

# 附录

## 正文未报告部分

### 附 1 数字技术识别的说明

本文通过文本技术对专利名称和摘要进行识别能一定程度上增加数字专利识别的完整性和准确性。具体说明和示例如下：

目前，关于数字技术相关领域专利的识别主要有两种方法：①国际专利IPC分类号（陶锋等，2023）；②文本分析法（黄勃等，2023；刘青和肖柏高，2023）。还有两种方法的综合应用，如Chen et al.（2019）先限定了专利分类号大类，再在其中通过文本分析的方法识别金融科技相关专利；罗佳等（2023）先根据数字技术关键词检索出所有相关专利并获得IPC分类号，再利用分类号识别数字专利。

仅利用IPC分类号识别存在较大偏误，尤其对于数字技术应用类的创新。例如“车联网”技术，据国家知识产权局依据《数字经济及其核心产业统计分类(2021)》制定的《数字经济核心产业分类与国际专利分类参照关系表(2023)》显示，该技术包含于大类“01 数字产品制造业”中的“010405 智能车载设备制造”，国际专利分类为“B60R\*”。如果仅在国家知识产权局中检索该分类号，会出现与数字技术不相关的技术，如公开号为“CN117022166A”的专利“一种汽车防撞梁”，其IPC分类号为“B60R19/34”，摘要表述为：

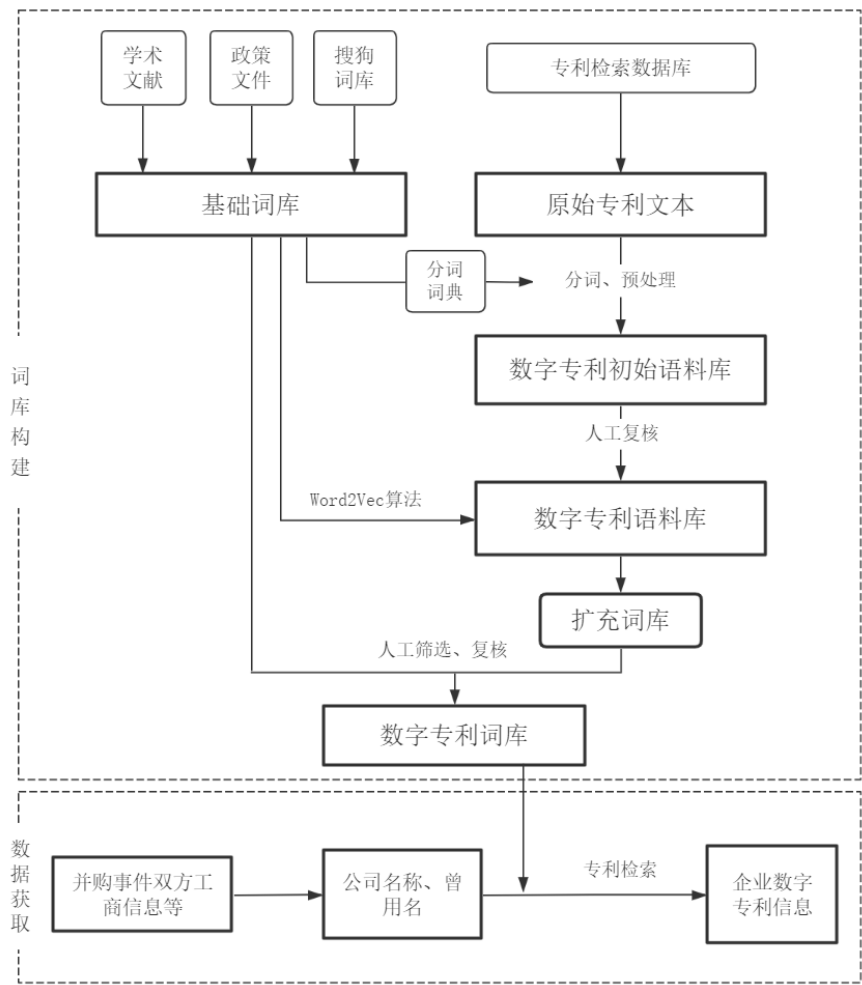
“本发明提供了一种汽车防撞梁，其包括防撞梁本体、两个吸能盒、两个安装板，所述防撞梁本体的后侧面的一端通过一个所述吸能盒与一个所述安装板相连，所述防撞梁本体的后侧面的另一端通过另一个所述吸能盒与另一个所述安装板相连；所述吸能盒包括上U型板和下U型板，所述上U型板的开口与所述下U型板的开口相对设置，所述上U型板的侧板的内壁面用于与所述下U型板的侧板的外壁面相贴合；所述上U型板的底板与所述下U型板的底板之间的竖直距离可调。本发明的汽车防撞梁具有较好地通用性，能够有效地降低车身的开发成本。”

这种情况下，利用专利文本进行识别可能更加精确，如刘青和肖柏高（2023）通过专利摘要语义判断出其是否属于劳动节约型技术。

此外，还有部分数字技术的IPC分类号并未纳入《数字经济核心产业分类与国际专利分类参照关系表(2023)》。例如，中国船舶工业股份有限公司（600150.SZ）于 2020 年 2 月 26 日公告并购的目标企业中拥有的专利“一种基于机器视觉的零件分拣方法及系统”（公开号CN110180799A，IPC分类号B07C5/34；B07C5/36），其分类号并未包含在文件中，但该专利的确涉及了数字技术。

附 2 数字技术创新词库构建步骤

数字技术创新词库构建和数据获取步骤如附图 1 所示：



附图 1：数字技术创新词库构建和数据获取流程

1. 基础词库构建

本文基础词库的构建基于三个渠道：学术文献、政策文件以及搜狗词库，各部分词库词汇数量见附表 3。基础词库具体构建步骤如下：

**步骤一：**参考一系列与文本相关的数字化主题文献（李春涛等，2020；吴非等，2021；张叶青等，2021），归纳出与数字技术相关的关键词。

**步骤二：**由于上述渠道词语来源存在对某些领域的偏向，可能会影响数字专利词汇拓展的全面性，本文还考虑了政府官方对数字技术的定义和描述。具体的，本文依据国家统计局颁布的《数字经济及其核心产业统计分类（2021）》、国务院制定的《“十四五”数字经济发展规划》以及国务院颁布的《数字中国建设整体布局规划》等三十余份数字经济政策的核心内容，经 Python 文本分词和人工标注的方式整理出数字技术相关词汇。

**步骤三：**在基础词库中纳入了数字技术相关的搜狗词库，主要包含“人工智能专业术语【官方推荐】”、“人工智能常用词库”、“区块链”、“深度学习”、“计算机名词”、“自然语言处理及计算语言学相关术语”、“网络工程”等词库。需要说明的是，在纳入前我们进行了人工核查筛选，剔除了一些可能对之后专利检索存在歧义的词汇（如“参数”、“表征”、

“知识库”等)，也删除了一些与数字专利不相关的词语（如一些互联网公司名称、特定应用软件名称等）。最后去重，得到该部分词库。

**步骤四：**对上述词汇进行进一步筛选。如果上述词汇未在语料当中出现，我们进行剔除。然后加入 Python 第三方软件包 jieba 自带的中文分词库，以增加分词的准确性。

整体来看，基础词库不同来源的词语类型具有不同特征，如附表 1 所示，学术文献的词语较为“概念化”，主要涉及数字技术应用领域；政策文件关注到了数字基础技术；搜狗词库拓展了各个领域更加详细的技术。三个来源起到很好的互补作用，为之后利用 Word2Vec 算法拓展词库提供了较为全面的“种子词”。

附表 1：基础词库示例

学术文献	政策文件	搜狗词库
人工智能、商业智能、大数据、异构数据、云计算、类脑计算、物联网、工业互联网、区块链、分布式计算等	射频芯片、无线通讯、通信系统、量子通信、工控机、中央处理器、GPU、半导体、光刻机、网络安全等	迁移学习、贪婪算法、遗传算法、潜在语义分析、公钥、哈希锁定、网络拓扑、边界网关、报文加密、通信协议等

## 2. 数字专利文本语料获取

拓展数字专利词需要合适的专业性文本，我们选择数字专利的名称和摘要作为文本语料。这些数字专利语料来自全球数字专利统计数据中的中国数字技术专利申请排名前十企业的专利文本，共计 415410 条专利文本。

为了拓展更精确的词汇，我们借鉴 Chen et al. (2019) 的方法，将专利 IPC 分类号的部限定在“G”和“H”，因为该类别涵盖了与数字技术相关领域的专利。然后，本文将专利名称和摘要作为数字专利词库构建的原始语料，对其进行分词处理，并且去除“一种”“及其”“方法”等无创新意义的停用词，形成处理后语料库。

## 3. 词库扩充

我们进一步利用 Word2Vec 算法从数字专利文本语料中训练词语向量并计算词语相似度，从而提取与前文基础词库词语高度相关的词语，并进行筛选，以此实现发现新词并扩充词库的目的。具体步骤和技术细节如下：

**步骤一：**Word2Vec 模型构建。Word2Vec 算法由 Mikolov et al. (2013) 提出，基于词嵌入（Word Embedding）方法，对文本的上下文中词汇向量化。该方法相较于传统词汇表示法的优势之一在于可以根据向量计算词语之间的相似度。借鉴已有研究，我们利用 Word2Vec 的 Skip-gram 模型对处理后的语料库中进行词向量训练。Mikolov et al. (2013) 和姜富伟等 (2021) 认为 Skip-gram 模型较 CBoW (Continuous Bag-of-words Mode) 具有更准确的预测力，同时考虑到专利文本中存在低频的专业术语，因此，本文使用 Word2Vec 的 Skip-gram 模型。已有研究中对长文本的 Word2Vec 模型参数设置较高，考虑到专利摘要属于短文本，我们将向量维度设置为 100 维（vector\_size=100），窗口大小为 5（window\_size=5）。

**步骤二：**词库扩展。本文拓展词库的方式为根据词语向量间的余弦相似度进行选择，余弦相似度越大说明两个词语越相关，对于 A、B 两个词语相似度计算方式为：

$$\text{similarity} = \frac{A \times B}{|A| \times |B|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

我们选择了与基础词库中各个词语相似度前 15 的词汇或相似度大于 0.8 的词汇，原因是专利摘要是短文本，选择过少的词汇可能导致词库不全面，选择过多数量的词汇可能

导致词库质量下降。表 2 展示了三个词相似度前 20 的例子。然后，我们将所有扩充后的词进行了去重和人工筛选，主要有两种情况，第一种是拓展词包含了扩充词，例如附表 2 中，与人工智能高度相似的词有“人工智能技术”、“AI 人工智能”，这类词需要剔除；第二种是不相关的词或存在歧义的词，例如“交互式”。在去重和筛选后，我们得到扩充词库。

附表 2：词库拓展示例

输入： 基础词	人工智能		射频芯片		迁移学习	
输出：	相似词汇	余弦相似度	相似词汇	余弦相似度	相似词汇	余弦相似度
1	AI	0.837203	射频天线	0.866116	监督学习	0.902624
2	人工智能技术	0.771487	NFC 芯片	0.814568	域适应	0.891347
3	AI 智能	0.745322	RFID 芯片	0.800473	半监督	0.891143
4	智能决策	0.740446	电子标签天线	0.792717	集成学习	0.874537
5	AI 人工智能	0.719849	RFID 天线	0.791958	度量学习	0.873881
6	云计算	0.713692	电子标签芯片	0.780426	多任务学习	0.865856
7	人工智能学习	0.705848	超高频 RFID	0.779220	增量学习	0.865472
8	深度学习	0.690341	UHFRFID	0.772641	深度神经网络	0.857353
9	交互式	0.685981	智能卡芯片	0.772231	深度学习	0.855221
10	智慧校园	0.685647	NFC 天线	0.762804	知识迁移	0.854692
11	信息技术	0.684155	无线射频	0.759624	代价敏感	0.852964
12	智慧养老	0.682082	蓝牙天线	0.757909	深度卷积	0.852886
13	智慧旅游	0.677765	读卡芯片	0.756841	深层卷积	0.850167
14	智慧医疗	0.675736	读写器	0.755695	神经网络学习	0.848617
15	智能机器人	0.674230	射频识别系统	0.755632	知识蒸馏	0.847053
16	智慧	0.673472	射频收发	0.755378	学习分类器	0.845729
17	人工智能语音	0.672773	超高频天线	0.754574	神经网络训练	0.844922
18	物联网	0.669994	RF 芯片	0.752185	人群计数	0.844192
19	机器学习	0.668300	天线射频	0.742171	语义分割	0.841953
20	健康管理	0.665619	射频识别	0.740012	字典学习	0.840782

**步骤三：整合词库。**本文将基础词库与扩充词库整合，再次去重整理，通过多次复核并经过专家复核后得到最终的数字专利词库，共包含 671 个词语。

附表 3：词库来源情况

来源		筛选后词语数	原词语数	筛选后词语数占 最终词库比例
基础词库	学术文献	78	91	11. 62%
	政策文件	109	221	16. 24%
	搜狗词库	116	373	17. 29%
拓展词库		368	4704	54. 85%
合计		671	5389	

附3 数字技术创新分类界定及方法

1. 数字技术创新类型界定

对数字技术的分类我们重点参考了：①国家政策文件，如《数字经济及其核心产业统计分类（2021）》数字经济活动的大类和中类技术分类（包含了5个大类32个中类）；②专业智库研究，如中国信息通信研究院每年发布的ICT深度观察和系列白皮书。其中，ICT深度观察报告将相关产业划分成了ICT产业、信息网络、先进计算、大数据与人工智能、网络安全等八大领域，并对具体技术进行了阐述。白皮书系列（包括历年《中国数字经济发展研究报告》、《云计算白皮书》、《大数据白皮书》等）也对相应技术进行了分类和描述；③相关学术研究，如吴非等（2021）将企业数字化内容分为了人工智能、区块链、大数据、云计算、数字技术应用五个类型。同时，结合中国具体实践，我们对数字技术进行了如下分类（附表4总结了这些数字技术类别与每个类别相关的关键技术）：

人工智能技术：包含支持智能制造、计算机视觉、自然语言处理、智能语音等领域的底层算法（如机器学习、深度学习等）、前端技术（知识图谱、语义处理等）和综合应用（智能机器人、实时翻译等）。

云计算技术：包含了从云存储（NAS存储、DAS存储等）到云服务（IaaS、PaaS、SaaS等）的云计算相关技术，以及其他相关的计算技术（分布式计算、网格计算、边缘计算等）。

安全技术：包含了加密技术（公/私钥、椭圆曲线加密算法、密文映射等）、网络安全（如DDoS攻击防御、报文过滤技术等）、区块链技术（公有链、共识算法、智能合约等）在内的相关技术及应用。

大数据技术：包含了从数据存储（除云技术存储之外的数据存储）、数据采集技术（网络爬虫、容器编排等）、到大数据分析技术（数据挖掘、数据可视化等）、再到大数据应用（个性化推荐、用户画像等）。

通信网络技术：包含了通信技术（3G/4G/5G、卫星通信、光纤通信、GPS定位等）、网络网联技术（互联网、IPv6、以太网、物联网、传感技术等）以及两者的结合技术（移动支付、车联网、智能家居、智能可穿戴设备等）。

电子芯片技术：包含了上述技术的硬件技术，如电子芯片（GPU、CPU、类脑芯片、基带芯片等）、半导体技术（离子注入、光刻技术、刻蚀技术等）、微电子制造技术（ASIC、CPLD、现场可编程逻辑器件等）等相关基础技术。

附表4：数字技术类别及其关键技术

类别	关键技术
人工智能技术	人工智能、工业机器人、智能机器人、机器学习、深度学习、机器视觉、自然语言处理、知识图谱、图像识别、生物识别技术等
云计算技术	云计算、云原生、云边协同、公有云、混合云、分布式存储、效用计算、并行计算、网格计算、类脑计算等
安全技术	区块链、共识机制、数字签名、公/私钥、加密/解密技术、密文映射、攻击预防技术、入侵检测技术等
大数据技术	数据采集、容器编排、数据挖掘、文本挖掘、数据分析、数据可视化、用户画像分析、数字营销、个性化推荐等
通信网络技术	移动通信、近场通信、射频技术（RFID）、卫星通信、移动支付、工业互联网、物联网、车联网、智能家居、智能可穿戴设备等
电子芯片技术	类脑芯片、基带芯片、专用集成芯片、CPU、GPU、半导体技术、刻蚀技术、光刻技术、现场可编程逻辑器件等



## 1. 数字技术创新类型分类方法

针对数字专利训练样本的分类具体步骤和说明如下：

**第一步，数字技术类型和训练样本创建。**本文在 Chen et al. (2019) 的研究设计，增加了训练样本比例进行人工标注，在所有数字专利中随机选择了 2200 条样本进行手动分类，构成了训练样本。本文人工标注参考类似研究的方法（林建浩等，2019；佟岩等，2020），首先，我们与专家讨论后，初步生成了一套分类标注逻辑。然后，我们进行了两人“背对背”标注，将标注结果交由第三人，由第三人对有分歧的专利进行重新判断，并向专家咨询有争议的分类结果，再进行修正。最后，我们将标注训练样本交给专家，专家进行抽样检查后未发现问题。

需要说明的是，大部分专利能明确被分类，但存在涉及多类技术的专利，此时我们的判断方式主要为根据该专利核心创新部分或最终的应用场景划分到相应的类别，同时辅以《数字经济核心产业分类与国际专利分类参照关系表(2023)》进行判断。

例 1：

本发明涉及一种用于核验公民身份信息的多功能移动警务终端，包括集成微处理器、无线通讯模块、指纹识别模块、二代身份证识读模块、音视频模块、安全验证模块、人机交互模块即触摸屏、USB 识别模块以及供电模块；无线通讯模块、指纹识别模块、二代身份证识读模块、音视频模块、安全验证模块、人机交互模块、USB 识别模块分别与集成微处理器、电源模块连接；安全验证模块与指纹识别模块、视频模块、无线通讯模块连接；无线通讯模块包括 GPS 通信模块、Wi-Fi 通讯模块和 4G 通讯模块，本发明实现二代身份证核验比对，指纹信息核验比对，公民头像信息核验，现场录音录像取证功能，具有部署灵活、使用方便，价格低廉的特点，有效提高警务活动效率，增加执法能力。

该专利摘要包含了多个类别技术，如人工智能技术（指纹识别）、通信网络技术（无线通讯）、电子芯片技术（集成微处理器），专利内容主要是应用包含的各模块的成熟技术，重点在于形成了移动终端，需要各模块之间网联技术，初步可以归为通信网络技术。同时，参考该专利的 IPC 主分类号大类在《数字经济核心产业分类与国际专利分类参照关系表(2023)》对应的数字经济核心产业各类名称和关键词。该专利 IPC 主分类号大类为“G06K”，对应文件中产业小类名称“可穿戴智能设备制造”、“记录媒介复制”、“应用软件开发”、“物联网技术服务”。可以看到文件中类别名称更加侧重于传感、物联网技术，我们将该技术归为通信网络技术。

对于 IPC 没有在《数字经济核心产业分类与国际专利分类参照关系表(2023)》出现的技术，我们先根据专利摘要初步判断类别，然后根据该专利的专利说明书内容核实其主要创新点，主要利用专利说明书中技术背景、内容和优势（若有）。例如公开号为 CN109731254A 的专利（IPC 分类号为 A62C3/00 A62C37/00 A62C35/68），摘要内容为：

本发明公开了一种地下管廊灭火系统，属于消防技术领域，其特征在于：包括灭火设备系统、数据传感监测系统、网关通信系统、后端数据处理平台；所述灭火系统包括若干的灭火节点，在灭火节点上设置灭火设备；所述数据传感监测系统包括若干监测点和监测终端，监测点上设置传感监测模块；所述传感监测模块通过网关通信系统连接监测终端，所述监测终端通过网关通信系统连接后端数据处理平台。本系统可大大降低消防设

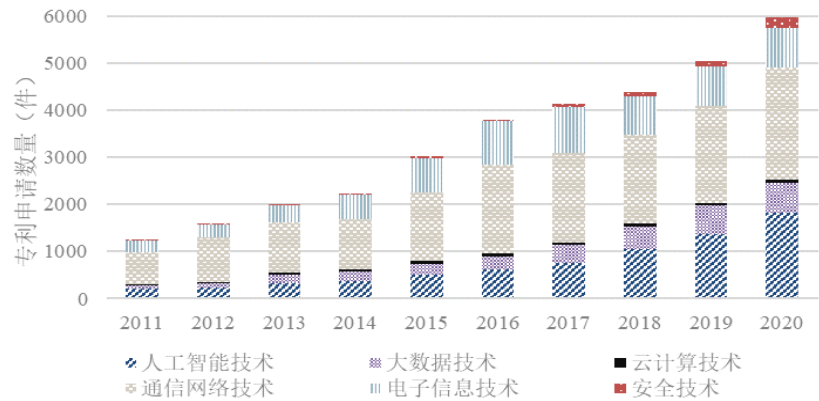
备的检查维护成本，且可快速反应灭火装置的性能，确保灭火系统正常工作。

该专利摘要信息包含了多个类别技术，核心创新点也在于各个模块通过网络技术联通，形成系统，初步判断为通信网络技术。在专利说明书中技术背景部分提出“现有市场上的城市地下管廊消防系统的监控都是基于人工的方式来进行的……二则监控管理所生成的数据没法实施上报，导致处理的时间有较大延迟……”。优势部分提出“本发明与现有技术相比，具有以下优点：1、本灭火系统整体结构设计巧妙合理，易于搭建实现及使用，制造成本较低，且智能化程度高。2、通过该系统的数据传感监测系统实时定点监测传感监测点位的传感数据，在发现数据异常后可技术处理，反应速度快，无需定期逐一排查消防设备，效率高且可大大降低消防设备的检查维护成本”。根据该专利的核心创新点在于数据传感监测系统的应用，我们将该类技术归为通信网络技术。

**第二步，训练分类算法。**首先，在训练机器学习模型之前，本文对训练样本的专利文本进行预处理，包括分词和删除停用词等常规操作，并使用词频-逆文档频率（TF-IDF）将每条处理后的专利文本进行数值向量化，作为机器学习模型的特征输入。参考已有利用机器学习进行文本分类的研究（Renault, 2017；钱宇等, 2020；范小云等, 2022），本文选择了 80%的训练样本作为训练集，用来训练机器学习模型，其余 20%作为测试集，用来评估模型样本外的预测效果。本文选取了多个在经管领域广泛使用的机器学习分类模型进行训练，模型主要包含：K 近邻（K-nearest Neighbor, KNN）、逻辑回归（Logistic Regression, LR）、朴素贝叶斯（Naïve Bayes, NB）、支持向量机（Support Vector Machines, SVM）、神经网络（Artificial Neural Network, ANN）、随机森林（Random Forest, RF）、梯度提升树（Gradient Boosting, GBDT）。调整好每个模型的超参数后，本文将测试集输入模型查看各个分类器的性能。7 个模型均用 Python scikit-learn 库的函数实现，具体超参数选择如表 5 所示：

附表 5：模型超参数

机器学习模型	超参数
KNN	n_neighbors=9
Logistic Regression	cv=4, scoring='accuracy'
Naïve Bayes	alpha=0.1
SVM	C=2, kernel='sigmoid'
ANN	activation='tanh',solver='adam',alpha=0.1,hidden_layer_sizes=(190,70,30)
Random Forest	n_estimators=200,oob_score=True, criterion='gini', bootstrap=True
GBDT	max_features=100,n_estimators=180, learning_rate=0.1, subsample=0.83



附图 2：2011-2020 年目标企业的数字专利申请情况

附 4 累积超额收益率计算和描述性统计

1. 累积超额收益率计算

首先，针对每个并购事件买方，利用模型（1）对其估计窗口内股票日收益率 $R_{i,t}$ 和市场收益率 $R_{m,t}$ 进行估计，得到 $\hat{\alpha}_i$ 和 $\hat{\beta}_i$ 。

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \epsilon_{i,t} \tag{1}$$

然后，使用方程（2）计算股票 i 在第 t 天的超额收益率 $AR_{i,t}$ 。最后，用模型（3）计算出窗口期 $(t_1, t_2)$ 的累积超额收益率 $CAR_i$ 。若 $CAR_i$ 值越高，表明该并购事件带来的短期市场价值越大。

$$AR_{i,t} = R_{i,t} - (\hat{\alpha}_i + \hat{\beta}_i R_{m,t}) \tag{2}$$

$$CAR_i(t_1, t_2) = \sum_{t=t_1}^{t_2} AR_{i,t} \tag{3}$$

2. 描述性统计

附表 6：描述性统计

变量名称	全样本（N=4221）					目标方有数字技术并购（N=546）				
	平均值	标准差	最小值	中位数	最大值	平均值	标准差	最小值	中位数	最大值
CAR	0.050	0.143	-0.324	0.019	0.413	0.076	0.162	-0.324	0.046	0.413
DPatentN	0.781	3.204	0.000	0.000	23.000	6.035	6.908	1.000	3.000	23.000
DPatent	0.204	0.605	0.000	0.000	3.178	1.574	0.822	0.693	1.386	3.178
DPatent_w	0.187	0.574	0.000	0.000	3.045	1.442	0.857	0.035	1.134	3.045
MKT	15.533	0.836	13.866	15.468	17.857	15.546	0.837	13.866	15.469	17.857
Age	2.737	0.387	1.609	2.773	3.466	2.761	0.395	1.609	2.773	3.466
LEV	0.381	0.204	0.041	0.358	0.916	0.383	0.200	0.041	0.349	0.916
Cashflow	0.043	0.066	-0.160	0.042	0.213	0.040	0.061	-0.160	0.042	0.213
Growth	0.221	0.474	-0.487	0.122	3.192	0.199	0.448	-0.487	0.132	3.192
Indep	0.375	0.051	0.333	0.364	0.556	0.380	0.055	0.333	0.364	0.556
R_D	0.036	0.045	0.000	0.030	0.227	0.046	0.046	0.000	0.037	0.227
Acq_Patent	2.758	1.639	0.000	2.944	6.365	3.547	1.488	0.000	3.555	6.365
Tar_Patent	0.850	1.315	0.000	0.000	5.094	2.307	1.542	0.000	2.250	5.094
Dealvalue	18.876	1.956	14.221	19.009	23.265	19.338	1.944	14.221	19.519	23.265
Relevance	0.397	0.489	0.000	0.000	1.000	0.456	0.499	0.000	0.000	1.000
MajorMA	0.274	0.446	0.000	0.000	1.000	0.350	0.477	0.000	0.000	1.000



附 5 基础结果稳健性检验结果

附表 7：数字技术创新对主并购方市场价值的影响

变量	(1) CAR	(2) CAR	(3) CAR
<i>DPatent</i>	0.018*** (4.747)	0.016*** (3.820)	0.014*** (3.291)
<i>MKT</i>		-0.032*** (-11.558)	-0.024*** (-7.276)
<i>Age</i>		-0.025*** (-4.543)	0.004 (0.703)
<i>Lev</i>		0.005 (0.385)	-0.002 (-0.193)
<i>Cashflow</i>		0.067** (2.035)	0.060* (1.732)
<i>Growth</i>		-0.008* (-1.676)	-0.007 (-1.404)
<i>Dealvalue</i>		0.009*** (6.379)	0.009*** (6.613)
<i>Relevance</i>		0.014*** (2.784)	0.015*** (3.054)
<i>MajorMA</i>		0.057*** (8.066)	0.046*** (6.656)
<i>Indep</i>		0.118*** (2.926)	0.121*** (2.965)
<i>Acq_Patent</i>		-0.003** (-2.366)	-0.003* (-1.803)
<i>Tar_Patent</i>		0.000 (0.177)	-0.000 (-0.162)
<i>R_D</i>		-0.034 (-0.705)	0.071 (1.286)
<i>_cons</i>	0.042*** (18.339)	0.388*** (8.850)	0.189*** (3.441)
年份固定效应	否	否	是
行业固定效应	否	否	是
样本量	4387	4221	4221
调整 <i>R</i> <sup>2</sup> 值	0.005	0.127	0.187

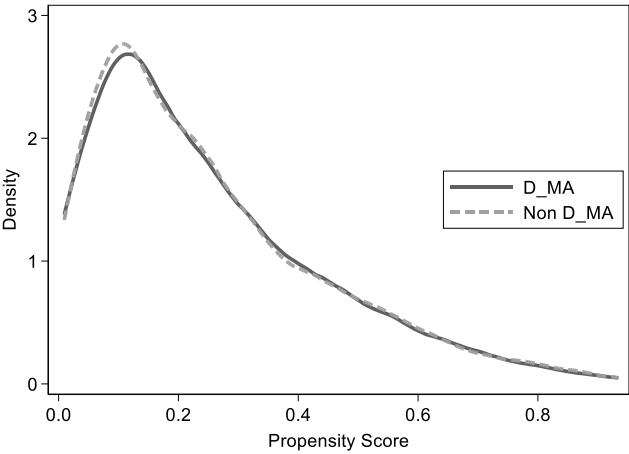
注： 本表对应原文表 5。

附表 8：稳健性检验

Panel A	更换事件研究法估计窗口					更换被解释变量	
变量	(1) CAR (-5,5)	(2) CAR (-10,10)	(3) CAR (-10,60)	(4) CAR (-10,90)	(5) CAR (-10,180)	(6) $\Delta ROA$	(7) $\Delta EPS$
<i>DPatent</i>	0.022*** (3.763)	0.035*** (4.675)	0.031*** (2.980)	0.046*** (3.616)	0.045** (2.314)	0.006** (2.375)	0.041** (2.014)
<i>_cons</i>	0.217*** (3.009)	0.396*** (4.392)	0.770*** (5.236)	1.090*** (5.835)	1.681*** (5.637)	0.175*** (4.547)	0.338 (1.011)
控制变量	是	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是	是
行业固定效应	是	是	是	是	是	是	是
样本量	4221	4221	4221	4221	4221	2980	2980
调整 <i>R</i> <sup>2</sup> 值	0.193	0.191	0.161	0.165	0.162	0.119	0.066

Panel B	更换解释变量		更严格的控制方法			倾向得分匹配法		Heckman
变量	(8) CAR	(9) CAR	(10) CAR	(11) CAR	(12) CAR	(13) CAR	(14) CAR	(15) CAR
<i>DPatent</i>			0.013*** (2.933)	0.013*** (2.946)	0.013*** (2.810)	0.017** (2.544)	0.019*** (2.975)	0.016*** (3.160)
<i>D_MA</i>	0.019** (2.444)							
<i>DPatent_ratio</i>		0.030** (2.140)						
<i>IMR</i>								-0.055 (-1.551)
<i>_cons</i>	0.185*** (3.352)	0.180*** (3.275)	0.273*** (4.687)	0.588*** (9.055)	0.627*** (5.452)	-0.127 (-0.767)	-0.125 (-0.771)	0.415*** (2.964)
控制变量	是	是	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是	是	是
行业固定效应	是	是	是	是	是	是	是	是
年份×行业 固定效应	否	否	是	否	是	否	否	否
年份×省份 固定效应	否	否	否	是	是	否	否	否
样本量	4221	4221	4221	4221	4221	841	740	2220
调整 <i>R</i> <sup>2</sup> 值	0.186	0.186	0.245	0.245	0.289	0.119	0.165	0.169

Panel C	重新定义 数字专利	限定经营 范围	更改变量定义				基于行业 调整的 CAR
变量	(16) CAR	(17) CAR	(18) CAR	(19) CAR	(20) CAR	(21) CAR	(22) CAR
<i>DPatent</i>	0.016*** (3.405)	0.016*** (3.352)	0.010*** (2.896)	0.009*** (2.852)	0.008*** (2.696)	0.008*** (2.790)	0.013*** (3.106)
<i>_cons</i>	0.194*** (3.515)	0.145** (2.274)	0.190*** (3.452)	0.190*** (3.450)	0.190*** (3.453)	0.191*** (3.462)	0.207*** (3.904)
控制变量	是	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是	Yes
行业固定效应	是	是	是	是	是	是	Yes
样本量	4190	3133	4221	4221	4221	4221	4221
调整 <i>R</i> <sup>2</sup> 值	0.188	0.183	0.186	0.186	0.186	0.186	0.187



附图 3：DPatent PSM 核密度图

附 6 DPatent\_w ①实证结果

附表 9：数字技术创新对主并购市场价值的影响

变量	(1) CAR	(2) CAR	(3) CAR
DPatent_w	0.020*** (4.994)	0.017*** (3.780)	0.015*** (3.286)
MKT		-0.032*** (-11.581)	-0.024*** (-7.289)
Age		-0.024*** (-4.506)	0.004 (0.725)
Lev		0.004 (0.370)	-0.003 (-0.213)
Cashflow		0.067** (2.045)	0.060* (1.733)
Growth		-0.008* (-1.654)	-0.007 (-1.378)
Dealvalue		0.009*** (6.317)	0.009*** (6.542)
Relevance		0.014*** (2.787)	0.015*** (3.048)
MajorMA		0.057*** (8.047)	0.046*** (6.639)
Indep		0.119*** (2.941)	0.121*** (2.982)
Acq_Patent		-0.003** (-2.299)	-0.003* (-1.782)
Tar_Patent_w		0.000 (0.088)	-0.000 (-0.158)
R_D		-0.035 (-0.708)	0.071 (1.288)
_cons	0.041*** (18.357)	0.389*** (8.890)	0.191*** (3.464)
年份固定效应	否	否	是
行业固定效应	否	否	是
样本量	4387	4221	4221
调整 R <sup>2</sup> 值	0.006	0.127	0.187

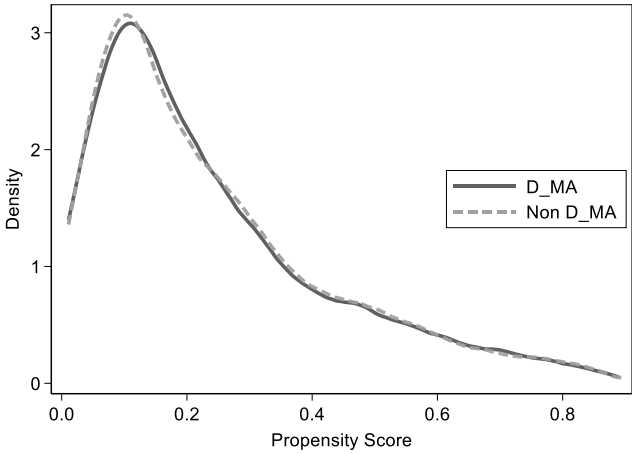
附表 10：稳健性检验

Panel A	更换事件研究法估计窗口					更换被解释变量	
变量	(1) CAR (-5,5)	(2) CAR (-10,10)	(3) CAR (-10,60)	(4) CAR (-10,90)	(5) CAR (-10,180)	(6) ΔROA	(7) ΔEPS
DPatent_w	0.022*** (3.592)	0.037*** (4.674)	0.032*** (2.911)	0.051*** (3.702)	0.051** (2.457)	0.007** (2.553)	0.044** (2.062)
_cons	0.218*** (3.027)	0.397*** (4.405)	0.771*** (5.245)	1.093*** (5.854)	1.684*** (5.646)	0.180*** (4.516)	0.330 (0.988)
控制变量	是	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是	是
行业固定效应	是	是	是	是	是	是	是
样本量	4221	4221	4221	4221	4221	2980	2980
调整 R <sup>2</sup> 值	0.193	0.191	0.161	0.165	0.162	0.118	0.066

① 考虑到虽然部分并购股权比例小于 100%，但主并方并购后属于目标企业的控股股东，我们将这部分样本权重设为 1。

Panel B	更严格的控制方法			倾向得分匹配法		Heckman
变量	(8) CAR	(9) CAR	(10) CAR	(11) CAR	(12) CAR	(13) CAR
<i>DPatent_w</i>	0.013*** (2.878)	0.013*** (3.059)	0.013*** (2.831)	0.015** (2.290)	0.013* (1.850)	0.016*** (3.034)
<i>IMR</i>						-0.055 (-1.564)
<i>_cons</i>	0.274*** (4.692)	0.587*** (9.044)	0.630*** (5.478)	0.518*** (2.586)	0.170 (1.045)	0.416*** (2.965)
控制变量	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是
行业固定效应	是	是	是	是	是	是
年份×行业 固定效应	是	否	是	否	否	否
年份×省份 固定效应	否	是	是	否	否	否
样本量	4221	4221	4221	826	769	2220
调整 $R^2$ 值	0.245	0.245	0.289	0.145	0.177	0.169

Panel C	重新定义 数字专利	限定经营 范围	更改变量定义				基于行业 调整的 CAR
变量	(14) CAR	(15) CAR	(16) CAR	(17) CAR	(18) CAR	(19) CAR	(20) CAR
<i>DPatent_w</i>	0.015*** (3.200)	0.017*** (3.244)	0.010*** (2.769)	0.010*** (2.852)	0.009*** (2.756)	0.009*** (2.873)	0.013*** (3.006)
<i>_cons</i>	0.200*** (3.620)	0.146** (2.297)	0.190*** (3.456)	0.191*** (3.463)	0.191*** (3.471)	0.191*** (3.473)	0.209*** (3.934)
控制变量	是	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是	是
行业固定效应	是	是	是	是	是	是	是
样本量	4176	3133	4221	4221	4221	4221	4221
调整 $R^2$ 值	0.188	0.183	0.186	0.186	0.186	0.186	0.187



附图 4：DPatent\_w PSM 核密度图

附表 11：数字技术创新的长期赋能效应

变量	(1) <i>TFP</i>	(2) <i>Acq_DPatent</i>	(3) <i>B_Model</i>
<i>DPatent_w</i>	0.063** (2.471)	0.215*** (4.149)	0.557** (2.343)
<i>_cons</i>	0.342 (0.725)	-7.944*** (-12.903)	-4.635** (-2.155)
控制变量	是	是	是
年份固定效应	是	是	是
行业固定效应	是	是	是
样本量	2960	3114	3114
调整 <i>R</i> <sup>2</sup> 值	0.446	0.513	0.127

附表 12：异质性分析

变量	(1) <i>CAR</i>	(2) <i>CAR</i>	(3) <i>CAR</i>
<i>DPatent_w</i>	0.002 (0.314)	-0.001 (-0.169)	0.014*** (3.115)
<i>IPP</i>	-0.000 (-0.068)		
<i>Coverage</i>		0.001 (0.238)	
<i>CTO</i>			-0.047* (-1.958)
<i>IPP</i> × <i>DPatent_w</i>	0.021** (2.473)		
<i>Coverage</i> × <i>DPatent_w</i>		0.023** (2.407)	
<i>CTO</i> × <i>DPatent_w</i>			0.054*** (2.816)
<i>_cons</i>	0.192*** (3.476)	0.191*** (3.470)	0.190*** (3.451)
控制变量	是	是	是
年份固定效应	是	是	是
行业固定效应	是	是	是
样本量	4221	4221	4221
调整 <i>R</i> <sup>2</sup> 值	0.189	0.189	0.188



附表 13：不同数字技术创新的价值

变量	(1)	机械设备 制造业	轻工制造 业	能源化工 业	信息技术 制造业	信息技术 服务业
		(2)	(3)	(4)	(5)	(6)
<i>AI</i>	-0.002 (-0.156)	0.088*** (4.104)	-0.035 (-0.979)	-0.033 (-0.824)	-0.019 (-0.931)	-0.016 (-0.461)
<i>Cloud_Computing</i>	0.063* (1.927)	0.169*** (2.927)	0.421*** (3.665)	0.278*** (4.554)	0.009 (0.127)	0.121* (1.657)
<i>Security</i>	0.013 (0.525)	-0.165 (-1.287)	0.335*** (4.624)	0.392 (1.507)	-0.025 (-0.900)	0.170*** (3.170)
<i>Big_Data</i>	0.004 (0.573)	0.007 (0.148)	0.112** (1.998)	-0.139 (-1.454)	0.002 (0.076)	-0.015 (-0.924)
<i>Commu_Network</i>	0.012* (1.687)	-0.009 (-0.581)	0.035 (0.896)	-0.002 (-0.058)	0.022* (1.801)	-0.032 (-1.257)
<i>E_Chip</i>	0.005 (0.571)	0.013 (0.428)	-0.038 (-0.860)	0.023 (0.961)	0.006 (0.526)	0.018 (0.328)
<i>_cons</i>	0.189*** (3.418)	0.204 (1.244)	0.207 (1.131)	0.023 (0.185)	0.290* (1.706)	0.338* (1.756)
控制变量	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是
行业固定效应	是	是	是	是	是	是
样本量	4221	649	361	960	770	477
调整 R <sup>2</sup> 值	0.187	0.167	0.334	0.233	0.183	0.248

附 7 行业划分的理论逻辑①

数据要素化是数字技术与实体经济深度融合的基础，也是微观主体进行数字化转型的关键。从物联网、传感等技术“赋能”数据采集，到互联网、5G 等技术“赋能”数据传输，再到云计算技术、加密技术“赋能”数据存储和数据安全，新一代信息技术正围绕“数据-算法-算力”，通过推进数字产业化和产业数字化，不断推动企业数字化转型、加速数据资产化、提升企业价值。然而，目前中国大部分企业数字化水平还处于基础阶段。工信部 2020 年数据显示，中国仅有 25% 的企业开展了数字化转型，超过 55% 的企业尚未完成基础设备的数字化改造，且大多数企业难以突破数字化转型的初级阶段（王海等，2023）。

分行业看，数字经济发展的特征存在较大差异。在处于产业数字化阶段的行业中，以电子商务为代表的服务业积极利用数字技术，如大数据、互联网技术，搭建数字化平台、进行数字化转型，其数字化程度领先于其他产业；制造业企业数字化基础能力逐步提升，在数据要素的驱动下，通过设备数字化、网联等方式实现智能化升级。但不同类型的制造业企业在数字化转型各个阶段的特征和所需技术仍具有一定的异质性。在**离散型制造业**②中，对于**机械、汽车、电子等产品设计和生产复杂、设备运维要求高的离散型行业**，除了需要在数字化转型初级阶段进行物理层面的数据采集和存储，还需利用工业机器人、人工智能等技术实现生产流程优化和质量实时监督等数字化与应用场景的融合，从而进一步走向企业内部和产业链间的互联互通、进阶企业数字能力聚合的数字化转型的中高级阶段（王永贵等，2023）；对于**纺织服饰、家居建材等需求个性化突出的离散型行业**，需要获

① 该部分内容主要根据中国信息通信研究院 2017 年-2023 年各年度《中国数字经济发展白皮书》、《企业数字化转型技术发展趋势研究报告（2023 年）》以及国家信息中心信息化和产业发展部《中国产业数字化报告 2020》进行总结。

② 离散型和流程型是制造业两种不同的生产方式。根据中国信息通信研究院的界定，离散型制造业主要包含机械、汽车、电子、家电、服装等制造业；流程型制造业主要包含化学制品、医药、石化等制造业。

取海量客户数据进行大数据分析，同时借助人工智能算法实现自主设计、自动排产和智能营销，进而打通生产现场与客户端，搭建基于互联网技术的智能生产-销售系统。在**流程型制造业**中，无论是对生产过程控制严格的**医药生物、基础化工等行业**，还是对**能源控制要求高的冶炼、石化等重化型行业**，均需在利用云技术进行数据采集存储的基础上引入智能化、联网化生产设备，通过网联、人工智能、大数据等技术进行生产全流程监测、精密控制，实现产能优化，并形成一体化的智能生产控制系统。

而对于数字产业化行业企业，则是以自身业务为主线，稳步推进服务数字化，支撑社会的数字化转型。在该产业中，**电信业和电子信息制造业**积极进行新型信息和通信基础设施建设，同时也拉动互联网数据服务（如数据中心业务、云计算业务等）实现快速发展；**软件和信息技术服务业**则加快了云化发展，云服务为信息技术服务业带来的收入持续高速增长。

## 附录参考文献

- [1] 黄勃,李海彤,刘俊岐等.数字技术创新与中国企业高质量发展——来自企业数字专利的证据[J].经济研究,2023,58(03):97-115.
- [2] 姜富伟,孟令超,唐国豪.媒体文本情绪与股票回报预测[J].经济学(季刊),2021,21(04):1323-1344.
- [3] 林建浩,陈良源,宋登辉.如何测度央行行长的口头沟通信息——一种基于监督学习的文本分析方法[J].统计研究,2019,36(08):3-18.
- [4] 刘青,肖柏高.劳动力成本与劳动节约型技术创新——来自 AI 语言模型和专利文本的证据[J].经济研究,2023,58(02):74-90.
- [5] 罗佳,张蛟蛟,李科.数字技术创新如何驱动制造业企业全要素生产率?——来自上市公司专利数据的证据[J].财经研究,2023,49(02):95-109+124.
- [6] 陶锋,朱盼,邱楚芝等.数字技术创新对企业市场价值的影响研究[J].数量经济技术经济研究,2023,40(05):68-91.
- [7] 佟岩,王茜,曾韵等.并购动因、融资决策与主并方创新产出[J].会计研究,2020,(05):104-116.
- [8] 王海,闫卓毓,郭冠宇等.数字基础设施政策与企业数字化转型:“赋能”还是“负能”?[J].数量经济技术经济研究,2023,40(05):5-23.
- [9] 王永贵,汪淋淋,李霞.从数字化搜寻到数字化生态的迭代转型研究——基于施耐德电气数字化转型的案例分析[J].管理世界,2023,39(08):91-114.
- [10] 吴非,胡慧芷,林慧妍等.企业数字化转型与资本市场表现——来自股票流动性的经验证据[J].管理世界,2021,37(07):130-144+10.
- [11] Chen, M. A., Wu, Q. and Yang, B. How Valuable Is Fintech Innovation? [J]. The Review of Financial Studies, 2019,32(5): 2062-2106.
- [12] Mikolov, T., Sutskever, I., Kai, C., Corrado, G., and Dean, J. Distributed Representations of Words and Phrases and their Compositionality [J]. Advances In Neural Information Processing Systems, 2013, 26, 3111-3119.

**注：该附录是期刊所发表论文的组成部分，同样视为作者公开发表的内容。如研究中使用该附录中的内容，请务必在研究成果上注明引文和下载附件出处。**

引用示例：

参考文献引用范例：

[1] 朱军. 技术吸收、政府推动与中国全要素生产率提升[J]. 中国工业经济. 2017, (1):5-24.

如果研究中使用了未在《中国工业经济》纸质版刊发、但在杂志网站上正式公开发表的数字内容（包括数据、程序、附录文件），请务必在研究成果正文中注明：

数据（及程序等附件）来自朱军（2017），参见在《中国工业经济》网站（<http://ciejournal.ajcass.org>）附件下载。