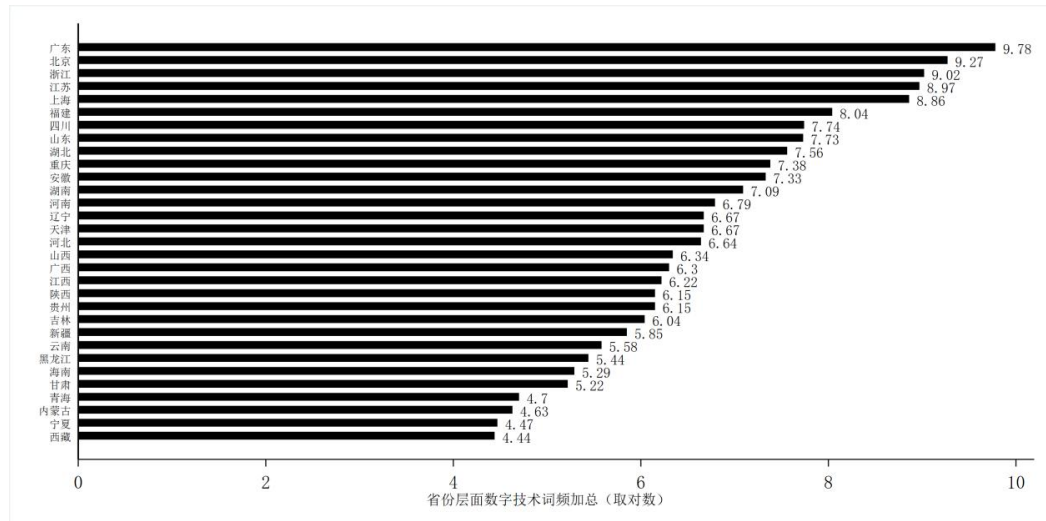


附录

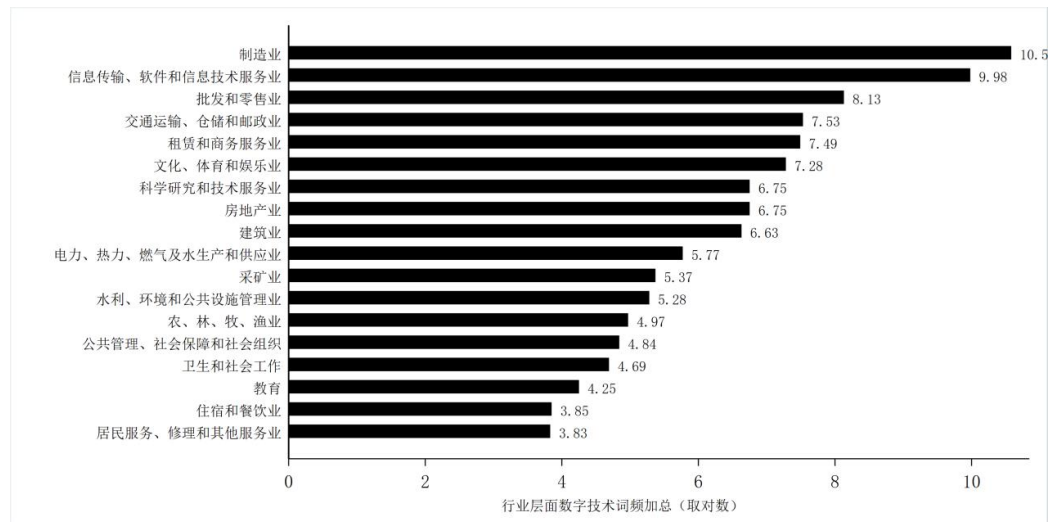
附录 A：企业数字技术应用水平指标的合理性与有效性说明

本文参考现有文献，采用机器学习的文本分析法，通过分析上市公司年报中披露的“数字技术”相关的信息来刻画企业数字技术应用水平（吴非等，2021；袁淳等，2021）。这种衡量方法的优点在于：根据中国证监会修订的《公开发行证券的公司信息披露内容与格式准则》，上市公司要对报告期内经营情况进行回顾并对未来发展进行展望。因此，如果上市公司将数字技术应用作为企业发展战略的重要组成部分，那么根据证监会的年报信息披露原则，企业需要在年报内容中对其进行披露。考虑到上市公司披露的年报是基于公司实际运营情况客观的陈述。因此，企业针对数字技术应用的重视和实施程度越高，越会倾向于披露“数字技术”相关信息。所以从上市企业年报中涉及“数字技术”的词频统计角度来刻画其数字技术应用程度，有一定的可行性和科学性（吴非等，2021）。并且，本文从企业对数字技术的“使用广度”和“应用深度”两个角度构建企业层面的数字技术应用指标，实现了对企业内部各环节（生产、管理和经营等）数字化程度较为细致地刻画，为后续关于企业数字技术应用的相关研究提供了有益借鉴。

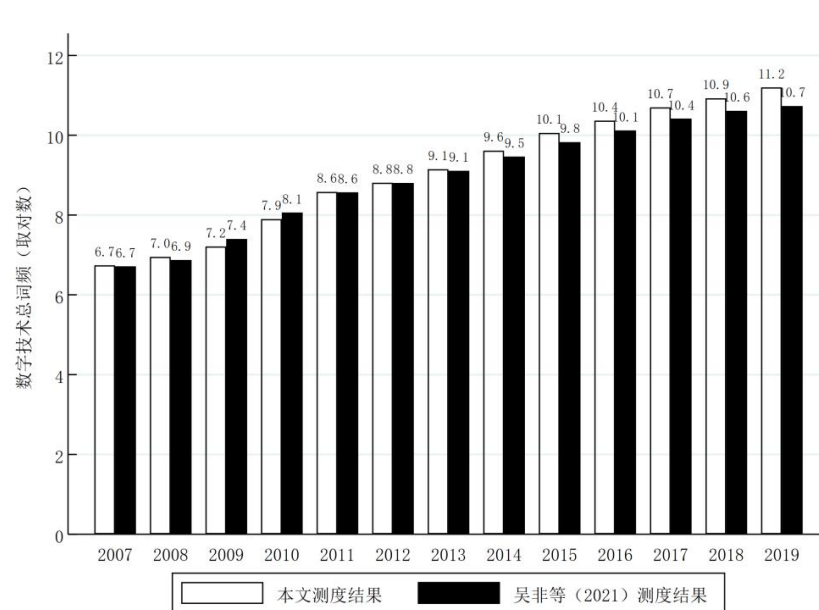
本文构建的企业数字技术应用水平指标具有如下特征：第一，从地区分布看，本文将2019年的企业数字技术应用关键词词频在省份层面加总后发现，位于一线城市和东南沿海省份（如北京、上海、广东、浙江、江苏和福建）的上市企业的数字技术应用程度更高，这与这些地区的数字经济发展程度有关。根据《中国数字经济白皮书（2020）》（下文简称为“《报告》”）披露的2019年各省市数字经济增加值总体规模数据，广东省、北京市、上海市、江苏省和浙江省处于国内数字经济发展的领先地位。可以看出，本文的测度结果与《报告》的统计结果基本保持一致（见附图1）。第二，从行业特征看，本文将2019年的企业数字技术应用关键词词频在行业层面加总后发现，数字技术应用排名前三的行业分别为制造业、信息传输、软件和信息技术服务业，以及批发和零售业；农业、卫生和社会工作、教育以及住宿和餐饮业的数字技术应用水平则较为落后，这可能是因为这些行业对劳动力的需求较大，并且多为非常规的复杂劳动，难以被数字技术替代（附图2）。第三，本文构建的企业数字技术应用指标与已有文献具有可比性。吴非等（2021）从“底层技术运用”和“技术实践应用”两个角度构建了企业数字化转型的指标。为此，本文统计了吴非等（2021）在指标构建中用到的关键特征词，将其加总后与本文的测度结果按照年份进行比对。附图3的结果表明，二者的测度结果也是较为接近的。因此，可以认为本文关于企业数字技术应用的测度结果是可靠的。



附图 1 2019 年 31 个省（市）数字技术总词频统计



附图 2 2019 年行业层面数字技术总词频统计



附图 3 2007—2019 年企业数字技术关键词总词频趋势图

附录 B：稳健性检验

1.更换企业数字技术应用的测度方法。

①考虑到企业数字技术应用水平在不同行业层面会存在差异，本文将核心解释变量作行业均值调整（*DigtechAdj*），以此反映不同行业下企业数字技术应用的相对水平；②本文将“数字技术使用广度”和“数字技术应用深度”指标下的“数字化生产”“数字化管理”“数字化营销”和“数字化产品”四个子指标标准化，然后使用主成分分析法重新构建数字技术应用指标（*DigtechPca*）。③本文使用《中国上市公司数字化转型研究数据库》中的企业数字化转型指数，重新衡量企业数字技术应用程度（*DigtechCsmar*），该指标来源于国泰安数据库。附表 1 报告了替换解释变量之后的检验结果，可以看出，企业数字技术应用对专利质量存在提升作用的结论始终成立。

附表 1 更换企业数字技术应用的测度方法			
	<i>PatQuality</i>		
	(1)	(2)	(3)
<i>DigtechAdj</i>	0.0087*** (0.0020)		
<i>DigtechPca</i>		0.0189*** (0.0034)	
<i>DigtechCsmar</i>			0.0016*** (0.0004)
控制变量	Yes	Yes	Yes
企业固定效应	Yes	Yes	Yes
年份固定效应	Yes	Yes	Yes
观测值	32056	32056	28152
组内 R ²	0.0277	0.0283	0.0307

注：括号内为企业层面聚类的标准误，***、** 和 * 分别表示在 1%、5%和 10%的水平上显著。下同。由于《中国上市公司数字化转型研究数据库》的数据起始时间为 2010 年，因此列（3）存在部分缺失值。

2.更换企业专利质量的测度方法。

①参考张杰和郑文平（2018）的方法，根据平均值将专利层面的知识宽度加总至企业层面，记为 *PatQualityMean*。②鉴于企业专利涉及的知识领域越多，其更多利用的是与已有知识相关度较低的新知识进行创新，专利质量可能越高。因此，本文使用企业发明专利分类号的大组数量代表企业在新知识领域的探索程度，记为 *ClassNum*。③采用企业当年申请的发明专利与专利申请总量的比例（*InvPatRatio*）来衡量企业专利的技术复杂程度，发明专利占比越高说明企业的专利质量越高。④采用企业专利的平均被引用次数（*PatCited*）衡量专利质量，平均被引用次数越高，说明专利质量越高。从附表 2 的回归结果可以看出，无论采用何种方式替换被解释变量，*Digtech* 系数始终至少在 5%的水平上正向显著。

附表 2 更换企业专利质量的测度方法				
	<i>PatQualityMean</i> (1)	<i>ClassNum</i> (2)	<i>InvPatRatio</i> (3)	<i>PatCited</i> (4)
<i>Digtech</i>	0.0061*** (0.0017)	0.0727*** (0.0121)	0.0124*** (0.0026)	0.0100** (0.0044)
控制变量	Yes	Yes	Yes	Yes
企业固定效应	Yes	Yes	Yes	Yes
年份固定效应	Yes	Yes	Yes	Yes
观测值	32056	32056	32056	32056
组内 R ²	0.0220	0.0625	0.0112	0.3064

3.排除企业策略性信息披露行为。

本文主要依据企业年报中披露的数字技术应用信息来分析其对专利质量的影响,然而这一结果可能会受到企业自身策略性信息披露行为的影响(袁淳等,2021)。例如,企业为了博取市场关注而夸大自身对数字技术的应用。为了排除这一影响,本文采用两种方法进行检验:①参考赵璨等(2020)的做法,构建“数字技术应用”信息披露程度的模型,估算出企业数字技术应用相关词频的正常披露次数,进而剔除存在夸大披露的企业样本。②删除深圳证券交易所信息披露质量的等级仅为合格和不合格的上市企业样本,因为这类企业可能更有动机进行策略性披露。附表3的第(1)和(2)列报告了剔除可能存在策略性披露行为企业后的回归结果,可以发现, *Digtech* 的系数依然在1%的水平上显著为正。

4.剔除特殊样本。

2009年10月《中华人民共和国专利法》执行了新的发明专利的授权标准,将原先的“相对新颖性”标准上升为“绝对新颖性”的高度。为此,本文删除了2010年之前的样本重新进行回归。从附表3的第(3)列回归结果中可以看出,在剔除特殊样本之后,企业数字技术应用对专利质量的提升效果依然显著。

附表3	排除企业策略性信息披露行为与剔除特殊样本		
	<i>PatQuality</i>		
	(1)	(2)	(3)
<i>Digtech</i>	0.0115*** (0.0044)	0.0084*** (0.0020)	0.0095*** (0.0020)
控制变量	Yes	Yes	Yes
企业固定效应	Yes	Yes	Yes
年份固定效应	Yes	Yes	Yes
观测值	18930	29626	29550
组内 R ²	0.0187	0.0286	0.0298

参考文献

[1] 吴非,胡慧芷,林慧妍,任晓怡.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(7):130-144+10.
[2] 袁淳,肖土盛,耿春晓,盛誉.数字化转型与企业分工:专业化还是纵向一体化[J].中国工业经济,2021(9):137-155.
[3] 赵璨,陈仕华,曹伟.“互联网+”信息披露:实质性陈述还是策略性炒作——基于股价崩盘风险的证据[J].中国工业经济,2020(3):174-192.