宏易	国瓷材料	187.50	人民币元	25.00	2008-08-26	A

4.数据库内容

数据模块	数据表	数据字段
基金基本信息	基金基本信息	基金公司编号、基金全称、基金简称、注册地点、组织形式、基金类型、资本类型、存续年限、基金状态、投资项目个数、管理公司地址、基金描述、募集开始时间、募集完成时间、募集目标规模、募集完成规模、管理公司名称、管理公司简称、投资项目金额
投资机构	投资机构基本信息	投资机构编号、投资机构名称、投资机构简称、所在国家、省份、城市、成立日期、募集基金个数、投资项目个数、公司地址、公司简介、所属行业、公司性质、法人代表、募集基金金额、投资项目金额
	投资机构管理人员	投资机构编号、投资机构名称、管理人员姓名、在本公司职位、任职公司名称、任职公司职位、在职时间、状态、个人简介
融资企业基本信息	公司工商信息	融资企业编号、融资企业名称、国民经济行业分类、注册资金、法定代表人、公司类型、所属省份、实缴资本、经营状态、实际控制人、营业期限、人员规模、参保人数、成立日期、核准日期、登记机关、办公地址、注册地址、经营范围、公司简介

数据模块	数据表	数据字段
融资企业 基本信息	公司股东信息	融资企业编号、融资企业名称、序号、股东名称、实际持股比例、认缴出资、实际出资
	公司对外投资	融资企业编号、融资企业名称、序号、企业名称、法定代表人、注册资本、登记状态、持股比例
	公司主要人员	融资企业编号、融资企业名称、序号、姓名、职务
投资事件	投资事件	投资事件 ID、投资方编号、投资方名称、融资方编号、融资方名称、融资方简称、投资金额、币种、获得股份数量、股权比例、融资金额、币种、融资总股权比例、融资时间、融资方式、融资轮次、事件描述
退出事件	管理层基本信息	退出事件 ID、退出方编号、退出方名称、公司编号、公司名称、公司简称、股票代码、交易所、上市募集资金、所属行业、上市/退出时间、退出方式、总投资金额、账面退出金额、币种、账面回报倍数、事件描述

5.数据范围

数据模块	数据起始时间
全部模块	1990 年

（四）中国金融理财研究数据库-CFSD

1.数据库简介

中国金融理财研究数据库 (*Chinese Financial Services Research Database, CFSD*) 是以研究为目的而设计的专业学术研究数据库,旨在为国内外从事金融理财产品研究的学者和机构从业人员提供精确、完整的研究数据。本数据库覆盖中国投资者主要关注的理财产品类型,如银行理财产品、保险理财产品、阳光私募基金产品和 P2P 网贷。本数据库不仅包含相关产品的基本信息,还综合汇总了产品的具体信息。具体地,本数据库包含:

低风险理财产品 分为 2 个子模块:银行理财、保险理财。该数据库汇总了低风险理财产品的基本信息。

阳光私募基金产品 分为 3 个子模块:阳光私募基金产品信息、阳光私募基金经理信息、阳光私募基金收益率信息。该数据库包含所有私募基金的评级以及私募基金经理擅长的投资策略,内容非常完整且全面。

P2P 理财产品 分为 4 个子模块:P2P 网贷平台基本信息、P2P 网贷平台所属公司信息、P2P 问题平台信息、P2P 网贷平台高管基本信息。该数据库为投资者和研究人员提供详尽的 P2P 网贷平台信息。

2.数据库特色

- 数据完整性

中国金融理财研究数据库囊括了国内投资者与研究者关注的几乎所有的理财产品与项目,不仅包括产品本身的各方面详细信息,还包括产品的投资人、委托管理人、借贷人等多方位的综合信息。

3.典型数据举例

(1) P2P 网贷平台基本信息

平台名称	平均收益	投资期限	注册资金	债权转让	投标保障	业务类型
懒虫网	18%	1 月标 (100.0%)	1000 万元	不可转让	本金保障	车贷 (99%)
天天赢	16%	1 月标 (60.0%)、 天标 (40.0%)	3000 万元	1 个月	100%保 本保息	车贷 (30%)、银 行过桥 (70%)
宏量财富	16%	1 月标 (100.0%)	2000 万元	不可转让	VIP 保本 保息	房贷 (100%)
龙商 e 贷	10%	2 月标 (35.0%)、 1 月标 (25.0%)	5000 万元	随时	风险保证 金制度	车贷 (35%)、房 贷 (35%)、供应 链 (30%)

(2) 阳光私募基金经理

姓名	学历/ 学位	从业年限	任职私募年限	任职日期	履历背景	基金公司	管理基 金数	擅长策略
范赞车	大专	20 年	1 年 103 天	2015-08-21	民间	象学资产	2 只	股票型
云峰	硕士	8 年	126 天	2016-04-26	期货	泰通资管	1 只	管理期货
何淼	硕士	10 年	266 天	2015-12-08	公募	相聚资本	1 只	其他
海乐	硕士	21 年	2 年 11 天	2014-08-20	券商	银领资产	19 只	定向增发

4.数据内容

数据模块	数据表	数据字段
低风险理财产品	银行理财产品基本信息	产品名称、银行性质、银行名称、收益类型、委托币种、起始金额、委托单位递增、预期收益率、与同期储蓄比、销售状态、风险评级、流动性评级、发行起始日、发行截止日、收益计起日、到期日、委托管理期限、是否能被质押、是否有权提前赎回、本息返还方式、产品管理费、投资标的、收益率说明、收益计算基础
	保险理财	产品名称、保险类别、投保年龄、发行机构、保障期限、主附险、缴费期限、保障范围、保险责任、产品特色
阳光私募基金产品	阳光私募基金产品信息	基金简称、基金类型、综合评级、起购资金、封闭期、浮动管理利率、基金公司、成立以来收益率、组织形式、单位净值、成立日期、结构形式、基金状态、托管银行、开放日期、备案号
	阳光私募基金经理	阳光私募基金经理姓名、学历/学位、从业年限、任职私募年限、任职日期、履历背景、基金公司、管理基金数、擅长策略、基金简称、经理介绍、公司介绍
	阳光私募基金收益率	基金名称、基金类型、成立以来收益率、数据更新时间、近一年收益率、近二年收益率、近三年收益率、近1月收益率、近3月收益率、近6月收益率、今年以来收益率
P2P理财产品	P2P网贷平台基本信息	平台名称、平台所属企业名称、平均收益、投资期限、注册资金、自动投标、债权转让、资金托管、投标保障、保障模式、担保机构、银行存管、风险准备金存管、业务类型、平台背景
	P2P网贷平台所属公司信息	平台名称、公司名称、公司法人、公司类型、股东结构、注册资本、实缴资本、注册地址、开业日期、核准日期、登记机关、统一社会信用代码
	P2P问题平台	平台名称、公司名称、所在地、发生问题时间、法定代表人、注册资本、利率范围、期限范围、上线时间、公司规模、问题原因
	P2P网贷平台高管基本信息	平台名称、P2P网贷平台高管姓名、职位、毕业学校、专业、高管简介

5.数据范围

数据模块	数据起始时间
低风险理财产品	2004年
阳光私募基金产品	2003年
P2P理财产品	2000年

（五）基金公司研究数据库-FCRD

1.数据库简介

基金公司研究数据库（*Fund Company Research Database, FCRD*）基于相关基金公司的年报信息和披露信息等而建立的专业数据库。*FCRD* 包含基金公司基本信息、基金管理信息、基金经理信息 3 个模块，内容涵盖基金公司概况、股东情况、持股公司、旗下基金、管理费用等多方面数据，此外，还对基金经理的信息进行了详细展示。

2.数据库特色

- 较为完善的基金公司和基金经理数据

FCRD 包含了基金公司股东构成、持股公司、旗下基金等多类重要数据。*FCRD* 不仅介绍基本个人信息，还包括了基金经理的教育背景、工作经历、管理基金等重要信息。

3.典型数据举例

基金持股公司

基金公司代码	基金公司名称	统计日期	持股公司代码	持有基金数	持股总量(股)	占该公司流通股比例 (%)
500038	华安基金管理有限公司	2003-12-31	600019	6	57994372	3.08973744
500038	华安基金管理有限公司	2003-12-31	600036	6	43984734	2.93231561
500038	华安基金管理有限公司	2003-12-31	600009	5	39830897	5.80345732

4.数据库内容

数据模块	数据表	数据字段
基金公司基本信息	基金公司概况	基金公司代码、基金公司名称、成立日期、公司性质、注册资本、法定代表人、总经理、注册地址、办公地址
	基金公司股东构成	基金公司代码、基金公司名称、股东名称、统计日期、持股数量、持股比例(%)、股东性质
基金管理信息	基金持股公司	基金公司代码、基金公司名称、统计日期、持股公司代码、持股公司名称、持有基金数、持股总量、持股占该公司流通股比例(%)、持股总市值、持股市值占基金净值比例(%)、持股市值占基金股票投资市值比例(%)
	基金公司旗下基金	基金公司代码、基金公司名称、基金代码、基金名称、基金类型、基金经理、基金成立日期
	基金管理费用	基金代码、基金名称、统计日期、基金公司代码、基金管理公司、基金份额、资产净值、管理费用
基金经理信息	基金经理基本信息	基金经理姓名、性别、出生年份、学历/学位、个人简介
	基金经理工作经历	基金经理姓名、基金代码、基金简称、任职机构名称、职位、入职日期、离职日期
	基金经理管理基金	基金经理姓名、基金代码、基金名称、基金公司名称、开始日期、结束日期
	基金经理教育背景	基金经理姓名、基金代码、基金简称、学历/学位、专业、毕业学校

5.数据范围

数据模块	数据起始时间
基金公司基本信息	1998 年
基金管理信息	2001 年
基金经理信息	截面数据

CNRDS

CNRDS

CNRDS

CNRDS

（六）券商投行数据库-CIBD

1.数据库简介

券商投行数据库 (*Chinese Investment Bank Database, CIBD*) 从券商年报和披露信息中提取了券商财务信息、大股东及其持股比例信息、大股东性质信息,董监高个人信息等独特数据,提供了丰富和准确的券商投行相关数据。

2.数据库特色

- 独有的券商相关信息

CIBD 包括资本市场主要券商的大股东、高管、财务数据等,是目前唯一提供此类数据的数据库。

3.典型数据举例

券商股东构成

券商名称	券商代码	会计报告日期	大股东名称	大股东持股比例	大股东性质	最终控制人
爱建证券有限责任公司	B00001	2014-12-31	上海陆家嘴金融发展有限公司	0.51136	地方国有	地方国资局
安信证券股份有限公司	B00002	2014-12-31	国家开发投资公司	0.5801	中央国有	国资委

4.数据库内容

数据模块	数据表	数据字段
券商基本信息	股东构成	券商名称、券商代码、统计日期、大股东名称、大股东占股比例(%)、大股东性质、最终控制人
	高管特征	统计日期、券商名称、券商代码、董监高姓名、职位类型、性别、出生年份、学历/学位、现任或者曾任、工作经历起止日期、工作单位、职位、是否全职、个人简介
	承销排名	券商名称、券商代码、会计年度、当年承销 IPO 总募集资金金额、当年承销 IPO 总募集资金排名……
	财务比率	券商名称、券商代码、合并会计报表或母公司会计报表、统计日期、券商净资本、营业收入、营业收入增长率(%)……
券商财务报表	资产负债表	券商名称、券商代码、合并会计报表或母公司会计报表、统计日期、货币资金、客户资金存款、交易性金融资产、衍生金融资产、买入返售金融资产……
	利润表	券商名称、券商代码、合并会计报表或母公司会计报表、统计日期、营业收入、营业利润、营业外收入、营业外支出、利润总额、所得税、净利润……
	现金流量表	券商名称、券商代码、合并会计报表或母公司会计报表、统计日期、经营活动产生的现金流量净额……

5.数据范围

数据模块	数据起始时间
券商基本信息	2003 年
券商财务报表	2006 年

(七) 保险机构研究数据库-IIRD

1. 数据库简介

保险机构研究数据库 (*Insurance Institution Research Database, IIRD*) 基于中国财险、寿险等公司的年报数据以及中国保险年鉴的相关数据而建立的专业数据库。*IIRD* 囊括了 2005 年至今相关保险公司的经营和财务信息, 数据内容丰富完整, 结构清晰, 极大满足了学术研究的需求。

2. 数据库特色

- 丰富的财报字段, 涵盖大多数保险公司

IIRD 数据来源于各大保险公司年报及中国保险年鉴, 因此展示了较为完整的财报字段和全面的财务数据。*IIRD* 通过多方数据来源, 收集整理了大部分保险公司的基本信息和财务信息。

3. 典型数据举例

保险公司基本信息

保险公司代码	保险公司名称	保险公司类型	成立日期	所在城市
800005	安邦财产保险股份有限公司	保险公司(财险)	2011-12-31	深圳市
800006	安邦养老保险股份有限公司	保险公司(养老险)	2013-12-31	北京市

4. 数据库内容

数据模块	数据表	数据字段
保险公司基本信息	保险公司基本信息	保险公司代码、保险公司名称、保险公司类型、成立日期、注册资本、所在城市、注册地址、经营范围、法定代表人、总经理
保险公司财务报表	资产负债表	保险公司代码、保险公司名称、统计日期、货币资金、交易性金融资产、衍生金融资产、买入返售金融资产、应收利息、应收代位追偿款、应收分保未到期责任准备金、应收分保未决赔款准备金、应收分保寿险责任准备金、应收分保长期健康险责任准备金、保户质押贷款、短期借款……
	利润表	保险公司代码、保险公司名称、统计日期、营业收入、已赚保费、保费总收入、分保费收入、分出保费、提取未到期责任准备金、营业支出、赔付支出、摊回赔付支出、提取保险合同准备金净额、摊回保险责任准备金、保单红利支出、分保费用、营业利润、营业外收入、营业外支出……
	现金流量表	保险公司代码、保险公司名称、统计日期、经营活动产生的现金流量、投资活动产生的现金流量、筹资活动产生的现金流量
保险公司分支机构	机构设立情况	机构编码、机构名称、机构住所、邮政编码、发证日期、批准日期、发证机关、业务范围
	机构失控情况	机构编码、机构名称、机构住所、邮政编码、发证日期、批准日期、发证机关、业务范围
	机构退出列表	机构编码、机构名称、机构住所、邮政编码、发证日期、批准日期、发证机关、业务范围、退出日期

5. 数据范围

数据模块	数据起始时间
全部模块	2005 年

（八）信托机构研究数据库-CTID

1.数据库简介

信托机构研究数据库（*Chinese Trust Institution Database, CTID*）主要根据中国资本市场上信托机构披露的年报数据整理而成的数据库。涵盖了信托公司基本信息、股东构成、高管特征、审计信息、员工与股东信息以及财务信息等多类重要信息，为学者研究信托行业提供了完善而丰富的数据。

2.数据库特色

- 市场独有的信托公司数据

CTID 囊括了企业工商信息、股东信息、人员信息及财务信息等，字段丰富而完善。

3.典型数据举例

信托公司基本信息

信托公司代码	信托公司名称	所在城市	成立日期	注册资本
200001	陕西省国际信托股份有限公司	西安市	1985-01-05	3964012846
200002	云南国际信托有限公司	昆明市	1997-09-03	1200000000

4.数据库内容

数据模块	数据表	数据字段
信托公司 基本信息	公司概况	信托公司代码、信托公司名称、所在城市、成立日期、注册地址、注册资本、经营范围、法定代表人、总经理
	股东构成	信托公司代码、信托公司名称、统计日期、序号、股东名称、持股金额、期末股份数、期末持股比例（%）、出资比例（%）、法定代表人、注册资本、注册地址、主要经营业务
	高管特征	信托公司代码、信托公司名称、统计日期、高管姓名、性别、年龄、学历/学位、专业、职位、选任日期、兼职单位名称、兼职单位职务、任职的股东单位名称、任职的股东单位职务、金融从业年限、所推举的股东名称、推举股东的持股比例（%）、个人简介
	审计信息	信托公司代码、信托公司简称、统计日期、审计日期、审计意见类型、审计师、会计师事务所名称、审计报告
	员工与股东信息	信托公司代码、信托公司名称、统计日期、股东总数、员工总数……
信托公司 财务报表	资产负债表	信托公司代码、信托公司名称、统计日期、资产总计、流动资产、长期投资、流动负债、长期负债、负债合计、实收资本、所有者权益
	利润表	信托公司代码、信托公司名称、统计日期、营业收入、营业支出、投资收益、利润总额、净利润
	其他指标	信托公司代码、信托公司名称、统计日期、信托资产运用合计、信托收入合计、信托费用合计、信托损益、资本收益率（%）、信托报酬率（%）、员工人数

5.数据范围

数据模块	数据起始时间
全部模块	2003 年

(九) 金融中介处罚数据库-FIPD

1. 数据库简介

金融中介处罚数据库 (*Financial Intermediary Penalty Database, FIPD*) 是基于监管机构对资本市场相关金融中介机构作出监管和处罚措施整理而成的专业数据库。*FIPD* 的金融中介包括会计师事务所、证券公司、基金公司和保险公司等, 均列示了处罚类型、监管机构、具体内容、处罚对象、处罚原因、处罚决定等重要关键字段。*FIPD* 解决了资本市场金融中介处罚数据缺失和分散的痛点, 为该领域研究提供了重要数据支撑。

2. 数据库特色

- 数据来源权威和全面

FIPD 的处罚数据全部来自各个类型的官方监管机构, 数据来源丰富、权威。

3. 典型数据举例

基金公司处罚

处罚类型	事件标题	发文日期	监管机构	受处罚机构	处罚原因	处罚决定
监管措施	关于对汇添富基金管理股份有限公司采取责令改正措施的决定	2021-03-10	北京证监局	万鼎国际股权投资基金管理(北京)有限公司	你公司存在发行未备案私募基金的行为	我局决定对你公司采取责令改正行政监管措施。
监管措施	关于对广发基金管理有限公司采取责令改正措施的决定	2021-03-17	北京证监局	品今(北京)投资基金管理有限公司	你公司存在发行未备案私募基金的行为	我局决定对你公司采取责令改正行政监管措施。

4. 数据库内容

数据模块	数据表	数据字段
会计师事务所处罚	会计师事务所处罚	处罚类型、事件标题、发文日期、监管机构、具体内容、受处罚机构、受处罚人员 1、受处罚人员 2、受处罚人员 3、受处罚人员 4、处罚原因、处罚决定、涉及的上市公司
证券公司处罚	证券公司处罚	处罚类型、事件标题、发文日期、监管机构、具体内容、受处罚机构、受处罚人员 1、受处罚人员 2、处罚原因、处罚决定
基金公司处罚	基金公司处罚	处罚类型、事件标题、发文日期、监管机构、具体内容、受处罚机构、受处罚人员 1、受处罚人员 2、受处罚人员 3、受处罚人员 4、处罚原因、处罚决定
保险公司处罚	保险公司处罚	处罚类型、事件标题、发文日期、监管机构、具体内容、受处罚机构、受处罚人员、处罚原因、处罚决定

5. 数据范围

数据模块	数据起始时间
会计师事务所处罚	2001 年
证券公司处罚	2008 年
基金公司处罚	2002 年
保险公司处罚	2005 年

（十）商业银行数字化数据库-CBDD

1.数据库简介

商业银行数字化数据库 (*Commercial Banks Digitalize Database, CBDD*) 基于中国商业银行年报文本整理而成的专业数据库。*CBDD* 根据几百家中国商业银行数据披露的年报文本, 采用较为通用的数字化关键词汇列表, 对这些年报文本进行了数字化词频统计, 用以衡量商业银行的数字化情况。此外, 为了提供更多维度数据, *CBDD* 还提供了商业银行年报文本基本信息、可读性等其他重要文本模块。希望 *CBDD* 的推出能为该领域提供更多数据支持。

2.数据库特色

- 数据处理精确、可靠

CBDD 的数据来源于商业银行在官网披露的年报, 我们对数据进行了精确的文本清洗、处理和计算, 可直接用于学术研究。

3.典型数据举例

数字化词频统计

银行名称	会计年度	词汇名称	词汇类别	词汇组别	词频统计	出现该词的句子数
北京农村商业银行股份有限公司	2012	数据中心	搜索词汇	29	2	2
北京银行股份有限公司	2012	互联网金融	搜索词汇	19	1	1
成都银行股份有限公司	2012	数据中心	搜索词汇	29	1	1

4.数据库内容

数据模块	数据表	数据字段
商业银行年报基本信息	商业银行年报基本信息	银行名称、会计年度、总字数 1、总字数 2、总词数 1、总词数 2、文件大小、总页数、句子数
商业银行数字化	数字化词频统计	银行名称、会计年度、词汇名称、词汇类别、词汇组别、词频统计、出现该词的句子数
	信息化词频统计	银行名称、会计年度、词汇名称、词汇类别、词汇组别、词频统计、出现该词的句子数
	智能化词频统计	银行名称、会计年度、词汇名称、词汇类别、词汇组别、词频统计、出现该词的句子数
	互联网化词频统计	银行名称、会计年度、词汇名称、词汇类别、词汇组别、词频统计、出现该词的句子数
商业银行年报可读性	商业银行年报可读性	银行名称、会计年度、总字数 1 (个)、总字数 2 (个)、总词数 1 (个)、总词数 2 (个)、句子数 (个)、分句中的平均字数 (个)、副词和连词的比例 (%)、每一千字中常用字字数 (个)、每一千字中次常用字字数 (个)

5.数据范围

数据模块	数据起始时间
全部模块	2007 年

公司债券信息研究

(一) 发债企业财务报表数据库-BCFS

1. 数据库简介

发债企业财务报表数据库 (*Financial Statements of Bond-Issuing Company, BCFS*) 是以研究为目的设计的, 面向国内外机构投资者和学术研究人员, 旨在提供最完整、最精确的发债企业财务数据。*BCFS* 的数据主要来源于中国货币网、中国债券信息网等披露的发债企业年报中的财务报表数据, 在数据处理过程中, 保留了研究中常用的财务报表数据字段, 并提供资产负债表、利润表和现金流量表三个模块的数据。*BCFS* 填补了债券领域研究财务数据的缺失, 为该领域研究提供了数据支持。

2. 数据库特色

- 来源权威、内容详细

BCFS 根据中国货币网及中国债券信息网披露的财务报告, 全面汇总了发债企业自 1990 年至今的财务报表数据。

3. 典型数据举例

资产负债表

发债企业	统计日期	货币资金 (元)	存货 (元)	固定资产 (元)
中国邮政集团有限公司	2016-12-31	1502858225463.29	6695961635.63	93679387880.20
中国邮政集团有限公司	2017-12-31	1724781288681.16	7461082313.83	97899088878.55
中国邮政集团有限公司	2018-12-31	1359659895933.56	7629022652.69	100679795670.32
安徽省皖能股份有限公司	2016-12-31	1215200380.39	235518332.00	12429097536.54
安徽省皖能股份有限公司	2017-12-31	907752289.53	144667132.01	13872358236.29

4. 数据库内容

数据模块	数据表	数据字段
资产负债表	资产负债表	发债企业、统计日期、货币资金、交易性金融资产、应收票据、应收账款、预付账款、其他应收款、存货、流动资产合计、长期应收款、长期股权投资、投资性房地产、固定资产、在建工程、无形资产、开发支出、商誉、递延所得税资产、非流动资产合计、资产总计……
利润表	利润表	发债企业、统计日期、营业总收入、主营业务收入、主营业务成本、营业税金及附加、销售费用、管理费用、财务费用……
现金流量表	现金流量表	发债企业、统计日期、销售商品、提供劳务收到的现金、收到的税费返还、购买商品、接受劳务支付的现金、支付给职工以及为职工支付的现金、支付的各项税费、经营活动产生的现金流量净额……

5. 数据范围

数据模块	数据起始时间
全部模块	1990 年

(二) 发债企业财务报表附注数据库-BNFS

1.数据库简介

发债企业财务报表附注数据库 (Notes to Financial Statements Database of Bond Issuing company, *BNFS*) 是基于发债企业年报数据, 对报表附注数据进行加工整理而成的专业数据库。根据报表附注中的披露项目, *BNFS* 设置了资产项目、应收项目、债务项目、费用明细、损益项目、其他项目等 6 个方面的财务数据, 可为财务报表使用者提供了完整、准确的报表附注数据支持。

2.数据库特色

- 丰富的报表附注字段

BNFS 包含多种类型的报表附注字段, 从资产、债务、债权和损益等多个方面体现了附注披露情况。

3.典型数据举例

应收账款

发债企业	统计日期	应收账款金额	应收账款比例	应收账款坏账准备
内江兴元实业集团有限责任公司	2012-12-31	139482704.4	43.32	
玉环市城市建设集团有限公司	2012-12-31	9765422.94	39.41	
吉林经济技术开发区经济技术开发总公司	2012-12-31	383105183.4	42.79	5079.47
秦皇岛城市发展投资控股集团有限公司	2012-12-31	12629575.02	59.03	258070.71
重庆开乾投资集团有限公司	2012-12-31	14603.06	9.12	

4.数据库内容

数据模块	数据表	数据字段
资产项目	存货	发债企业、统计日期、原材料、在产品、产成品、低值易耗品、包装物、已完工未结算资产
	资产减值准备	发债企业、统计日期、坏账准备合计、坏账准备_应收账款、坏账准备_其他应收款、存货跌价准备合计、存货跌价准备_库存商品、存货跌价准备_原材料、存货跌价准备_产成品、存货跌价准备_开发成本、存货跌价准备_在产品、存货跌价准备_开发产品、存货跌价准备_自制半成品、长期投资减值准备合计、长期投资减值准备_长期股权投资、固定资产减值准备合计、固定资产减值准备_房屋、建筑物、固定资产减值准备_机器设备、固定资产减值准备_专用设备、固定资产减值准备_运输工具、固定资产减值准备_电子设备、固定资产减值准备_其他设备、无形资产减值准备合计、无形资产减值准备_专利权、无形资产减值准备_商标权、委托贷款减值准备、贷款呆账准备
	商誉	发债企业、统计日期、商誉减值准备、商誉账面价值
应收项目	应收账款	发债企业、统计日期、应收账款金额、应收账款比例、应收账款坏账准备
债务项目	短期负债	发债企业、统计日期、短期应付债券
	应交税费	发债企业、统计日期、增值税、营业税、企业所得税、城市维护建设税、房产税、个人所得税、消费税、土地使用税、印花税、车船使用

		税、土地增值税、资源税、教育费附加、地方教育费附加、其他税项
费用明细	研发费用	发债企业、统计日期、研发支出合计、研发支出总额占营业收入比例、本期资本化研发支出、研发费用明细_工资薪酬、研发费用明细_折旧摊销费用、研发费用明细_租赁费、研发费用明细_其他
损益项目	资产减值损失	发债企业、统计日期、坏账损失、存货跌价损失、商誉减值损失、可供出售金融资产减值损失、持有至到期投资减值损失、固定资产减值损失
	政府补助	发债企业、统计日期、政府补助
其他项目	审计意见	发债企业、统计日期、审计单位、签字注册会计师、当期实付审计费用、审计意见类别
	客户及供应商	发债企业、统计日期、主要客户名称、主要客户销售额合计、主要客户销售额合计占比、主要供应商名称、主要供应商采购额合计、主要供应商采购额合计占比

5.数据范围

数据表	数据起始时间
存货	2007 年
资产减值准备	2007 年
商誉	2007 年
应收账款	2007 年
短期负债	2018 年
应交税费	2007 年
研发费用	2007 年
资产减值损失	2007 年
政府补助	2007 年
审计意见	2007 年
客户及供应商	2007 年

(三) 上市公司发债研究数据库-BIRD

1.数据库简介

上市公司发债研究数据库 (*Bond Issuance Research Database of Listed Company, BIRD*) 是基于上市公司发行一般债券和可转债公告整理而成的专业数据库。*BIRD* 涵盖上市公司一般债券发行情况和预案情况, 以及债券及发行主体的评级情况, 总共分三大模块, 全方位展示了上市公司的债券市场融资情况。

2.数据库特色

- 内容详细完善

BIRD 不仅包括了上市公司债券的发行信息和预案信息, 还提供了债券的评级信息, 为研究提供了多维度数据支持。

3.典型数据举例

上市公司可转债发行预案

股票代码	预案公告日	方案进度	发行方式	股东大会公告日	发行规模(亿元)	发行期限(年)	利率类型	利率(%)	付息频率
000034	2022-12-27	董事会修改方案		2023-01-10	13.49	6			1年1次
000062	2019-06-12	董事会取消方案	非公开发行						
000155	2022-05-30	董事会批准	非公开发行			3	递进利率	0.5	1年1次
000430	2018-04-03	停止实施		2018-04-25	6.1	6			1年1次
000537	2001-09-25	股东大会否决	优先配售, 网上发行, 网下发行	2001-10-26	6	5			1年1次
000601	2001-07-28	股东大会否决	网上发行, 优先配售, 网下发行	2001-09-13	6.8	5	累进利率	1	1年1次
000623	2001-08-30	股东大会批准	交易所网上发行, 非公开发行	2001-10-09	4	5	固定利率	1.2	1年1次

4.数据库内容

数据模块	数据表	数据字段
上市公司一般债券发行	上市公司一般债券发行	股票代码、股票简称、发行年份、发行日期、债券名称、债券类型、发行对象、发行规模(亿元)、债券发行面值、期限(年)、债券形式、利率类型、债券发行票面利率(%)、债券票面利率备注、计息方式、计息基准、付息方式、付息频率、付息日说明、上市日期、上市地点、发行币种、起息日期、到期日期、计息截止日、摘牌日期

数据模块	数据表	数据字段
	上市公司一般债券发行预案	股票代码、股票简称、董事会预案公告日、债券类型、方案进度、发行数量(亿元)、数量说明、期限说明、利率说明、主承销商、是否向股东配售、向股东配售安排、决议有效期、股东大会公告日、发改委获准公告日、证监会批准公告日期、发审委通过公告日、实施公告日
上市公司可转债发行	上市公司可转债发行	股票代码、股票简称、转债代码、转债名称、信用评级、发行方式、发行起始日、发行规模(亿元)、发行期限(年)、票面利率(%)、补偿利率(%)、付息说明、利率类型、付息频率、上市日期、起息日期、到期日期、转股起始日、初始转股价(元)、担保方式、担保范围、评估机构、主承销商
	上市公司可转债发行预案	股票代码、股票简称、预案公告日、方案进度、发行方式、股东大会公告日、发行规模(亿元)、发行期限(年)、利率类型、利率(%)、付息频率
上市公司债券评级	上市公司债券信用评级	债券名称、发行时债项评级、发行时主体评级、发行时主体评级展望、发行人委托评级机构
	债券发行主体历史信用评级	债券名称、发布日期、评级类型、信用评级、评级机构

5.数据范围

数据模块	数据起始时间
上市公司一般债券发行	2007年
上市公司一般债券发行预案	2003年
上市公司可转债发行	2002年
上市公司可转债发行预案	2001年
上市公司债券信用评级	截面数据
债券发行主体历史信用评级	2005年

社会经济组织研究

(一) 政府审计数据库-CGAD

1. 数据库简介

政府审计数据库 (Government Audit Database, **CGAD**) 是基于中国审计年鉴、国家审计署等官方数据整理而成的专业数据库。该数据库全面记录了各省、市、县在各年度的政府审计情况、违规处理信息等。**CGAD** 共包括两张表, 分别是地方政府审计信息、审计工作综合信息。

2. 数据库特色

- 字段丰富

CGAD 除了包括各地市每年的审计结果信息, 还包括审计署及其分派机构、地方审计单位的详细审计情况。

3. 典型数据举例

地方政府审计信息

省份名称	地市名称	会计年度	审计机关	完成审计项目(个)	应上缴财政(万元)
安徽省	宿州市	2013	宿州市本级	26	68
安徽省	宿州市	2013	萧县审计局	17	48
安徽省	宿州市	2013	灵璧县审计局	36	50
安徽省	宿州市	2013	泗县审计局	13	1209
安徽省	淮北市	2013	淮北市本级	52	3847

4. 数据库内容

数据模块	数据表	数据字段
地方政府 审计信息	地方政府 审计信息	省份名称、地市名称、会计年度、审计机关、完成审计项目、审计查出主要问题金额、应上缴财政、应减少财政拨款或补贴、应归还原渠道资金、应调账处理金额、移送处理事项、出具审计报告数量、提出审计建议、提交审计信息
审计工作 综合情况	审计工作 综合情况	指标名称、计量单位、会计年度、合计、审计署本级汇总、审计署派出局汇总、审计署特派办汇总、地方审计机关汇总、北京市、天津市、河北省、山西省、内蒙古自治区、辽宁省、吉林省、黑龙江省、上海市、江苏省、浙江省、安徽省、福建省……

5. 数据范围

数据模块	数据起始时间
全部模块	1999年

(二) 中国高校科技统计数据库-USTS

1. 数据库简介

中国高校科技统计数据库 (*Science and Technology Statistics Database of Chinese University, USTS*) 主要摘录和整理了来自中华人民共和国教育部科学技术司《高等学校科技统计资料汇编》的重要数据。该数据库记录了按学校和按地区统计的中国各大高校的科研情况，详细反映了高校的科技活动、研发情况和科技成果等情况。

2. 数据库特色

- 数据权威可靠

USTS 的主要数据来自教育部，因此统计数字翔实可靠，字段丰富，可直接为研究者使用。

3. 典型数据举例

高校科技活动分学校统计情况

学校名称	会计年度	教研人员总数	科技经费总额	科技课题总数	专著数量	成果授奖
北京大学	2012	12601	2342999	3815	31	51
清华大学	2012	5293	3892514	5135	45	115
北京交通大学	2012	1418	718431	3311	17	29
北京工业大学	2012	1796	640682	2275	14	31
北京邮电大学	2012	1393	407970	1388	22	7

4. 数据库内容

数据模块	数据表	数据字段
高校科技活动	高校科技活动分学校统计情况	学校名称、学校类型、会计年度、教学与科研人员总数、其中：科学家与工程师人数 a、其中：具有高级职称的科学家与工程师人数 a、研究与发展人员总数、其中：科学家与工程师人数 b、其中：具有高级职称的科学家与工程师人数 b、全时当量研究与发展人员、科技经费总额、科技经费当年拨入金额、其中：政府资金拨入、其中：当年企事业单位委托拨入（进入学校财务）、其中：当年其他拨入、当年科技经费内部支出金额、科技课题总数、科技课题当年投入人数、科技课题当年拨入经费、科技课题当年支出经费、专著数量、专著字数、学术论文……
	高校科技活动分地区统计情况	地区名称、会计年度、教学与科研人员总数、其中：科学家与工程师人数 a、其中：具有高级职称的科学家与工程师人数 a、研究与发展人员总数、其中：科学家与工程师人数 b、其中：具有高级职称的科学家与工程师人数 b、全时当量研究与发展人员、科技经费总额、科技经费当年拨入金额、其中：政府资金拨入、其中：当年企事业单位委托拨入（进入学校财务）、其中：当年其他拨入、当年科技经费内部支出金额、科技课题总数、科技课题当年投入人数、科技课题当年拨入经费、科技课题当年支出经费、专著数量、专著字数、学术论文、其中：国外及全国性刊物发表……

数据模块	数据表	数据字段
高校研发情况	高校研发机构分地区统计情况	地区名称、会计年度、机构数、从事研究与发展人员合计数、其中：具有高级职称人数、其中：具有中级职称人数、其中：具有初级职称人数、其中：具有其他职称人数
	高校研发项目分地区统计情况	地区名称、会计年度、研发项目合计数、研发项目合计当年投入人员、研发项目合计参与研究生、研发项目合计当年拨入经费、研发项目合计当年支出经费、基础研究项目数、基础研究当年投入人员、基础研究参与研究生、基础研究当年拨入经费、基础研究当年支出经费、应用基础项目数、应用基础当年投入人员、应用基础参与研究生、应用基础当年拨入经费、应用基础当年支出经费、试验发展项目数、试验发展当年投入人员、试验发展参与研究生、试验发展当年拨入经费、试验发展当年支出经费
高校科技成果	高校科技成果分地区统计情况	地区名称、会计年度、出版科技著作数量、出版科技著作字数、发表学术论文合计、其中：发表在国外学术刊物的学术论文数、国家级项目验收合计数、其中：与其他单位合作的国家级项目验收、验收项目来源于 973 计划的数量、验收项目来源于科技攻关计划的数量、验收项目来源于 863 计划的数量……
	高校技术转让分地区统计情况	地区名称、会计年度、签订技术转让合同合计数、与国有企业签订的技术转让合同数、与外资企业签订的技术转让合同数、与民营企业签订的技术转让合同数、与其他机构签订的技术转让合同数、签订技术转让合同合计金额、与国有企业签订的技术转让合同金额……
	高校科技获奖分地区统计情况	地区名称、会计年度、国家最高科学技术奖特等奖、国家自然科学奖二等奖、国家发明奖一等奖、国家发明奖二等奖、国家科技进步奖特等奖、国家科技进步奖一等奖、国家科技进步奖二等奖、国务院各部门科技进步奖、省、自治区、直辖市科技进步奖
其他	高校科技人力分地区统计情况	地区名称、会计年度、学校数、教学与科研人员合计、其中：科学家和工程师 a、研究与发展人员合计、其中：科学家和工程师 b、研究与发展全时人员合计、……
	高校科技经费分地区统计情况	地区名称、会计年度、学校数、拨入经费合计、其中：拨入科研事业费、其中：拨入主管部门专项费、其中：拨入其他政府部门专项费、其中：拨入企事业单位委托经费、其中：各种收入中转为科技经费、其中：拨入其他经费、支出经费合计、其中：支出科研人员费、其中：支出业务费、其中：支出固定资产购置费、其中：支出上缴税金、其中：支出其他经费、其中：支出转拨给外单位经费
	高校国际科技交流分地区统计情况	地区名称、会计年度、国际合作研究派遣、国际合作研究接受、国际学术会议出席人员、国际学术会议交流论文、国际学术会议特邀报告、国际学术会议主办

5.数据范围

数据模块	数据起始时间
全部模块	2003 年

(三) 中国非营利组织数据库-CNGO

1. 数据库简介

中国非营利组织数据库 (*Chinese Non-Government Organization Database, CNGO*) 是面向广大高校教授和学者以及实务界人士, 研究中国非营利组织机构特征、财务状况、人员概况以及捐赠情况的数据库。目前该数据库主要包括我国所有公募和非公募慈善基金会的独特数据, 是反映中国慈善基金会发展状况较为全面的学术数据库。

2. 数据库特色

- 全面

在时间跨度上, *CNGO* 包括所有中国慈善基金会 2005 年至今的所有基本信息, 学术研究提供较长的时间序列数据; 在空间跨度上, 除了个别未公开披露、数据缺失的基金会, *CNGO* 收录了中国大部分的基金会信息, 这为学术研究提供了大量的横截面数据。

- 唯一

CNGO 的高管模块收集了基金会大部分理事和监事在其他单位的任职背景, 这为考察慈善基金会的各种行为提供了不同视角。

3. 典型数据举例

(1) 基金会信息

基金会代码	基金会名称	基金会类型	基金会范围	成立时间	原始基金数额
10001	爱佑慈善基金会	非公募	全国	2008/5	2000 万
10002	北京大学教育基金会	非公募	全国	1995/7	2000 万
10293	北京当代艺术基金会	公募	地方	2008/12	400 万

(2) 基金会财务概况

基金会名称	会计年度	年度总收入 (元)	捐赠收入 (元)	年度总支出 (元)	公益事业支出 (元)
北京大学教育基金会	2014	786623183	567967683	309490517	302509385
北京大学教育基金会	2013	1489656769	1395841805	175105873	209366023
北京大学教育基金会	2012	428954733	383523157	267340159	259211412

4. 数据库内容

数据模块	数据表	数据字段
基本信息	基金会信息	基金会代码、基金会名称、基金会地址、所属省份、宗旨、基金会类型、基金会范围、志愿者数量、评估等级、业务主管单位、成立时间、原始基金数额
	基金会关联方	基金会代码、基金会名称、会计年度、关联方、与基金会的关系
	审计信息	基金会代码、基金会名称、会计年度、审计意见、审计机构、注册会计师_1、注册会计师_2
财务信息	财务概况	基金会代码、基金会名称、会计年度、净资产、年度总收入、捐赠收入、投资收入、服务收入、政府补助收入、其他收入、年度总支出、用于公益事业的支出、工作人员工资福利支出、行政办公支出、其他支出

数据模块	数据表	数据字段
财务信息	业务活动表	基金会代码、基金会名称、会计年度、捐赠收入_非限定性、会费收入_非限定性、提供服务收入_非限定性、商品销售收入_非限定性、政府补助收入_非限定性、投资收益_非限定性、其他收入_非限定性、收入合计_非限定性、业务活动成本_非限定性、管理费用_非限定性、筹资费用_非限定性、其他费用_非限定性、费用合计_非限定性……
	资产负债表	基金会代码、基金会名称、会计年度、货币资金、短期投资、应收款项、预付账款、待摊费用、一年内到期的长期债权投资、其他流动资产、流动资产合计、长期股权投资、长期债权投资、长期投资合计、固定资产原价、累计折旧、固定资产净值、在建工程、文物文化资产、固定资产清理、固定资产合计、无形资产……
捐赠情况	公益支出情况	基金会代码、基金会名称、会计年度、基金会类型、上年度基金余额（非公募）/上年度实际收入合计(公募)、本年度总支出、本年度用于公益事业的支出、工作人员工资福利支出、行政办公支出
	接受捐赠情况	基金会代码、基金会名称、会计年度、本年度捐赠收入、来自境内的捐赠、来自境内自然人的捐赠、来自境内法人或者其他组织的捐赠、来自境外的捐赠、来自境外自然人的捐赠、来自境外法人或者其他组织的捐赠
	大额捐赠情况	基金会代码、基金会名称、会计年度、捐赠方、捐赠金额、捐赠用途
	重大公益项目	基金会代码、基金会名称、会计年度、项目名称、大额支付对象、支付金额、占年度公益总支出比例（%）、用途
	业务活动情况	基金会代码、基金会名称、会计年度、项目名称、项目本年度收入、项目本年度支出、项目介绍
高管信息	理事背景	基金会代码、会计年度、基金会名称、理事姓名、理事会职务、性别
	监事背景	基金会代码、会计年度、基金会名称、监事姓名、性别

5.数据范围

数据模块	数据起始时间
全部模块	2005年

（四）中国商会研究数据库-CCCD

1.数据库简介

中国商会研究数据库（*Chamber of Commerce Research Database, CCCD*）中所涉及的商会主要指异地商会，是指由异地同一籍自然人或法人在某级行政辖区内投资兴办并自愿发起组成，以原籍地行政区域名称命名为基本特征的非营利性社会团体。具体来说，*CCCD*数据库包含商会基本信息、理事会成员、主要会员企业 3 个模块，共 4 张表。商会基本信息模块主要包括商会自身的一些基本信息，例如，名称、原籍地、活动地域、成立时间等；理事会成员包括理事的个人信息，如姓名、性别、最高学历、毕业学校、籍贯等信息；主要会员企业信息包括会员企业基本信息以及股东出资信息情况。

2.数据库特色

- 唯一

CCCD 是目前国内唯一一个关注商会研究的数据库，弥补了该领域研究的数据缺陷。

3.典型数据举例

商会会员基本信息

商会代码	商会名称	会员企业名称	公司类型	法定代表人	注册资本	成立日期
A00001	上海市浙江商会	上海复星高科技（集团）有限公司	有限责任公司	郭广昌	480000 元	2005-03-08
A00001	上海市浙江商会	万丰奥特控股集团	有限责任公司	陈爱莲	12000 元	1998-03-04
A00001	上海市浙江商会	上海均瑶（集团）有限公司	有限责任公司	王均金	80000 元	2001-02-14

4.数据库内容

数据模块	数据表	数据字段
商会基本信息	基本信息	商会名称、原籍地、活动地域、简介、章程、成立日期
理事会成员	理事会成员	商会名称、理事职位、理事姓名、性别、最高学历/学位、毕业学校、籍贯、政治面貌、理事所在单位及职位、理事其他社会职务、理事个人简介
主要会员企业	会员基本信息	商会名称、会员企业名称、经营状态、公司类型、成立日期、法定代表人、注册资本、注册资本单位、营业期限、登记机关、发照日期、所属行业、企业地址、经营范围、公司简介
	股东与出资	商会名称、公司名称、股东、认缴出资金额、认缴出资金额单位、认缴出资时间、实缴出资金额、实缴出资金额单位、实缴出资时间

5.数据范围

数据模块	数据起始时间
全部模块	截面数据

（五）高新技术企业研究数据库-HTED

1.数据库简介

高新技术企业研究数据库 (*High-tech Enterprises Research Database, HTED*) 是基于各省市科技厅、科技委员会披露的高新技术企业认定名单整理而成的专业数据库。*HTED* 囊括了高新技术企业的认定情况、更名情况和复审情况等重要信息。此外, *HTED* 还统计了高新技术企业专利申请及授权情况。

2.数据库特色

- 权威、全面的数据内容

HTED 数据来源于中国所有重要省市科技厅等官方网站公布的认定名单, 数据权威且全面。

3.典型数据举例

高新技术企业认定情况

省份	年份	批次	企业编号	企业名称	证书编号	地区
安徽省	2008	1	GX00113820	合肥阳光电源有限公司	GR200834000001	合肥市
安徽省	2008	1	GX00002011	安徽合力股份有限公司	GR200834000003	合肥市
安徽省	2008	1	GX00113577	合肥通用机械研究院	GR200834000004	合肥市
安徽省	2008	1	GX00001763	安徽广成科技有限责任公司	GR200834000005	合肥市

4.数据库内容

数据模块	数据表	数据字段
高新技术企业认定情况	高新技术企业认定情况	省份、年份、批次、序号、企业编号、企业名称、证书编号、发证日期、地区
高新技术企业更名情况	高新技术企业更名情况	省份、年份、批次、序号、变更前企业编号、变更前企业名称、变更后企业编号、变更后企业名称、更名方式、证书编号、发证日期、地区
高新技术企业复审情况	高新技术企业复审情况	省份、年份、批次、序号、企业编号、企业名称、证书编号、发证日期、地区
高新技术企业专利申请与授权情况	高新技术企业专利申请与授权情况	企业编号、企业名称、年份、当年独立申请的发明数量、当年独立申请的实用新型数量……
高新技术企业基本信息	高新技术企业基本信息	企业编号、企业名称、登记状态、法定代表人、注册资本、成立日期、核准日期、所属省份、所属城市、所属区县、参保人数、企业类型、所属行业、企业地址、经营范围

5.数据范围

数据模块	数据起始时间
高新技术企业认定情况	2005 年
高新技术企业更名情况	2008 年
高新技术企业复审情况	2008 年
高新技术企业专利申请与授权情况	1985 年
高新技术企业基本信息	截面数据

(六) 全国高校专利研究数据库-CUID

1. 数据库简介

全国高校专利研究数据库 (*Innovation of Chinese Universities Research Database, CUID*) 是基于教育部公布的全国普通高校名单, 以及国家知识产权局专利数据, 从而建立的研究中国高校专利申请及授权情况的专业数据库。*CUID* 从全国普通高校统计、高校专利申请与授权、高校绿色专利、高校数字经济专利四个模块对我国高校专利研究情况进行展示。

2. 数据库特色

- 多维度展示高校专利情况

CUID 不仅统计了高校历年专利的申请和授权情况, 还统计了高校绿色专利及数字经济专利情况。

3. 典型数据举例

高校绿色专利

高校名称	会计年度	当年独立授权的绿色发明数量	当年独立授权的绿色实用新型数量	当年联合授权的绿色发明数量	当年联合授权的绿色实用新型数量	当年授权的绿色专利总数
常州大学	2016	28	4	4	1	37
东北大学	2016	18	6	4	0	28
东华大学	2016	21	1	8	0	30

4. 数据库内容

数据模块	数据表	数据字段
全国普通高校统计	全国普通高校统计	统计日期、省份、学校数量、高校名称、学校标识码、主管部门、所在地、办学层次、备注
高校专利申请与授权	高校专利申请与授权情况	高校历史关系编码、高校名称、会计年度、当年独立申请的发明数量、当年独立申请的实用新型数量、当年独立申请的外观设计数量……
	高校专利申请后授权情况	高校历史关系编码、高校名称、会计年度、更新日期、专利申请个数、专利申请_授权个数、专利申请_当年授权个数、专利申请_第二年授权个数……
	高校专利授权年限情况	高校历史关系编码、高校名称、会计年度、专利授权、专利授权_平均年限、发明专利授权、发明专利授权_平均年限、发明专利授权_公开平均年限、实用新型授权、实用新型授权_平均年限、外观设计授权、外观设计授权_平均年限
高校绿色专利	高校绿色专利	高校历史关系编码、高校名称、会计年度、当年独立申请的绿色发明数量、当年独立申请的绿色实用新型数量……
高校数字经济专利	高校数字经济专利	高校历史关系编码、高校名称、会计年度、当年独立申请的数字经济发明数量、当年独立申请的数字经济实用新型数量……

5. 数据范围

数据模块	数据起始时间
全国普通高校统计	2007 年
高校专利相关模块	1985 年

（七）战略联盟研究数据库-SARD

1.数据库简介

战略联盟研究数据库 (*Strategic Alliance Research Database, SARD*) 是基于专利研发联盟和上市公司战略合作等相关数据研发的专业数据库。*SARD*包括四个模块：企业研发联盟、校企研发联盟、高校研发联盟以及公司战略合作，其中企业研发联盟是指企业与企业之间联合申请与获得专利；校企研发联盟是指企业与高校之间联合申请与获得专利；高校研发联盟是指高校之间联合申请与获得专利；公司战略合作则包含了上市公司公告的与其他相关方进行的战略合作情况。

2.数据库特色

- 多维度展示公司战略联盟情况

SARD 不仅展示了企业及高校之间的研发联盟情况，还囊括了上市公司战略合作情况。

3.典型数据举例

企业研发联盟_发明公布

专利公告号	专利类型	专利分类号	专利权人数	专利申请年度	研发联盟类型
CN85100648	发明公布	C23C28/00	2	1985	企业研发联盟
CN85108248	发明公布	H04M1/27	2	1985	企业研发联盟

4.数据库内容

数据模块	数据表	数据字段
企业研发联盟	企业研发联盟_发明公布	专利公告号、专利类型、专利权人、专利分类号、专利权人数、专利申请年度、研发联盟类型
	企业研发联盟_发明授权/实用新型/外观设计	专利公告号、专利类型、专利权人、专利分类号、专利权人数、专利申请年度、专利授权公告年度、研发联盟类型
校企研发联盟	校企研发联盟_发明公布	专利公告号、专利类型、专利权人、专利分类号、专利权人数、专利申请年度、研发联盟类型
	校企研发联盟_发明授权/实用新型/外观设计	专利公告号、专利类型、专利权人、专利分类号、专利权人数、专利申请年度、专利授权公告年度、研发联盟类型
高校研发联盟	高校研发联盟_发明公布	专利公告号、专利类型、专利权人、专利分类号、专利权人数、专利申请年度、研发联盟类型
	高校研发联盟_发明授权/实用新型/外观设计	专利公告号、专利类型、专利权人、专利分类号、专利权人数、专利申请年度、专利授权公告年度、研发联盟类型
公司战略合作	上市公司战略合作	股票代码、公司简称、公告标题、公告时间、战略合作公司、战略合作内容、备注

5.数据范围

数据模块	数据起始时间
专利研发联盟相关模块	1985年
上市公司战略合作	2006年

（八）排污许可研究数据库-DPRD

1.数据库简介

排污许可研究数据库（Discharge Permit Research Database, **DPRD**）基于生态环境部公布的各地排污许可公开信息整理而成的专业数据库。包含三个模块：排污许可企业、限期整改企业、碳排放企业，其中，排污许可企业展示了几十万条各省市的排污企业情况；限期整改企业是在排污过程中存在违规行为而受到处罚的企业信息公示；碳排放企业则是在生态环境部公示的重点碳排放企业名单。**DPRD**全方位展示了我国企业的排污监管情况，为该领域研究提供了数据支持。

2.数据库特色

- 详细的排污企业信息

DPRD 展示了取得排污许可企业的经营场所、排放明细、排放标准等详细信息。

3.典型数据举例

碳排放企业

省份名称	地市名称	区县名称	单位名称
广东省	惠州市	惠州大亚湾经济技术开发区	广东惠州天然气发电有限公司
贵州省	遵义市	赤水市	贵州赤天化纸业股份有限公司
河北省	石家庄市	栾城区	河北宏源热电有限责任公司
河北省	邢台市	宁晋县	河北健民淀粉糖业有限公司
河南省	安阳市	林州市	大唐林州热电有限责任公司
湖北省	孝感市	汉川市	汉川市福星热电有限公司

4.数据库内容

数据模块	数据表	数据字段
排污许可企业	排污许可企业	省份名称、地市名称、区县名称、排污许可证编号、排污单位名称、行业类别、有效期限、发证日期、发证机关、业务类型、版本、主要污染物类别、大气主要污染物种类、大气污染物排放执行标准、废水主要污染物种类、废水污染物排放执行标准
限期整改企业	限期整改企业	省份名称、地市名称、区县名称、许可证编号、排污整改单位名称、行业类别、下达时间、序号、整改问题、整改措施、整改时限、整改计划、整改通知书编号、业务类型、版本、办结日期、有效期限
碳排放企业	碳排放企业	省份名称、地市名称、区县名称、单位名称

5.数据范围

数据模块	数据起始时间
排污许可企业	2017 年
限期整改企业	2017 年
碳排放企业	截面数据

（九）小巨人企业数据库-SGED

1.数据库简介

小巨人企业数据库（*Small Giant Enterprise Database, SGED*）是根据工信部发布的专精特新“小巨人”企业名单整理而成的专业数据库。专精特新“小巨人”企业是由工信部认定的专注于细分市场、创新能力强、市场占有率高、掌握关键核心技术、质量效益优的排头兵企业。*SGED*搜集了2019年以来工信部公布的多个批次的小巨人企业名单，并在此基础上匹配了这些公司的上市情况、专利申请授权情况和工商信息等，期望能为该领域研究提供数据支持。

2.数据库特色

- 数据可靠、精确

SGED 的数据来源于工信部公布的名单，并精准匹配了上市情况和专利情况。

3.典型数据举例

专精特新小巨人企业认定情况

省份	年份	批次	企业编号	企业名称	地区	经度	纬度	是否上市公司	上市板块	与上市公司关系
天津市	2019	第一批	1006	瑞普（天津）生物药业有限公司	东丽区	117.376900	39.068340	参控股公司		子公司
天津市	2019	第一批	1008	天津久日新材料股份有限公司	北辰区	117.126420	39.259860	A股	A股	上市公司本身

4.数据库内容

数据模块	数据表	数据字段
认定情况	专精特新小巨人企业认定情况	省份、年份、批次、序号、企业编号、企业名称、地区、经度、纬度、是否上市公司、上市板块、与上市公司关系
更名情况	专精特新小巨人企业更名情况	省份、年份、批次、序号、企业编号、变更前企业名称、变更后企业名称、地区
专利申请与授权情况	专精特新小巨人企业专利申请与授权情况	企业编号、企业名称、年份、当年独立申请的发明数量、当年独立申请的实用新型数量、当年独立申请的外观设计数量、当年联合申请的发明数量、当年联合申请的实用新型数量、当年联合申请的外观设计数量……
基本信息	专精特新小巨人企业基本信息	企业编号、企业名称、登记状态、法定代表人、注册资本、成立日期、核准日期、所属省份、所属城市、所属区县、参保人数……

5.数据范围

数据模块	数据起始时间
认定情况	2019年
更名情况	2019年
专利申请与授权情况	1985年
基本信息	截面数据

三、CNRDS 经济特色数据库

随着中国成为国际上举足轻重的经济体，中国的经济发展模式和经济奇迹受到世界各国的关注，因此对于中国经济问题的研究逐渐成为国内外学术界的研究热点和前沿，而这些研究过程中的数据成为学者们的关键需求。**CNRDS**正是基于国内外学者在研究中国经济问题过程中的数据需求和数据质量要求，推出了一系列重要的经济特色数据库。其特色之处在于这些数据库非常符合中国经济问题研究过程中学者所关注的热点问题，虽获取难度大，但数据价值高且各有特色，可满足不同研究领域的独特数据需求。目前，**CNRDS**根据已有研究，结合经济热点问题，推出了以下六个系列数据：

宏观经济研究系列 主要包括中国经济发展过程中具有一定特色的宏观经济相关数据，例如夜间灯光数据、高铁航线数据等。

区域经济研究系列 主要包括中国区域经济发展的数据，例如中国城市统计数据、中国农村统计数据等。

产业经济研究系列 主要分不同产业建立了相应数据库，例如，中国能源数据、房地产数据等。

对外经济研究系列 囊括中国经济对外开放过程中的相关数据，包括对外直接投资、海外并购、一带一路等重要数据。

人文社科研究系列 主要包括科教文卫等领域的的数据，例如，儒家文化数据、科技统计数据、旅游统计数据等。

财政金融研究系列 主要包括财政、税收、金融和保险等经济领域数据，例如，税收数据、保险数据、PPP数据等。

总之，经济特色库的推出不仅对具体数据库进行了清晰、正确的归类，更方便了经济研究领域学者的使用。这些经济特色数据库都自成体系，几乎囊括了经济研究该领域可能用到的关键数据，这大大缩短了繁琐的数据寻找、合并过程，真正减少了学者处理数据的时间成本。

CNRDS经济特色数据库具有以下三个特点：

- 数据多数来自官方渠道，权威可靠
- 深度整合源数据，数据结构简约
- 多维度经济数据，满足不同领域研究

宏观经济研究系列

(一) 全球夜间灯光数据库-GNLD

1. 数据库简介

全球夜间灯光数据库 (*Global Night-time Light Database, GNLD*) 是基于美国国家海洋与大气的管理局(NOAA)的 *DMSP/OLS* 影像数据和 *VIIRS/DNB* 影像数据开发而成的, 目前主要反映中国各省、地级市夜间灯光数据和一带一路沿线国家的夜间灯光数据情况。*DMSP/OLS* 影像灯光数据目前已经应用于经济类研究, 并且已有较多文献发表在诸多高水平杂志上, 因此逐渐受到学者关注。*VIIRS/DNB* 影像灯光数据在国内文献中还未被广泛应用, 但基于其具有许多优点因此逐渐受到学者关注。由于学者通常只能获得 NOAA 上面的原始图片资料, 将这些图像资料转化为可用数据需要用到较为复杂的计算机和编程技术, 因此, 为了方便学者使用该数据, 促进中国问题研究, 我司联合相关领域专家开发了全球夜间灯光数据库。具体的开发步骤请见数据库说明书。

2. 数据库特色

- 丰富多样的数据字段

GNLD 既提供了连续时间序列的校正后的 DN 均值和灯光指数, 还提供不同时期不同卫星测定的影像数据, 方便学者进行深度研究。

- 提供地级市灯光数据

GNLD 不仅包括省份灯光数据, 还包括中国所有地级市及县区的灯光数据, 为学者拓宽研究思路提供了数据支持。

- 提供一带一路沿线国家夜间灯光数据

GNLD 是市场上唯一提供一带一路沿线国家的夜间灯光数据, 对该领域研究进行了扩展。

3. 典型数据举例

(1) *DMSP* 中国各地级市灯光数据 (校正后)

省份名称	城市名称	年度	DN 均值
江苏省	常州市	1998	10.3544
江苏省	常州市	1999	10.6286
江苏省	常州市	2000	10.7672

(2) *DMSP* 一带一路沿线国家灯光数据 (校正后)

国家名称	国家简称	年度	DN 均值 (不含油田)	DN 均值 (含油田)
阿尔巴尼亚	ALB	2013	4.6405	4.6405
阿富汗	AFG	2013	0.2613	0.2613
阿联酋	ARE	2013	19.7063	20.7325
阿曼	OMN	2013	3.3275	4.1676
阿塞拜疆	AZE	2013	4.3688	4.426
埃及	EGY	2013	3.9639	4.0588

4.数据库内容

数据模块	数据表	数据字段
DMSP/OLS 全球灯光数据	DMSP 中国各省份灯光数据（校正后）	省份名称、年度、DN 均值
	DMSP 中国各省份灯光数据（原始）	省份名称、年度、所属卫星、DN 均值
	DMSP 中国各地级市灯光数据（校正后）	省份名称、城市名称、年度、DN 均值
	DMSP 中国各地级市灯光数据（原始）	省份名称、城市名称、年度、所属卫星、DN 均值
	DMSP 中国各县区灯光数据（校正后）	省份名称、城市名称、县区名称、年度、DN 均值
	DMSP 中国各县区灯光数据（原始）	省份名称、城市名称、县区名称、年度、所属卫星、DN 均值
VIIRS/DNB 全球灯光数据	VIIRS 中国各省份灯光数据	省份名称、年度、月份、DN 均值
	VIIRS 中国各地级市灯光数据	省份名称、城市名称、年度、月份、DN 均值
	VIIRS 中国各县区灯光数据	省份名称、城市名称、县区名称、年度、月份、DN 均值
一带一路沿线国家夜间灯光数据	DMSP 一带一路沿线国家灯光数据（校正后）	国家名称、国家简称、年度、DN 均值（不含油田）、DN 均值（含油田）
	DMSP 一带一路沿线国家灯光数据（原始）	国家名称、国家简称、年度、所属卫星、DN 均值（不含油田）、DN 均值（含油田）

5.数据范围

数据模块	数据起始时间
DMSP/OLS 全球灯光数据	1992 年
VIIRS/DNB 全球灯光数据	2013 年
DMSP 一带一路沿线国家灯光数据（原始）	1992 年

(二) 中国高铁航线数据库-CRAD

1.数据库简介

中国高铁航线数据库 (*Chinese High-speed Rail and Airline Database, CRAD*) 主要包含国内各省市所开通的高铁线路、高铁火车站点、飞机航线和航站的情况。其中, 高铁列车信息包括车站所在省市、列车车次、出发站、到达站、运行时间、里程等信息; 高铁站开通时间数据主要来自国家铁路总公司公布的自 2003 年以来高铁线路建设和开通情况, 该表提供各高铁站名称、开通时间、所属地区和所属高铁线路等信息; 高铁线路信息包括近百条已开通高铁的开通时间、站点和运营时速等信息; 飞机航线信息包括机场所在城市、机场名称、航班信息等。

2.数据库特色

CRAD 详细记录了中国各省、市、县所开通的列车站和飞机场的情况, 尤其是, 是否开通高铁站信息以及高铁站开通时间, 可以在一定程度上反映一个城市对外沟通和交流的便利程度, 可为学者研究提供更多的数据支持。

3.典型数据举例

高铁列车信息

列车车次	出发站	出发站类型	到达站	车型	运行时间	里程
G9464	南京南	过路车	徐州东	高速铁路	01 小时 38 分	331
D5641	南京南	始发车	芜湖	动车组	00 小时 43 分	20
G1113	南京南	始发车	广州南	高速动车	07 小时 35 分	1583
G7041	南京南	始发车	上海	高速动车	02 小时 09 分	311

4.数据库内容

数据模块	数据表	数据字段
高铁列车信息	高铁列车信息	列车车次、出发站、出发站所属地级市、出发站所属省份、出发站类型、到达站、到达站所属地级市、到达站所属省份、到达站类型、车型、开车时间、到站时间、运行时间、里程
	高铁站开通时间	高铁站名称、开通时间、所在省份、所在城市、所属线路名称、备注
	高铁线路信息	高铁线路名称、起点名、终点名、开通时间、线路长度(km)、设计速度(km/h)、沿途主要车站
飞机航线信息	飞机航线信息	起点城市、起点城市所属地级市、起点城市所属省份、起点机场、终点城市、终点城市所属地级市、终点城市所属省份、终点机场、航空公司、更新日期、航班、出发时间、到达时间、准点率、班次_周一、班次_周二、班次_周三、班次_周四、班次_周五、班次_周六、班次_周日

5.数据范围

数据模块	数据起始时间
全部模块	截面数据

(三) 中国环境数据库-CEDS

1.数据库简介

中国环境数据库 (*Chinese Environment Database, CEDS*) 对中国环境状况进行全方面的考察与总结, 共包含气象、环境、自然灾害、城市地理信息、自然生态、环境治理、环境影响评价、环评工作人员全国各省、地级市 1998 年以来各年年均 $PM_{2.5}$ 等 9 个模块。(关于 1998 年以来各年年均 $PM_{2.5}$ 模块详细数据处理流程可见说明书附录)

2.数据库特点

- 独特性

CEDS 对中国环境各方面的数据都进行了梳理总结, 不仅包含气象、环境和自然灾害数据, 还汇总了各保护区、自然生态、环评机构与环评工作人员的信息。

- 多维性

CEDS 不仅包括基础的大气环境与自然灾害方面的信息, 还包含有城市地理位置、环评机构与人员、国家污染排放重点监控企业名单等环境相关的内容。

- 完整性

CEDS 包括大气与灾害、自然生态、环境影响评价等信息, 数据库内容丰富。

3.典型数据举例

空气质量信息

城市名称	日期	质量等级	AQI 指数	当天 AQI 排名	二氧化硫	二氧化氮	一氧化碳	PM2.5	PM10
上海	2015/12/01	良	90	200	16	75	1.03	66	52
上海	2015/12/02	良	97	259	20	54	0.98	71	75
上海	2015/12/03	良	97	332	29	54	0.92	54	142

4.数据库内容

数据模块	数据表	数据字段
气象数据	天气	城市编号、所属省份、所属地级市、城市名称、日期、星期时间、最高气温、最低气温、天气情况、风向、风力
环境数据	空气质量	城市编号、城市名称、所属省份、日期、AQI 指数、质量等级、当天 AQI 排名、 $PM_{2.5}$ 24 小时均值、 PM_{10} 24 小时均值、二氧化硫 24 小时均值、二氧化氮 24 小时均值、一氧化碳 24 小时均值、臭氧指标 24 小时均值
	污染物排放	省级名称、年度、指标种类、排放量指标数值、排放量指标单位、子指标、具体污染指标
	水体环境	水系名称、点位名称、断面情况、溶解氧 (mg/l)、氨氮含量 (mg/l)、高锰酸钾化学需氧量 (mg/l)、PH 值、河流或湖泊名称、上周水质、本周水质、主要污染物指标、报告年份、报告周数
自然灾害	台风灾害	台风编号、台风中文名称、台风英文名称、年度、台风数据来源、台风中心风速 (m/s)、时间、台风中心风力、台风中心气压、7

数据模块	数据表	数据字段
		级风力半径 (km)、类型、10 级风力半径 (km)、台风中心经度、台风中心纬度
	地震灾害	发震日期、发震时间、震源纬度、震源经度、震源深度、震级、震源位置
城市地理信息	城市经纬度	城市编码、城市级别、省级名称、地市级名称、县级名称、区号、邮政编码、经度、纬度
	城市海拔	城市编号、城市级别、省级名称、地市级名称、县级名称、区号、邮政编码、海拔
自然生态	全国自然保护区名录	保护区编号、保护区名称、所属省份、所属地级市、所属区县、总面积、主要保护对象、类型、级别、始建时间、主管部门
	国家级生态乡镇名单	年度、省级名称、地市级名称、县级名称、乡镇名称
自然生态	国家生态工业示范园区	示范园区编号、示范园区名称、类别、批准文号、批准时间
	环境标志获证企业	企业名称、企业编号、所属省份、产品种类、批准时间、有效期限、证书编号、证书状态
环境治理	污染物排放国家重点监控企业名单	行政区划代码、法人代码、公司名称、监控类别、所属省份、所属年份
环境影响评价	环境影响评价机构资质信息	所属省份、所属地级市、机构名称、机构资质等级、机构资质证书号、甲级报告书评价范围、乙级报告书评价范围、报告表评价范围、有效期
	环境影响评价机构基本信息	所属省份、所属地级市、机构名称、成立时间、首次发证时间、经济类型、通讯地址、邮编
	环境影响评价机构诚信信息	所属省份、所属地级市、机构名称、奖惩时间、奖惩部门、文号、奖惩类型、具体情况
环评工作人员	环境影响评价工程师	环评工程师姓名、登记单位、登记证号、登记类别、登记有效期起始日期、登记有效期终止日期、职业资格证书号
	岗位证书持有人员	岗位持证人姓名、岗位证书号、所在单位、所属省份
1998 年以来各年年均 PM2.5	省级 1998 年以来各年年均 PM2.5	年度、省级名称、地区编码、PM2.5 年均浓度 (微克/立方)
	地市级 1998 年以来各年年均 PM2.5	年度、省级名称、地市级名称、地区编码、PM2.5 年均浓度 (微克/立方)

5.数据范围

数据模块	数据起始时间
气象数据	2011 年
环境数据	2009 年
自然灾害	2002 年
城市地理信息	截面数据
自然生态	2007 年
环境治理	截面数据
环境影响评价	2005 年
环评工作人员	2013 年
1998 年以来各年年均 PM2.5	1998 年

（四）中国商品交易市场统计数据库-CEMD

1.数据库简介

中国商品交易市场统计数据库（*Statistical Database of China Commodity Exchange Market, CEMD*）是基于国家统计局数据和中国商品交易市场统计年鉴数据整理开发的专业数据库。*CEMD* 收集整理了中国近十几年各地区交易市场的成交额、摊位数、市场数量等数据内容，为学术界进行相关研究提供了重要数据支持。

2.数据库特色

- 数据来源权威

CEMD 数据来源于国家统计局以及历年商品交易市场统计年鉴，指标覆盖完整，数据权威。

- 结构的一致性

CEMD 在数据整理后的模块结构与表格形式上与近几年的年鉴格式保持一致，便于进行数据比对。

3.典型数据举例

市场综合_历年商品交易市场情况

会计年度	市场数(个)	年末出租摊位数(个)	营业面积(平方米)	成交额(万元)
2014	5023	3534757	295679201	1003099019
2015	4952	3468638	300657328	1001337865
2016	4861	3457899	300233736	1021396989

4.数据库内容

数据模块	数据表
市场综合	历年商品交易市场情况；历年商品交易市场变动；市场综合商品交易市场总体情况；市场综合商品交易市场情况（按营业状态分）……
省、自治区、直辖市	省、自治区、直辖市商品交易市场总体情况；省、自治区、直辖市商品交易市场情况（按市场类别分）；省、自治区、直辖市商品交易市场情况（按营业状态分）；省、自治区、直辖市商品交易市场情况（按经营方式分）；省、自治区、直辖市商品交易市场情况（按经营环境分）；省、自治区、直辖市商品交易市场成交情况（按摊位分）
东中西部及东北地区	东中西部及东北地区商品交易市场总体情况；东中西部及东北地区商品交易市场情况(按市场类别分)；东中西部及东北地区商品交易市场情况（按营业状态分）；东中西部及东北地区商品交易市场情况（按经营方式分）；东中西部及东北地区商品交易市场情况（按经营环境分）……
36 城市	36 城市商品交易市场总体情况；36 城市商品交易市场情况(按市场类别分)；36 城市商品交易市场情况（按营业状态分）；36 城市商品交易市场情况（按经营方式分）……
三大地带	三大地带商品交易市场总体情况；三大地带商品交易市场情况(按市场类别分)；三大地带商品交易市场情况（按营业状态分）……

5.数据范围

数据模块	数据起始时间
全部模块	2000 年

（五）中国人口和就业统计数据库-PESD

1.数据库简介

中国人口和就业统计数据库（*Chinese Population and Employment Statistics Database, PESD*）收集了全国和各省、自治区、直辖市人口就业统计的主要数据，同时收录了世界部分国家和地区的相关数据，全面反映了我国人口和就业状况。**PESD**数据主要来源于国家统计局发布的《中国人口和就业统计年鉴》，根据年鉴中的数据内容，将其归类为六大模块：综合数据、劳动力抽样调查、城镇单位就业人员、全国户籍统计人口、计划生育统计人口、世界人口与就业。

2.数据库特色

- 准确的数据归类

PESD根据《中国人口和就业统计年鉴》，将其归类六大模块，直观反映我国人口和就业状况。

- 简洁的数据结构

PESD整合了多个年度数到一张表格中，数据结构简洁，下载方便，节约了数据搜寻和整理时间。

3.典型数据举例

综合数据_人口_按性别分人口数

会计年度	总人口 (年末)	男性人口数	男性人口比重	女性人口数	女性人口 比重
2005	130756	67375	51.53	63381	48.47
2006	131448	67728	51.52	63720	48.48

4.数据库内容

数据模块	数据表
综合数据	按性别分人口数；人口年龄结构和抚养比；各地区人口出生率、死亡率和自然增长率；流动人口数；各地区人口平均预期寿命；全国历年人口密度……
劳动力抽样调查	全国按就业身份、性别分的就业人员年龄构成；城镇按就业身份、性别分的就业人员年龄构成；城镇按行业、性别分的就业人员年龄构成……
城镇单位就业人员	各地区分行业国有单位就业人员数；各地区分行业城镇集体单位就业人员数；各地区分行业其他单位就业人员数；各地区国有单位就业人员数；各地区城镇集体单位就业人员数；各地区其他行业就业人员数
全国户籍统计人口	各地区总户数、总人口；各地区市总户数、总人口；各地区镇总户数、总人口；各地区县总户数、总人口；各地区非农业、农业人口；各地区市非农业、农业人口；各地区镇非农业、农业人口……
计划生育统计人口	各地区分孩次出生政策符合率与上年同期比较；各地区已婚育龄妇女领证情况、避孕率与上年同期比较；各地区采用各种节育措施人数……
世界人口与就业	各国人口出生率、死亡率、自然增长率；各国人口年龄构成；各国人口指标；人口主要指标(粗)；人口主要指标……

5.数据范围

数据模块	数据起始时间
全部模块	2006年

（六）脱贫数据库-LPOP

1.数据库简介

脱贫数据库 (*Database of Lift People Out of Poverty, LPOP*) 是基于中国农村贫困监测报告以及我国贫困县戴帽、摘帽及其社会经济发展等情况而建立的专业数据库。党的十八届五中全会提出全面建成小康社会新的目标要求,其中之一便是要在我国现行标准下,贫困县全部摘帽,解决区域性整体贫困。以此为契机, *LPOP* 深入考察了改革开放以来,中国农村各地区逐步脱贫的情况以及 800 多个贫困县的基本情况、交通发展情况、经济产业状况、社会保障状况等多方面内容,多维度展示了近年来我国脱贫攻坚取得的成果,为该领域的学术研究提供数据支持。

2.数据库特色

- 全方位展示脱贫情况

LPOP 涵盖了分地区脱贫情况和贫困县具体脱贫情况,对与脱贫相关的多个方面数据进行了展示。

- 数据来源权威、完整

LPOP 根据国务院扶贫开发领导小组办公室公布的贫困县名单确定贫困县,并根据这些贫困县列表整理了来自国家统计局的各方面数据。

3.典型数据举例

贫困县脱贫情况_贫困县名单与摘帽情况

省份	所属地级市	贫困县	地区编码	是否为重点贫困县	贫困县是否摘帽
云南	昆明市	东川区	530113	1	0
云南	昆明市	禄劝彝族苗族自治县	530128	1	0
云南	昆明市	寻甸回族彝族自治县	530129	1	0
云南	曲靖市	师宗县	530323	0	0
云南	曲靖市	罗平县	530324	0	1

4.数据库内容

数据模块	数据表
农村脱贫情况	历年全国农村贫困状况、全国分地区农村贫困人口规模、全国分地区农村贫困发生率、全国分地区农村常住居民人均可支配收入、全国分地区农村常住居民人均消费支出、全国农村常住居民收入消费增长情况、全国农村常住居民人均可支配收入及构成、全国农村常住居民消费支出及构成、全国农村常住居民年末每百户主要耐用消费品拥有量、贫困地区经济社会发展情况、贫困地区居民生产生活情况、贫困地区文化教育、医疗保健情况、贫困地区农村贫困人口变化情况、贫困地区农村常住居民收入情况……
贫困县脱贫情况	贫困县名单与摘帽情况、贫困县基本情况、贫困县综合经济、贫困县农业、工业及投资、贫困县教育、卫生和社会保障、贫困县是否开通高铁、贫困县上市公司基本信息

5.数据范围

数据模块	数据起始时间
全部模块	2001 年

(七) 中国固定资产投资统计数据库-CIFA

1.数据库简介

中国固定资产投资统计数据库 (*Statistical Database of Chinese Investment in Fixed Assets, CIFA*) 是为全面反映中国固定资产投资情况而进行整理的专业数据库, 其内容涵盖了全国、各省、自治区、直辖市、国民经济各行业等地区 and 行业有关固定资产投资的统计数据。*CIFA* 整体分为三个部分: 全社会固定资产投资、固定资产投资(不含农户)、农户固定资产投资, 其中固定资产投资(不含农户)又分为两个模块, 分别是: 固定资产投资(不含农户)和房地产开发。

2.数据库特色

- 覆盖重点统计指标

CIFA 囊括了按地区、国民经济各行业、隶属关系、经济类型、建设定制等分组的本年完成投资, 总投资规模、新增固定资产、投资到位资金、项目个数及有关建筑面积等重点指标数据。

3.典型数据举例

全社会固定资产投资主要指标

指标名称	年份	国有经济	集体经济	私营个体经济	联营经济
房屋建筑面积(万平方米)-竣工面积	2012	42298.72	14993.15	175360.50	485.75
投资总额(亿元)-按产业分-第三产业	2012	77847.33	10286.52	44112.66	309.94
投资总额(亿元)-按产业分-第二产业	2012	24442.41	2822.87	53772.21	227.75
投资总额(亿元)-按产业分-第一产业	2012	2485.79	765.41	5126.15	44.52

4.数据库内容

数据模块	数据表
全社会固定资产投资	全社会固定资产投资主要指标及增长速度; 全社会固定资产投资主要指标; 按结构分全社会固定资产投资情况; 各地区全社会固定资产投资建设规模; 各地区全社会按经济类型分固定资产投资建设规模; ……
固定资产投资(不含农户)	固定资产投资(不含农户)主要指标; 各地区固定资产投资(不含农户)建设规模; 国民经济行业大类固定资产投资(不含农户)建设规模; 各地区固定资产投资(不含农户)和新增固定资产……
房地产开发	房地产开发投资主要指标; 各地区按登记注册类型分的房地产开发单位个数; 各地区房地产开发固定资产投资建设规模; 各地区房地产开发投资和新增固定资产; 各地区按登记注册类型分的房地产开发投资……
农户固定资产投资	各地区农村农户固定资产投资增长情况; 各地区按构成分农村农户投资; 各地区按项目分农村农户投资; 各地区按主要行业分农村农户投资; 各地区农村农户投资实际到位资金; 各地区农村农户房屋建筑面积和投资; 各地区农村农户住宅建筑面积和投资

5.数据范围

数据模块	数据起始时间
全部模块	1996年

（八）最低工资标准数据库-WAGE

1.数据库简介

最低工资标准数据库 (*Minimum Wage Standard Database, WAGE*) 是基于中国各省市人力资源和社会保障厅各年度所公布的最低工资标准文件整理而成的, 反映我国各省市最低工资标准情况的专业数据库。*WAGE* 从各个渠道搜集了 2001 年以来, 各省市县的最低工资标准数据, 需要说明的是, 由于各省份并非在每个年度都会对最低工资标准进行调整, 因此, 为了提供完整的时间序列数据, 在未公布最低工资调整文件的年份, *WAGE* 使用上一次调整的最低工资标准作为当年度值; 此外, 由于某些省份可能在省级层面发布了最低工资标准文件, 可是文件中并没有对这些最低工资标准适用的城市进行列示, 需要从城市层面发布的最低工资标准文件进行数据搜集, 因此, 存在部分省份的某些城市无法搜集的情况。

2.数据库特色

- 数据来源权威

WAGE 的数据来源于各省市人力资源和社会保障厅官方网站所披露数据, 数据来源权威。

- 多年度数据

WAGE 包含了 2001 年以来各省市县的最低工资标准数据, 提供了丰富的时间序列数据。

3.典型数据举例

最低工资标准

会计年度	省份名称	地级市名称	区县名称	调整时间	最低月工资 (元)	最低小时工资 (元)
2019	北京市	市辖区	东城区	2019/7/1	2200	24
2019	北京市	市辖区	西城区	2019/7/1	2200	24
2019	贵州省	贵阳市	开阳县	2019/12/1	1670	17.5
2019	贵州省	贵阳市	息烽县	2019/12/1	1570	16.5
2019	贵州省	六盘水市	钟山区	2019/12/1	1790	18.6

4.数据库内容

数据模块	数据字段
最低工资标准	省份名称、城市名称、区县名称、统计年度、调整时间、最低月工资、最低小时工资

5.数据范围

数据模块	数据起始时间
全部模块	2001 年

（九）中国住户调查数据库-HSLD

1.数据库简介

中国住户调查数据库 (*Database of Chinese Household Survey, HSLD*) 是基于国家统计局展开的住户调查资料而建立的, 反映我国城乡居民收支、生产和生活状况的专业数据库。*HSLD* 主要收录了历年全国及分城乡居民收支和生活状况数据, 以及分省、分地区及按收入等份分组和按 4 个经济区域分组的住户调查基本情况及收支数据, 还包括农村住户固定资产投资情况, 以及住户调查其他数据。

2.数据库特色

- 两种统计口径

HSLD 根据 2012 国家统计局实施的住户调查一体化改革, 将数据分成 2 种口径进行收录。

3.典型数据举例

分地区居民可支配收入

分类	地区	单位	年份	可支配收入	工资性收入	经营净收入	财产净收入
全体居民	河北	元/人	2013	15189.6	8858.9	2462.3	1048.3
全体居民	山西	元/人	2013	15119.7	9424	2338.3	739
全体居民	内蒙古	元/人	2013	18692.9	10004.9	4729.8	1075.2

4.数据库内容

数据模块	数据表
全国及分城乡居民收支与生活状况 (新口径)	全国居民家庭基本情况; 全国居民可支配收入及构成; 全国居民现金可支配收入及构成; 全国居民按收入五等份分组的人均可支配收入……
分地区居民收支与生活状况 (新口径)	分地区居民可支配收入; 分地区居民可支配收入构成; 分地区居民现金可支配收入; 分地区居民现金可支配收入构成; 分地区居民消费支出; 分地区居民消费支出构成; 分地区居民现金消费支出……
历年城镇和农村住户调查主要数据 (老口径)	城镇居民现金可支配收入和指数; 城镇居民现金收入来源及构成; 城镇居民现金消费支出及增长率; 城镇居民现金消费支出及构成; ……
分地区城镇和农村住户调查收支数据 (老口径)	分地区全体居民可支配收入及构成; 分地区全体居民现金可支配收入及构成; 分地区全体居民消费支出及构成; 分地区全体居民现金消费支出及构成; 分地区城镇居民现金收入来源……
全国及分地区农村住户固定资产投资情况	全国农村住户固定资产投资情况; 分地区农村住户固定资产投资结构情况; 分地区农村住户固定资产投资完成额; 分地区农村住户固定资产投资投向情况……
各地区城镇、农村居民家庭情况 (老口径)	各地区城镇居民家庭人口就业情况; 各地区城镇居民家庭全年收入来源; 各地区城镇居民家庭总支出及构成; 各地区城镇居民家庭现金消费支出及构成……
住户调查其他数据	全国居民人均可支配收入基尼系数; 农村贫困状况; 农民工数量及收入水平

5.数据范围

数据模块	数据起始时间
全部模块	1978 年

(十) 碳中和研究数据库-CNLD

1.数据库简介

碳中和研究数据库 (Carbon Neutral Research Database, **CNLD**) 是根据近年来中国提出为全球气候治理自主贡献新目标和碳中和愿景, 坚定贯彻新发展理念、建设清洁美丽世界的决心而开发的专业数据库。**CNLD** 主要从碳排放权交易、各地区排污费、各地区环境概况、环境保护与评价、环境监督与法制、污染物排放、污染治理投资、工业污染治理、生活污染治理和环保机构情况等十多个方面展示了历史以来我国为生态环保和实现碳中和所付出的努力和取得的成效。

2.数据库特色

- 数据来源权威

CNLD 的数据来源比较多样, 碳排放权交易数据来自于中国八大碳排放权交易所公布的数据, 各地区排污费和其他数据均来自中国环境年鉴和生态环境部。

- 数据内容丰富多样

CNLD 从碳排放权交易、排污费、环保情况、环保治理、环保投资等多个维度展示了我国的碳中和实际情况。

3.典型数据举例

各地区排污费

地区	年份	排污费收入总额(万元)	缴纳排污费单位(个)
安徽	2014	57449.78	9076
北京	2014	27844.01	3913
福建	2014	33135.68	14408
甘肃	2014	20273.43	3840
广东	2014	87763.22	48475

4.数据库内容

数据模块	数据表
碳排放权交易	碳排放权交易
各地区排污费	各地区排污费
环境基本情况	各地区城市环境情况、各地区生态环境基本情况、重点城市空气质量情况、各地区突发环境事件
环境保护与评价	各地区自然生态保护与建设情况、各地区环境污染控制与管理情况、各地区环境环境污染源控制与管理情况、各地区污染源自动监控情况、各地区建设项目“三同时”执行情况、各地区环境影响评价情况、各地区环境宣教情况、各地区建设项目竣工环境保护验收情况、各地区环境保护档案工作情况
环境监督与法制	各地区环境执法情况、各地区环境法制工作情况、各地区环境信访工作情况、各地区建设项目环境影响评价执行情况、各地区环境信访与环境法制情况、各地区人大建议、政协提案情况

数据模块	数据表
环境治理投资	各地区工业污染防治投资情况、各地区环境污染治理投资情况、重点城市工业污染防治投资情况、湖泊水库工业污染防治投资情况、各地区工业污染治理投资完成情况、各地区城市环境基础设施建设投资情况、各地区工业污染治理项目建设情况、重点城市工业污染治理项目建设情况、各地区环境保护能力建设投资情况、各地区环境科技工作情况
污染物排放	各地区主要污染物排放情况、各地区废气、废水中主要污染物排放情况、各地区农业污染排放情况、各地区机动车污染排放情况、重点城市生活及其他污染情况、城市主要工业污染物单位工业总产值排放强度、重点城市农业污染排放情况、重点城市机动车污染排放情况
工业污染治理	各地区工业废气排放及处理情况、重点城市工业废气排放及处理情况、各地区工业废水排放及处理情况、重点城市工业废水排放及处理情况、各地区固体废物处理利用情况、重点城市工业固体废物产生及处置利用情况、各地区工业企业“三废”治理效率、重点城市工业企业“三废”治理效率、各地区工业危险废物产生及处置利用情况、重点城市工业危险废物产生及处置利用情况
生活污染治理	各地区城镇生活污染排放及处理情况、重点城市城镇生活污染排放及处理情况、各地区生活垃圾处理情况、重点城市生活垃圾处理情况、各地区城镇污水处理情况、重点城市污水处理情况、各地区危险(医疗)废物集中处置情况、各地区城市生活垃圾清运和处理情况、重点城市危险(医疗)废物集中处置情况
环保机构情况	各地区环保机构情况、各地区环境保护系统各级机构人员数、各地区环境保护系统年末机构总数、各地区环境保护系统年末实有人数

5.数据范围

数据模块	数据起始时间
全部模块	1990年

（十一）影子银行研究数据库-SBRD

1.数据库简介

影子银行研究数据库 (*Shadow Banking Research Database, SBRD*) 是基于对影子银行行为深入考察而建立的专业数据库,旨在为学术研究者提供权威的 SBR 数据。通过考察相关权威文献, *SBRD* 从微观的商业银行层面和上市公司层面,到宏观的社会融资层面,对不同领域的影子银行数据进行了全面展示。其中,商业银行层面的影子银行数据主要来自银行年报;上市公司层面的影子银行数据来自上市公司年报及公告;社会融资层面的影子银行数据则根据中国人民银行披露的相关数据进行了总结和归纳。

2.数据库特色

- 全面且多样的数据内容

SBRD 包括了商业银行、上市公司和社会融资三个维度的统计数据,并进行多方位信息来源的综合,且详细列示了与研究相关的大量字段。

3.典型数据举例

上市公司层面的测量_信用中介类影子银行规模

股票代码	统计日期	委托贷款(元)	委托理财(元)	其他应收款(元)
000016	2020-12-31	20000000	639160000	2095679366.13
000055	2020-12-31	20000000	790297400	162145236.85
000609	2020-12-31	21000000	40000000	183801460.59
000672	2020-12-31	500000000	995150000	26304125.80
000811	2020-12-31	135000000	1420630000	82380513.18

4.数据库内容

数据模块	数据表
商业银行层面的测量	商业银行层面的测量
上市公司层面的测量	信用中介类影子银行规模、信用链条类影子银行规模
社会融资层面的测量	全国社会融资规模存量统计、全国社会融资规模增量统计、地区社会融资规模增量统计

5.数据范围

数据模块	数据起始时间
商业银行层面的测量	1998 年
上市公司层面的测量	2012 年
社会融资层面的测量	2012 年

（十二）绿色金融研究数据库-GFRD

1.数据库简介

绿色金融研究数据库（*Green Finance Research Database*, **GFRD**）基于我国坚持绿色发展理念，推进生态文明建设，考察金融领域的绿色发展情况而构建的专业数据库。**GFRD** 囊括四大模块：绿色保险、绿色财政、绿色信贷和绿色债券。其中，绿色保险主要展示了各省、地级市和保险公司的农业保险情况；绿色财政则从地方财政环保支出、环境治理投资以及国家节能减排财政政策展示了财政领域对于绿色发展的支持；绿色信贷包括主要上市银行、银保监会和人民银行对于绿色信贷的披露和统计情况；最后，我们根据债券所投资行业，总结了当前中国绿色债券的发行情况。

2.数据库特色

- 全面的数据内容

GFRD 包括了绿色保险、绿色财政、绿色信贷和绿色债券四个方面的数据，并进行多方位信息来源的综合，且详细列示了与研究相关的大量字段。

- 来源权威

GFRD 数据来源于中国人民银行、国家统计局和上市公司年报等权威来源。

3.典型数据举例

绿色信贷_各银行绿色信贷

银行名称	统计日期	绿色贷款余额（亿元）
中国银行股份有限公司	2012-12-31	2274.80
中国银行股份有限公司	2013-12-31	2587.59
中国银行股份有限公司	2014-12-31	3010.43
中国银行股份有限公司	2015-12-31	4123.15

4 数据库内容

数据模块	数据表
绿色保险	各省农业保险、各地级市农业保险、主要保险公司农业保险
绿色财政	环境保护支出、环境治理投资完成额、国家节能减排财政政策综合示范城市名单
绿色信贷	各银行绿色信贷、两高一剩行业统计、银行信贷投向（银保监会口径）、银行信贷投向（人民银行口径）
绿色债券	绿色债券

5.数据范围

数据模块	数据起始时间
绿色保险	2005 年
绿色财政	2004 年
绿色信贷	2007 年
绿色债券	2010 年

(十三) 国家统计局标准数据库-CNSD

1. 数据库简介

国家统计局标准数据库 (*Chinese National Statistical Standard Database, CNSD*) 是根据国家统计局所公布的统计标准整理而成的专业数据库。*CNSD* 囊括了自 2009 年起的统计用区划和城乡划分、共计 4 个版本的国民经济行业分类以及统计用产品分类。可为相关部门及各界人员提供研究使用的完善、权威的数据支持。

2. 数据库特色

- 数据来源权威

CNSD 数据来源国家统计局, 直观体现各年度、各版本数据变化。

- 全面的字段内容

CNSD 遵照国家统计局的原始字段进行了细致整理, 在保留字段的基础上提供更详细的字段释义。

3. 典型数据举例

国民经济行业分类和代码

年份	代码门类	代码大类	代码中类	代码小类	类别名称
1984	1	05	051	0511	海水养殖业
1984	1	05	051	0512	海洋捕捞业
1984	1	05	052		淡水渔业
1984	1	05	052	0512	淡水养殖业

4 数据库内容

数据模块	数据表	数据字段
国民经济行业分类	国民经济行业分类和代码	年份、代码门类、代码大类、代码中类、代码小类、类别名称、说明
	国民经济行业分类新旧类目对照	对比年份、行业类别代码(新)、行业类别名称(新)、行业类别代码(旧)、行业类别名称(旧)、变动说明
	国民经济行业分类与国际标准行业分类对照	年份、国民经济行业分类代码、国民经济行业分类名称、国际标准产业分类代码、国际标准产业分类名称
统计用区划和城乡划分	统计用区划和城乡划分	年份、省份、地级市、地级市代码、区县、区县代码
统计用产品分类	统计用产品分类	主类别代码、主类别名称、一级分类代码、一级分类名称、二级分类代码、二级分类名称、三级分类代码、三级分类名称、四级分类代码、四级分类名称、说明

5. 数据范围

数据模块	数据起始时间
国民经济行业分类	截面数据
统计用区划和城乡划分	2009 年
统计用产品分类	截面数据

（十四）环境处罚研究数据库-EPRD

1.数据库简介

环境处罚研究数据库 (*Environmental Penalty Research Database, EPRD*) 是基于全国各个城市生态环境局公布的环境处罚公告整理而成的专业数据库。近年来, 保护生态环境, 推动绿色发展成为我国社会发展的重要方向, *EPRD* 手工收集了来自 300 多个地级市的环境处罚公告数据, 对原始公告数据进行了清洗与分析, 提供诸如处罚原因、处罚内容、处罚金额等重要字段。希望 *EPRD* 能为该领域研究提供多维度的学术数据支持。

2.数据库特色

- 提供丰富的环境处罚数据

EPRD 不仅提供全国多个城市的环境处罚数据, 还提供了多个年度的时间序列数据, 为学术研究提供重要数据支持。

3.典型数据举例

环境处罚

省份名称	城市名称	披露日期	处罚对象	处罚原因	罚没金额
浙江省	丽水市	2022/1/11	浙江耀水家具销售有限公司	未建立危险废物管理台账	10
浙江省	丽水市	2022/1/24	浙江陕鼓能源开发有限公司	超标排放废气	15
浙江省	丽水市	2022/1/24	浙江陕鼓能源开发有限公司	未按规定贮存危险废物	13.9
浙江省	丽水市	2022/1/25	丽水市丽农生态农牧有限公司	超标污水通过雨水沟外排	

4 数据库内容

数据模块	数据表	数据字段
环境处罚	环境处罚	省份名称、城市名称、披露日期、处罚对象、处罚原因、处罚内容、罚没金额、处理机构

5.数据范围

数据模块	数据起始时间
环境处罚	2008 年

(十五) 共同富裕研究数据库-CPRD

1.数据库简介

共同富裕是社会主义的本质要求,是中国式现代化的重要特征。基于党和国家提出的“正确认识和把握实现共同富裕的战略目标和实践途径”的重大理论和实践问题,共同富裕研究数据库(*Common Prosperity Research Database, CPRD*)参考近期前沿性研究成果,从发展性、共享性和可持续性三个大模块总共 11 个维度构建了中国各省市共同富裕发展情况的相关研究数据库。其中发展性包括富裕度、共同度和基尼系数三个子模块;共享性包括文化教育、医疗健康、社会保障、公共设施、信息化五个子模块;可持续性则包含科技创新、社会治理和生态环境三个子模块。希望 *CPRD* 的推出能为该领域研究提供重要数据支持。

2.数据库特色

- 多维度展示共同富裕的各个方面

CPRD 从发展性、共享性和可持续三个方面,科学构建了研究共同富裕的多维度指标体系。

3.典型数据举例

共享性_信息化各项指标

地区	年份	电话普及率(包括移动电话)	移动电话普及率	互联网普及率(%)
安徽省	2010	65.73	45.65	22.7
安徽省	2011	75.6	54.72	26.6
安徽省	2012	78.8	60.5	31.3
安徽省	2013	81.85	65.65	35.9

4 数据库内容

数据模块	数据表	数据字段
发展性	富裕度各项指标	地区、年份、全体居民人均可支配收入、城镇居民人均可支配收入、农村居民人均可支配收入、人均地区生产总值、人均可支配收入占人均 GDP 比重(%)、劳动报酬占 GDP 比重(%)、全体居民人均消费支出、城镇居民人均消费支出、农村居民人均消费支出、人均社会消费品零售总额、人均居民人民币储蓄存款、恩格尔系数_城镇、恩格尔系数_农村
	共同度各项指标	地区、年份、城乡居民收入倍差、地区人均可支配收入差异系数_城镇、地区人均可支配收入差异系数_农村、地区人均 GDP 差异系数、城镇平均低保标准、农村平均低保标准、城镇化率(%)
	基尼系数计算指标	地区、年份、调查户数_城镇、平均每户家庭人数_城镇、调查总人数_城镇、居民全年人均可支配收入总平均_城镇、最低收入户的居民人均年可支配收入_城镇、困难户的居民人均年可支配收入_城镇、低收入户的居民人均年可支配收入_城镇、中低收入户的居民人均年可支配收入_城镇、中等收入户的居民人均年可支配收入_城镇、中高收入户的居民人均年可支配收入_城镇、高收入户的居民人均年可支配收入_城镇、最高收入户的居民人均年可支配收入_城镇、更高收入户的居民人均年可支配收入_城镇、调查户数_农村、平均每户家庭人数_农村、调

		查总人数_农村、居民全年人均可支配收入总平均_农村、低收入户的居民人均年可支配收入_农村、中低收入户的居民人均年可支配收入_农村、中等收入户的居民人均年可支配收入_农村、中高收入户的居民人均年可支配收入_农村、高收入户的居民人均年可支配收入_农村
共享性	文化教育各项指标	地区、年份、人均教育经费、小学生师比、初中生师比、普通高中生师比、人均拥有公共图书馆藏书/册/人、每万人拥有公共图书馆建筑面积/平方米
	医疗健康各项指标	地区、年份、每万人医疗机构床位数、城市每万人医疗机构床位数、农村每万人医疗机构床位数、每万人卫生人员数、每万人卫生技术人员数、每万人执业(助理)医师数、每万人执业医师数、每万人注册护士数、每万人药师数、每万人乡村医生和卫生员数
	社会保障各项指标	地区、年份、城乡居民社会养老保险参保人数、城乡居民社会养老保险实际领取待遇人数、社会保障支出占 GDP 之比(%)
	公共设施各项指标	地区、年份、城市污水处理率(%)、公园数、建成区绿化覆盖率(%)、城市用水普及率(%)、城市燃气普及率(%)、每万人拥有公共交通工具、人均城市道路面积、人均公园绿地面积、每万人拥有公共厕所数
	信息化各项指标	地区、年份、电话普及率(包括移动电话)、移动电话普及率、互联网普及率(%)
可持续性	科技创新各项指标	地区、年份、国内专利申请受理量、国内专利申请授权量、各地区研究与试验发展(R&D)经费投入强度(%)、R&D 经费内部支出占国内生产总值的比重(%)、城镇登记失业率(%)、人均 GDP 增长率(%)
	社会治理各项指标	地区、年份、每万人拥有登记社会组织数、慈善捐赠占 GDP 比重(%)
	生态环境各项指标	地区、年份、单位地区生产总值能耗(等价值)/吨标准煤、单位地区生产总值能耗(等价值)_同比增长(%)、单位工业增加值能耗(规模以上当量值)/吨标准煤、单位工业增加值能耗(规模以上当量值)_同比增长(%)、单位地区生产总值电耗(等价值)/千瓦小时、单位地区生产总值电耗(等价值)_同比增长(%)、林业用地面积、森林面积、人工林面积、森林覆盖率(%)、活立木总蓄积量、森林蓄积量

5.数据范围

数据模块	数据起始时间
富裕度各项指标	2001 年
共同度各项指标	2001 年
基尼系数计算指标	2001 年
文化教育各项指标	2001 年
医疗健康各项指标	2001 年
社会保障各项指标	2001 年
公共设施各项指标	2011 年
信息化各项指标	2010 年
科技创新各项指标	2001 年
社会治理各项指标	2008 年
生态环境各项指标	2004 年

(十六) 中国人口普查分县数据库-CDPC

1. 数据库简介

中国人口普查分县数据库 (*Chinese County-level Data on Population Census, CDPC*) 是基于中国每十年一次的人口普查资料整理而成的专业数据库。*CDPC* 涵盖了全国各县级单位的人口分布、结构以及家庭居住状况等多项指标, 是比较系统、详实的县级大型人口数据资料。全国人口普查采用了长、短两种普查表, 本数据库有关婚姻、生育、家庭居住条件等指标取自普查长表。希望 *CDPC* 的推出能为该领域研究提供重要数据支撑。

2. 数据库特色

- 数据来源权威

CDPC 的数据来源于官方发布的人口普查数据, 数据来源丰富、权威。

3. 典型数据举例

人口基本状况(全部数据)

年份	地区	人口数-合计 (人)	人口数-男 (人)	人口数-女 (人)	人口数-性别 比(女=100)	户籍人口 (人)	少数民族人 口比重(%)
2000	密云县	420019	210890	209129	100.84	417605	5.644030389
2000	怀柔县	296002	152357	143645	106.06	266590	7.802649982
2000	平谷县	396701	197646	199055	99.29	388985	0.676075936
2000	大兴县	671444	350691	320753	109.33	533127	3.829060949

4 数据库内容

数据模块	数据表	数据字段
人口基本状况 (全部数据)	人口基本状况 (全部数据)	年份、地区、人口数-合计(人)、人口数-男(人)、人口数-女(人)、人口数-性别比(女=100)、户籍人口(人)、少数民族人口比重(%)、非农业户口人口比重(%)、城镇人口(人)、乡村人口(人)、城市城区人口(人)、家庭户-户数(户)、家庭户-人口数(人)、家庭户-户规模(人/户)、家庭户-其中:一人户(户)、家庭户类别-一代户(户)、家庭户类别-二代户(户)、家庭户类别-三代户(户)、家庭户类别-四代以上户(户)
分年龄、性别 的人口(全部 数据)	分年龄、性别 的人口(全部 数据)	年份、地区、0岁-男(人)、0岁-女(人)、1-4岁-男(人)、1-4岁-女(人)、5-9岁-男(人)、5-9岁-女(人)、10-14岁-男(人)、10-14岁-女(人)、15-19岁-男(人)、15-19岁-女(人)、20-24岁-男(人)、20-24岁-女(人)、25-29岁-男(人)、25-29岁-女(人)、30-34岁-男(人)、30-34岁-女(人)、35-39岁-男(人)……
各年龄组人口 比重、有老年 人口的户数 (全部数据)	各年龄组人口 比重、有老年 人口的户数 (全部数据)	年份、地区、各种年龄组人口占总人口比重(%)、0-14岁、各种年龄组人口占总人口比重(%)、15-64岁、各种年龄组人口占总人口比重(%)、65岁及以上、各种年龄组人口占总人口比重(%)、15-49岁育龄妇女、有65岁及以上老年人口的户数(户)-合计、有65岁及以上老年人口的户数(户)-有一个老人、有65岁及以上老年人口的户数(户)-有二个老人、有65岁及以上老年人口的户数(户)-有三个及以上老人、人口自然变动(%)、出生率、人口自然变动(%)、死亡率、人口自然变动(%)、自然增长率……

人口受教育状况(全部数据)	人口受教育状况(全部数据)	年份、地区、6岁及以上各种受教育程度人口(人)-未上过学-男、6岁及以上各种受教育程度人口(人)-未上过学-女、6岁及以上各种受教育程度人口(人)-扫盲班-男、6岁及以上各种受教育程度人口(人)-扫盲班-女、6岁及以上各种受教育程度人口(人)-学前教育-男、6岁及以上各种受教育程度人口(人)-学前教育-女、6岁及以上各种受教育程度人口(人)-小学-男……
各职业大类人口和未工作人口(长表数据)	各职业大类人口和未工作人口(长表数据)	年份、地区、16岁及以上人口(人)、各种职业人口总计(人)、各种职业人口总计-国家机关、党群组织、企业、事业单位负责人(人)、各种职业人口总计-专业技术人员(人)、各种职业人口总计-办事人员和有关人员(人)、各种职业人口总计-商业、服务业人员(人)、各种职业人口总计-农、林、牧、渔、水利业生产人员(人)……
各行业门类人口及三次产业人口比重(长表数据)	各行业门类人口及三次产业人口比重(长表数据)	年份、地区、15岁及以上人口-总计(人)、各行业人口合计(人)、各行业人口合计-农林牧渔业(人)、各行业人口合计-采掘业(人)、各行业人口合计-制造业(人)、各行业人口合计-电力、煤气及水的生产和供应业(人)、各行业人口合计-建筑业(人)、各行业人口合计-地质勘查业、水利管理业(人)、各行业人口合计-交通运输、仓储及邮电通信业(人)、各行业人口合计-批发和零售贸易……
人口婚姻状况和妇女生育状况(长表数据)	人口婚姻状况和妇女生育状况(长表数据)	年份、地区、15岁及以上人口-合计(人)、15岁及以上人口-男(人)、15岁及以上人口-女(人)、未婚-小计(人)、未婚-男(人)、未婚-女(人)、有配偶-小计(人)、有配偶-男(人)、有配偶-女(人)、离婚-小计(人)、离婚-男(人)、离婚-女(人)、丧偶-小计(人)、丧偶-男(人)、丧偶-女(人)、15—50岁妇女-平均活产子女数(人)、15—50岁妇女-平均存活子女数(人)、15—64岁妇女-平均活产子女数(人)、15—64岁妇女-平均存活子女数(人)
家庭户住房状况(全部数据、长表数据)	家庭户住房状况(全部数据、长表数据)	年份、地区、全部数据-家庭户(户)、全部数据-平均每户住房间数(间/户)、全部数据-人均住房建筑面积(平方米/人)、长表数据(户)-家庭户、长表数据(户)-住房内有厨房、长表数据(户)-住房内无厨房、长表数据(户)-住房内有管道自来水、长表数据(户)-住房内无管道自来水、长表数据(户)-住房内有洗澡设施、长表数据(户)-住房内无洗澡设施、长表数据(户)-住房内有厕所、长表数据(户)-住房内无厕所、长表数据(户)-住房所在建筑有电梯、长表数据(户)-拥有家用汽车、长表数据(户)-住房来源-自建、长表数据(户)-住房来源-购买、长表数据(户)-住房来源-租赁、长表数据(户)-住房来源-其他、长表数据(户)-住房来源-继承或赠予

5.数据范围

数据模块	数据起始时间
全部模块	2000年

(十七) 数字经济政策数据库-DEPD

1. 数据库简介

数字经济政策数据库 (*Digital Economy Policy Database, DEPD*) 是基于全国各个省份及地级市人民政府官网公布的政府工作报告整理而成的专业数据库。结合已有研究文献成果, 我们建立了数字经济政策相关的关键字列表, 使用该列表分析了政府工作报告文字表述中对数字经济的关注程度。*DEPD* 不仅列示了数字经济关注程度的统计词频结果, 还展示了每个关键词的出现频率, 希望 *DEPD* 的推出能为该领域研究提供相应的数据支持。

2. 数据库特色

- 数据来源权威

DEPD 的政府工作报告文本信息来自各省份及地级市人民政府官网, 关键词列表参照权威期刊文献, 数据来源丰富、权威。

3. 典型数据举例

省级数字经济政策统计表

省名称	年份	总字数 1	总字数 2	总词数 1	总词数 2	句子数	数字经济政策词频汇总
安徽	2011	24128	20907	12559	8490	498	27
安徽	2012	15692	13518	8160	5550	378	13
安徽	2013	14256	12076	7469	4958	387	24
安徽	2014	14761	12732	7669	5274	321	25

4 数据库内容

数据模块	数据表	数据字段
省级数字经济政策统计表	省级数字经济政策统计表	省名称、年份、总字数 1、总字数 2、总词数 1、总词数 2、句子数、数字经济政策词频汇总
省级数字经济政策词频表	省级数字经济政策词频表	省名称、年份、词汇名称、词频统计、句子示例、出现该关键字的句子数
地级市数字经济政策统计表	地级市数字经济政策统计表	省名称、地市级名称、年份、总字数 1、总字数 2、总词数 1、总词数 2、句子数、数字经济政策词频汇总
地市级数字经济政策词频表	地市级数字经济政策词频表	省名称、地市级名称、年份、词汇名称、词频统计、句子示例、出现该关键字的句子数

5. 数据范围

数据模块	数据起始时间
省级数字经济政策统计表	1979 年
省级数字经济政策词频表	1979 年
地级市数字经济政策统计表	1986 年
地市级数字经济政策词频表	1986 年

（十八）环境注意力数据库-GEAD

1.数据库简介

环境注意力数据库（*Government Environmental Attention Database, GEAD*）涵盖了我国 31 个省市和近 300 个地级市的政府工作报告文件基本信息，并根据政府环境注意力定义及相关文献研究，从保护和改善环境、防治污染和其他公害、资源节约、协同发展与环境共治、推进生态文明建设、促进经济社会可持续发展 6 个维度，设计了 140 个环境注意力理论关键词，利用智能文本分析系统，对各省省市及其地级市政府工作报告文件中关键词词频进行分析统计。**GEAD** 数据统计起始于 1979 年，旨在为环境注意力相关研究提供全面、具体的数据支持，极大满足了学术研究的需求。

2.数据库特色

- 内容设计科学

GEAD 参考了已有学术研究，并结合政府环境注意力定义，科学地设计 6 个维度与 140 个理论关键词来测量政府环境注意力。

- 数据结构清晰

GEAD 分别列示各省市和各地级市的政府工作报告基本信息统计、各维度关键词词频统计，数据结构清晰简洁，便于研究者进行数据使用和匹配。

3.典型数据举例

省市-保护和改善环境

省份	年份	词汇名称	词汇类别	词汇组别	词频统计	出现该关键字的句子数
安徽	2012	环境保护	搜索词汇	1	2	2
广东	2012	环境保护	搜索词汇	1	5	5
广西	2012	环境保护	搜索词汇	1	3	3
海南	2012	环境保护	搜索词汇	1	1	1
湖北	2012	环境保护	搜索词汇	1	4	4

4.数据库内容

数据模块	数据表	数据字段
省市环境注意力	省市-基本信息统计	省份、年份、总字数 1、总字数 2、总词数 1、总词数 2、句子数
	省市-保护和改善环境	省份、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、出现该关键字的句子数
	省市-防治污染和其他公害	省份、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、出现该关键字的句子数
	省市-资源节约	省份、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、出现该关键字的句子数
	省市-协同发展与环境共治	省份、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、出现该关键字的句子数
	省市-推进生态	省份、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、

数据模块	数据表	数据字段
	文明建设	出现该关键字的句子数
	省市-促进经济社会可持续发展	省份、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、出现该关键字的句子数
地级市环境注意力	地级市-基本信息统计	省份、地级市、年份、总字数 1、总字数 2、总词数 1、总词数 2、句子数
	地级市-保护和改善环境	省份、城市、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、出现该关键字的句子数
	地级市-防治污染和其他公害	省份、城市、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、出现该关键字的句子数
	地级市-资源节约	省份、城市、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、出现该关键字的句子数
	地级市-协同发展与环境共治	省份、城市、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、出现该关键字的句子数
	地级市-推进生态文明建设	省份、城市、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、出现该关键字的句子数
	地级市-促进经济社会发展	省份、城市、年份、词汇名称、词汇类别、词汇组别、词频统计、句子示例、出现该关键字的句子数

5.数据范围

数据模块	数据起始时间
省市环境注意力	1979 年
地级市环境注意力	1986 年

（十九）金融科技研究数据库-FTCH

1.数据库简介

金融科技研究数据库（Fintech Research Database, FTCH）基于不同方法构建了多维度的金融科技发展情况，用以测度我国地区和银行业的金融科技发展水平。FTCH 分为两个模块：地区金融科技情况、银行金融科技情况。其中，地区金融科技情况主要从各地区的金融科技数量以及该地区对于金融科技相关关键词的网络搜索指数进行衡量；银行金融科技情况则根据银行申请金融科技相关专利情况进行数据构建。希望 FTCH 的推出能为学者深入研究金融科技提供更多数据支持。

2.数据库特色

- 多元的测度方法

FTCH 从不同角度衡量我国各地区以及银行业的金融科技发展水平，有利于满足学者多方面的研究需求。

3.典型数据举例

各地区金融科技搜索指数

省份	地级市	统计日期	金融科技搜索指数
山东	全省	2022-01-01	3333
新疆	阿拉尔	2022-01-01	0
河南	新乡	2022-01-01	612
广东	广州	2022-01-01	2133
青海	全省	2022-01-01	645

4.数据库内容

数据模块	数据表	数据字段
地区金融科技发展水平	各地区金融科技 公司数	统计年份、省份、地级市、地级市代码、当年新注册金融科技数量、累计注册金融科技数量
	各地区金融科技 搜索指数	省份、地级市、统计日期、金融科技搜索指数
银行金融科技发展水平	银行金融科技创 新专利	申请公告号、专利名称、申请公告日、申请日、申请人、分类号、专利类型

5.数据范围

数据模块	数据起始时间
地区金融科技发展水平	截面数据
银行金融科技发展水平	2003 年

(二十) 工业机器人研究数据库-IRRD

1. 数据库简介

工业机器人研究数据库 (*Industrial Robots Research Database, IRRD*) 基于中国海关提供的各省市工业机器人数据以及国家知识产权局的工业机器人专利数据, 从多个维度展示了我国工业机器人进出口及其专利研发情况。*IRRD* 分为四个表格, 各省份工业机器人进口和出口的月度数据、上市公司及参控股公司工业机器人专利数据、各省市所申请工业机器人专利数据。希望 *IRRD* 的推出能为学者深入研究工业机器人相关课题提供多维度的数据支持。

2. 数据库特色

- 多维度数据研究

IRRD 研究了我国各地区和上市公司及其参控股公司的工业机器人发展情况, 有利于满足学者多方面的研究需求。

- 全面的数据展示

IRRD 展示了工业机器人进出口情况和工业机器人相关专利申请情况, 数据从不同角度全面展示了我国工业机器人发展情况。

3. 典型数据举例

(1) 各省市工业机器人专利

年份	省份	地级市	发明公布	发明授权	实用新型	外观设计
2014	海南省	海口市	0	0	0	0
2015	湖北省	孝感市	2	2	0	0
2016	黑龙江省	哈尔滨市	16	11	9	0
2022	浙江省	杭州市	53	2	42	4

(2) 上市公司工业机器人专利

股票代码	上市公司名称	会计年度	申请日	参控股公司名称	参控关系	专利名称	分类号	专利类型
000039	中集集团	2019	2019.01.23	深圳前海瑞集科技有限公司	子公司	一种前端角柱焊接设备	B23K37/047(2006.01)I;B23K37/02(2006.01)I	实用新型
000100	TCL 科技	2016	2016.11.29	武汉华星光电技术有限公司	子公司	一种液晶玻璃基板搬运机器人	B25J9/00(2006.01)I;B25J9/12(2006.01)I;B25J15/02(2006.01)I	发明公布
000584	ST 工智	2019	2019.01.17	苏州哈工易科机器人有限公司	子公司	外部 PLC 变位机系统与焊接机器人速度匹配方法	B23K37/047(2006.01)I;B23K37/02(2006.01)I	发明授权

601238	广汽集团	2021	2021.06.24	广汽菲亚特克莱斯勒汽车有限公司	合营公司	一种增压机节能控制系统	G05B19/042(2006.01)I	实用新型
--------	------	------	------------	-----------------	------	-------------	----------------------	------

4.数据库内容

数据模块	数据表	数据字段
各省份工业机器人出口	各省份工业机器人出口	年份、月份、省份、省份编码、出口数量、出口金额
各省份工业机器人进口	各省份工业机器人进口	年份、月份、省份、省份编码、进口数量、进口金额
各省市工业机器人专利	各省市工业机器人专利	年份、省份、地级市、发明公布、发明授权、实用新型、外观设计
上市公司工业机器人专利	上市公司工业机器人专利	股票代码、上市公司名称、会计年度、申请日、参控股公司名称、参控关系、专利名称、分类号、专利类型

5.数据范围

数据模块	数据起始时间
各省份工业机器人出口	2015 年
各省份工业机器人进口	2015 年
各省市工业机器人专利	2012 年
上市公司工业机器人专利	2015 年

区域经济研究系列

(一) 城市统计数据库-CCSD

1. 数据库简介

中国城市统计数据库 (*Chinese City Statistics Database, CCSD*) 涵盖了我国 600 多个城市 (含地级以上城市和县级城市) 社会经济发展方面的主要统计数据。*CCSD* 数据主要来源于国家统计局发布的《中国城市统计年鉴》，根据年鉴中的数据内容，将其归类为 11 个大模块，分别是行政区划、人口就业、公共设施、土地农业、教育教学、工业生产、财政投资、环境环保、经济金融、邮政电信、其他。这些模块既包括了地级以上城市的统计资料，又囊括县级城市统计资料。根据学术研究需要，*CCSD* 在横向上从数据类别对数据进行归纳总结，在纵向上从时间序列对数据进行多年度整合，极大满足了学术研究的需求。

2. 数据库特色

- 准确的数据归类

CCSD 根据《中国城市统计年鉴》各个表格数据内容，准确地将其归类为 11 个大模块，直观体现了城市社会经济发展方面的主要统计数据。

- 简洁的数据结构

CCSD 不仅将地级市和县级市数据分别列示，并整合了多个年度的相同数据到一张表格中，数据结构简洁，下载方便，节约了数据搜寻和整理时间。

3. 典型数据举例

在岗职工人数及工资状况_地级市

省份名称	城市名称	会计年度	在岗职工平均人数	在岗职工工资总额	职工平均工资	统计范围
河北省	石家庄市	1999	96	758877	7938	全市
河北省	唐山市	1999	88	640579	7288	全市
河北省	秦皇岛市	1999	33	272117	8314	全市
河北省	邯郸市	1999	69	425401	6169	全市
河北省	邢台市	1999	41	255322	6255	全市

4. 数据库内容

数据模块	数据表
行政区划	行政区划和区域分布
人口状况	人口及户数；人口变动情况（截止 2019 年）；人口密度（截止 2015）
资源环境	行政区域土地面积及水资源总量_地级市；城市建设用地状况_市辖区；绿地面积及建成区绿化覆盖面积_市辖区；工业废水、颗粒物排放量和二氧化硫及氮氧化物排放量_全市；工业烟(粉)尘排放量和工业固体废物综合利用率_全市（截止 2019 年）；可吸入细颗粒物年平均浓度和污水及生活垃圾处理率_全市
经济发展	地区生产总值_地级市；地区生产总值构成_地级市；公共财政收支_地级市；年末金融机构存贷款余额_地级市；固定资产投资情况_地级市（截止 2019 年）；商品

数据模块	数据表
经济发展	房销售面积；对外经济贸易_地级市；规模以上工业企业数_地级市；规模以上工业企业主要财务指标_地级市；社会消费品零售总额及批发零售贸易业情况_地级市；规模以上工业总产值_地级市（截止 2016 年）
科技创新	科技创新情况_全市
人民生活	劳动力就业及收入状况_地级市；按产业划分的年末单位从业人员_地级市（截止 2019 年）；按产业划分的年末单位从业人员构成(比重)_地级市（截止 2019 年）；按行业分组的单位从业人员_地级市（截止 2019 年）；在岗职工人数及工资状况_地级市（截止 2019 年）
公共服务	学校数量_地级市；专任教师数_地级市；在校学生数_地级市；文化体育设施_地级市；医院、床位和医生数_地级市；社会保障主要指标_地级市
基础设施	市政公用事业_市辖区；道路面积及公共汽车、出租车拥有情况_市辖区；按运输方式分类的客运量_全市；按运输方式分类的货运量_全市；邮政、电信业发展情况_地级市；供水、用水及用电情况_地级市（截止 2019 年）；煤气及液化石油气供应及利用情况_市辖区
县级城市	人口状况_县级市；劳动力就业状况_县级市（截止 2019 年）；行政区划土地面积_县级市；地区生产总值_县级市；公共财政收支_县级市；年末金融机构存贷款余额_县级市；规模以上工业企业情况_县级市；在校学生数_县级市；学校数量_县级市（截止 2014 年）；专任教师数_县级市（截止 2014 年）；医院、卫生院床位数及医护人员数_县级市（截止 2014 年）；固定资产投资情况_县级市（截止 2016 年）；社会消费品零售总额及批发零售贸易业情况_县级市（截止 2014 年）；工资水平_县级市（截止 2012 年）；主要农产品产量_县级市（截止 2009 年）

5.数据范围

数据模块	数据起始时间
全部模块	1999 年

（二）中国县域统计数据库-COTY

1.数据库简介

中国县域统计数据库 (*Chinese County Statistics Database, COTY*) 是基于中国县域统计年鉴中各县市社会经济指标数据而开发的专业数据库。*COTY* 反映了我国 2000 多个县域单位的基本情况、综合经济、农业、工业、基本建设、教育、卫生、社会保障等方面情况。为了更为直观地展示各县域单位的社会经济情况, *COTY* 将各类指标经整理后区分为四个模块: 基本情况; 综合经济; 农业、工业及投资; 教育、卫生和社会保障 4 个部分, 希望本数据库能为相关领域研究提供重要数据支持。

2.数据库特色

- 数据来源权威

COTY 数据来源于历年中国县域统计年鉴, 指标覆盖完整, 数据权威。

- 数据分表与整理

COTY 根据县域统计年鉴各年份指标与数据特性, 将其归纳整理为四个不同内容的表格, 结构更清晰、直观。

3.典型数据举例

区县基本情况

指标	单位	年份	省份	地区	地区编码	数值
行政区域面积	平方公里	2001	北京市	大兴县	110115	1030
乡(镇)个数	个	2001	北京市	大兴县	110115	14
村民委员会个数	个	2001	北京市	大兴县	110115	544
年末总户数	户	2001	北京市	大兴县	110115	175791
年末总户数其中: 乡村户数	户	2001	北京市	大兴县	110115	117951

4.数据库内容

数据模块	数据表	数据字段
区县基本情况	区县基本情况	指标、单位、年份、省份、地区、地区编码、数值
区县综合经济	区县综合经济	指标、单位、年份、省份、地区、地区编码、数值
区县农业、工业及投资	区县农业、工业及投资	指标、单位、年份、省份、地区、地区编码、数值
区县教育、卫生和社会保障	区县教育、卫生和社会保障	指标、单位、年份、省份、地区、地区编码、数值

5.数据范围

数据模块	数据起始时间
全部模块	2001 年

(三) 中国区域经济统计数据库-REGN

1. 数据库简介

中国区域经济统计数据库 (*China Statistical Database for Regional Economy, REGN*) 是基于国家统计局数据, 提供了 1999 年以来全国多个经济区域、省市、地级市和县域的主要社会经济统计数据。主要内容涵盖自然资源、人口与就业、国民核算、工业、农业、建筑业、金融、教育等社会经济发展的各个方面。

2. 数据库特色

- 数据精准划分

REGN 将分省份数据和分地级市数据中的诸多表格, 按照关联性进行了分组, 并整合在不同模块中, 极大便利了使用者进行数据查找、浏览和下载。

3. 典型数据举例

东部 10 省(市) 主要指标

指标	单位	会计年度	合计	占全国比重(%)
土地面积	万平方公里	2012	91.6	9.5
年底总人口	万人	2012	51460.9	38.2
城镇就业人员	万人	2012	7225.3	47.4
国内(地区)生产总值	亿元	2012	295892	51.3
国内(地区)生产总值_第一产业	亿元	2012	18339.6	35

4. 数据库内容

数据模块	数据子模块	数据表
分经济区域数据	分经济区域数据	东部 10 省(市)主要指标; 中部 6 省主要指标; 西部 12 省(区、市)主要指标; 东北 3 省主要指标; 民族自治地方主要指标; 长江三角洲经济区主要指标; 珠江三角洲经济区主要指标; 海峡西岸经济区主要指标; 北部湾经济区主要指标; 关中-天水经济区主要指标
分省份数据	综合经济	全国行政区划; 国内生产总值; 支出法国内生产总值; 地区生产总值收入法构成项目; 财政收入; 财政支出; 进出口商品总值; 电力消费量; 亿元以上商品交易市场基本情况; 接待入境旅游人数和国际旅游外汇收入; 实际利用外商直接和其他投资额; 外商投资企业进出口商品总值; 外商投资企业年末登记情况
	人口就业	总人口和出生率、死亡率、自然增长率; 人口的城乡构成和出生率、死亡率、自然增长率; 按城乡分的就业人员数; 按三次产业分的年底就业人员数; 按行业分城镇单位就业人员数(年底数); 乡村从业人员; 城镇登记失业人员及失业率; 职业介绍工作情况; 农村基层组织情况
	基础设施	城市设施水平; 城市建设情况; 城市公共交通情况; 城市园林绿化情况; 城市市政工程情况; 城市煤气、液化石油气、天然气情况; 城市自来水情况; 城市集中供热情况……
	农林牧渔	农、林、牧、渔业总产值及指数; 主要农产品产量; 主要林产品产量; 畜产品产量; 水产品产量; 主要农产品单位面积产量……

数据模块	数据子模块	数据表
分省份数据	教育医疗	普通高校基本情况；普通高中基本情况；中等职业学校(机构)基本情况；普通初中基本情况；职业初中基本情况；普通小学基本情况；幼儿园基本情况；特殊教育情况……
	工业生产	主要工业产品产量；大中型工业企业主要指标；大中型工业企业主要经济效益指标；国有及国有控股工业企业主要指标；国有及国有控股工业企业主要经济效益指标……
	固定资产投资	资本形成总额及构成；固定资产投资价格指数；按城乡分全社会固定资产投资；全社会固定资产投资实际到位资金……
	建筑地产	城市房屋建筑及住房情况；全社会施工、竣工房屋面积和价值；建筑业劳动生产率；房屋建筑面积；房地产开发企业(单位)建设房屋建筑面积和造价；总承包建筑业企业主要经济指标……
	环境资源	自然保护工作基本情况；主要城市空气质量指标；主要城市道路交通噪声监测情况；工业固体废物产生及处理利用情况；废气排放及处理情况；废水排放及处理情况；环境污染与破坏事故情况；造林面积；湿地面积；水资源情况；森林资源情况；主要有色金属、非金属矿产基础储量；主要能源、黑色金属矿产基础储量；能源消耗指标
	运输邮电	全社会客运量；全社会旅客周转量；货运量；货物周转量；货物进出口总额；运输线路长度；民用车辆拥有量；公路营运汽车拥有量；私人汽车拥有量；邮电业务量；电信主要通信能力；邮政业网点及邮递线路
	餐饮零售	连锁零售企业基本情况；连锁餐饮企业基本情况；限额以上批发零售贸易、餐饮业产业活动单位数；限额以上批发零售贸易……
	文化科技	文化事业机构数；博物馆基本情况；图书、杂志和报纸出版情况；录像制品出版情况；录音制品出版情况；图书纯销售情况；技术市场成交额；国有企事业单位专业技术人员数；三种专利申请受理数和授权数
人民生活	居民消费水平；最终消费及构成；居民消费价格分类指数；居民消费价格指数和商品零售价格指数；城镇居民家庭平均每人全年消费性支出……	
分地级市数据	综合经济	地区生产总值及指数；财政收入和支出；全社会固定资产投资和房屋销售；货物进出口总额和利用外资
	农业工业	农村和农业生产基本情况；播种面积和农产品产量；规模以上内资工业企业总产值；规模以上工业企业单位数、总产值和主要经济指标；建筑业情况
	第三产业	交通运输业、邮电业情况；社会消费品零售总额和批发零售业、住宿餐饮业情况；金融机构存款和贷款；旅游业情况
	科教文卫	普通高等学校情况；普通中等学校情况；幼儿园和小学情况；文化事业情况；卫生机构、床位数和人员数
	居民生活	土地面积和人口情况；就业、失业和工资；城乡居民家庭收入和支出；城市供气、供水和交通；城市排水、绿化和垃圾处理
分县域数据	分县域数据	分县域数据

5.数据范围

数据模块	数据起始时间
全部模块	1999年

（四）中国城乡建设统计数据库-URCS

1.数据库简介

中国城乡建设统计数据库（China Urban-Rural Construction Statistical Database, **URCS**）是基于住房和城乡建设部的年鉴数据，反映我国城乡市政公用设施建设与发展状况的专业数据库。**URCS**包括城市建设、县城建设和村镇建设三个大模块，每个大模块根据各自的内容分为不同的子模块。其中城市建设模块和县城建设模块包括道路交通、公用设施、供气供热、供水用水、固定资产投资、环境绿化、建设资金和人口土地八个子模块；村镇建设模块包括基本信息、公用设施、供水用水、燃气供热道路、环境绿化、建设资金和住宅房屋七个模块。

2.数据库特色

- 模块重新划分

URCS 将原始年鉴中的表格进行了重新划分，并建立相应的子模块，便于研究人员寻找所需数据。

3.典型数据举例

城市建设_城市道路和桥梁

地区名称	会计年度	道路长度 (公里)	道路面积 (万平方米)	桥梁数 (座)	桥梁数_立交桥 (座)	道路照明灯盏数 (盏)
北京	2013	8107	13834	2244	433	248695
天津	2013	7275	13144	869	110	309857
河北	2013	12859	30113	1378	259	624580
山西	2013	7107	14470	640	132	436661

4.数据库内容

数据模块	数据子模块	数据表
城市建设	道路交通	全国历年城市轨道交通情况；全国历年城市道路和桥梁情况；城市轨道交通(建成)；城市轨道交通(在建)；城市道路和桥梁
	公用设施	全国历年城市市政公用设施水平；全国城市市政公用设施水平；全国城市市政公用设施建设施工规模和新增生产能力(或效益)
	供气供热	全国历年城市燃气情况；全国历年城市集中供热情况；城市天然气；城市人工煤气；城市液化石油气；城市集中供热
	供水用水	全国历年城市供水情况；全国历年城市节约用水情况；全国历年城市排水和污水处理情况；城市供水；城市供水(公共供水)；城市供水(自建设施供水)；城市节约用水；城市排水和污水处理
	固定资产投资	按行业分全国历年城市市政公用设施建设固定资产投资；按资金来源分全国历年城市市政公用设施建设固定资产投资；按行业分全国城市市政公用设施建设固定资产投资；按资金来源分全国城市市政公用设施建设固定资产投资
	环境绿化	全国历年城市市容环境卫生情况；全国历年城市园林绿化情况；城市园林绿化；城市市容环境卫生；国家级风景名胜区
	建设资金	全国历年城市维护建设资金收入；全国历年城市维护建设资金支出；全国城市维护建设资金(财政性资金)收支
	人口土地	全国历年城市数量及人口、面积情况；全国城市人口和建设用地

数据模块	数据子模块	数据表
县城建设	道路交通	全国历年县城道路和桥梁情况；县城道路和桥梁
	公用设施	全国历年县城市政公用设施水平；全国县城市政公用设施水平；全国县城市政公用设施建设施工规模和新增生产能力(或效益)
	供气供热	全国历年县城燃气情况；全国历年县城集中供热情况；县城人工煤气；县城天然气；县城液化石油气；县城集中供热
	供水用水	全国历年县城供水情况；全国历年县城节约用水情况；全国历年县城排水和污水处理情况；县城供水；县城供水(公共供水)；县城供水(自建设施供水)；县城节约用水；县城排水和污水处理
	固定资产投资	按行业分全国历年县城市政公用设施建设固定资产投资；按资金来源分全国历年县城市政公用设施建设固定资产投资；按行业分全国县城市政公用设施建设固定资产投资；按资金来源分全国县城市政公用设施建设固定资产投资
	环境绿化	全国历年县城市容环境卫生情况；全国历年县城园林绿化情况；县城市容环境卫生；县城园林绿化
	建设资金	全国历年县城维护建设资金收入；全国历年县城维护建设资金支出；全国县城维护建设资金(财政性资金)收支
	人口土地	全国历年县城数量及人口、面积情况；全国县城人口和建设用地
村镇建设	基本情况	全国历年村庄基本情况；乡基本情况；建制镇基本情况；镇乡级特殊区域基本情况；村庄人口及面积；村庄规划及整治
	公用设施	全国历年乡市政公用设施情况；全国历年建制镇市政公用设施情况；乡市政公用设施水平；建制镇市政公用设施水平；镇乡级特殊区域市政公用设施水平；村庄公共设施
	供水用水	乡供水；建制镇供水；镇乡级特殊区域供水；乡排水和污水处理；建制镇排水和污水处理；镇乡级特殊区域排水和污水处理
	燃气供热道路	乡燃气、供热、道路桥梁；建制镇燃气、供热、道路桥梁；镇乡级特殊区域燃气、供热、道路桥梁
	环境绿化	乡园林绿化及环境卫生；建制镇园林绿化及环境卫生；镇乡级特殊区域园林绿化及环境卫生
	建设资金	乡建设投入；村庄建设投入；建制镇建设投入；镇乡级特殊区域建设投入；村庄建设财政性资金投资；乡建设财政性资金投资；建制镇建设财政性资金投资；镇乡级特殊区域建设财政性资金投资
	住宅房屋	全国历年乡及住宅基本情况；全国历年建制镇及住宅基本情况；村庄房屋；乡房屋；建制镇房屋；镇乡级特殊区域房屋

5.数据范围

数据模块	数据起始时间
全部模块	1978年

（五）中国农村统计数据库-CRSD

1.数据库简介

中国农村统计数据库（*China Rural Statistical Database, CRSD*）是基于国家统计局数据整理而成的一个用于反映我国农村经济发展情况的特色数据库。*CRSD* 主要包括综合和概要，农村基本情况与农业生产条件，农业生态与环境，农村投资，农林牧渔业总产值、中间消耗及增加值，主要农产品种植（养殖）面积与产量，农村市场与物价，农产品进出口，收入与消费，农村文化、教育、卫生及社会服务，各地区主要农村经济指标排序等十一个大模块。

2.数据库特色

- 精简的数据划分

CRSD 将关联度较大的表合并整理成一张表格，极大地节省了数据使用者的下载时间。

- 数据质量较高

CRSD 对年鉴原文中的缺失数据进行了多来源数据的补充，大大提高了数据质量。

3.典型数据举例

农村基本情况与农业生产条件_各地区农用地情况

地区	会计年度	农用地数量 (千公顷)	耕地(千公 顷)	园地(千公 顷)	林地(千公 顷)	草地(千公 顷)	其他(千公 顷)
全国	2016	645126.6	134921	14266.3	252908.1	219359.2	23672.2
北京	2016	1145.6	216.3	133.5	739.7	0.2	55.8
河北	2016	13069.2	6520.5	834.4	4599	401.3	714.1
山西	2016	10027	4056.8	406.4	4856	33.8	674.1
内蒙古	2016	82882.4	9258	56.6	23228	49519.9	820
辽宁	2016	11531.5	4974.5	468.2	5615.9	3.2	469.7

4.数据库内容

数据模块	数据表
综合和概要	农村经济主要指标；人均主要农产品产量；农村经济在国民经济中的地位；各地区农村经济在国民经济中的地位；各地区社会消费品零售额；各地区城乡居民收入水平；各地区城乡居民消费水平；主要农产品供需情况
农村基本情况与农业生产条件	全国乡村人口和乡村就业人员情况；各地区乡村人口数；主要农业机械年末拥有量；各地区主要农业机械年末拥有量；农村电力、灌溉面积、化肥施用量情况；农村电力和农田水利建设情况；农用化肥、农膜、柴油和农药使用量；各地区农村电力和农田水利建设情况；各地区农用化肥施用量（按折纯法计算）；各地区农用塑料薄膜使用量；各地区农用柴油和农药使用量；各地区农用地情况；各地区耕地面积构成
农业生态与环境	全国自然保护区情况；农村环境情况；各地区自然保护基本情况；各地区农村改厕情况；各地区农村可再生资源利用情况；全国林业重点生态工程历年完成造林面积；各地区林业重点生态工程建设情况；灌区、水库、除涝、治水情况；各地区水利设施和除涝、治水面积……

数据模块	数据表
农村投资	国家财政用于农林水各项支出；农村住户固定资产投资情况；各地区农村住户固定资产投资情况；农村住户固定资产投资和建房情况；分地区农村住户固定资产投资和建房情况
农林牧渔业总产值、中间消耗及增加值	农林牧渔业增加值和指数；各地区分部门农林牧渔业增加值(按当年价格计算)；各地区分部门农林牧渔业增加值构成(按当年价格计算)；各地区分部门农林牧渔业增加值率(以该部门总产值为 100)；各地区分部门农林牧渔业中间消耗(按当年价格计算)……
主要农产品种植（养殖）面积与产量	各地区农作物总播种面积；各地区粮食播种面积；各地区粮食播种面积（按季节分）；各地区粮食播种面积（按品种分）；各地区油料播种面积；各地区棉花和麻类播种面积；各地区糖料播种面积；各地区烟叶和药材种植面积；各地区蔬菜、瓜果类和青饲料播种面积……
农村市场与物价	农村主要物价总指数（以上年价格为 100）；各地区农村商品零售价格分类指数(上年价格=100)；各地区农村居民消费价格分类指数（以上年价格为 100）；各地区农业生产资料价格分类指数（以上年价格为 100）；各地区主要农产品分品种生产者价格指数（以上年价格为 100）
农产品进出口	海关出口主要农产品数量；海关进口主要农产品数量；海关出口农副产品及加工品数量；海关出口农副产品及加工品金额；海关进口农副产品及加工品数量；海关进口农副产品及加工品金额
收入与消费	农村居民可支配收入及构成；农村居民消费支出及构成；农村居民现金消费支出及构成；农村居民主要食品消费量；农村居民年末主要耐用消费品拥有量；分地区农村居民可支配收入构成；分地区农村居民消费支出；分地区农村居民消费支出构成；分地区农村居民现金消费支出；分地区农村居民现金消费支出构成
农村文化、教育、卫生及社会服务	乡村教育情况；农民成人教育基本情况；农村乡（镇）卫生院情况；各地区农村村卫生室和人员情况；农村养老服务和文化机构情况；各地区农村社会救济情况
各地区主要农村经济指标排序	粮食总产量和人均占有量；油料总产量和人均占有量；糖料总产量与人均占有量；肉类总产量与人均占有量；水产品总产量和人均占有量；水果总产量与人均占有量；奶类总产量与人均占有量；各地区农村居民人均可支配收入位次

5.数据范围

数据模块	数据起始时间
全部模块	1990 年

（六）区域碳排放数据库-RCED

1.数据库简介

区域碳排放数据库 (Regional Carbon Emission Database, RCED) 利用 IPCC 法核算各省份 (除西藏外) 每年温室气体的碳排放量 [IPCC (Intergovernmental Panel on Climate Change) 即联合国政府间气候变化专门委员会, 其主要目的是研究自然气候变化, 以及人与自然相处过程中的关系问题。IPCC 法是目前各方公认的区域温室气体核算方法。]。RCED 参考联合国环境规划署公布的《城市温室气体排放国际标准》和相关领域文献研究, 并基于各统计年鉴公布的数据, 分别从能源活动、工业生产过程、农业活动、土地利用变化和林业以及废弃物处理五个部门计算各地区的碳排放量。其中, 碳排放量均以排放的温室气体中碳元素重量 (万吨) 表示, 由各活动数据乘以《省级温室气体清单编制指南 (试行)》报告中给出的对应排放因子折算并加总得到。RCED 的推出旨在为学者深入研究区域碳排放和气候变化等提供更多数据支持。各部门的碳排放核算方法请见本文附录。

2.数据库特色

- 科学的测算方法

RCED 根据目前各方公认的 IPCC 法, 并参考相关文献对温室气体的碳排放进行核算, 数据来源可靠, 核算方法科学。

- 清晰的数据结构

RCED 分别列示了我国各省份五个部门的碳排放量, 并折算加总出各省碳排放总量, 数据详尽, 结构清晰。

3.典型数据举例

土地利用变化和林业碳排放

省份	年份	活立木总蓄积量 (万立方米)	生长率 (%)	消耗率 (%)	基本木材密度 (吨/m ³)	加权平均生物量转换系数	固碳量 (万吨)	土地利用变化和林业碳排放量 (万吨)	数据来源
广西	2012	51056.78	8.94	5.9	0.43	1.819	607.0132	-607.0132	《中国环境统计年鉴》
广东	2012	32160.74	8.24	7.18	0.474	1.915	154.7209	-154.7209	《中国环境统计年鉴》
湖南	2012	38177.2	9.9	6.38	0.394	1.712	453.2280	-453.2280	《中国环境统计年鉴》
湖北	2012	23121.55	8.29	4.94	0.459	1.848	328.5083	-328.5083	《中国环境统计年鉴》
河南	2012	18051.16	11.68	6.86	0.488	1.74	369.3952	-369.3952	《中国环境统计年鉴》
山东	2012	8627.99	15.28	9.51	0.412	1.774	181.9308	-181.9308	《中国环境统计年鉴》

4.数据库内容

数据模块	数据表	数据字段
区域碳排放	能源活动碳排放	省份、年份、原煤/焦炭/原油/汽油/煤油/柴油/燃料油/天然气(万吨)、原煤/焦炭/原油/汽油/煤油/柴油/燃料油/天然气平均低位发热量(KJ/kg)、原煤/焦炭/原油/汽油/煤油/柴油/燃料油/天然气单位热值含碳量(吨碳/TJ)、原煤/焦炭/原油/汽油/煤油/柴油/燃料油/天然气碳氧化率、原煤/焦炭/原油/汽油/煤油/柴油/燃料油/天然气/电力碳排放量(万吨)、电力(亿 KWh)、电力排放因子(kg 二氧化碳/KW.h)、电力碳排放量(万吨)、能源活动碳排放量(万吨)、数据来源
	工业生产过程碳排放	省份、年份、水泥(万吨)、水泥排放因子(吨二氧化碳/吨熟料)、水泥生产过程碳排放量(万吨)、平板玻璃(万重量箱)、平板玻璃排放因子(吨二氧化碳/吨玻璃)、平板玻璃生产碳排放量(万吨)、工业生产过程碳排放量(万吨)、数据来源
	农业活动碳排放	省份、年份、非奶牛/奶牛/马/驴/骡(万头)、非奶牛/奶牛/马/驴/骡甲烷排放因子(kg/头/年)、非奶牛/奶牛/马/驴/骡肠道发酵碳排放量(万吨)、单季稻/双季早稻/双季晚稻(千公顷)、单季稻/双季早稻/双季晚稻甲烷排放因子(kg/公顷)、单季稻/双季早稻/双季晚稻碳排放量(万吨)、农业活动碳排放量(万吨)、数据来源
	土地利用变化和林业碳排放	省份、年份、活立木总蓄积量(万立方米)、生长率(%)、消耗率(%)、基本木材密度(吨/m ³)、加权平均生物量转换系数、固碳量(万吨)、土地利用变化和林业碳排放量(万吨)、数据来源
	废弃物处理碳排放	省份、年份、废水化学需氧量排放总量(万吨)、最大甲烷生产能力、废水碳排放量(万吨)、垃圾焚烧量(万吨)、废弃物含碳比例、矿物碳比例、垃圾焚烧碳排放量(万吨)、垃圾填埋量(万吨)、垃圾填埋甲烷排放因子、垃圾含水率、垃圾填埋碳排放量(万吨)、废弃物处理碳排放量(万吨)、数据来源
	各省碳排放总量	省份、年份、能源活动碳排放量(万吨)、工业生产过程碳排放量(万吨)、农业活动碳排放量(万吨)、土地利用变化和林业碳排放量(万吨)、废弃物处理碳排放量(万吨)、各省碳排放总量(万吨)

5.数据范围

数据模块	数据起始时间
全部模块	2005 年

产业经济研究系列

(一) 产业政策数据库-IPRD

1. 数据库简介

产业政策研究数据库 (*Industrial Policy Research Database, IPRD*) 是基于我国中央政府和各省级政府每五年发布一次的国民经济和社会发展五年规划纲要, 并结合已有产业政策方面的学术研究文献, 整理归纳“五年规划”中提及的相关产业制作而成的数据库。*IPRD* 不仅提取了五年规划纲要中提及的相关产业及规划内容, 还依据文中对提及产业的描述语气, 归纳了国家及各省对提及的相关产业的政策态度, 分为鼓励、中性和抑制三类。此外, 还根据证监会发布的“上市公司行业分类指引 2001”对五年规划大纲中提及的产业进行标准化分类。

2. 数据库特色

- 丰富的产业政策相关字段

IPRD 不仅提取了五年规划中提及的相关产业及规划内容, 还对这些产业进行了标准化分类, 并对产业政策态度进行了区分, 为该领域研究提供更多直观可用的数据。

- 囊括中央和地方产业政策

除了主流研究使用的国家产业政策, *IPRD* 还提供各省级政府产业政策, 进一步拓展了研究范围。

3. 典型数据举例

中央产业政策

五年规划	批准日期	规划提及产业	行业代码 2001	政策态度	是否重点支持产业
十五计划	2001-03-15	旅游业	K34	鼓励	是
十五计划	2001-03-15	零售业和餐饮业	H11; K30	鼓励	否
十五计划	2001-03-15	职业培训产业	K20	鼓励	否
十一五规划	2006-03-14	建材业	C61	抑制	否

4. 数据库内容

数据模块	数据表	数据字段
产业政策	中央产业政策	五年规划、批准日期、规划提及产业、规划内容、所属行业大类、行业分类 2001、行业代码 2001、政策态度、是否重点支持产业
	省级产业政策	所属省份、五年规划、批准日期、规划提及产业、规划内容、所属行业大类、行业分类 2001、行业代码 2001、政策态度、是否重点支持产业、是否国家重点支持产业

5. 数据范围

数据模块	数据起始时间
中央产业政策	2001 年
省级产业政策	2006 年

(二) 中国房地产统计数据库-RES D

1. 数据库简介

中国房地产统计数据库 (*Real Estate Statistics Database, RES D*) 是基于国家统计局数据和中国房地产统计年鉴数据整理开发的专业数据库。*RES D* 收集整理了中国近十几年来的房地产开发企业的基本信息、房地产企业经营财务数据以及全国 35 个大中城市的房地产经济统计数据等, 提供了全面的数据和丰富的字段, 为学术界对房地产领域的研究提供了重要数据支持。

2. 数据库特色

- 数据来源权威

RES D 数据来源于国家统计局以及历年房地产统计年鉴, 指标覆盖完整, 数据权威。

- 结构的一致性

RES D 经整理后的模块结构与表格形式上与近几年的年鉴格式保持一致, 便于进行数据比对。

3. 典型数据举例

35 个大中城市按用途分房地产开发企业商品房平均销售价格

地区名称	会计年度	商品房均价 (元/平方米)	住宅 (元/平方米)	别墅、高档公寓 (元/平方米)
北京	2009	13799	13224	20385
北京	2010	17782	17151	28680
北京	2011	16852	15518	28543

4. 数据库内容

数据模块	数据表
房地产开发企业基本概况	房地产开发企业主要指标完成情况; 各地区按登记注册类型分房地产开发企业个数; 各地区按资质等级分房地产开发企业个数; 各地区按隶属关系分房地产开发企业个数; 各地区按登记注册类型分房地产开发企业从业人数; 各地区按资质等级分房地产开发企业从业人数; 各地区按隶属关系分房地产开发企业从业人员
房地产开发投资与资金状况	各地区房地产开发企业投资规模与完成情况; 各地区按登记注册类型分房地产开发企业计划总投资; 各地区按资质等级分房地产开发企业计划总投资……
房屋开竣工面积、商品房销售及土地情况	各地区按用途分的房地产开发企业房屋施工面积; 各地区按资质等级分房地产开发企业房屋施工面积; 各地区按资质等级分的房地产开发企业住宅施工面积……
房地产开发企业主要财务状况	各地区房地产开发企业主要财务指标; 各地区按登记注册类型分房地产开发企业资产总计; 各地区按资质等级分房地产开发企业资产总计……
35 个大中城市房地产综合情况	35 个大中城市按登记注册类型分房地产开发企业个数; 35 个大中城市按资质等级分房地产开发企业个数……

5. 数据范围

数据模块	数据起始时间
全部模块	1999 年

（三）中国能源统计数据库-CESD

1.数据库简介

中国能源统计数据库（*Chinese Energy Statistics Database, CESD*）涵盖了多个能源品种、多个行业和城市的能源生产和消费数据。*CESD* 数据主要来源于国家统计局发布的《中国能源统计年鉴》，根据年鉴中的数据内容，将其归类为 5 个大模块，分别是综合、能源建设、能源生产、能源消费、能源平衡表。这些模块既包括能源总量和分能源品种的数据，还包括分行业分地区、分地区分能源品种等数据，在横向上从数据类别对数据进行归纳总结，在纵向上从时间序列对数据进行多年度整合。

2.数据库特色

- 准确的数据归类和简洁的数据结构

CESD 根据《中国能源统计年鉴》各个表格数据内容，在整合归纳的基础上，将其准确地归类为 5 个大模块，直观体现了中国能源发展的主要统计数据。

3.典型数据举例

能源生产与消费增长速度

会计年度	能源生产增长速度	能源消费增长速度	能源生产弹性系数	能源消费弹性系数
2012	3.2%	3.9%	0.40	0.49
2013	2.2%	3.7%	0.28	0.47
2014	0.9%	2.1%	0.12	0.29

4.数据库内容

数据模块	数据表
综合	能源生产与消费增长速度；能源经济主要指标；平均每万元国内生产总值能源消费量；能源加工转换效率；人均能源生产和消费量；人均生活用能量；交通运输设备拥有量；主要能源品种进、出口量；主要高耗能产品进、出口量；分地区主要污染物排放情况 1994~2010；分地区主要污染物排放情况 2011~2016
能源建设	国有经济能源工业分地区行业固定资产投资；能源工业分地区分行业投资；能源工业分行业投资
能源生产	一次能源生产量和构成；分地区分品种能源生产供应
能源消费	能源消费总量和构成；工业分行业终端能源消费量(实物量)；工业分行业终端能源消费量(标准量)；分行业分品种能源消费量；分地区分品种能源消费量
能源平衡表	中国能源平衡表(实物量)；中国能源平衡表(标准量)；分品种能源平衡表；分地区能源平衡表

5.数据范围

数据模块	数据起始时间
全部模块	1980 年

（四）中国工业统计数据库-CISD

1.数据库简介

中国工业统计数据库 (*Chinese Industrial Statistics Database, CISD*) 是基于历年中国工业统计年鉴整理而成的数据库, 系统地收录了全国各经济类型、各工业行业和各省、自治区、直辖市等工业经济各方面的统计数据。全库分为三个模块: 综合数据、分行业数据及分地区数据, 主要反映中国工业经济的发展、中国工业经济的行业结构、中国工业经济的地区布局, 可供学者了解和研究我国工业经济发展情况。

2.数据库特色

- 数据整合和归类

CISD 根据《中国工业统计年鉴》各表数据内容, 准确地将其归类为三个大模块, 并对属于类似内容的表格进行了合并, 数据结构更为简洁和直观。

3.典型数据举例

工业统计_工业企业单位数

年度	全国总计(万个)	国有控股(万个)	私营(万个)	外商投资和港澳台商投资(万个)
2011	32.56	1.71	18.06	5.72
2012	34.38	1.79	18.93	5.69
2013	36.98	1.86	20.84	5.74
2014	37.79	1.88	21.38	5.52

4.数据库内容

数据模块	数据表
综合数据	工业企业单位数; 工业企业主要经济指标; 分登记注册类型规模以上工业企业主要经济指标; 分行业规模以上工业企业主要效益指标; 分地区规模以上工业企业主要效益指标; 历年主要工业产品产量
分行业数据	分行业规模以上工业企业主要经济指标; 分行业分企业性质工业企业主要经济指标; 分行业大中型工业企业主要经济指标
分地区数据	分地区规模以上工业企业主要经济指标; 分地区分企业形式工业企业主要经济指标; 分地区分企业性质工业企业主要经济指标; 分地区分企业规模工业企业主要经济指标; 分地区采矿业-合计数; 分地区采矿业主要经济指标; 分地区制造业-合计数; 分地区制造业主要经济指标; 电力、热力、燃气及水生产和供应业; 分地区主要工业产品产量

5.数据范围

数据模块	数据起始时间
全部模块	1999 年

(五) 中国高技术产业统计数据库-HISD

1.数据库简介

中国高技术产业统计数据库 (*Chinese High-tech Industry Statistical Database, HISD*) 基于历年中国高技术产业统计年鉴整理而成, 系统地收录了中国高技术产业的生产经营情况、研发和投资情况。全库分为四个模块, 通过分行业、分地区和分企业性质等进行分类统计, 分别展示了高技术产业的生产经营情况、R&D 及相关活动情况、固定资产投资情况和国际比较情况。*HISD* 可为学者研究中国高技术产业的发展提供数据支持和参考。

2.数据库特色

- 数据整合和归类

HISD 根据《中国高技术产业统计年鉴》各表数据内容, 准确地将其归类为四个大模块, 并对属于类似内容的表格进行了合并, 数据结构更为简洁和直观。

3.典型数据举例

生产经营情况_高技术产业和制造业生产经营情况

会计年度	指标	高技术产业	制造业
2016	主营业务收入(亿元)	153796	1047711
2016	利润总额(亿元)	10302	65281
2016	出口交货值(亿元)	52445	117590

4.数据库内容

数据模块	数据表
生产经营情况	高技术产业和制造业生产经营情况、按行业分高技术产业生产经营情况、各地区高技术产业生产经营情况、按行业和企业规模分高技术产业生产经营情况、按行业分国有及国有控股企业高技术产业生产经营情况、按行业和登记注册类型分高技术产业生产经营情况、按地区和企业规模分高技术产业生产经营情况……
R&D 及相关活动情况	大中型高技术产业和制造业企业 R&D 及相关活动情况、分行业大中型企业 R&D 及相关活动情况、大中型国有及国有控股企业分行业 R&D 及相关活动情况、大中型内资企业分行业 R&D 及相关活动情况、大中型港澳台资企业分行业 R&D 及相关活动情况、大中型外资企业分行业 R&D 及相关活动情况……
固定资产投资情况	高技术产业和制造业固定资产投资基本情况、分行业固定资产投资情况、国有及国有控股企业固定资产投资情况、按地区和登记注册类型分高技术产业投资基本情况、按行业和登记注册类型分高技术产业投资基本情况、各地区国有及国有控股企业高技术产业投资基本情况、按地区和行业分高技术产业投资基本情况
国际比较情况	高技术产业出口总额及其占制造业出口的比重

5.数据范围

数据模块	数据起始时间
全部模块	2009 年

(六) 中国第三产业统计数据库-TISD

1.数据库简介

中国第三产业统计数据库(*Chinese Tertiary Industry Statistics Database, TISD*)是基于国家统计局出版的《中国第三产业统计年鉴》整理而成的一个专门反映我国第三产业经济发展情况的特色数据库。全库严格按照年鉴最新的模块划分对以往年度的相关模块进行了排列重组,形成了较为合理的体系。同时,在模块内对关联度较大的行业数据进行了整合,确保了数据的简洁性。

2.数据库特色

- 统一的模块划分

TISD 严格按照最新的年鉴模块分布对以前年度年鉴中分布杂乱的模块进行了整合。

- 简洁的数据归类

TISD 将具有相同字段名的各明细行业进行了统一整合,大大缩短了使用者搜寻相关数据的时间。

3.典型数据举例

第三产业单位数_第三产业按行业、东中西部以及东北地区分组的法人单位数

行业	会计年度	合计(个)	东部(个)	中部(个)	西部(个)	西北(个)
道路运输业	2012	122318	60152	28041	23949	10176
城市公共交通业	2012	9473	3543	2415	2319	1196
水上运输业	2012	10960	6783	2271	1396	510

4.数据库内容

数据模块	数据表
第三产业单位数	分地区的第三产业法人单位数及所占比重; 第三产业按行业、东中西部以及东北地区分组的法人单位数……
第三产业按行业门类分就业人员数	第三产业就业人员数及比重; 第三产业按行业按登记注册类型分就业人员数; 各地区按行业门类与登记注册类型分的第三产业就业人员数; 各地区按企业性质分的第三产业就业人员数
第三产业增加值	第三产业增加值及所占比重; 第三产业分行业增加值与构成……
第三产业固定资产投资	各地区按行业门类分第三产业固定资产投资; 按行业门类分第三产业投资占社会投资比重; 按地区按登记注册类型分第三产业全社会固定资产投资……
第三产业双向投资与服务贸易进出口情况	第三产业双向投资情况; 服务进出口分类金额情况
第三产业能源消费情况	分品种的能源平衡表
第三产业分行业主要指标	第三产业分行业分地区企业法人单位主要指标; 规模以上服务业企业法人单位分地区主要指标; 规模以上服务业企业法人单位分行业主要指标……
派生产业情况	国家旅游及相关产业、文化及相关产业增加值情况; 国内旅游情况; 历年入境游客数及其增长速度; 按国籍分入境外国游客……

5.数据范围

数据模块	数据起始时间
全部模块	2008年

（七）数字经济研究数据库-DERD

1.数据库简介

数字经济研究数据库 (*Digital Economy Research Database, DERD*) 是基于数字经济发展和相关研究所开发的专业数据库。*DERD* 从数字创新、数字产业和数字平台三个模块对我国数字经济发展情况进行展示, 其中, 数字创新包括各省、地级市、各行业的数字专利申请与授权情况, 上市公司数字经济申请与授权情况; 数字产业包括我国信息传输、计算机服务和软件行业的主要经济指标情况; 数字平台则展示了各省和地级市电信及互联网等数字经济基础设施的发展情况。*DERD* 从三个大维度 14 张表全方位展示了我国数字经济发展的情况, 为该领域研究提供了良好的数据支持。

2.数据库特色

- 多维度展示数字经济情况

DERD 从数字产业、数字创新和数字平台三个维度, 展示了我国数字经济发展情况, 字段和数据量均相当丰富。

- 提供独特的数字经济专利数据

DERD 根据相关部门发布的权威文件, 对专利数据进行了精准的匹配和处理, 提供了独有的数字经济专利统计数据。

3.典型数据举例

各省市各行业数字经济专利

省份	城市	会计年度	国民经济行业代码	当年申请的数字经济相关发明数量	当年授权的数字经济相关发明数量	当年申请的数字经济相关实用新型数量	当年授权的数字经济相关实用新型数量
江苏省	南京市	2012	3562	261	89	159	138
江苏省	南京市	2012	3563	251	91	149	141
江苏省	苏州市	2013	3912	829	108	238	282
江苏省	苏州市	2013	3913	466	58	225	269
江苏省	无锡市	2014	3973	706	196	371	311
江苏省	无锡市	2014	3974	570	149	324	291

4.数据库内容

数据模块	数据表	数据字段
数字创新	各省市数字经济专利	省份、地市、会计年度、当年申请的数字经济相关发明数量、当年授权的数字经济相关发明数量、当年申请的数字经济相关实用新型数量、当年授权的数字经济相关实用新型数量
	各行业数字经济专利	会计年度、国民经济行业代码、国民经济行业名称、当年申请的数字经济相关发明数量、当年授权的数字经济相关发明数量、当年申请的数字经济相关实用新型数量、当年授权的数字经济相关实用新型数量
	各省市各行业数字经济专利	省份、地市、会计年度、国民经济行业代码、国民经济行业名称、当年申请的数字经济相关发明数量、当年授权的数字经济相关发明数量、当年申请的数字经济相关实用新型数量、当年授权的数字经济相关实用新型数量

	上市公司数字经济专利	股票代码、会计年度、公司类型、上市前后、当年独立申请的数字经济发明数量、当年独立授权的数字经济发明数量、当年独立申请的数字经济实用新型数量、当年独立授权的数字经济实用新型数量、当年联合申请的数字经济发明数量、当年联合授权的数字经济发明数量、当年联合申请的数字经济实用新型数量、当年联合授权的数字经济实用新型数量
	上市公司数字经济专利引用基本信息	股票代码、专利类型、被引用专利号、被引用专利申请日、专利所有人、专利引用数量、非专利引用数量、被引用次数
	上市公司数字经济专利引用数量	股票代码、专利类型、被引用专利号、被引用专利申请日、引用年度、公司类型、发明类型、各年被引用次数、剔除自引用的各年被引用次数、各年累计被引用次数、剔除自引用的各年累计被引用次数
数字产业	各省市软件和信息 技术服务业主要经济 指标	地区、年份、软件业务收入、软件产品收入、信息技术服务收入、信息安全收入、嵌入式系统软件收入、软件业务出口
	各省市信息传输、 计算机服务和软件 业法人单位数	地区、年份、法人单位数、信息传输、计算机服务和软件业法人单位数
	各省市信息传输、 计算机服务和软件 业固定资产投资	地区、年份、固定资产投资(不含农户)、信息传输、计算机服务和软件业固定资产投资(不含农户)
	各地级市信息传 输、计算机服务和 软件业从业人员数	省份名称、城市名称、会计年度、信息传输，计算机服务和软件业从业人员数
数字平台	各省互联网发展主 要指标	省份、年份、网站数量、互联网上网人数、网页数、互联网拨号用户、互联网宽带接入用户、城市宽带接入用户、农村宽带接入用户、移动互联网用户、互联网宽带接入端口、移动互联网接入流量、域名数量、其中 CN 域名数量、其中中国域名数量、各省 IPv4 比例
	各省电信基础设施 情况	地区、年份、固定长途电话交换机容量、局用交换机容量、移动电话交换机容量、光缆线路长度、长途光缆线路长度
	各省电信通信服务 水平	地区、年份、电话普及率(包括移动电话)、移动电话普及率、每千人拥有公用电话数、开通互联网宽带业务的行政村比重(%)、互联网普及率(%)
	各地级市电信产业 发展情况	省份名称、城市名称、会计年度、电信业务收入、移动电话年末用户数、互联网宽带接入用户数

5.数据范围

数据模块	数据起始时间
数字创新	1985 年
数字产业	2003 年
数字平台	1998 年

(八) 工业企业专利数据库-IIED

1. 数据库简介

工业企业专利数据库 (*Innovation of Industrial Enterprise Database, IIED*) 是基于国家统计局的工业企业数据匹配国家知识产权局专利数据整理而成的专业数据库。*IIED* 分为四张表格, 分别为工业企业历年专利情况、工业企业专利情况 (1998-2013)、工业企业基本信息和工业企业绿色专利。其中, 工业企业历年专利情况提供了 1985 年以来工业企业的所有专利信息; 工业企业专利情况 (1998-2013) 则完全根据工业企业 1993-2013 年名单进行专利的一一匹配; 工业企业基本信息提供了来自国家企业信用信息公示系统的基本工商信息; 工业企业绿色专利提供了专利分类号符合《国际专利分类表》中绿色专利定义的工业企业专利情况。*IIED* 不仅考虑了发明、实用新型和外观设计三类专利的历年申请及授权情况, 其中还区分了诸如专利独立申请还是联合申请等重要信息, 方便学者进行深度研究。

2. 数据库特色

- 全面的工业企业数据

IIED 不仅统计了 1998-2013 年间工业企业的专利申请及授权情况, 还统计了这些工业企业 1985 至今所有的专利统计情况。

- 唯一工业企业法人单位

由于工业企业数据本身在不同年度存在样本匹配混乱、指标存在缺失等问题, 因此, 参照 Brandt 等(2012)的做法, *IIED* 对不同年度的工业企业名称进行了名称统一, 在不同年度提供了统一法人名称字段。

3. 典型数据举例

工业企业专利情况 (1998-2013)

工业企业法人名称	年份	当年独立申请的发明数量	当年独立申请的实用新型数量	当年独立申请的外观设计数量	当年独立授权的发明数量	当年独立授权的实用新型数量	当年独立授权的外观设计数量
中国重汽集团青岛重工有限公司	2011	1	4	1	0	24	1
青岛啤酒股份有限公司	2011	3	0	4	1	2	2
山东五征集团有限公司	2011	2	7	6	1	29	23
兖矿集团有限公司	2011	8	21	0	3	4	0
中国一拖集团有限公司	2011	8	59	4	7	94	16

4.数据库内容

数据模块	数据表	数据字段
工业企业历年专利情况	工业企业历年专利情况	工业企业法人名称、年份、当年独立申请的发明数量、当年独立申请的实用新型数量、当年独立申请的外观设计数量、当年联合申请的发明数量、当年联合申请的实用新型数量、当年联合申请的外观设计数量、当年申请专利总数、当年独立授权的发明数量、当年独立授权的实用新型数量、当年独立授权的外观设计数量、当年联合授权的发明数量、当年联合授权的实用新型数量、当年联合授权的外观设计数量、当年授权专利总数
工业企业专利情况 (1998-2013)	工业企业专利情况 (1998-2013)	法人代码、企业名称、工业企业法人名称、年份、当年独立申请的发明数量、当年独立申请的实用新型数量、当年独立申请的外观设计数量、当年联合申请的发明数量、当年联合申请的实用新型数量、当年联合申请的外观设计数量、当年申请专利总数、当年独立授权的发明数量、当年独立授权的实用新型数量、当年独立授权的外观设计数量、当年联合授权的发明数量、当年联合授权的实用新型数量、当年联合授权的外观设计数量、当年授权专利总数
工业企业绿色专利	工业企业绿色专利	工业企业法人名称、年份、当年独立申请的绿色发明数量、当年独立申请的绿色实用新型数量、当年联合申请的绿色发明数量、当年联合申请的绿色实用新型数量、当年独立授权的绿色发明数量、当年独立授权的绿色实用新型数量、当年联合授权的绿色发明数量、当年联合授权的绿色实用新型数量、当年绿色专利申请总数、当年绿色专利授权总数
工业企业基本信息	工业企业基本信息	工业企业法人名称、登记状态、法定代表人、注册资本、成立日期、核准日期、所属省份、所属城市、所属区县、参保人数、企业类型、所属行业、曾用名、企业地址、经营范围

5.数据范围

数据模块	数据起始时间
工业企业历年专利情况	1985 年
工业企业专利情况 (1998-2013)	1998 年
工业企业绿色专利	1985 年
工业企业基本信息	截面数据

（九）电子商务研究数据库-ECRD

1.数据库简介

电子商务研究数据库 (*E-commerce Research Database, ECRD*) 是在数字经济大发展背景下而开发的专业数据库, *ECRD* 主要整理了与电子商务相关的数据, 涵盖了电子商务进农村综合示范县、跨境电子商务企业、电子商务交易额、电子商务指数等多个方面的指标, 从多个维度体现了我国目前电子商务的整体情况。希望 *ECRD* 的推出能为学者研究提供更多数据支持。

2.数据库特色

- 数据全面

ECRD 包括了电子商务研究多个方面的数据, 有助于减少学者的数据搜集时间和成本。

3.典型数据举例

电子商务指数

年份	总指数绝对量	总指数同比 (%)	企业成长指数绝对量	企业成长指数同比 (%)	渗透指数绝对量	渗透指数同比 (%)	消费者参与指数绝对量	消费者参与指数同比 (%)	环境指数绝对量	环境指数同比 (%)
2017 第一季度	126.8	8.5	107.9	2.4	152.7	17.1	128.7	7.4	118.8	5.8
2017 第二季度	130.7	6.3	108.9	2	164.4	10.6	130.4	6.2	120	5.5
2017 第三季度	129.9	5.2	109.1	1.9	159.7	9.4	130.2	3.4	122.1	5.5
2017 第四季度	135.3	6.2	110.9	2.3	174.6	13.1	133.2	2.5	123.9	5
2018 第一季度	135.8	7.1	110.1	2	175.5	14.9	134.5	4.5	124.6	4.9
2018 第二季度	139.3	6.6	110.9	1.8	185.9	13.1	136.5	4.7	125.4	4.5
2018 第三季度	138	6.3	111.3	2	179.5	12.4	136.1	4.6	126.8	3.9
2018 第四季度	140.5	3.8	112.6	1.5	186.9	7.1	136	2.1	128.2	3.5

4.数据库内容

数据模块	数据表	数据字段
电子商务综合发展	分地区的企业信息化基本情况、应用情况及电子商务应用情况	省份、年份、企业数 (个)、使用计算机的企业 (个)、使用计算机的企业比重 (%)、使用互联网的企业 (个)、使用互联网的企业比重 (%)、有网站的企业 (个)、有网站的企业比重 (%)、采用信息化管理的企业 (个)、采用信息化管理的企业比重 (%)、通过互联网开展有关生产经营活动的企业 (个)、通过互联网开展有关生产经营活动的企业比重 (%)、通过互联网进行宣传和推广的企业 (个)、通过互联网进行宣传和推广的

数据模块	数据表	数据字段
		企业比重 (%)、有电子商务交易的企业 (个)、有电子商务交易的企业比重 (%)、有电子商务销售的企业 (个)、有电子商务销售的企业比重 (%)、有电子商务采购的企业 (个)、有电子商务采购的企业比重 (%)、电子商务销售额 (亿元)、电子商务采购额 (亿元)
	分行业的企业信息化基本情况、应用情况及电子商务应用情况	项目、年份、企业数 (个)、使用计算机的企业 (个)、使用计算机的企业比重 (%)、使用互联网的企业 (个)、使用互联网的企业比重 (%)、有网站的企业 (个)、有网站的企业比重 (%)、采用信息化管理的企业 (个)、采用信息化管理的企业比重 (%)、通过互联网开展有关生产经营活动的企业 (个)、通过互联网开展有关生产经营活动的企业比重 (%)、通过互联网进行宣传和推广的企业 (个)、通过互联网进行宣传和推广的企业比重 (%)、有电子商务交易的企业 (个)、有电子商务交易的企业比重 (%)、有电子商务销售的企业 (个)、有电子商务销售的企业比重 (%)、有电子商务采购的企业 (个)、有电子商务采购的企业比重 (%)、电子商务销售额 (亿元)、电子商务采购额 (亿元)
	通信快递业	年份、移动互联网用户数绝对值 (亿户)、移动互联网用户数同比 (%)、5G 移动电话用户数绝对值 (亿户)、5G 移动电话用户数同比 (%)、固定宽带用户数绝对值 (亿户)、固定宽带用户数同比 (%)、户均移动互联网接入流量绝对值 (GB)、户均移动互联网接入流量同比 (%)、快递业务量绝对值 (亿件)、快递业务量同比 (%)
	电子商务指数	年份、总指数绝对量、总指数同比 (%)、企业成长指数绝对量、企业成长指数同比 (%)、渗透指数绝对量、渗透指数同比 (%)、消费者参与指数绝对量、消费者参与指数同比 (%)、环境指数绝对量、环境指数同比 (%)
电子商务示范地区	国家电子商务示范基地	省份、基地名称、批次、发布日期
	省级电子商务示范基地	省份、基地名称、发布日期
	电子商务进农村综合示范县	年份、示范区县、区县代码、所属省份、所属地级市、地级市代码
电子商务交易情况	电子商务交易额	年份、全国电子商务交易额绝对值 (万亿元)、全国电子商务交易额同比 (%)、全国网上零售额绝对值 (亿元)、全国网上零售额同比 (%)、全国实物商品网上零售额绝对值 (亿元)、全国实物商品网上零售额同比 (%)、实物商品网上零售额占社会消费品零售总额的比重 (%)、实物商品网上零售额对社会消费品零售总额增长的贡献率 (%)、重点监测在线餐饮销售额同比 (%)、重点监测在线旅游销售额同比 (%)、农村网络零售额同比 (%)、全国农产品网络零售额同比 (%)、跨境电商进出口交易额同比增速 (%)、跨境电商进口交易额同比增速 (%)、跨境电商出口交易额同比增速 (%)
	重点监测按卖家所在地零售额占比同比情况	年份、地区、重点监测按卖家所在地分网络零售额占比 (%)、重点监测按卖家所在地分网络零售额同比 (%)、重点监测按卖家所在地分实物网络零售额占比 (%)、重点监测按卖家所在地分实物网络零售额同比 (%)
	重点监测分商品网络零售额	年份、指标、重点监测分商品网络零售额占比 (%)、重点监测分商品网络零售额同比 (%)

数据模块	数据表	数据字段
	占比同比情况	
跨境电子商务企业	跨境电子商务企业	企业名称、统一社会信用代码、企业经营类别、跨境电子商务企业类型、信用等级、法定代表人、登记状态、注册资本、实缴资本、成立日期、企业类型、所属行业、人员规模、所属省份、所属地区、注册地址、参保人数、登记机关、经营范围

5.数据范围

数据表	数据起始时间
分地区的企业信息化基本情况、应用情况及电子商务应用情况	2012 年
分行业的企业信息化基本情况、应用情况及电子商务应用情况	2012 年
通信快递业	2017 年
电子商务指数	2015 年
国家电子商务示范基地	截面数据
省级电子商务示范基地	截面数据
电子商务进农村综合示范县	2014 年
电子商务交易额	2017 年
重点监测按卖家所在地零售额占比同比情况	2017 年
重点监测分商品网络零售额占比同比情况	2017 年
跨境电子商务企业	截面数据

对外经济研究系列

(一) 中国对外直接投资数据库-OFDI

1.数据库简介

中国对外直接投资数据库 (*Outward Foreign Direct Investment Database, OFDI*) 是基于商务部统计公报以及国内对外投资相关新闻报道等信息整理而成的专业数据库。*OFDI* 数据库包括中国境外投资企业、对外直接投资事件、对外直接投资情况 (分国家地区)、对外直接投资情况 (分省市) 四种类型的表格, 不仅分年度展示了中国对世界各地和国家的投资情况, 还列明了重要投资事件的具体信息, 包括投资年份、投资方、交易方、投资事件、投资金额、持股比例、投资方所属行业、交易方所属行业、投资国家或地区等重要字段。

2.数据库特色

- 数据来源权威

OFDI 基于政府部门或公益性组织等统计数据, 来源权威可靠。

- 丰富的统计数据 and 事件信息

OFDI 不仅包括一定时期内中国对各国家和地区直接投资的总体数据, 还列明了该时期内的重要事件具体信息, 为学术研究提供了重要数据支持。

3.典型数据举例

对外直接投资事件

年份	月份	投资方	交易方	投资事件	投资金额	持股比例	投资国家
2016	1	海信集团有限公司	欧洲杯	海信将成为 2016 年欧洲杯的顶级赞助商			法国
2016	1	航天科技控股集团股份有限公司	Hiwinglux S.A.	向益圣国际有限公司发行股份购买 Hiwinglux S.A.100%股权	28.78	100%	美国
2016	1	航天科技控股集团股份有限公司	Naviligh	以现金购买国新国际投资有限公司持有的 Navilight S.a.r.l.股权	21.1	100%	美国
2016	1	海尔集团	General Electric	拟以 54 亿美元购买通用电气及其子公司所持有家电业务资产	5400		美国
2016	1	两家中资公司	阿尔及利亚政府	签约成立合资公司, 共同兴建阿尔及利亚境内最大港口	3300		阿尔及利亚

4.数据库内容

数据模块	数据表	数据字段
中国境外投资企业	中国境外投资企业	境外投资企业 (机构) 名称、境内投资者名称、统计日期、投资国家 (地区)
对外直接投资事件	对外直接投资事件	年份、月份、投资方、交易方、投资事件、投资金额、持股比例、投资方所属行业、交易方所属行业、投资国家 (地区)

数据模块	数据表	数据字段
对外直接投资情况 (分国家地区)	对外直接投资流量情况 (分国家地区)	国家(地区)、序号、年份、投资金额
	对外直接投资存量情况 (分国家地区)	国家(地区)、序号、年份、投资金额
对外直接投资情况 (分省市)	对外直接投资流量情况 (分省市)	地区、序号、年份、投资金额
	对外直接投资存量情况 (分省市)	地区、序号、年份、投资金额

5.数据范围

数据模块	数据起始时间
中国境外投资企业	截面数据
对外直接投资事件	2000年
对外直接投资情况(分国家地区)	2003年
对外直接投资情况(分省市)	2003年

（二）一带一路研究数据库-BRRD

1.数据库简介

一带一路研究数据库 (*The Belt and Road Research Database, BRRD*) 是根据我国提出的“一带一路”发展战略, 充分调查和研究了一带一路沿线 65 个国家的政治、经济、人口、环境、社会等信息开发而成的专业数据库。为了给学者提供准确、详细、权威的数据字段, *BRRD* 借鉴了多方数据, 例如, 世界银行、国际货币基金组织、联合国、国际劳工组织等多个机构的独有数据, 涵盖了财政税收、对外贸易、发展援助、国际收支、基础设施、教育教学、经济金融、企业经营调查、人工就业、土地农业、外部债务、医疗卫生、资源环境等 13 个重要方面。*BRRD* 根据 65 个国家所在区域分为七个子库, 分别为: 东南亚 11 国、东亚-蒙古、南亚 8 国、中亚 5 国、西亚 17 国、中东欧 16 国、独联体 7 国。

2.数据库特色

- 1000 多个特色指标

BRRD 囊括了 1000 多个特色指标, 从多个维度反映了一带一路各国的基本情况。

- 数据来源权威

BRRD 整合了来自多个渠道的各国数据, 包括世界银行、国际货币基金组织、联合国等。

3.典型数据举例

菲律宾_财政税收

国家名称	会计年度	指标名称	数据值
菲律宾	2015	一般政府最终消费支出占 GDP 百分比	10.96
菲律宾	2015	一般政府最终消费支出 (年增长率)	7.766
菲律宾	2015	企业与税务官员见面的平均次数	1
菲律宾	2015	企业其他应缴税种 (占商业利润的百分比)	13.9

4.数据库内容

数据模块	数据模块
财政税收	企业经营调查
对外贸易	人口就业
发展援助	土地农业
国际收支	外部债务
基础设施	医疗卫生
教育教学	资源环境
经济金融	

5.数据范围

数据模块	数据起始时间
全部模块	1960 年

（三）海外并购数据库-COMA

1.数据库简介

海外并购数据库 (*Overseas M&A Database, COMA*) 是基于公司公告、新闻报道、商务部官方网站等数据建立的专业数据库。*COMA* 包括 1988 年以来中国公司 (包括上市公司和非上市公司) 进行海外并购的事件数据, 主要包括四大表格: 并购事件、并购方、被并购方以及被并购方财务信息。

2.数据库特色

- 整合海外并购事件, 提供丰富字段

COMA 提供了较为全面的海外并购事件信息, 并囊括研究所用到的常用字段, 弥补了该领域研究数据匮乏的缺陷。

3.典型数据举例

海外并购事件

并购 ID	首次公告日	并购生效日	转让股权比例 (%)	转让后并购方持股比例 (%)	转让股权估值 (美元)	现金支付比例 (%)
331144	1992/12/3	1992/12/3	30.1	30.1	21071000	100
331283	1998/3/4	1998/7/16	85	85	20000000	100
331422	2001/9/12	2001/10/1	29	29	17372000	46.05
333349	2013/5/29	2013/9/26	100	100	4752384000	100

4.数据库内容

数据模块	数据表	数据字段
海外并购	并购事件	并购 ID、并购方名称、被并购方名称、首次公告日、并购生效日、转让股权比例 (%)、转让后并购方持股比例 (%)、转让股权估值、货币类型、现金支付比例 (%)、竞买方数量、完成状态
	并购方	并购 ID、并购方名称、并购方英文名称、并购方企业性质、并购方所在交易所、并购方股票代码、并购方上市状态、并购方所在地区、并购方所属行业
	被并购方	并购 ID、被并购方名称、被并购方所在国、被并购方所属行业、被并购方企业性质
	被并购方财务信息	并购 ID、被并购方名称、首次公告日、并购生效日、货币类型、被并购方资产负债率、被并购方净销售额、被并购方 EBIT、被并购方净资产、被并购方总资产、被并购方 EPS、被并购方并购公告后一周股价、被并购方并购公告后 180 天股价、被并购方并购公告后四周股价、被并购方并购公告后 90 天股价、被并购方并购公告当天股价、被并购方并购公告前一天股价、被并购方并购公告前一周股价、被并购方并购公告前四周股价

5.数据范围

数据模块	数据起始时间
全部模块	1988 年

(四) 中国贸易外经统计数据库-TEES

1. 数据库简介

中国贸易外经统计数据库 (China Trade and External Economic Statistical Database, **TEES**) 是基于国家统计局数据整理而成的一个用于反映中国国内贸易、对外经济贸易和旅游业发展情况的特色数据库。**TEES** 主要包括综合、消费品市场、批发业和零售业、住宿和餐饮业、国际收支、对外贸易、利用外资、对外投资与经济合作、旅游等 9 个大模块。

2. 数据库特色

- 合理整合数据

TEES 将关联度较大的表合并整理成一张表格, 极大地节省了数据使用者的下载时间。

3. 典型数据举例

批发和零售业_各地区限额以上批发业企业基本情况

地区	年份	批发业法人单位数(个)	批发业年末从业人数(人)	国有控股批发业法人单位数(个)	国有控股批发业年末从业人数(人)
河北	2009	805	82140	167	42300
山西	2009	668	94393	248	62883
内蒙古	2009	461	42829	92	18080

4. 数据库内容

数据模块	数据字段
综合	国内生产总值指数(不变价, 上年=100); 国内生产总值构成(按当年价格计算); 三次产业贡献率(按不变价格计算)……
消费品市场	历年社会消费品零售总额; 各地区社会消费品零售总额; 各地区社会消费品零售总额构成; 分地区网上零售额; 亿元以上商品交易市场基本情况; 亿元以上商品交易市场摊位分类情况; 各地区亿元以上商品交易市场基本情况
批发和零售业	限额以上批发和零售业企业基本情况; 各地区限额以上批发业企业基本情况……
住宿和餐饮业	限额以上住宿和餐饮业企业基本情况; 各地区限额以上住宿业企业基本情况; 各地区限额以上住宿业企业基本情况(按登记注册类型分)……
国际收支	国际收支概况; 国际收支平衡表_2015年以后; 国际收支平衡表_2014年以前……
对外贸易	进出口额(按人民币计算); 进出口额(按美元计算); 进出口额指数(按美元计算)……
利用外资	利用外资情况; 分方式外商直接投资; 外商直接投资行业分布; 分国别(地区)外商直接投资; 历年外商投资企业登记概况……
对外投资与经济合作	按主要国别(地区)分对外直接投资; 按行业分对外直接投资; 对外经济合作概况; 分国别(地区)对外承包工程合同数量及金额; 分国别(地区)对外承包工程营业额……
旅游	历年入境游客; 历年入境游客增长速度; 历年入境过夜游客和国际旅游收入; 入境外国游客分组构成; 按国别分入境外国游客……

5. 数据范围

数据模块	数据起始时间
全部模块	1978年

（五）中国海关贸易数据库-CCTD

1.数据库简介

中国海关贸易数据库（*Chinese Customs Trade Database, CCTD*）是基于中华人民共和国海关总署数据整理而形成的一个用于反映国家经济发展规律及发展趋势的数据库，可以帮助研究者获得全面的进出口贸易数据，掌握国际贸易动向。

CCTD 自 2014 年起以月度的形式，根据进出口商品的类型、构成情况、目的地与货源地、贸易方式、外商投资企业、特定地区等不同类别进行表格设定及发布金额总值情况，包括有：进出口商品年度总值表、进出口商品月度总值表、进出口商品国别（地区）总值表、进出口商品构成表、进出口商品类章总值表、进出口商品贸易方式总值表、出口商品贸易方式企业性质总值表、进口商品贸易方式企业性质总值表、进出口商品收发货人所在地总值表、进出口商品境内目的地/货源地总值表、进出口商品关别总值表、特定地区进出口总值表、外商投资企业进出口总值表、出口主要商品量值表、进口主要商品量值表、对部分国家（地区）出口商品类章金额表、自部分国家（地区）进口商品类章金额表、部分出口商品主要贸易方式量值表、部分进口商品主要贸易方式量值表，共计 19 张表格。

2.数据库特色

- 数据来源权威

CCTD 数据来源于中华人民共和国海关总署，准确地整理数据信息，直观体现了中国海关方面的主要统计数据。

3.典型数据举例

进出口商品总值表_进出口商品年度总值表

年度	进出口额（亿元人民币）	出口额	进口额	贸易差额
2012	244160	129359	114801	14558
2013	258169	137131	121037	16094
2014	264242	143884	120358	23526
2015	245503	141167	104336	36831

4.数据库内容

数据模块	数据字段
进出口商品年度总值表	年度、单位、进出口额、出口额、进口额、贸易差额、进出口比去年同期±%、出口比去年同期±%、进口比去年同期±%
进出口商品月度总值表	日期、单位、进出口额、出口额、进口额、贸易差额
进出口商品国别总值表	进口原产国（地）出口最终目的国（地）、日期、单位、进出口额、出口额、进口额、进出口累计比去年同期±%、出口累计比去年同期±%、进口累计比去年同期±%
进出口商品构成表	商品构成（按 SITC 分类）、日期、单位、当月出口额、当月进口额、出口累计比去年同期±%、进口累计比去年同期±%
进出口商品类章总值表	类章、日期、单位、当月出口额、当月进口额、出口累计比去年同期±%、进口累计比去年同期±%

进出口商品贸易方式总值表	贸易方式、日期、单位、当月进出口额、当月出口额、当月进口额
出口商品贸易方式企业性质总值表	贸易方式、日期、单位、合计金额、国有企业金额、外商投资企业金额小计、中外合作金额、中外合资金额、外商独资金额、私营企业金额、除国有企业、外商投资企业、中外合作企业、中外合资企业、外商独资企业、私营企业外，其他出口商品贸易金额
进口商品贸易方式企业性质总值表	贸易方式、日期、合计金额、国有企业金额、外商投资企业金额小计、中外合作金额、中外合资金额、外商独资金额、私营企业金额、其他金额
进出口商品收发货人所在地总值表	经营单位所在地、日期、单位、当月进出口额、当月出口额、当月进口额、进出口累计比去年同期±%、出口累计比去年同期±%、进口累计比去年同期±%
进出口商品境内目的地@货源地总值表	境内目的地/货源地、日期、单位、当月进出口额、当月出口额、当月进口额、进出口累计比去年同期±%、出口累计比去年同期±%、进口累计比去年同期±%
进出口商品关别总值表	关别、日期、单位、当月进出口额、当月出口额、当月进口额、进出口累计比去年同期±%、出口累计比去年同期±%、进口累计比去年同期±%
特定地区进出口总值表	特定经济地区、日期、单位、当月进出口额、当月出口额、当月进口额、进出口累计比去年同期±%、出口累计比去年同期±%、进口累计比去年同期±%
外商投资企业进出口总值表	外商投资企业、日期、单位、当月进出口额、当月出口额、当月进口额、进出口累计比去年同期±%、出口累计比去年同期±%、进口累计比去年同期±%
出口主要商品量值表	商品名称、日期、单位、计量单位、当月数量、当月金额、数量比去年同期±%、金额比去年同期±%、数量累计比去年同期±%、金额累计比去年同期±%
进口主要商品量值表	商品名称、日期、单位、计量单位、当月数量、当月金额、数量比去年同期±%、金额比去年同期±%、数量累计比去年同期±%、金额累计比去年同期±%
对部分国家(地区)出口商品类章金额表	类章、日期、单位、缅甸、中国香港、印度、印度尼西亚、伊朗、日本、中国澳门、马来西亚、阿曼、巴基斯坦、菲律宾、沙特阿拉伯、新加坡、韩国、泰国、土耳其、阿拉伯联合酋长国、越南、中国台湾、哈萨克斯坦、南非、欧洲联盟、比利时、丹麦、英国、德国、法国、意大利、荷兰、西班牙、奥地利、芬兰、瑞典、罗马尼亚、瑞士、俄罗斯联邦、乌克兰、阿根廷、巴西、智利、加拿大、美国、澳大利亚、新西兰
自部分国家(地区)进口商品类章金额表	类章、日期、单位、缅甸、中国香港、印度、印度尼西亚、伊朗、日本、中国澳门、马来西亚、阿曼、巴基斯坦、菲律宾、沙特阿拉伯、新加坡、韩国、泰国、土耳其、阿拉伯联合酋长国、越南、中国台湾、哈萨克斯坦、南非、欧洲联盟、比利时、丹麦、英国、德国、法国、意大利、荷兰、西班牙、奥地利、芬兰、瑞典、罗马尼亚、瑞士、俄罗斯联邦、乌克兰、阿根廷、巴西、智利、加拿大、美国、澳大利亚、新西兰
部分出口商品主要贸易方式量值表	商品、日期、单位、计量单位、一般贸易当月数量额、一般贸易当月金额、加工贸易当月数量额、加工贸易当月金额、保税物流当月数量额、保税物流当月金额、其他贸易当月数量额、其他贸易当月金额
部分进口商品主要贸易方式量值表	商品、日期、单位、计量单位、一般贸易当月数量额、一般贸易当月金额、加工贸易当月数量额、加工贸易当月金额、保税物流当月数量额、保税物流当月金额、在其他贸易的贸易方式下部分进口商品的当月数量额、其他贸易当月金额

5.数据范围

数据模块	数据起始时间
全部模块	2014年

（六）外国在华专利研究数据库-FCID

1.数据库简介

外国在华专利研究数据库（*Innovation of Foreign Countries Research Database*, **FCID**）是结合国家知识产权局、中国科技统计年鉴和 WIPO 知识产权统计数据中心专利数据，基于不同外国国家在中国的专利申请和授权情况建立的专业数据库。**FCID** 包括外国在华专利情况、外国在华专利统计和按技术分类的外国在华专利三个模块。其中，外国在华专利情况模块统计了外国在华专利申请与授权情况、外国在华专利申请后授权情况和外国在华专利授权年限情况，此外，**FCID** 还提供了专利分类号符合《国际专利分类表》中绿色专利定义、以及符合国家统计局 2021 年发布的《数字经济及其核心产业统计分类》中数字经济专利定义的外国在华专利情况；外国在华专利统计模块统计了按国别(地区)分国外三种专利申请受理、授权及有效数；按技术分类的外国在华专利模块统计了按技术分类的外国在华发明申请和授权情况。

2.数据库特色

- 多维度展示外国在华专利情况

FCID 统计了外国在华历年专利申请与授权的情况、专利申请后的授权情况、各类专利的授权年限，以及外国在华的绿色专利及数字经济专利等多个数据。

- 分模块展示不同数据来源

FCID 分模块统计了三个数据来源的外国在华专利情况，其中，外国在华专利情况模块来自国家知识产权局专利数据，外国在华专利统计模块来自中国科技统计年鉴数据，按技术分类的外国在华专利模块来自 WIPO 知识产权统计数据中心数据。

3. 典型数据举例

（1）外国在华专利申请与授权情况

国家(地区)	会计年度	外国在华当年独立申请的发明数量	外国在华当年独立申请的实用新型数量	外国在华当年独立申请的外观设计数量	外国在华当年联合申请的发明数量	外国在华当年联合申请的实用新型数量	外国在华当年联合申请的外观设计数量	外国在华当年申请专利总数
阿根廷	2012	3	0	0	6	0	0	9
爱尔兰	2012	205	0	4	8	0	0	217
爱沙尼亚	2012	5	0	3	0	0	0	8
安道尔	2012	1	0	0	0	0	0	1

（2）外国在华绿色专利

国家(地区)	会计年度	外国在华当年独立申请的绿色发明数量	外国在华当年独立申请的绿色实用新型数量	外国在华当年联合申请的绿色发明数量	外国在华当年联合申请的绿色实用新型数量	外国在华当年申请绿色专利总数
--------	------	-------------------	---------------------	-------------------	---------------------	----------------

阿根廷	2012	0	0	1	0	1
爱尔兰	2012	12	0	1	0	13
爱沙尼亚	2012	2	0	0	0	2
奥地利	2012	30	0	1	0	31
澳大利亚	2012	40	1	3	0	44
巴巴多斯	2012	2	0	0	0	2

(3) 按国别(地区)分国外三种专利申请受理

国家(地区)	年份	合计(件)	发明(件)	实用新型(件)	外观设计(件)
百慕大群岛	2012	129	59	31	39
巴西	2012	159	107	4	48
加拿大	2012	1352	1111	54	187
开曼群岛	2012	805	712	40	53
智利	2012	21	21		
哥伦比亚	2012	10	9	1	
克罗地亚	2012	5	4	1	

4. 数据库内容

数据模块	数据表	数据字段
外国在华专利情况	外国在华专利申请与授权情况	国家(地区)、会计年度、外国在华当年独立申请的发明数量、外国在华当年独立申请的实用新型数量、外国在华当年独立申请的外观设计数量、外国在华当年联合申请的发明数量、外国在华当年联合申请的实用新型数量、外国在华当年联合申请的外观设计数量、外国在华当年申请专利总数、外国在华当年独立授权的发明数量、外国在华当年独立授权的实用新型数量、外国在华当年独立授权的外观设计数量、外国在华当年联合授权的发明数量、外国在华当年联合授权的实用新型数量、外国在华当年联合授权的外观设计数量、外国在华当年授权专利总数
	外国在华专利申请后授权情况	国家(地区)、会计年度、更新日期、外国在华专利申请个数、外国在华专利申请_授权个数、外国在华专利申请_当年授权个数、外国在华专利申请_第二年授权个数、外国在华专利申请_第三年授权个数、外国在华专利申请_第四年授权个数、外国在华发明申请个数、外国在华发明申请_授权个数、外国在华发明申请_当年授权个数、外国在华发明申请_第二年授权个数、外国在华发明申请_第三年授权个数、外国在华发明申请_第四年授权个数、外国在华实用新型申请个数、外国在华实用新型申请_授权个数、外国在华实用新型申请_当年授权个数、外国在华实用新型申请_第二年授权个数、外国在华实用新型申请_第三年授权个数、外国在华实用新型申请_第四年授权个数、外国在华外观设计申请个数、外国在华外观设计申请_授权个数、外国在华外观设计申请_当年授权个数、外国在华外观设计申请_第二年授权个数、外

数据模块	数据表	数据字段
		国在华外观设计申请_第三年授权个数、外国在华外观设计申请_第四年授权个数
	外国在华专利授权年限情况	国家(地区)、会计年度、外国在华专利授权、外国在华专利授权_平均年限、外国在华发明授权、外国在华发明授权_平均年限、外国在华发明授权_公开平均年限、外国在华实用新型授权、外国在华实用新型授权_平均年限、外国在华外观设计授权、外国在华外观设计授权_平均年限
	外国在华绿色专利	国家(地区)、会计年度、外国在华当年独立申请的绿色发明数量、外国在华当年独立申请的绿色实用新型数量、外国在华当年联合申请的绿色发明数量、外国在华当年联合申请的绿色实用新型数量、外国在华当年申请绿色专利总数、外国在华当年独立授权的绿色发明数量、外国在华当年独立授权的绿色实用新型数量、外国在华当年联合授权的绿色发明数量、外国在华当年联合授权的绿色实用新型数量、外国在华当年授权绿色专利总数
	外国在华数字经济专利	国家(地区)、会计年度、外国在华当年独立申请的数字经济发明数量、外国在华当年独立申请的数字经济实用新型数量、外国在华当年联合申请的数字经济发明数量、外国在华当年联合申请的数字经济实用新型数量、外国在华当年申请数字经济专利总数、外国在华当年独立授权的数字经济发明数量、外国在华当年独立授权的数字经济实用新型数量、外国在华当年联合授权的数字经济发明数量、外国在华当年联合授权的数字经济实用新型数量、外国在华当年授权数字经济专利总数
外国在华专利统计	按国别(地区)分国外三种专利申请受理	国家(地区)、年份、合计(件)、发明(件)、实用新型(件)、外观设计(件)
	按国别(地区)分国外三种专利授权	国家(地区)、年份、合计(件)、发明(件)、实用新型(件)、外观设计(件)
	按国别(地区)分国外三种专利有效数	国家(地区)、年份、合计(件)、发明(件)、实用新型(件)、外观设计(件)
按技术分类的外国在华专利	按技术分类的外国在华发明申请情况	申请地区、主管地区、年份、电气机械、电气装置、电能、音像技术、电信、数字通信、基本通信过程、计算机技术、用于管理的信息技术方法、半导体、光学、测量、生物材料分析、控制、医疗技术、有机精细化学、生物技术、药品、高分子化学、聚合物、食品化学、基础材料化学、材料、冶金、表面技术、涂层、微观结构技术和超微技术、化学工程、环境技术、搬运、机床、发动机、泵、涡轮机、纺织机械和纸张机械、其他专用机械、热工过程和器具、机械元件、运输、家具、游戏、其他消费品、土木工程、未知
	按技术分类的外国在华发明授权情况	申请地区、主管地区、年份、电气机械、电气装置、电能、音像技术、电信、数字通信、基本通信过程、计算机技术、用于管理的信息技术方法、半导体、光学、测量、生物材料分析、控制、医疗技术、有机

数据模块	数据表	数据字段
		精细化学、生物技术、药品、高分子化学、聚合物、食品化学、基础材料化学、材料、冶金、表面技术、涂层、微观结构技术和超微技术、化学工程、环境技术、搬运、机床、发动机、泵、涡轮机、纺织机械和纸张机械、其他专用机械、热工过程和器具、机械元件、运输、家具、游戏、其他消费品、土木工程、未知

5.数据范围

数据表	数据起始时间
外国在华专利申请与授权情况	1985 年
外国在华专利申请后授权情况	1985 年
外国在华专利授权年限情况	1985 年
外国在华绿色专利	1985 年
外国在华数字经济专利	1985 年
按国别(地区)分国外三种专利申请受理	1997 年
按国别(地区)分国外三种专利授权	1997 年
按国别(地区)分国外三种专利有效数	2009 年
按技术分类的外国在华发明申请情况	1985 年
按技术分类的外国在华发明授权情况	1993 年

（七）专利国际化研究数据库-PIRD

1.数据库简介

专利国际化研究数据库（*Patent Internationalization Research Database, PIRD*）是结合 WIPO 知识产权统计数据中心专利数据，基于不同专利类型和不同国际专利申请途径建立的专业数据库。

PIRD 包括四个模块的数据：

（1）国际发明专利申请与授权模块提供了中国历年国际发明专利的申请与授权情况，包含直接申请的发明专利申请与授权情况、进入 PCT 国家阶段的发明专利申请与授权情况、按技术分类的发明公布情况和按技术分类的发明授权情况四张数据表。该模块包括自 1980 年以来每年中国在世界各主管局申请和获得授权的国际发明专利的数量情况，以及提出申请的发明专利在各个技术领域的分布情况。

（2）国际实用新型申请与授权模块提供了中国历年国际实用新型的申请与授权情况，包含直接申请的实用新型申请与授权情况、进入 PCT 国家阶段的实用新型申请与授权情况两张数据表。该模块包括自 1980 年以来每年中国在世界各主管局申请和获得授权的国际实用新型的数量情况。

（3）国际外观设计申请与注册模块提供了中国历年国际外观设计的申请与注册情况，包含直接申请的外观设计申请与注册情况、通过海牙体系申请的外观设计申请与注册情况两张数据表，该模块包括自 1980 年以来每年中国在世界各主管局申请和注册的国际外观设计的数量情况。

（4）PCT 申请模块提供了中国向海外提出的 PCT 专利申请情况，包含按申请日开列的 PCT 申请情况、按 IPC 分类开列的 PCT 公布情况两张数据表。

2.数据库特色

- 全面展示不同申请途径的国际专利情况

PIRD 展示了历年我国直接申请、通过 PCT 申请的国际发明专利和实用新型的申请与授权情况，以及直接申请、通过海牙体系进行申请的国际外观设计的申请与注册情况，全面展示了我国在知识产权领域的创新活动和保护水平。

- 提供不同技术领域的中国海外专利布局规模数据

PIRD 统计了历年我国国际发明专利在不同技术领域的申请与授权情况，以及世界知识产权组织公开发布的 PCT 申请在各技术领域的分布情况数据。

4. 典型数据举例

直接申请的发明专利申请与授权情况

主管地区	主管地区代码	来源地	年份	直接申请的发明专利申请	直接申请的发明专利授权
埃及	EG	中国	2012	1	2
埃塞俄比亚	ET	中国	2012		
爱尔兰	IE	中国	2012		
巴拿马	PA	中国	2012	1	

4.数据库内容

数据模块	数据表	数据字段
国际发明专利申请与授权	直接申请的发明专利申请与授权情况	主管地区、主管地区代码、来源地、年份、直接申请的发明专利申请、直接申请的发明专利授权
	进入 PCT 国家阶段的发明专利申请与授权情况	主管地区、主管地区代码、来源地、年份、进入 PCT 国家阶段的发明申请、进入 PCT 国家阶段的发明授权
	按技术分类的发明公布情况	主管地区、主管地区代码、来源地、年份、电气机械、电气装置、电能、音像技术、电信、数字通信、基本通信过程、计算机技术、用于管理的信息技术方法、半导体、光学、测量、生物材料分析、控制、医疗技术、有机精细化学、生物技术、药品、高分子化学、聚合物、食品化学、基础材料化学、材料、冶金等……
	按技术分类的发明授权情况	主管地区、主管地区代码、来源地、年份、电气机械、电气装置、电能、音像技术、电信、数字通信、基本通信过程、计算机技术、用于管理的信息技术方法、半导体、光学、测量、生物材料分析、控制、医疗技术、有机精细化学、生物技术、药品等……
国际实用新型申请与授权	直接申请的实用新型申请与授权情况	主管地区、主管地区代码、来源地、年份、直接申请的实用新型申请、直接申请的实用新型授权
	进入 PCT 国家阶段的实用新型申请与授权情况	主管地区、主管地区代码、来源地、年份、进入 PCT 国家阶段的实用新型申请、进入 PCT 国家阶段的实用新型授权
国际外观设计申请与注册	直接申请的外观设计申请与注册情况	主管地区、主管地区代码、来源地、年份、直接申请的外观设计申请、直接申请的外观设计注册
	通过海牙体系申请的外观设计申请与注册情况	主管地区、主管地区代码、来源地、年份、通过海牙体系进行申请的外观设计申请、通过海牙体系进行申请的外观设计注册
PCT 申请	按申请日开列的 PCT 申请情况	受理地区、受理地区代码、来源地、年份、PCT 申请
	按 IPC 分类开列的 PCT 公布情况	IPC 分类、来源地、来源地代码、年份、PCT 公布

5.数据范围

数据表	数据起始时间
国际发明专利申请与授权	1980 年
国际实用新型申请与授权	1980 年
国际外观设计申请与注册	1980 年
PCT 申请	1995 年

人文社科研究系列

（一）中国文化研究数据库-CCRD

1. 数据库简介

中国语言类别众多，官话、方言、少数民族方言差异巨大，语言的差异体现着文化的差异；宗教的种类、分布，庙宇的数量也在一定程度上体现出各地区人们的宗教信仰程度，以及各宗教在中国的传播状况；中国不同地区因历史、文化和风俗等差异，对于购买彩票的态度也存在差异，彩票文化的地区差异在一定程度上反应了各地冒险精神或赌博偏好。中国文化研究数据库（*Chinese Culture Research Database, CCRD*）通过对中国语言地图信息、中国宗教庙宇信息、中国彩票文化信息的全面考察和整理，为学者进行文化研究提供了详实的数据来源。

中国语言地图信息 包含各省、市、县对应的语言片区，包括官话、汉语方言、少数民族方言的大类、小类与细分信息。

中国宗教庙宇信息 包含中国所有寺庙、道观、清真寺和教堂的多维度信息，如名称、建立日期、负责人姓名、所属地区、是否为国家文物保护单位、具体地址、邮编以及简介等。

中国彩票文化信息 包含各省、直辖市自 2008 年以来的地区不同类型彩票销售额、增长率信息，数据主要来自财政部官网相关内容披露。

2. 数据库特色

- 独特性

CCRD 对中国各地区方言分布、中国宗教庙宇分布进行了综合梳理和总结，对中国彩票文化信息数据进行了整理，对研究中国文化传统和文化遗产具有重要参考价值。

- 完整性

CCRD 中国语言地图数据涵盖了中国现存的所有方言分布信息，中国宗教庙宇信息包含五个宗教所有的庙宇信息，中国彩票文化信息包括 2008 年至今的彩票销售情况。

- 多维性

CCRD 的各个模块都包含详细的字段内容。其中，中国语言地图信息模块将官话、汉语方言、少数民族方言全部进行片区与大小类的细分；中国宗教庙宇信息对基督教与天主教各个教堂的礼拜与弥撒时间都详细标出，对佛教、道教、伊斯兰教、基督教、天主教，各教的建筑状况均做出标识与介绍。

3. 典型数据举例

（1）汉语方言片区

区域代码	省级名称	市级名称	县级名称	官话区/汉语方言区	汉语方言 (大类)	汉语方言 (小类)	汉语方言 (细分)
350182	福建省	福州市	长乐市	汉语方言区	闽语	闽东区	侯官片
350205	福建省	厦门市	海沧区	汉语方言区	闽语	闽南区	泉漳片
350421	福建省	三明市	明溪县	汉语方言区	客家话	客家话	汀州片
350423	福建省	三明市	清流县	汉语方言区	客家话	客家话	汀州片

(2) 佛寺信息

寺庙名称	支派类别	建立日期	所属省份	所属地市	所属区县	景区级别
白居寺	藏语系	公元 1427 年	西藏	日喀则市	江孜县	国家 AAAA 级旅游景区
白马寺	汉语系	公元 68 年	河南省	洛阳市	洛龙区	国家 AAAA 级旅游景区
风穴寺	汉语系	公元 190 年	河南省	平顶山市	汝州市	国家 AAA 级旅游景区

4. 数据库内容

数据模块	数据表	数据字段
中国语言地图信息	官话片区	省级名称、地市级名称、县级名称、区域代码、官话区/汉语方言区/少数民族方言区、官话（大类）、官话（片区）、官话（小片区）
	汉语方言片区	省级名称、地市级名称、县级名称、区域代码、官话区/汉语方言区/少数民族方言区、汉语方言（大类）、汉语方言（小类）、汉语方言（细分）
	少数民族方言片区	省级名称、地市级名称、县级名称、区域代码、官话区/汉语方言区/少数民族方言区、少数民族方言（语族）、少数民族方言（语族大类）、少数民族方言（语族小类）、少数民族方言（语族细分 1）、少数民族方言（语族细分 2）
中国宗教庙宇信息	佛寺	寺庙名称、支派类别、负责人姓名、建立日期、所属省份、所属地级市、所属区县、具体地址、佛寺简介、是否为国内名寺、是否为国家文物保护单位、景区级别
	道观	道观名称、支派类别、负责人姓名、建立日期、所属省份、所属地级市、所属区县、具体地址、道观简介、是否为国内名观、是否为国家文物保护单位、景区级别
	清真寺	清真寺名称、建立日期、所属省份、所属地级市、所属区县、具体地址、清真寺简介、是否为国内名寺、是否为国家文物保护单位、景区级别
	基督教堂	教堂名称、建立日期、所属省份、所属地级市、所属区县、具体地址、礼拜时间、基督教堂简介
	天主教堂	教堂名称、教堂主管、所属省份、所属地级市、所属区县、具体地址、弥撒时间、天主教堂简介、是否为国内朝圣地、是否为国家文物保护单位
中国彩票文化信息	全国各地区彩票销售情况表	地区、年份、月份、福利彩票本月销售额、福利彩票本月销售额比上年同期增长（%）、福利彩票本年累计销售额、福利彩票本年累计销售额比上年同期增长（%）、体育彩票本月销售额、体育彩票本月销售额比上年同期增长（%）、体育彩票本年累计销售额、体育彩票本年累计销售额比上年同期增长（%）、本月合计销售额、本月合计销售额比上年同期增长（%）、本年累计销售额、本年累计销售额比上年同期增长（%）、销量累计排序
	全国各类型彩票销售情况表	类型、年份、月份、本月销售额、上年同期销售额、本月销售额比上年同期增长（%）、该月环比增长（%）、本年累计销售额、上年累计销售额、本年累计销售额比上年同期增长（%）

5. 数据范围

数据模块	数据起始时间
中国语言地图信息	截面数据
中国宗教庙宇信息	截面数据
中国彩票文化信息	2008 年

（二）儒家文化数据库-CFCN

1.数据库简介

儒家文化数据库（*Confucian Culture Database, CFCN*）是基于中国传统儒家文化及由此衍生出的各类反映儒家文化发展情况的历史事件、人物和文物等进行统计整理而成的，反映中国各地方儒家文化发展情况的专业数据库。*CFCN* 查阅了诸多经典权威文献，并结合网络搜索等，从明清进士、书院数量和孔庙遗存三个方面对中国各地区的儒家文化发展情况进行了展示。

2.数据库特色

- 权威和全面的儒家文化数据

CFCN 考察和借鉴了诸多权威书籍文献，并且结合网络搜索等渠道进行补充，提供了该领域研究的重要数据。

- 多维度儒家文化衡量指标

CFCN 包含从进士人数、书院数量和孔庙数量三个维度体现了各地区的儒家文化发展程度，为研究提供了多维度支持。

3.典型数据举例

明清进士人数

省份名称	城市名称	区县名称	明清各省进士人数	明清各市进士人数	明清各县区进士人数
安徽省	安庆市	宿松县	2178	439	26
安徽省	安庆市	太湖县	2178	439	58
安徽省	安庆市	桐城市	2178	439	213

4.数据库内容

数据模块	数据表	数据字段
明清进士	明清进士人数	省份名称、城市名称、区县名称、明清各省进士人数、明清各城市进士人数、明清各县区进士人数
	明朝进士人数	省份名称、城市名称、区县名称、明朝各省进士人数、明朝各城市进士人数、明朝各县区进士人数
	清朝进士人数	省份名称、城市名称、区县名称、清朝各省进士人数、清朝各城市进士人数、清朝各县区进士人数
	明清历代进士人数	朝代名称、省份名称、城市名称、各朝代进士人数、各朝代省份进士人数、各朝代城市进士人数
书院数量	书院数量	省份名称、地级市名称、区县名称、省份书院数量、地级市书院数量、区县书院数量
孔庙遗存	孔庙遗存	省份名称、城市名称、区县名称、孔庙名称、孔庙地址、始建年代、保护级别、经度、纬度

5.数据范围

数据模块	数据起始时间
全部模块	截面数据

（三）中国旅游统计数据库-CTSD

1.数据库简介

中国旅游统计数据库（*Chinese Tourism Statistics Database, CTSD*）是基于中国旅游统计年鉴相关数据而建立的专业数据库，全面反映了这几十年来我国旅游业的发展情况。*CTSD* 囊括了 1978 年以来中国出入境游、旅游景区、旅行社和旅游饭店等重要旅游数据。

2.数据库特色

- 模块重新整合

CTSD 根据研究需求，重新整合了旅游年鉴中的数据，模块设计更为简洁。

3.典型数据举例

国内游_全国国内旅游基本情况

居民类型	会计年度	总人次(亿人次)	出游率(%)	总花费(亿元)	人均花费(元/人·次)
全国总计	2017	50.01	361.6	45660.77	913
城镇居民	2017	36.77	463.7	37673.03	1024.6
农村居民	2017	13.24	224.4	7987.74	603.3

5.数据库内容

数据模块	数据表
入境游	中国入境过夜旅游者人数和国际旅游(外汇)收入的世界排名；入境旅游人数；各月入境旅游人数；各月入境旅游人数(按入境方式分)；各月入境外国游客人数(按国籍分)；入境外国旅游者人数(按国籍分)……
国内游	全国国内旅游基本情况；城镇居民国内游客人次构成(按旅游目的分)；城镇居民国内游客人均每次花费(按旅游目的分)；城镇居民国内游客人次构成(按旅游方式分)……
星级饭店	全国星级饭店基本情况(按经济类型、规模和星级分)；全国星级饭店基本情况(按地区分)；全国各地区星级饭店数(按经济类型分)；全国各地区星级饭店数(按规模分)；全国各地区星级饭店数(按星级分)；星级饭店主要经济指标(按地区分)……
旅行社	旅行社总数；旅行社外联、接待入境旅游者情况；旅行社组织、接待国内游客情况；全国旅行社主要经济指标(按地区分)；主要城市旅行社主要经济指标
旅游景区	旅游景区单位数；旅游景区基本情况；全国旅游景区主要经济指标(按地区分)；主要城市旅游景区主要经济指标
旅游企业	全国旅游企业主要经济指标(按行业分)；全国旅游企业主要经济指标(按地区分)；主要城市旅游企业主要经济指标(按行业分)；主要城市旅游企业主要经济指标(按城市分)；分地区旅游企业资料；分城市旅游企业资料
旅游从业人员及培养	旅游从业人员；旅游院校情况

5.数据范围

数据模块	数据起始时间
全部模块	1999 年

（四）中国科技统计数据库-CSTD

1.数据库简介

中国科技统计数据库（*China Statistical Database on Science and Technology, CSTD*）是为反映我国科技活动情况而进行整理统计的数据库，收录了全国 31 个省、自治区、直辖市及相关部门的科技活动、企业创新活动、高新技术发展、国家科技计划等多方位数据资料。

2.数据库特色

- 数据架构清晰

CSTD 数据列明 10 个模块，结构布局明确，由企业到院校机构等全方面展现了历年国内科技发展在各界的表现情况。

3.典型数据举例

研究与试验发展(R&D)人员

项目	年份	R&D 人员(人)	R&D 人员_女性(人)	R&D 人员_全时人员(人)	R&D 人员_博士毕业(人)	R&D 人员_硕士毕业(人)	R&D 人员_本科毕业(人)
内蒙古	2018	41182	12710	23183	2490	6906	19130
辽宁	2018	153303	45703	99511	16389	28923	70884
吉林	2018	64190	25453	37352	13071	20816	19409

4.数据库内容

数据模块	数据表
综合	研究与试验发展(R&D)人员、全国研究与试验发展(R&D)人员全时当量……
工业企业	规模以上工业企业的科技活动基本情况、按企业规模及登记注册类型分规模以上工业企业数量情况、按行业分规模以上工业企业数量情况……
研究与开发机构	研究与开发机构基本情况、按隶属关系和学科分研究与开发机构 R&D 人员、各地区研究与开发机构 R&D 人员……
高等学校	高等学校科技活动情况、各地区高等学校 R&D 人员……
高技术产业	高技术产业生产经营情况、各地区高技术产业生产经营情况……
企业创新活动	规模(限额)以上企业创新活动总体情况、规模(限额)以上企业产品和工艺创新分布情况、规模(限额)以上企业产品或工艺创新活动类型……
国家科技计划	国家主要科技计划基本情况、国家科技支撑计划除以国家科技攻关计划中央财政拨款、国家重点基础研究发展计划(973 计划)中央财政拨款……
科技活动成果	国内专利申请数、国内专利授权数、国内有效专利数、国内、外三种专利申请数……
科技服务	全国非油气地质勘查分地区情况、全国气象部门基本情况、地震台、网基本情况、海洋观测预报单位机构、人员情况……
国际比较	研究与试验发展(R&D)经费及占国内生产总值的比重、研究与试验发展(R&D)活动的国际比较、按 ESI 论文数量排序的前 20 个国家……

5.数据范围

数据模块	数据起始时间
全部模块	1980 年

（五）全国景点数据库-ATRA

1.数据库简介

全国景点数据库 (*Chinese Attractions Database, ATRA*) 是基于中国各省市文化和旅游厅所公布的 A 级旅游景区名录整理而成的专业数据库。在获得景区名称的基础上, *ATRA* 对此进行了进一步加工, 列示了景区被评级的批准时间、景区所属城市和县区、景区的地址、景区的经纬度等重要字段。需要注意的是, 由于大部分省市的文化和旅游厅在近年才公布 A 级景区名录, 因此, *ATRA* 所搜集的也是近几年 A 级景区的截面数据, 并且通过“截止时间”字段注明了数据的发布时间。

2.数据库特色

- 数据来源权威

ATRA 的数据来源于官方网站所披露数据, 数据来源权威。

- 字段丰富

ATRA 包含了与景区相关的诸多字段, 大大便利了学术研究。

3.典型数据举例

全国景点

景区名称	所属省份	所属城市	所属县区	景区级别	批准时间
故宫博物院	北京市	北京市	东城区	5A	2007
天坛公园	北京市	北京市	东城区	5A	2007
圆明园遗址公园	北京市	北京市	海淀区	5A	2020
颐和园	北京市	北京市	海淀区	5A	2007
十三陵	北京市	北京市	昌平区	5A	2011

4.数据库内容

数据模块	数据表	数据字段
全国景点	全国景点	景区名称、所属省份、所属城市、所属县区、景区地址、景区级别、批准时间、截止时间、经度、纬度

5.数据范围

数据模块	数据起始时间
全国景点	截面数据

（六）中国民政统计数据库-CASD

1.数据库简介

中国民政统计数据库（China Civil Affairs Statistical Database, CASD）是基于民政部发布的统计年鉴而建立的，全面反映我国社会服务发展情况的专业数据库。CASD 主要从 4 个方面反映了我国社会服务发展情况，分别是社会服务主要图表数据、社会服务综合统计资料、社会服务历年统计资料、社会服务当年分省统计资料。

2.数据库特色

- 数据来源权威、结构清晰简洁

CASD 主要根据国家统计局、地方统计机构以及统计年鉴等多个渠道数据整合而成。CASD 不仅整合了多个年度相同数据到一张表格中，同时对每个指标的单位也进行了标注。

3.典型数据举例

各服务机构床位数

指标	年份	床位数（万张）	增长率（%）	每千人口床位数（张）
养老服务床位数	2009	293.5	9.8	17.56
养老服务床位数	2010	316.1	7.7	17.79
养老服务床位数	2011	369.2	16.8	19.96
养老服务床位数	2012	416.5	12.8	21.48

4.数据库内容

数据模块	数据子模块	数据表
社会服务主要数据	综合	市、区、县、乡镇与街道；六十岁以上老年人口占全国总人口比重；人口年龄结构……
	社会工作	各服务机构床位数；老年人福利；家庭收养……
	成员组织和其他社会服务	社会组织概况；自治组织概况；结婚登记概况……
社会服务综合统计	综合	“七五” — “十三五” 时期社会服务发展速度……
	社会工作	社会工作类机构财务状况……
	成员组织和其他社会服务	社会组织财务状况；社会组织分类情况与上年比较……
社会服务历年统计	综合	县及以上行政区划；乡镇级行政区划；全国人口情况……
	社会工作	提供住宿的社会服务机构；……
	成员组织和其他社会服务	社会组织；自治组织；结婚登记；离婚登记；殡葬服务
社会服务当年分省统计	行政区划和行政机关	省级行政区划；地级与县级行政区划……
	社会服务总体情况	社会服务总表；社会服务发展综合指数……
	社会工作	社会工作总表；提供住宿的社会服务机构总表……
	成员组织和其他社会服务	成员组织总表；社会组织总表；社会团体 2000-2005……

5.数据范围

数据模块	数据起始时间
全部模块	1978 年

(七) 中国卫生健康统计数据库-CHSD

1. 数据库简介

中国卫生健康统计数据库 (*China Health Statistical Database, CHSD*) 是基于国家卫健委发布的《中国卫生健康统计年鉴》整理而成的, 反映我国卫生健康事业发展情况和居民卫生健康状况的特色数据库。*CHSD* 主要包括医疗卫生机构, 卫生人员, 卫生设施, 卫生经费, 医疗服务, 基层医疗卫生服务, 中医药服务, 妇幼保健与计划生育, 人民健康水平, 疾病控制与公共卫生, 居民病伤死亡原因, 食品安全与卫生健康监督, 医疗保障, 人口数及构成这 14 个大模块。

2. 数据库特色

- 简洁的数据划分

CHSD 对原始年鉴中杂乱无规律且不连贯的数据进行了精简, 使得数据划分更为精简。

3. 典型数据举例

医疗卫生机构数

地区	年份	单位	医院_小计	医院_综合 医院	医院_中医 医院	医院_中西医 结合医院	医院_专科 医院
山西	2016	家	1393	693	210	26	463
内蒙古	2016	家	720	402	110	10	134
辽宁	2016	家	1190	694	139	10	339

4. 数据库内容

数据模块	数据表
医疗卫生机构	医疗卫生机构数; 各地区医疗卫生机构数……
卫生人员	卫生人员数; 各类医疗卫生机构人员数……
卫生设施	医疗卫生机构床位数; 各类医疗卫生机构床位数……
卫生经费	卫生总费用; 各地区卫生总费用; 政府卫生支出; 政府卫生支出所占比重……
医疗服务	医疗卫生机构诊疗人次数; 各类医疗卫生机构门诊服务情况……
基层医疗卫生服务	基层医疗卫生机构医疗服务量; 各地区基层医疗卫生机构工作情况……
中医药服务	中医类医疗机构诊疗人次; 其他机构中医类临床科室诊疗人次……
妇幼保健与计划生育	监测地 5 岁以下儿童和孕产妇死亡率……
人民健康水平	人口出生率、死亡率与自然增长率……
疾病控制与公共卫生	甲乙类法定报告传染发病数及死亡数排序……
居民病伤死亡原因	城市居民主要疾病死亡率及构成; 城市居民年龄别疾病别死亡率……
食品安全与卫生健康监督	各类致病因素食源性疾病暴发报告情况……
医疗保障	各地区城镇居民和职工基本医疗保险情况; 各地区生育保险情况……
人口数及构成	人口数及构成; 流动人口数; 人口基本情况; 各地区总人口……

5. 数据范围

数据模块	数据起始时间
全部模块	1978 年

（八）中国环境统计数据库-ENVS

1.数据库简介

中国环境统计数据库（China Statistical Database on Environment, ENVS）是基于国家统计局和生态环境部提供的资料而建立的，全面反映我国各省、自治区、直辖市环境分领域基本数据和主要年份全国环境统计数据的专业数据库。ENVS 共分为 13 个模块，从自然状况、水环境、大气环境等多个方面反映了我国各地区、各行业的状况。

2.数据库特色

- 权威且完整性

ENVS 反映了我国 2004 年以来环境各领域的基本情况，数据来源于国家统计局及生态环境部。

3.典型数据举例

全国水环境情况

年份	水资源总量（亿立方米）	水资源总量：地表水资源量（亿立方米）	水资源总量：地下水资源量（亿立方米）	降水量（亿立方米）	人均水资源量（立方米/人）	供水总量（亿立方米）
2002	28261	27243	8697	62610	2207.2	5497.3
2003	27460	26251	8299	60416	2131.3	5320.4
2004	24130	23126	7436	56876	1856.3	5547.8
2005	28053	26982	8091	61010	2151.8	5633

4.数据库内容

数据模块	数据表
自然状况	自然状况；土地状况；主要山脉基本情况；主要河流基本情况；主要城市气候情况；主要矿产基础储量；各地区主要能源、黑色金属矿产基础储量；各地区主要有色金属、非金属矿产基础储量；全国水环境情况
水环境	各流域水资源情况；各流域节水灌溉面积；各流域供水和用水情况；各地区水资源情况；各地区节水灌溉面积；各地区供水和用水情况；流域分区河流水质状况评价结果（按评价河长统计）；主要水系水质状况评价结果（按监测断面统计）；大型淡水湖泊、水库水质状况；重点评价湖泊水质状况；各地区废水排放及处理情况 2004-2010……
海洋环境	全国海洋环境情况；全海域未达到第一类海水水质标准的海域面积；海区废弃物倾倒及石油勘探开发污染物排入海情况；全国主要海洋产业增加值；海洋资源利用情况；全国近岸海域海水水质评价结果（按海水类别分）；全国近岸海域海水水质评价结果（按海域污染程度分）；全海域海水水质评价结果；海洋石油、原油及天然气
大气环境	全国废气排放及处理情况；各地区废气排放及处理情况 2004-2010；各地区废气排放及处理情况 2011 以后；各行业工业废气排放及处理情况 2004-2010；各行业工业废气排放及处理情况 2011 以后；主要城市工业废气排放及处理情况；沿海城市工业废气排放及处理情况；全国历年废气排放及处理情况
固体废物	全国工业固体废物产生、排放和综合利用情况；各地区固体废物产生和排放情况；各行业固体废物产生和排放情况 2004-2010；各行业固体废物产生和排放情况 2011 以后；主要城市固体废物产生和排放情况；沿海城市固体废物产生和排放情况……

数据模块	数据表
自然生态	全国自然生态情况；各地区自然保护基本情况；海洋类型自然保护区建设情况；各地区草原建设利用情况；各地区地质公园建设情况；各地区矿山环境保护情况；各流域除涝和水土流失治理情况；各地区除涝和水土流失治理情况；各地区湿地情况；各地区荒漠化和沙化土地情况 2004-2009；各地区沙化土地情况
土地利用	全国土地利用情况；各地区土地利用情况；各地区耕地变动情况
林业	全国造林情况；各地区森林资源情况；各地区造林情况；各地区天然林资源保护情况；各地区退耕还林情况；三北、长江流域等重点防护林体系工程建设情况；京津风沙源治理工程建设情况；各地区林业系统野生动植物保护及自然保护区工程建设情况；重点地区速生丰产用材林基地工程建设情况
自然灾害及突发事件	全国自然灾害情况；各地区自然灾害损失情况；地质灾害及防治情况；地震灾害情况；海洋灾害情况；各地区森林火灾情况；各地区森林有害生物防治情况；各地区草原灾害情况；各地区突发环境事件情况 2008-2010；各地区突发环境事件情况 2011 以后；各地区环境污染与破坏事故情况
环境投资	全国环境污染治理投资情况；各地区环境污染治理投资情况；各地区城镇环境基础设施建设投资情况；各地区工业污染治理投资完成情况；各地区林业投资完成情况；各地区环保验收项目情况；各地区林业投资资金来源情况；各地区工业污染治理投资来源情况；各地区排污费征收情况；各地区建设项目“三同时”执行情况
城市环境	全国城市环境情况；环保重点城市空气质量指标；各地区城市面积和建设用地情况 2004-2011；各地区城市面积和建设用地情况 2012 以后；各地区城市市政设施情况；各地区城市供水和用水情况；各地区城市节约用水情况；各地区城市污水排放和处理情况；各地区城市市容环境卫生情况；各地区城市燃气情况 2004；各地区城市燃气情况 2005 以后；各地区城市集中供热情况；各地区城市园林绿化情况；各地区城市公共交通情况；环保重点城市道路交通噪声监测情况；环保重点城市区域环境噪声监测情况；主要城市区域环境噪声声源构成情况；各地区城市环境综合整治情况
农村环境	全国农村环境情况；各地区农村改厕情况 2004-2014；各地区农村改厕情况 2015 以后；各地区农村可再生能源利用情况；各地区耕地灌溉面积和农用化肥施用情况；各地区农用塑料薄膜和农药使用量情况；各地区农村改水、改厕投资情况；各地区农村非商品能源生活消费情况
附录	人口资源环境主要统计指标；东中西部地区水资源情况；东中西部地区环境污染治理投资情况；东中西部地区城市环境情况；东中西部地区农村环境情况；东中西部地区废水排放情况；东中西部地区废气排放情况；东中西部地区固体废物产生及处理情况 2005-2010；东中西部地区固体废物产生及处理情况 2011 以后

5.数据范围

数据模块	数据起始时间
全部模块	2000 年

（九）宗族文化数据库-CLAN

1.数据库简介

宗族文化数据库（*Clan Culture Database, CLAN*）是基于中国各地方的家谱数据整理而成的，反映中国宗族文化的专业数据库。*CLAN*的家谱资料来自上海图书馆编著的《中国家谱总目》，该书收录的家谱条目高达五万多部，覆盖了全国 608 个姓氏，并按 2003 年行政区划著录了每部家谱谱主的姓氏和谱籍、家谱收藏机构等信息。

2.数据库特色

- 精确的数据处理过程

CLAN 严格按照已有学术文献的处理方法，对家谱总目数据进行严谨的数据整理。

- 全面的数据字段和详情

CLAN 列示了姓氏、谱籍所在省市县以及纂修年代等详细字段数据，并展示了统计之前每个谱籍的详细情况，便于学者进行多维度研究。

3.典型数据举例

家谱总目情况表

族谱序号	姓氏	家族所属省份	家谱所属地市	家谱所属县区	家谱年份
124-0097	李氏	江苏省	南京市	江宁区	1887
124-0098	李氏	江苏省	南京市	江宁区	1902
124-0099	李氏	江苏省	南京市	溧水区	1948
124-0101	李氏	江苏省	徐州市	丰县	1917

4.数据库内容

数据模块	数据表	数据字段
家谱总目情况表	家谱总目情况表	族谱序号、姓氏、家谱所属省份、家谱所属地市、家谱所属县区、家谱年份
家谱总目统计表	家谱总目统计表	家谱所属省份、家谱所属地市、地市级家谱数量统计值、地区年末总人口（万人）、市区年末总人口（万人）、地区每百万人拥有家谱数、市区每百万人拥有家谱数

5.数据范围

数据模块	数据起始时间
全部模块	截面数据

(十) 中国火炬统计数据库-TORCH

1.数据库简介

中国火炬统计数据库 (*China Torch Statistical Database, TORCH*) 是基于对《中国火炬统计年鉴》相关指标的深入考察,建立的详细、准确和全面的专业数据库。该库共分为十大模块,主要展示了2007年以来中国火炬计划、技术市场、全国生产力促进中心等相关数据。

2.数据库特色

- 准确的数据模块划分

TORCH 将全国各省市数据中的诸多表格按照关联性进行分组,并整合在不同模块中,极大便利了使用者进行数据查找、浏览和下载。

3.典型数据举例

全国高新技术企业主要经济指标

年份	入统企业数(个)	年末从业人员(万人)	工业总产值(亿元)	净利润(亿元)	上缴税额(亿元)	出口创汇(亿美元)
2014	62556	1914.8	211335.9	14399.2	10674.8	5068.6
2015	76141	2045.2	189757.5	14894.8	11052.1	4768.7
2016	100012	2360.7	212268.8	18859.7	13159.1	4694.9

4.数据库内容

数据模块	数据表
国家高新技术产业开发区	各国家高新区企业 R&D 活动与科技活动情况;各高新区企业主要经济指标……
全国高新技术企业	副省级城市高新技术企业 R&D 活动与科技活动情况……
科技企业孵化器	各地区国家级科技企业孵化器当年在孵企业情况……
国家火炬计划项目	火炬计划产业化类项目主要经济指标(按单位性质分类);火炬计划产业化类项目主要经济指标(按技术来源分类)……
国家大学科技园	各地区国家大学科技园毕业企业情况……
国家火炬软件产业基地	软件产业基地场地情况;软件产业基地出口和利税情况……
国家火炬特色产业基地	各地区火炬特色产业基地经济指标……
创新型产业集群	创新型产业集群主要服务机构情况;创新型产业集群主要经济指标;创新型产业集群主要科技活动成果情况;创新型产业集群主要情况
全国技术市场	东北地区技术交易情况;东部地区技术交易情况……
全国生产力促进中心	各省、自治区、直辖市国家级示范生产力促进中心服务情况;各省、自治区、直辖市国家级示范生产力促进中心服务业绩情况……

5.数据范围

数据模块	数据起始时间
全部模块	2007年

(十一) 教育经费统计数据库-EFSD

1. 数据库简介

教育经费统计数据库 (*Educational Finance Statistical Database, EFSD*) 是根据《中国教育经费统计年鉴》整理而成的, 反映全国教育经费来源和使用情况的专业数据库。*EFSD* 包括全国教育经费收支、各地区按来源分类教育经费收入、各地区各级各类教育机构教育经费收入情况、各地区各级各类教育机构教育经费支出明细、各地区各级各类教育机构财政补助支出明细、各地区各级各类教育机构公共预算教育事业费和基本建设支出明细等六个模块。

2. 数据库特色

- 数据来源权威

EFSD 的数据来源于《中国教育经费统计年鉴》等权威渠道, 数据准确和可靠。

3. 典型数据举例

各级各类教育机构财政补助支出明细(地方)

学校类别	年份	总计(千元)	个人部分(千元)	公用部分(千元)	基本建设支出(千元)
总计_小学	2014	809660510	534527367	264694765	10438377
总计_中学_成人中学	2014	511166	260389	250777	
总计_中学_普通中学_普通初中_农村	2014	311465987	203693621	101726350	6046017
总计_中学_普通中学_普通初中	2014	498062490	317418510	170872228	9771753

4. 数据库内容

数据模块	数据表
全国教育经费收支	各级各类教育机构财政补助支出明细(地方); 各级各类教育机构财政补助支出明细(地方教育和其他部门)……
各地区按来源分类教育经费收入	分地区教育经费总收入; 港澳台及海外捐赠……
各地区各级各类教育机构教育经费收入情况	教育经费收入情况(成人高等学校); 教育经费收入情况(成人小学)……
各地区各级各类教育机构教育经费支出明细	教育经费支出明细(成人高等学校); 教育经费支出明细(成人小学)……
各地区各级各类教育机构财政补助支出明细	财政补助支出明细(成人高等学校); 财政补助支出明细(成人小学)……
各地区各级各类教育机构公共预算教育事业费和基本建设支出明细	一般公共预算教育事业费和基本建设支出明细(成人高等学校); 一般公共预算教育事业费和基本建设支出明细(成人小学)……

5. 数据范围

数据模块	数据起始时间
全部模块	1995 年

(十二) 中国价格统计数据库-CPSD

1. 数据库简介

中国价格统计数据库 (China Price Statistical Database, CPSD) 是根据《中国价格统计年鉴》整理而成的, 反映我国价格变动情况的专业数据库。CPSD 包括综合、生产者价格指数、流通消费价格指数、70 个大中城市商品住宅销售价格指数四大模块, 反映了生产、流通、消费等环节的价格变动趋势和变动幅度。

2. 数据库特色

- 数据来源权威

CPSD 的数据来源于《中国价格统计年鉴》等权威渠道, 数据准确和可靠。

3. 典型数据举例

全国商品零售价格分类指数

项目	年份	全国(上年价格=100)	城市(上年价格=100)	农村(上年价格=100)	城市_36 个大中城市(上年价格=100)
商品零售价格总指数	2012	102	101.9	102.2	101.8
一、食品	2012	104.8	105.1	104	105.6
1.粮食	2012	103.8	103.9	103.7	104.4
2.淀粉及制品	2012	103.7	103.9	103.3	105
3.干豆类及豆制品	2012	102	102	102	102.9
4.油脂	2012	105.1	105.6	104.1	105.6

4. 数据库内容

数据模块	数据表
综合	各种价格总指数; 各种价格定基指数
生产者价格指数	全国工业生产者出厂价格分类指数; 全国按部门分工业生产者出厂价格指数; 全国按行业分工业生产者出厂价格指数……
流通消费价格指数	全国居民消费价格分类指数; 各地区居民消费价格总指数; 各地区居民消费价格分类指数; 各地区城市居民消费价格分类指数……
70 个大中城市商品住宅销售价格指数	70 个大中城市住宅销售价格指数

5. 数据范围

数据模块	数据起始时间
综合	1951 年
生产者价格指数	1985 年
流通消费价格指数	1978 年
70 个大中城市商品住宅销售价格指数	2012 年

（十三）中国知识产权研究数据库-IPRR

1.数据库简介

中国知识产权研究数据库 (*Intellectual Property Rights Research Database, IPRR*) 是根据《中国知识产权年鉴》中与知识产权相关的统计数据部分整理而成的专业数据库。*IPRR* 包含专利；商标；版权；知识产权司法保护；农业植物新品种；林草植物新品种；集成电路布图设计；海关知识产权保护；地理标志等模块，从多个维度展示了近年来我国知识产出情况以及知识产权保护情况。希望 *IPRR* 的推出能为该领域研究提供更多数据支持。

2.数据库特色

- 数据来源权威

IPRR 的数据来源于《中国知识产权年鉴》等权威渠道，数据准确和可靠。

3.典型数据举例

中国历年国内外三种专利申请受理年度状况

地区	年份	合计(单位:件)	发明(单位:件)	实用新型(单位:件)	外观设计(单位:件)
合计	2005	476264	173327	139566	163371
国内	2005	383157	93485	138085	151587
国外	2005	93107	79842	1481	11784

4.数据库内容

数据模块	数据表
专利	中国历年国内外三种专利申请受理年度状况；国内三种专利申请受理量……
商标	商标申请和注册概况；各省、自治区、直辖市商标申请与注册……
地理标志	主要林产品地理标志登记量
版权	全国版权合同登记情况；全国作品自愿登记情况统计表；版权引进地汇总……
集成电路布图设计	集成电路布图设计登记申请；集成电路布图设计登记发证
农业植物新品种	农业植物新品种权申请和授权（根据植物种类划分）；农业植物新品种权申请（根据单位性质划分）；农业植物新品种权申请（根据地区划分）；国外植物新品种申请情况
林草植物新品种	历年林业植物新品种申请量和授权量；林业授权品种中不同申请国家的授权量；各国授权品种的属(种)授权量……
海关知识产权保护	海关扣留货物的进出口流向；海关知识产权执法；中国海关查获侵权商品；海关保护的知识产权权利人来源情况……
知识产权司法保护	法院受理和审结各类知识产权案件；检察机关办理侵犯知识产权案件情况

5.数据范围

数据模块	数据起始时间
全部模块	1979 年

(十四) 中国劳动经济数据库-CLSD

1. 数据库简介

中国劳动经济数据库 (*China Labour Statistical Database, CLSD*) 是根据《中国劳动统计年鉴》整理而成的, 全面反映我国劳动经济情况的专业数据库。*CLSD* 分为 10 个模块, 分别为: 综合、就业与失业、城镇单位就业人员和工资总额、国有单位就业人员和工资总额、城镇集体单位就业人员和工资总额、其他单位就业人员和工资总额、职业培训与技能鉴定、劳动关系、社会保障、工会工作。希望 *CLSD* 的推出能为该领域研究提供数据支持。

2. 数据库特色

- 数据来源权威

CLSD 的数据来源于《中国劳动统计年鉴》等权威渠道, 数据准确和可靠。

3. 典型数据举例

综合_城镇按户口性质、性别分的就业人员工作时间构成

户口性质	单位	年份	合计	1-8 小时	9-19 小时	20-39 小时
总计	%	2004	100	0.2	0.6	6.8
总计: 户口在本省: 非农业	%	2004	100	0.3	0.3	4.1
总计: 户口在本省: 农业	%	2004	100	0.2	1.2	12.6
总计: 户口在外省: 非农业	%	2004	100	0.3	0.3	2.7
总计: 户口在外省: 农业	%	2004	100	0.1	0.6	2.5
总计: 其他	%	2004	100		0.4	4.8

4. 数据库内容

数据模块	数据表
综合	城镇按户口性质、性别分的就业人员工作时间构成; 城镇按就业身份、性别分的就业人员工作时间构成; 城镇按就业身份、性别分的就业人员年龄构成; 城镇按就业身份、性别分的就业人员受教育程度构成……
就业与失业	各地区城镇登记失业人员情况; 各地区公共就业服务工作情况; 各地区年末城镇登记失业人数及登记失业率; 公共就业服务工作情况; 年末城镇登记失业人数及登记失业率
城镇单位就业人员和工资总额	分行业城镇单位就业人员和工资总额; 各地区分登记注册类型城镇单位工资总额; 各地区分登记注册类型城镇单位年末人数; 各地区分登记注册类型城镇单位平均工资; 各地区分行业城镇单位就业人员和工资总额 1; 各地区分行业城镇单位就业人员和工资总额 2; 各地区分行业城镇单位其他就业人员和平均工资; 各地区分行业城镇单位在岗职工人数和工资
国有单位就业人员和工资总额	分行业国有单位就业人员和工资总额; 各地区分行业国有单位就业人员和工资总额 1998-2002; 各地区分行业国有单位就业人员和工资总额 2003 以后; 各地区分行业国有单位其他就业人员和平均工资 1998-2002; 各地区分行业国有单位其他就业人员和平均工资 2003 以后; 各地区分行业国有单位在岗职工人数和平均工资 1998-2002; 各地区分行业国有单位在岗职工人数和平均工资 2003 以后
城镇集体单位就业人	分行业城镇集体单位就业人员和工资总额; 各地区分行业城镇集体单位就业人员

数据模块	数据表
员和工资总额	和工资总额 1998-2002；各地区分行业城镇集体单位就业人员和工资总额 2003 以后；各地区分行业城镇集体单位其他就业人员和平均工资 1998-2002；各地区分行业城镇集体单位其他就业人员和平均工资 2003 以后；各地区分行业城镇集体单位在岗职工人数和平均工资
其他单位就业人员和工资总额	分行业其他单位就业人员和工资总额；各地区分行业其他单位就业人员和工资总额 1998-2002；各地区分行业其他单位就业人员和工资总额 2003 以后；各地区分行业其他单位其他就业人员和平均工资；各地区分行业其他单位在岗职工人数和平均工资
职业培训与技能鉴定	各地区技工学校综合情况；各地区就业训练中心综合情况；各地区民办职业培训机构综合情况；各地区职业技能鉴定综合情况；技工学校综合情况；历年全国职业技能鉴定综合情况
劳动关系	各地区劳动争议处理情况；劳动保障监察案件结案情况；劳动保障监察工作情况；历年劳动人事争议仲裁情况
社会保障	分地区因工死亡人员工伤认定情况；各地区城乡居民基本养老保险情况；各地区城镇居民基本医疗保险基本情况；各地区城镇职工基本养老保险情况；各地区工伤保险基本情况；各地区工伤认定情况；各地区劳动能力鉴定情况……
工会工作	分地区基层单位建立职工代表大会制度情况；各地区工会会员人数 1992-2000；各地区工会会员人数 2001 以后；各地区基层工会参与调节劳动争议工作情况；各地区基层工会开展合理化建议和劳动竞赛活动情况；各地区基层工会组织数；各地区基层以上工会职业介绍机构情况；各地区基层以上工会职业培训机构情况

5.数据范围

数据模块	数据起始时间
全部模块	1949 年

(十五) 人口老龄化数据库-PARD

1. 数据库简介

人口老龄化数据库 (Population Aging Research Database, PARD) 是基于国家统计局、各省市统计局等相关养老和老龄化数据而设计的专业数据库。PARD 从 5 个方面展示了我国人口老龄化情况, 分别是基本情况、老龄生活、养老机构、养老保险以及智慧健康养老示范名单。PARD 围绕中国人口老龄化构建的这些不同模块和丰富字段, 为该领域的学术研究提供了重要数据支持。

2. 数据库特色

- 数据层次丰富

PARD 从微观到宏观, 提供了多个层次的人口老龄化的数据颗粒度, 为该领域研究提供更多支持。

3. 典型数据举例

各省市老年人福利

地区	年份	老年人福利: 享受高龄补贴的老年人人数 (人)	老年人福利: 享受护理补贴的老年人人数 (人)	老年人福利: 享受养老补贴的老年人人数 (人)	老年人福利: 享受综合补贴的老年人人数 (人)
甘肃省	2021	7734	13562	93582	6040
甘肃省	2020	9523	20579	67738	7102
海南省	2021	225686	7947	1925	22
海南省	2020	232955	4530	5097	16

4. 数据库内容

数据模块	数据表
基本情况	各地区各年龄组人口比重、有老年人口的户数(普查); 各省市老龄事业发展情况; 各省市人口年龄构成和抚养比
老龄生活	各省市居民人均可支配收入; 各省市老年人福利; 各省市养老病床与护理
养老机构	各省市社区养老机构和设施; 各省市养老服务机构; 全国养老机构信息
养老保险	各省市城乡居民养老保险; 各省市基本养老保险
智慧健康养老示范名单	全国智慧健康养老示范企业名单; 全国智慧健康养老示范街道(乡镇)名单; 全国智慧健康养老示范基地名单

5. 数据范围

数据模块	数据起始时间
全部模块	截面数据

财政金融研究系列

（一）中国税收研究数据库-CTRD

1. 数据库简介

中国税收研究数据库 (*Chinese Tax Research Database, CTRD*) 基于对中国国税、地税以及上市公司税收的深入考察, 建立的详细、准确和全面的中国税收研究专业数据库。该库不仅包括中国税务部门分地区、分税种、分企业类型、分产业等的税收收入情况, 详细的分市县各税种税收收入, 还包括上市公司税收信息。

CTRD 包含中国税收收入信息、地税收入详细信息和上市公司税收信息三个大模块。其中, 中国税收收入信息模块分为全国税收收入、国税收入和地税收入三个小模块, 这些小模块又分为全部税收收入和具体税收收入两个部分, 包含了这三类税收收入各种情况下的收入信息以及细分为具体税收收入情况下的收入信息; 地税收入详细信息模块包含各个省份市级和县级的各税种税收情况; 上市公司税收信息包括我国上市公司各个年度的税收信息。

2. 数据库特色

- 唯一性与多维度

CTRD 是目前市场上唯一提供中国税收收入详细信息的数据库, 弥补了我国税收研究数据的空缺。该数据库从中国税收、地税收入和上市公司税收三个方面, 分地区、分税种、分企业类型、分产业等多个维度详细记录了我国税收的发展情况。

3. 典型数据举例

各省地税部门分税种分市、县收入情况

所在省级名称	所在市级名称	所在县级名称	会计年度	地税部门组织收入总计 (万元)	一、税收收入 (万元)	二、基金费收入合计 (万元)	三、税务部门其他罚没收入 (万元)
安徽省	合肥市	合肥市合计	2009	1722736	1192646	529463	627
安徽省	芜湖市	繁昌县	2010	64391	43258	21131	2
安徽省	宣城市	宣州区	2011	121581	88805	32771	5
安徽省	滁州市	来安县	2012	97850	72550	25299	1
安徽省	合肥市	市区	2013	3673069	2497914	1174956	199
安徽省	亳州市	亳州市合计	2014	725812	551119	174669	24

4. 数据库内容

数据模块	数据表
全国税收收入-全国税收收入	税收收入分税种分产业情况、税收收入分税种分地区情况、税收收入分税种分级次情况、税收收入分税种分企业类型情况、税收收入分地区分产业情况、税收收入分地区分企业类型情况

数据模块	数据表
全国税收收入-具体税收收入	各类税收分地区分企业收入情况、增值税收入分项目分企业类型收入情况、资源税分项目分企业类型收入情况、营业税分行业分企业类型收入情况、营业税分税目分企业类型收入情况、涉外税收分地区分税种收入情况、涉外税收分地区分项目收入情况、涉外税收分项目分税种收入情况、企业所得税分地区分行业收入情况、企业所得税分项目分企业类型收入情况、个人所得税分地区分项目收入情况
国税收入-全部税收收入	国家税务总局税收收入分税种分产业情况、国家税务总局税收收入分税种分地区情况、国家税务总局税收收入分税种分级次情况、国家税务总局税收收入分税种分企业类型情况、国家税务总局税收收入分地区分产业情况、国家税务总局税收收入分地区分企业类型情况
国税收入-具体税收收入	国家税务总局营业税分地区分企业类型收入情况、国家税务总局资源税分地区分企业类型收入情况、国家税务总局涉外税收分地区分税种收入情况、国家税务总局企业所得税分地区分企业类型收入情况
地税收入-全部税收收入	地方税务局税收收入分税种分产业情况、地方税务局税收收入分税种分地区情况、地方税务局税收收入分税种分级次情况、地方税务局税收收入分税种分企业类型情况、地方税务局税收收入分地区分产业情况、地方税务局税收收入分地区分企业类型情况
地税收入-具体税收收入	地方税务局营业税分地区分企业类型收入情况、地方税务局资源税分地区分企业类型收入情况、地方税务局涉外税收分地区分税种收入情况、地方税务局企业所得税分地区分企业类型收入情况
地税收入详细信息	安徽省、北京市、福建省、广东省、广西壮族自治区、黑龙江省、湖北省、山东省、云南省、浙江省、重庆市、河南省、湖南省、四川省、陕西省
上市公司税收信息	上市公司税收信息

5.数据范围

数据模块	数据起始时间
中国税收收入信息	2001 年
地税收入详细信息	1994 年
上市公司税收信息	2001 年

（二）中国保险统计数据库-INS D

1.数据库简介

中国保险统计数据库 (*Chinese Insurance Statistics Database, INS D*) 是基于中国保险年鉴相关数据而建立的专业数据库。*INS D* 囊括了 2005 年以来中国各保险公司的业务和财务信息, 各地区保险业务情况, 保险公司从业人员、再保险和保险院校等情况。*INS D* 数据内容丰富完整, 对人身、财产保险公司都进行了详细统计, 囊括了全国主要保险公司和各地区的保险统计数据, 极大满足了学术研究的需求。

2.数据库特色

- 详细的业务统计数据

INS D 数据来源于中国保险年鉴, 业务统计表不仅细分了财险及人寿公司的数据, 还分别从全国、公司、地方等 3 个方面进行统计。

- 多维度的保险行业数据

INS D 除了包含保险机构本身信息, 还展示了与之相关中介机构及从业人员信息和保险院校数据。

3.典型数据举例

各财产保险公司业务统计表 (全国版)

年份	保险公司名称	单位	保费收入合计	赔付支出	赔案件数 (万件)	未决赔款
2008	人保财险	百万元	101762.95	61430.87	1398.15	30750.62
2008	国寿财险	百万元	5343.20	1511.89	61.57	684.98
2008	大地	百万元	9424.49	6187.67	207.89	263.46
2008	太平保险	百万元	4269.34	2170.55	82.52	1087.28
2008	中国信保	百万元	3498.00	2619.00	0.08	2550.00

4.数据库内容

数据模块	数据表
各保险公司业务统计表	各财产保险公司业务统计表 (全国版); 各财产保险分公司业务统计表 (公司版); 各财产保险分公司业务统计表 (地方版); 各人身保险公司业务统计表 (全国版); 各人身保险分公司业务统计表 (公司版); 各人身保险分公司业务统计表 (地方版)
各保险公司财务报表	各保险公司资产负债表; 各保险公司损益表
各地区保险业务情况	各省、自治区、直辖市、计划单列市经济和社会发展情况; 各省、自治区、直辖市、计划单列市保险业务情况; 各省、自治区、直辖市、计划单列市财产保险公司市场份额表; 各省、自治区、直辖市、计划单列市人身保险公司市场份额表
其他	各保险公司再保险业务统计表; 各保险公司人员结构情况; 各保险中介公司基本情况; 全国中高等院校保险专业和保险院校情况

5.数据范围

数据模块	数据起始时间
全部模块	2005 年

(三) 中国行政审批数据库-CAAD

1.数据库简介

中国行政审批数据库 (*Chinese Administrative Approval Database, CAAD*) 是基于全国 300 多个地级市行政审批中心数据而建立的数据库。随着中国经济改革的深入, 行政审批中心的设立可以实现审批项目“一站式”办理, 从而反作用于经济发展。因此, 为了支持学者对这一领域的深入研究, *CAAD* 从多个维度提供了行政审批中心数据, 重要字段包括是否设立行政审批中心、审批中心成立时间、进驻部门或单位、进驻事项和进驻窗口等。由于中国行政审批制度改革并未停止, 近年来在行政审批中心基础上, 多个地方纷纷设立行政审批局, 启用行政审批专用章, 从而实现了“一颗印章管审批”。因此, 本文基于此还设立了是否组建行政审批局、行政审批局成立时间等重要字段, 扩展了学者的研究思路。

2.数据库特色

- 数据唯一, 字段丰富

近年来, 行政审批效率研究已然成为学术研究的重要议题, 但是这类研究数据的搜集需要学者耗费大量时间和精力, 尚无公开数据库提供该类数据, *CAAD* 弥补了这类公开数据缺失的不足。*CAAD* 基于全国 300 多个地级市审批中心数据, 提供成立时间、进驻事项、进驻窗口等近 20 个重要字段。

3.典型数据举例

行政审批中心

省级名称	地级名称	行政审批单位名称	设立时间	进驻事项	进驻窗口	进驻工作人员
河北省	唐山市	唐山市行政服务中心	2002-11-20	278	263	400
山西省	太原市	太原市政务服务中心	2003-01-00	450	70	160
内蒙古	包头市	包头市政务服务中心	2002-00-00	600	150	400
吉林省	长春市	长春市政务中心	2002-10-00	277	183	345
吉林省	松原市	松原市政务服务中心	2002-09-00	302	94	127

4.数据库内容

数据模块	数据表	数据字段
行政审批中心	行政审批中心	省级名称、地级名称、是否设立行政审批单位、行政审批单位名称、设立时间、单位介绍、是否组建行政审批局、行政审批局成立时间、进驻部门或单位、审批部门、服务部门、职能科室、公共服务单位、公共资源部门、内设机构、进驻事项、行政许可事项、进驻窗口数量、进驻工作人员

5.数据范围

数据模块	数据起始时间
------	--------

全部模块

2000 年

（四）政府和社会资本合作数据库-PPPD

1. 数据库简介

政府和社会资本合作数据库（*Public Private Partnerships Database, PPPD*）是基于财政部政府和社会资本合作中心的全国 PPP 综合信息平台项目管理库相关数据整理而成的专业数据库，旨在为学术研究者提供权威的 PPP 数据。PPPD 分为 PPP 管理项目、PPP 储备项目、PPP 相关机构等模块。

2. 数据库特色

- 全面的数据内容

PPPD 数据库中的 PPP 管理项目模块不仅披露了项目的评分指标及结果，同时对预算数额进行了展示；PPP 相关机构模块不仅有详细的咨询信息及人员信息，还包含了 2015 年以来的业绩情况。

3. 典型数据举例

管理项目基本信息表

项目名称	所在省份	所在区县	项目总投资 (万元)	合作 期限	批准日期
“美丽新昌”小城镇环境综合整治项目	浙江省	新昌县	188156.00	12 年	2017-06-27
“智慧岑巩”基础设施建设 PPP 项目	贵州省	岑巩县	188879.18	20 年	2018-01-24
“中国稀金谷”核心区基础设施 PPP 项目	江西省	赣县区	82601.00	10 年	2017-11-03
002 省道南京浦口段建设项目	江苏省	浦口区	79751.00	10 年	2017-09-19
2015 年正安县城北片区棚户区改造 PPP 项目	贵州省	正安县	21173.68	25 年	2020-05-27

4. 数据库内容

数据模块	数据表	数据字段
PPP 管理项目	管理项目基本信息表	项目名称、项目编号、所在省份、所在市、所在区县、所属行业、项目总投资、所处阶段、发起时间、项目示范级别/批次、回报机制……
	项目评分表	项目名称、项目编号、序号、指标名称、权重、评分标准、评分结果
	项目预算表	项目名称、项目编号、支出年度、本项目一般公共预算支出数额……
PPP 储备项目	储备项目基本信息表	项目名称、项目编号、所在省份、所在市、所在区县、所属行业、项目总投资、所处阶段、发起时间、项目示范级别/批次……
PPP 相关机构	机构基本信息	机构名称、机构简介、经营范围、注册成立时间、通讯地址、机构荣誉、机构资质
	机构项目咨询	机构名称、项目序号、项目名称、项目阶段、所在省份、所在市、所在区县、项目行业、项目总投资、示范级别、咨询服务业绩
	机构业绩	机构名称、年份、项目总数、项目总投资额、示范项目数、示范项目总投资额

数据模块	数据表	数据字段
	机构退库情况	机构名称、退库类型、退库日期、备注、退库原因、是否列入黑名单

5.数据范围

数据模块	数据起始时间
全部模块	2009 年

(五) 中国地方债数据库-CLGD

1.数据库简介

中国地方债数据库 (*Chinese Local Government Debt, CLGD*) 是基于中国地方政府发行的地方债、城投债等数据整理而成的专业债券研究数据库。*CLGD* 共分为三个模块, 分别为城投公司基本信息、城投债基本信息、地方债基本信息, 其中, 城投公司的确定采用银保监会公布的城投公司名单, 同时提供这些城投平台的基本工商信息; 进一步地, 基于银保监会城投平台公司名单匹配出城投债发行信息; 地方债信息则来自于官方发布的地方政府债券信息网站。

2.数据库特色

- 数据架构清晰

CLGD 依据城投公司、城投债、地方债三类分设三表, 结构布局简洁、清晰。

- 数据来源权威

CLGD 使用研究常用的银保监会城投平台名单, 并且地方债直接来自官方披露来源。

3.典型数据举例

地方债基本信息

发行人名称	债券简称	发行日期	兑付日期	票面利率 (%)	发行价格 (元)
江苏	20 江苏 22	2020-10-30	2035-11-02	3.8800	100.0000
江苏	20 江苏债 23	2020-12-04	2030-12-07	3.5300	100.0000
江苏	20 江苏 19	2020-10-16	2027-10-19	3.5300	100.0000
江苏	20 江苏债 18	2020-08-31	2030-09-01	3.2900	100.0000

4.数据库内容

数据模块	数据表	数据字段
城投公司 基本信息	城投公司 基本信息	企业名称、登记状态、法定代表人、注册资本、成立日期、核准日期、所属省份、所属城市、所属区县、电话、邮箱、统一社会信用代码、纳税人识别号、注册号、组织机构代码、参保人数、企业类型、所属行业、曾用名、企业地址……
城投债基 本信息	城投债基 本信息	代码、债券简称、发行日期、到期日、发行总额、利率、期限、发行价格、发行时主体评级、发行时债券评级、评级机构、增信情况、票面品种、付息利率品种、发行对象、发行人、债券类型
地方债基 本信息	地方债基 本信息	发行人名称、票面利率、基本利差、债券全称、债券期限、当期基础利率、债券简称、分期还本情况、首次划款日、债券代码、是否为续发行、上市流通日、发

	行日期、兑付日期、起息日、发行价格、计划发行量、到期日……
--	-------------------------------

5.数据范围

数据模块	数据起始时间
城投公司基本信息	截面数据
城投债基本信息	1999 年
地方债基本信息	2009 年

CNRDS

CNRDS

CNRDS

CNRDS

（六）财政预决算数据库-FBFA

1.数据库简介

财政预决算数据库 (*Financial Budget and Final Accounts Database, FBFA*) 是基于财政部省级预决算公开专栏以及各省市财政厅等提供的预决算资料而建立的专业数据库。*FBFA* 共分为财政预算和财政决算两个大模块, 每个模块均从一般公共财政、政府性基金、国有资本经营、社会保险基金四个方面反映了各省市的预算或决算情况。

2.数据库特色

- 数据来源权威

FBFA 数据来源于财政部以及各省市财政厅公布的公告, 权威可靠。

3.典型数据举例

财政预算_一般公共财政预算_一般公共财政收入预算

省份	收入项目	年份	预算数 (万元)
江苏省	其他收入	2015	420000
江苏省	国有资源 (资产) 有偿使用收入	2015	2070000
江苏省	罚没收入	2015	1220000

4.数据库内容

数据模块	数据表	数据字段
财政预算	一般公共财政收入预算	省份、收入项目、年份、预算数 (万元)
	一般公共财政支出预算	省份、支出项目、年份、预算数 (万元)
	政府性基金收入预算	省份、收入项目、年份、预算数 (万元)
	政府性基金支出预算	省份、支出项目、年份、预算数 (万元)
	国有资本经营收入预算	省份、收入项目、年份、预算数 (万元)
	国有资本经营支出预算	省份、支出项目、年份、预算数 (万元)
	社会保险基金收入预算	省份、收入项目、年份、预算数 (万元)
	社会保险基金支出预算	省份、支出项目、年份、预算数 (万元)
财政决算	一般公共财政收入决算	省份、收入项目、年份、决算数 (万元)
	一般公共财政支出决算	省份、支出项目、年份、决算数 (万元)
	政府性基金收入决算	省份、收入项目、年份、决算数 (万元)
	政府性基金支出决算	省份、支出项目、年份、决算数 (万元)
	国有资本经营收入决算	省份、收入项目、年份、决算数 (万元)
	国有资本经营支出决算	省份、支出项目、年份、决算数 (万元)
	社会保险基金收入决算	省份、收入项目、年份、决算数 (万元)
	社会保险基金支出决算	省份、支出项目、年份、决算数 (万元)

5.数据范围

数据模块	数据起始时间
全部模块	2013 年

四、CNRDS 基础数据库

基础数据是指在研究过程中必须用到的一系列基本数据，包括上市公司股价数据、财务数据、股权结构、公司治理数据等。这些数据为研究提供了各种原始资料，其质量高低可能会影响到研究结果的可靠性。因此，在有关中国问题的研究过程中，基础数据的涵盖范围、数据质量是学者关注的一个重点。**CNRDS** 参照国际标准并结合中国资本市场数据结构，基于 **CNRDS** 基础数据库整合了经济管理领域学术研究过程中可能用到的大部分研究基础数据。参阅各领域重点文献，依托已有的学术研究成果，**CNRDS** 平台根据研究需求对各类数据进行整合和重构，归纳了以下四个系列的基础数据：

上市公司股票基础数据 主要包括中国股票市场相关数据，例如股价数据、融资融券数据、股权融资数据。

上市公司财务基础数据 主要包括中国上市公司相关财务数据，例如财务报表数据、盈利预测数据、财务指标数据等。

上市公司治理基础数据 包括与公司治理相关的上市公司数据，例如关联交易、内部人交易、高管薪酬等重要数据。

经济研究基础数据 主要涵盖了经济研究所需要的基本研究数据，例如来自国家统计局的基本宏观经济数据和区域经济数据等。

CNRDS 基础数据库从数据生产、数据整理和数据呈现均有一套严格的标准和程序。首先，基础数据库的字段构成必须由相关领域的专家学者讨论通过；其次，数据库体系必须符合学术研究需求和学者浏览、下载数据的习惯；最后，基础数据的数据质量必须经过多次核查方可最终上线。

CNRDS 基础数据库具有以下三个特点：

- 数据库设计和布局合理，契合使用习惯
- 数据严格审核，数据质量较高
- 多维度基础数据，满足不同研究需求

上市公司股票基础数据

（一）中国上市公司股价数据库-CNSP

中国上市公司股价数据库 (*Securities Prices of Chinese Public Firms, CNSP*) 是以研究为目的而设计的, 对来自上海证券交易所和深圳证券交易所中国上市公司全部交易数据进行全面深入研究, 旨在提供完整精确的股价信息。本数据库提供准确、真实、全面的中国上市公司股价信息, 并结合中国股票市场自身特点进行校正。目前, 本数据库主要包含了两个数据模块, 即个股回报率和市场回报率。

个股回报率 分为三个子模块: 个股日回报率、个股周回报率、个股月回报率。该模块主要提供中国上市公司来源于权威信息源的股票信息, 严格按照 *CRSP* 国际标准数据库的调整技术计算个股回报率, 以日、周、月进行详细划分, 有利于对任一股票的行情做出合理判断。

市场回报率 分为三个子模块: 市场日回报率、市场周回报率、市场月回报率。该模块主要提供中国上市公司来源于权威信息源的股票信息, 严格按照 *CRSP* 国际标准数据库的调整技术计算市场回报率, 以日、周、月进行详细划分, 有利于对整个市场股票的行情做出合理判断。

（二）中国股票异常交易数据库-CAST

中国股票异常交易数据库 (*Chinese Abnormal Stock Trading Database, CAST*) 是以研究为目的而设计的专业学术研究数据库, 旨在为国内外进行中国股票市场交易情况研究的学者与研究人员提供专业、准确的学术研究数据, 本数据库收集整理了 2006 年以来的中国股票市场异常交易的数据情况。

本数据包含两个数据模块, 分别是: 交易公开信息列表与交易公开信息明细共两张表格, 22 个字段内容, 其中包括进行买卖行为的交易方以及交易成交的金额与数量情况。

（三）中国特殊处理股票数据库-CSTS

中国特殊处理股票数据库 (*Chinese Special Treatment Stock Database, CSTS*) 收集整理了 1998 年以来中国沪深两市受到特殊处理的股票数据。数据库包括股票代码、公司简称、特殊处理前的公司简称、特殊处理后的公司简称、处理类别、处理类别编号、信息发布日期、变动原因、以及处理实施日期等多个字段。

（四）中国融资融券数据库-CMTD

中国融资融券数据库 (*Chinese Margin Trading Database, CMTD*) 是基于对我国证券市场融资融券行为的分析而建立的专业数据库。该库不仅包括个股融资融券的详细信息, 还包括市场统计、券商统计、标的物 and 担保物等多方统计信息。

CMTD 包含个券交易明细、个券交易市场汇总、券商交易信息、标的物 and 担保物、市场交易统计 5 个模块, 7 张表, 共 50 多个字段。其中, 个券交易明细包括每天各个证券的融资融券具体情况; 个

券交易市场汇总是每天不同交易所融资融券情况的汇总；券商交易信息统计了每个月每个券商的融资融券情况；标的物 and 担保物列示了融资融券过程中涉及的标的物 and 担保物信息；最后，该数据库还统计了沪深两市的整体融资情况和融券情况。*CMTD* 数据库包括 2010 年至今沪深交易所的所有融资融券信息。

（五）中国公司 IPO 综合数据库-CIPO

IPO 综合数据库 (*IPO Integrated Database of Chinese Companies, CIPO*) 是基于对公司 IPO 上市全过程的深入考察和了解建立的详细、准确和全面的公司 IPO 研究专业数据库。该库几乎涵盖了中国公司上市前、上市时和上市后的全部活动和参与者信息，不仅包括已经成功上市的 IPO 公司情况，还涉及多次报审以及审核未通过的公司。

（六）中国上市公司增发与配股数据库-CSEO

中国上市公司增发与配股数据库 (*SEOs and Share Allotment Database of Chinese Companies, CSEO*) 收集整理了 1991 年以来中国上市公司增发股票与配股的数据，建立的详细、准确和全面的上市公司权益再融资研究学术数据库。该库涵盖了中国上市公司增发股票所涉及的几乎所有重要指标。

上市公司财务基础数据

（一）上市公司业绩预告数据库-CEPD

上市公司业绩预告数据库 (*Estimated Performance Database of Listed Companies, CEPD*) 是基于上市公司业绩预告和业绩快报公告整理而成的专业数据库。**CEPD** 包含上市公司业绩预告、上市公司业绩快报两张表, 包含了上市公司正式披露季报、半年报、年报之前, 交易所对上市公司提前预告经营业绩和披露业绩快报情况的披露, 反映了上市公司基本面的最新重大变化, 为该领域研究提供了全面、高质量的数据支持。

（二）财务报告披露时间数据库-FRDT

财务报告披露时间数据库 (*Financial Report Disclosure Time Database, FRDT*) 包含了 1999 年以来中国 A 股和 B 股上市公司的年报、中报和季报的预披露时间、更正时间和实际披露时间等多个重要时间字段。

（三）中国公司财务报表数据库-CNFS

中国公司财务报表数据库 (*Financial Statements of Chinese Firms, CNFS*) 是以研究为目的而设计的, 对 186 个重要会计科目进行全面深入研究, 旨在提供完整精确的财务信息。**CNFS** 以 2006 年财政部出台的企业会计准则为基本框架, 包含了中国自 1990 年以来公布的各公司全部财务信息, 数据库的设计基本遵循企业会计准则规定的合并报表列报格式, 适用于所有行业, 并能往前兼容历史数据。目前, 本数据库包含三个数据模块, 即资产负债表、利润表和现金流量表。

（四）上市公司财务报表附注数据库-NFSD

上市公司财务报表附注数据库 (*Notes to Financial Statements Database, NFSD*) 是基于上市公司年报数据, 对报表附注数据进行加工整理而成的专业数据库。根据报表附注中的披露项目, **NFSD** 设置了资产项目、应收项目、债务项目、费用明细、损益项目、其他项目等 6 个方面的财务数据, 可为财务报表使用者提供完整、准确的报表附注数据支持。

（五）上市公司财务指标数据库-CNFI

上市公司财务指标数据库 (*Financial Indicator Database of Listed Companies, CNFI*) 是为研究中国上市公司财务状况和经营成果而建立的专业数据库。本数据库的数据建立在 **CNDRS** 中国公司财务报表数据库的基础上, 包含了中国自 1991 年以来公布的各上市公司财务指标信息, 数据库的设计基本遵循企业会计准则规定的合并报表列报格式, 根据相关的会计制度、会计准则做必要的调整, 适用于所有行业, 并能往前兼容历史数据。**CNFI** 提供了偿债能力、盈利能力、成长能力、每股指标、

杜邦分析、现金流分析、经营能力等 7 个方面的财务指标。

（六）中国上市公司盈利预测数据库-CEFD

中国上市公司盈利预测数据库 (*Earnings Forecast Database of Chinese Companies, CEFD*) 是以研究为目的设计开发的学术数据库。

本数据库共由两个模块组成：分析师荐股评级与分析师盈利预测。两个模块的具体字段内容均是提取自 30 多万份分析师报告，包括有具体的分析师标准化荐股意见以及各相关主要指标与比率的预测情况，字段内容丰富、数据质量高、信息全面，便于研究者与学者的分析使用。

上市公司治理基础数据

（一）上市公司基本信息数据库-CBID

1.数据库简介

上市公司基本信息数据库 (*Basic Information Database of Listed Companies, CBID*) 包含了 1990 年以来在中国 A 股和 B 股市场上市或者曾经上市的公司信息。*CBID* 提供了这些上市公司的股票代码、经营范围、注册资本、所属行业、曾用名等多个维度的重要基本信息。

（二）中国公司治理数据库-CCGD

中国公司治理数据库 (*Chinese Corporate Governance Database, CCGD*) 是基于对中国公司治理研究的深入考察而建立的专业数据库, 包括了研究中国公司治理问题几乎所有的指标数据, 能够为学者提供全面、准确的研究数据。*CCGD* 包括 3 个模块, 分别为股东治理、管理层治理、外部治理。

（三）审计意见与费用数据库-AUDIT

审计意见与费用数据库 (*Audit Opinion & Expenses Database of Chinese Companies, AUDIT*) 包括审计师事务所对上市公司出具的审计意见与收取的审计费用等信息。具体包括上市公司聘请的注册会计师事务所及审计师姓名信息、出具的审计意见和支付的审计费用等; *AUDIT* 收录了 1990 年至今所有中国上市公司的审计意见与审计收费信息。

（四）中国上市公司管理层变更数据库-MTDB

中国上市公司管理层变更数据库 (*Management Turnover Database of Chinese Companies, MTDB*) 系统全面地统计了与上市公司管理层岗位变更和持股变更相关的重要数据信息。数据库涵盖了高管离职时间、离职日期, 高管持股变动日期、关联人、变动股份等。

（五）公司与高管违规处罚数据库-VPCE

公司与高管违规处罚数据库 (*Violation and Punishment of Chinese Companies & Executives, VPCE*) 主要基于资本市场监管部门对上市公司、上市公司高管和股东等的违规处罚公告而设计的专业数据库。*VPCE* 主要包括两张表, 上市公司违规与处罚表主要涵盖监管部门对上市公司违规行为做出处罚的数据, 上市公司高管违规与处罚则主要包括因公司违规而受处罚的高管, 以及违规不涉及公司而单纯属于高管个人违规的数据。在制作过程中, 我们对违规事件进行了编号, 可以在这两张表格中进行交叉核验。此外, *VPCE* 还对被处罚高管进行了编号, 该编号可以在其他人物数据库进行匹配, 便于研究。*VPCE* 收录了 1997 年至今所有的违规处罚数据。

（六）高管薪酬与激励数据库-ECEI

高管薪酬与激励数据库 (*Executive Compensation & Equity Incentive Database, ECEI*) 系统全面地统计了与上市公司管理层薪酬相关的重要数据信息。数据库不仅统计了管理层薪酬信息，同时还 将公司股权激励方案与股权激励实施明细模块纳入数据库内。

（七）中国上市公司关联交易数据库-CRTD

中国上市公司关联交易数据库 (*Related Transactions Database of Chinese Companies, CRTD*) 搜集整理了 2000 年以来的上市公司关联方交易数据。该数据库包括了关联交易的公司简称、会计年度、公告日期、关联方和关联交易金额等相关信息。

（八）中国公司股权研究数据库-CERD

中国公司股权研究数据库 (*Equity Research Database of Chinese Company, CERD*) 是基于对中国上市公司股权结构和股东信息的深入考察，从而建立的详细、准确和全面的公司股权研究学术数据库。**CERD** 包含股权结构、股权变动、股份质押与冻结、股东持股、股份回购 5 个模块，共 9 张表。其中，股权结构包括股本结构、实际控制人；股权变动包括股本变动、重要股权交易；股份质押与冻结包括相应股份的质押与冻结情况；股东持股包括大股东持股、机构持股；最后，股份回购模块展示了公司的股份回购情况。

（九）机构投资者持股研究数据库-IORD

机构投资者持股研究数据库 (*Institutional Ownership Research Database, IORD*) 包括机构持股与机构重仓股两大模块。机构持股模块主要包括公司基本信息、各大机构持股数量、各大机构持股比例以及持股机构的数量等字段。

（十）中国公司股利分红数据库-CCDD

中国公司股利分红数据库 (*Dividend Database of Chinese Companies, CCDD*) 基于对上市公司股利公告的全面收录，结合学术研究领域的需求建立了完整和详细的专业数据库。该库搜集整理了 1991 年至今中国上市公司全部的利润分配数据，涵盖了会计年度、预案内容、实施进度以及股本基准日、股权登记日、除权除息日、派息日、红股上市日等多维度数据。

（十一）中国上市公司诉讼仲裁数据库-CLAD

中国上市公司诉讼仲裁数据库 (*Litigation & Arbitration Database of Chinese Companies, CLAD*) 是以研究为目的而设计的, 旨在提供完整、精确的中国上市公司诉讼仲裁数据。该数据库提供上市公司的诉讼仲裁信息, 不仅详细罗列了一审、二审的诉讼情况, 还对后续执行情况进行了跟踪, 保障了信息的完整性和有效性。

（十二）中国上市公司内部人交易数据库-CITD

中国上市公司内部人交易数据库 (*Insider Trading Database of Chinese Companies, CITD*) 搜集整理了 1996 年以来的上市公司内部人股权交易数据。该数据库由 19 个字段组成, 包括内部人交易的公司简称、公告日期、内部人交易涉及的股东和高管信息以及内部人股权交易的具体信息。

经济研究基础数据

（一）中国宏观经济研究数据库（年度）-MACRO

中国宏观经济研究数据库（年度）（*Chinese Macro-economic Research Database, MACRO*）是以研究为目的设计的专业学术研究数据库，旨在为国内外从事宏观经济研究的学者和研究人员提供完整、可靠的学术研究数据。本数据库包含综合；国民经济核算；人口；就业人员和工资；固定资产投资和房地产；对外经济贸易；能源；财政；价格指数；人民生活等 28 个方面的经济数据指标。

（二）中国宏观经济研究数据库（季度）-MACROQ

中国宏观经济研究数据库（季度）（*Chinese Macro-economic Research Database (Quarterly), MACROQ*）包含：国民经济核算、农业、工业、建筑业、人民生活、价格指数、国内贸易 7 个方面的经济数据。

（三）中国宏观经济研究数据库（月度）-MACROM

中国宏观经济研究数据库（月度）（*Chinese Macro-economic Research Database (Month), MACROM*）包含以下 14 个方面的经济数据指标：价格指数、工业、能源、固定资产投资（不含农户）、服务业生产指数、城镇调查失业率、房地产、国内贸易、对外经济、交通运输、邮电通信、采购经理指数、财政、金融。

（四）中国区域经济研究数据库-CRED

中国区域经济研究数据库（*Chinese Regional Economy Database, CRED*）从 28 个方面反映了我国 31 个省市的区域经济与社会发展状况，囊括了各种类型的区域经济数据指标。*CRED* 包含：综合；国民经济核算；人口；固定资产投资和房地产等 28 个方面的经济数据指标，涉及 31 个省市。

（五）中国债券研究数据库-BOND

中国债券研究数据库（*Chinese Bond Research Database, BOND*）涵盖债券市场各方面信息，不仅收录债券及债券发行主体的信息，还对债券发行时的重要信息进行汇总，同时也整理了债券的信用评级及债券发行主体的财务报表信息。

五、最新学术动态

(一) 模块简介

为了使广大学者便于跟踪国内外权威学术期刊最新论文等信息，**CNRDS** 开辟了“最新学术动态”一栏。该栏目将定期及时更新国内外顶级财经期刊的文章、国内外学术会议信息。

(二) 模块功能

1.及时更新国内外各类权威期刊学术动态

CNRDS 可为您提供便于浏览的学术文章题目、摘要、链接地址等重要信息。该栏目将及时更新来自 20 本会计金融类国际顶级期刊上的最新论文列表，以及来自 13 本中国财经类权威期刊的最新论文列表。

2.总结近期国际期刊上的中国问题研究文章

来自 20 本会计金融类国际顶级期刊上的中国问题研究文章，可为学者浏览、查阅最前沿的学术论文，扩展研究思路提供有益参考。

六、数据定制

(一) 定制项目

1. 数据咨询

该服务主要为海内外高校学者免费提供中国学术研究所需数据的来源以及数据质量的咨询和建议。根据您提出的问题，我们将比较中国已有数据库之间的差异，找出最适合您需求的数据库；针对某些您不确定在中国是否存在的数据或者数据来源不明确的数据，我们将为您提供免费的数据咨询服务。我们的咨询团队都是由国内一流的数据和财经领域的专家组成，他们不仅长期关注中国资本市场的发展，还与证监会、银监会、财政部等中国资本市场监管部门保持良好的交流与联系。他们所提供的咨询服务能为您节约研究成本，提高研究效率。

2. RA 服务

我们可为国内外学者提供收集与整理数据服务，我司专业、高效的 RA 团队能为您及时提供准确的研究数据。您既可以将您需要搜集的数据项目告诉我们，也可以将您已有的数据项目提供给我们，我们将秉承准确性、全面性、及时性、保密性原则，在第一时间为您收集和整理出电子化的标准格式数据。

3. 学术数据定制

该服务将根据您在研究和实务中对独特数据的需要，为您提供专业的财经类数据查找、搜集、整理等定制服务。我们拥有长期、稳定的数据采集渠道，高效、可靠的数据整理团队，采用网页爬虫、数据抓取、大数据标注、人工搜集和整理等手段和方法，能够在第一时间获取您所需的定制数据。

(二) 定制流程



1. 提交数据定制需求

您可通过电话、邮件等方式提交您的数据定制需求。联系邮箱：customerservice@cnrds.com；联系固话：021-66181082。

2. 评估与反馈

我司对您的数据定制需求进行评估，并会在 3 个工作日内把数据可获得性、数据样本、报价以及完成所需时间等信息反馈给您。

3. 签订合同

我们双方之间就数据、价格、交付时间等方面达成一致后，我司拟定数据定制电子版合同电邮给您。最终确定好合同各项条款后，双方通过邮寄或扫描等方式签订正式的数据定制合同。

4.付款

我司可以配合您的工作单位科研经费的报销流程，出具合同、发票等财务报销相关资料。

5.交付数据

我司按合同约定的时间、交付方式交付数据。

6.售后跟踪

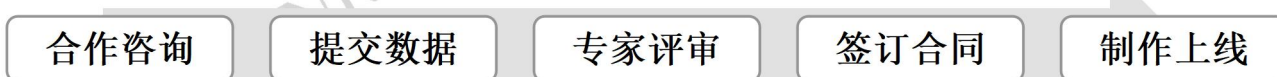
交付数据后，为了了解您对定制数据的使用情况，我司会通过邮件、电话或实地拜访等对您进行回访，以便及时、完整地解答您使用数据过程中所遇到的问题。

七、数据合作

(一) 数据合作

我们诚邀各专业领域的数据库公司、相关领域的专家和学者与 **CNRDS** 平台进行数据合作，我们相信每一位学者在支持和推动中国研究基础数据和资源构建上迈出的小小一步，都将成为中国学术研究工作的一大步。

(二) 合作流程



1. 合作咨询

若您有需要商业化的特色学术数据，可与我们联系，并咨询我司数据研发部工作人员，联系邮箱：liuchaoran@cnrds.com

2. 提交数据

我们双方确定好合作意向后，您可通过邮件等方式发送合作数据的样本给我司。

3. 专家团队评审

我方邀请相关领域的专家和学者对您的数据合作项目进行匿名审核。我方会在 10 个工作日内把专家团队的匿名评审结果反馈给您。

4. 签订合同

我们双方之间就数据合作各项细节达成一致后，我方拟定数据合作合同电邮给您，最终确定好合同各项条款后，双方通过邮寄或扫描等方式签订正式的数据合作合同。

5. 制作数据库并上线

您有以下两种数据呈现形式：第一，数据内容由贵研究团队独立制作，我方提供技术支持和形式审核以及后续维护；第二，贵研究团队进行顶层设计，我方进行数据库内容搜集和整理，并提供制作和后续维护。最后，我司将制作好的数据库发布在 **CNRDS** 平台上。