

2017

中国软件开发者调查报告

Survey Report on Software Developers in China

# CSDN学院 IT实战派

CSDN学院是CSDN在2015年初期推出的教育平台，是面向中国软件开发者和IT专业人士的全新技术服务平台，利用互联网优势，通过视频学习、线上答疑、作业辅导、实战演练等方式，让学习更加方便和高效。同时也为各大企业产品与技术提供更为有效的传播和推广途径。

- 企业培训：为企业提供一站式人才培养方案。
- 走进高校：覆盖高校2000+，助力大学生快速成长。

## 公开课

技术直播，与大咖面对面交流！

## 精品课

海量学习资源  
随时随地在线学！

## 技术峰会

名企一线工程师，实践案例深度剖析！

## 特训班

系统学习，实战演练，掌握专业技能，高薪就业

## 覆盖领域



CSDN学院服务号

咨询电话：13522267206

CSDN学院：<http://edu.csdn.net/>

# 2017年 中国软件开发者白皮书

## 版权声明

《2017年中国软件开发者白皮书》版权属于CSDN，并受法律保护。转载、摘编或利用其他方式使用本白皮书文字或观点，应注明“来源CSDN”。

# 目录 CONTENTS

第一章 报告概述	4
1.1 背景和方法	4
1.2 数据处理方法	5
1.3 样本处理方法	5
1.4 特邀专家顾问	6
第二章 重要发现	7
第三章 软件开发者现状分析	9
3.1 软件开发者基本特征	9
3.2 软件开发者薪资状况分析	12
3.3 软件开发者学习特点分析	18
3.4 软件开发工具使用特点分析	22
3.5 软件开发者大学毕业就业特点分析	27
第四章 大数据技术应用现状分析	29
4.1 企业大数据平台构建及应用情况	29
4.2 企业处理数据及平台集群规模情况	32
4.3 企业构建大数据平台所用技术情况	34
第五章 企业使用云服务现状分析	39
5.1 企业使用云服务方式分析	39

5.2 企业云平台部署特点	42
5.3 企业构建云服务平台技术选型分析	44
第六章 区块链技术应用现状分析	49
6.1 区块链技术应用现状分析	49
6.2 区块链技术开发特点	53
6.3 区块链应用前景展望	56
第七章 人工智能技术应用现状分析	59
7.1 企业人工智能应用状况及面临挑战	60
7.2 人工智能技术应用现状	63
7.3 人工智能行业应用及展望	66
第八章 物联网技术应用现状分析	69
8.1 物联网技术应用状况	69
8.2 物联网技术开发现状	71
8.3 业务涉及工业物联网的企业调查	73
附录：中国城市线具体划分	77

# 第一章 报告概述

## 1.1 背景和方法

本白皮书是基于一年一度的CSDN软件开发者大调查数据分析结果形成的。CSDN最早从2004年开始针对中国软件开发者进行大规模调查，是迄今为止覆盖国内各类软件开发者人群数量最多，辐射地域、行业分布最广的调查活动。该调查旨在全面和深入地了解中国软件开发者群体整体现状、应用开发技术以及开发工具、平台的状况和发展趋势等，它是各相关行业了解中国软件开发者群体以及软件开发服务领域市场的重要参考资料。

2017年中国软件开发者大调查针对软件开发技术、应用开发领域等方面对调查问卷进行了重新设计，同时还新增加了区块链、物联网等新兴软件开发领域的相关调查，更能体现中国IT业、尤其是软件开发领域的发展现状。本次调查具有以下特点：

- 调查内容 — 紧扣技术热点、把握知识趋势

我们根据2016-2017年的技术趋势和热点以及市场上出现的新产品、新技术，对调研选项进行了调整和补充，力求准确、及时地反映出2016-2017年中国软件开发技术发展变化趋势。

- 调查报告 — 提供定制方案服务、打造调研精品

本次报告依旧在以往简版报告、完整报告分类基础上，提供定制报告分析服务。根据客户的需求，我们以调查的原始数据为基础，进行不同维度的数据交叉分析和专家点评，为企业客户制定市场及产品开发策略提供更有针对性的决策依据。

- 专家介入 — 行业资深团队、全面洞悉业界

本次大调查特别成立了专家顾问团队，针对调查中涉及的技术领域，特邀了相关领域的技术专家，共同参与调查项的设计并对报告中的部分数据进行专业点评。本次调查数据的获取采用大规模线上问卷数据调查和社区平台数据挖掘相结合的方式，除了利用CSDN自身平台资源之外，还联合部分合作伙伴渠道进行大力推广以增加数据采集的全面性，力求调查结果更加客观、准确。此次调查在短时间内获得了涉及47个行业、200多个地区的上万名软件开发者的积极参与。

## 1.2 数据处理方法

### ● 数据处理流程



### ● 统计分析工具

利用目前流行的大规模数据统计分析工具SPSS，根据分析目的采用了描述性统计、交叉分析等数据处理方法。

## 1.3 样本处理方法

本次调查采取样本有效性系统方法来保证样本的有效性：包括重复填写问卷排除法、答题逻辑矛盾排除法、不完整问卷排除法等。

## 1.4 特邀专家顾问

开发平台及工具领域

**孙 玄** 转转公司架构算法部负责人

大数据领域

**毕洪宇** 饿了么大数据平台总监

云计算领域

**李明宇** OStorage创始人兼CTO

人工智能领域

**智 亮** TinyMind总经理

区块链领域

**王 玮** 北京志顶科技创始人

物联网领域

**唐翊国** GE资深技术经理

感谢以上专家对白皮书的指导与支持!

## 第二章 重要发现

我们在本次调研过程中有如下主要发现：

1. 30岁以下软件开发者人数占比超7成，全国近半数的开发者工作在一线城市
  - 互联网、软件、IT服务三个技术领域涵盖了国内65%以上的开发者
  - 30岁以上的开发者以高级工程师、项目经理、架构师、技术专家等职位居多
  - 40岁以上的开发者从事架构设计岗位的比例最高
  - 全国有4成以上开发者月均收入过万，其中超5成来自一线城市
  - 8成架构设计师月薪过万，超3成的架构师平均月薪超过2万元
2. 软件开发者学习热情高涨，9成开发者通过技术媒体、技术社区、技术论坛等网站了解行业最新动态、获取新知识
  - 超9成开发者计划通过学习新技术来提升自己的事业，6成开发者每周学习6小时以上
  - 74%的人群愿意通过购买图书杂志方式进行学习。超5成人员愿意为线上课程付费。愿为线下培训付费的仅占39%
  - 35%的开发者未来半年有培训计划。超过8成的开发者参加培训的预算来自个人
3. 近7成开发者在使用Java语言，近6成开发者近期最想学Python语言
  - 开发者的云/容器使用率仅11%，近6成开发者在使用Notepad++文本编辑器
  - Eclipse是最受欢迎的IDEs，有43%的开发者在使用。Frameworks方面，jQuery、Spring Framework较受欢迎，3成以上开发者在使用这两者
  - 在自动构建工具方面，Maven较受欢迎，有近4成开发者在使用
4. 超7成软件开发者主要来自二本以上院校，超五成毕业于计算机与软件专业
  - 超6成开发者毕业前就找到工作，超过6成开发者未曾参加过针对大学生的就业培训
  - 学生就业情况良好，培训机构的学生就业也较乐观。77%的软件开发者所在公司招聘大学生，46%的开发者所在公司招聘培训机构的学生

5. 大数据技术应用开始普及，但大数据应用场景相对单一

- 78%的企业在运用大数据技术，这一比例与2016年（24%）提升较大
- 目前大57%的企业对大数据的应用更多的体现在统计分析、报表及数据可视化上
- Spark、Redis和Kafka正在成为企业大数据平台通用技术组件

6. 83%的企业正在使用云服务，超7成企业利用云平台已经开发应用或正在开发应用

- 安全是企业选择公有云服务时最大的顾虑，也是进行云技术相关开发时的最大问题
- 企业在虚拟机、网络存储、负载均衡方面应用云计算较为普遍，使用率分别是48%、39%、33%
- 基于Docker或OpenStack是当前云平台部署的两种主流框架

7. 区块链技术近两年刚刚兴起，仅1成开发者有应用或准备应用，近7成开发者尚处于初步了解阶段，2成开发者完全不了解

- 比特币和以太坊是当前两种主流的区块链开发平台
- Python和Java是区块链的主流开发语言

8. 人工智能技术日益受到企业和市场的关注，应用前景广阔但任重道远

- 4成开发者表示其技术领域正在尝试使用AI/机器学习/深度学习
- 68%的开发者通过自学方式学习AI/机器学习/深度学习技术。但有四成开发者通过工作中实践、线上培训课程来学习AI技术

9. 物联网产业链上企业最多的在云平台环节，智能家居是物联网最大的应用领域

- 目前企业工业智能化水平参差不齐
- 智能家居是物联网最大的应用领域

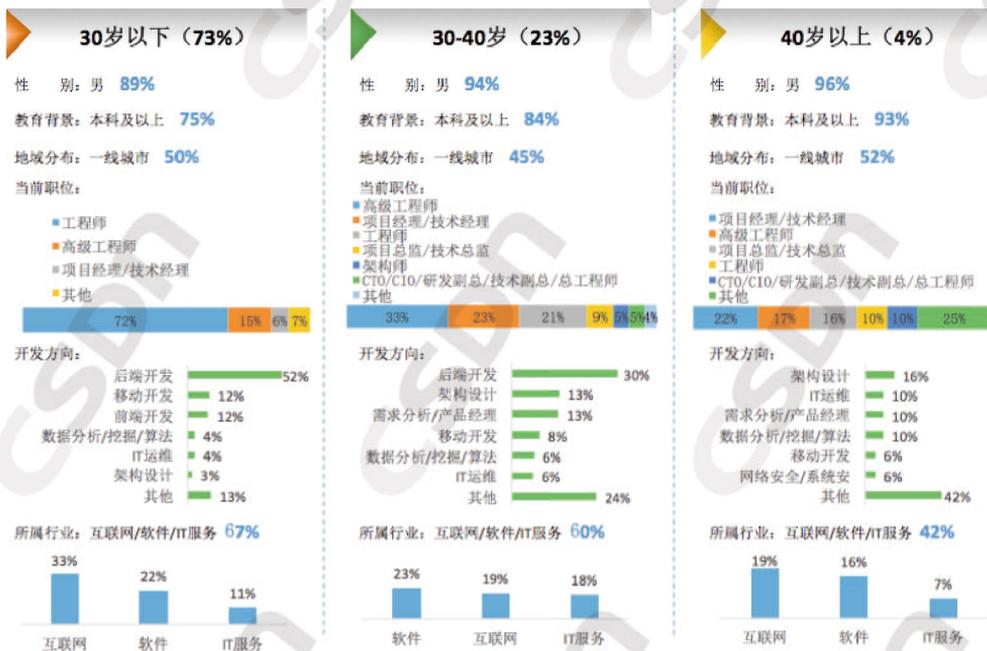
## 第三章 软件开发者现状分析

### 3.1 软件开发者基本特征

软件开发者基本画像：男性，本科，一线城市，开发/架构设计

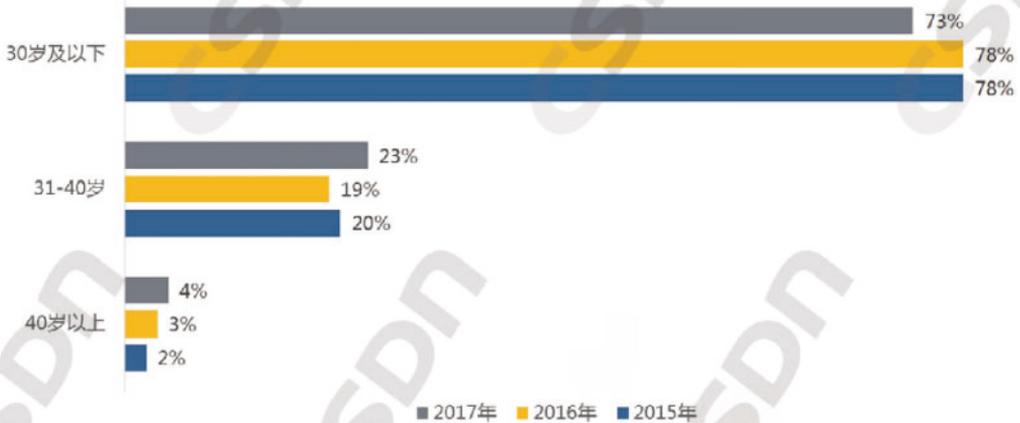
我们将开发者根据年龄范围划分成三类，不同年龄段开发者群体进行特征分析后呈现出如下特点：

- 30岁以下软件开发者人数超7成，全国近半数的开发者工作在一线城市（北京、上海、广州、深圳、天津），本科及以上学历占近8成，91%的开发者是男性
- 互联网、软件、IT服务三个技术领域涵盖了国内65%以上的开发者
- 30岁以上的开发者以高级工程师、项目经理、架构师、技术专家等职位居多
- 40岁以上的开发者从事架构设计岗位的比例最高



## 开发者群体年轻化比例增高，30岁以下占7成

和国外开发者年龄分布趋势大概一致，国内的软件开发者群体一直呈现出年轻化的特点。从我们2015年到2017年的调研数据来看，30岁以下的开发者人群占比都在7成以上，一直是软件开发领域的主力军。

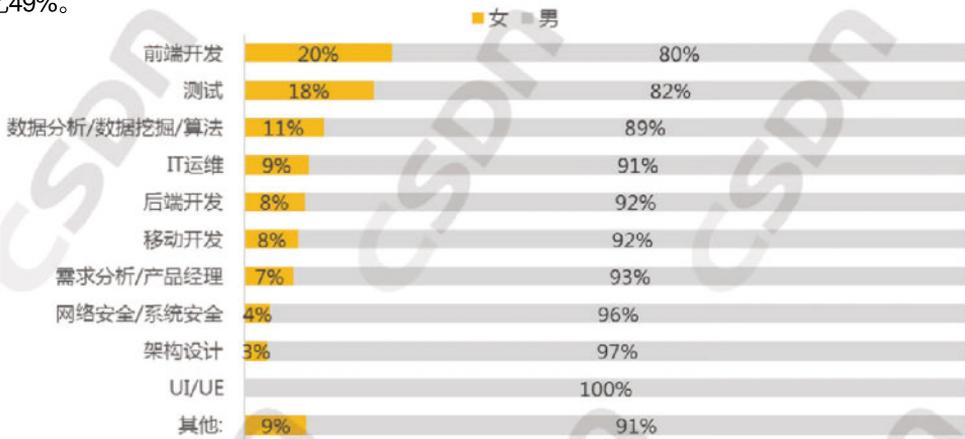


2017年度·中国软件开发者白皮书

图表 2：软件开发者年龄占比对比

## 女性开发者岗位TOP 3分别是：前端开发、测试、数据分析

从具体图表中可以看出，近一半的女性开发者相集中在了前端开发、测试及数据分析相关职位，合计占比49%。

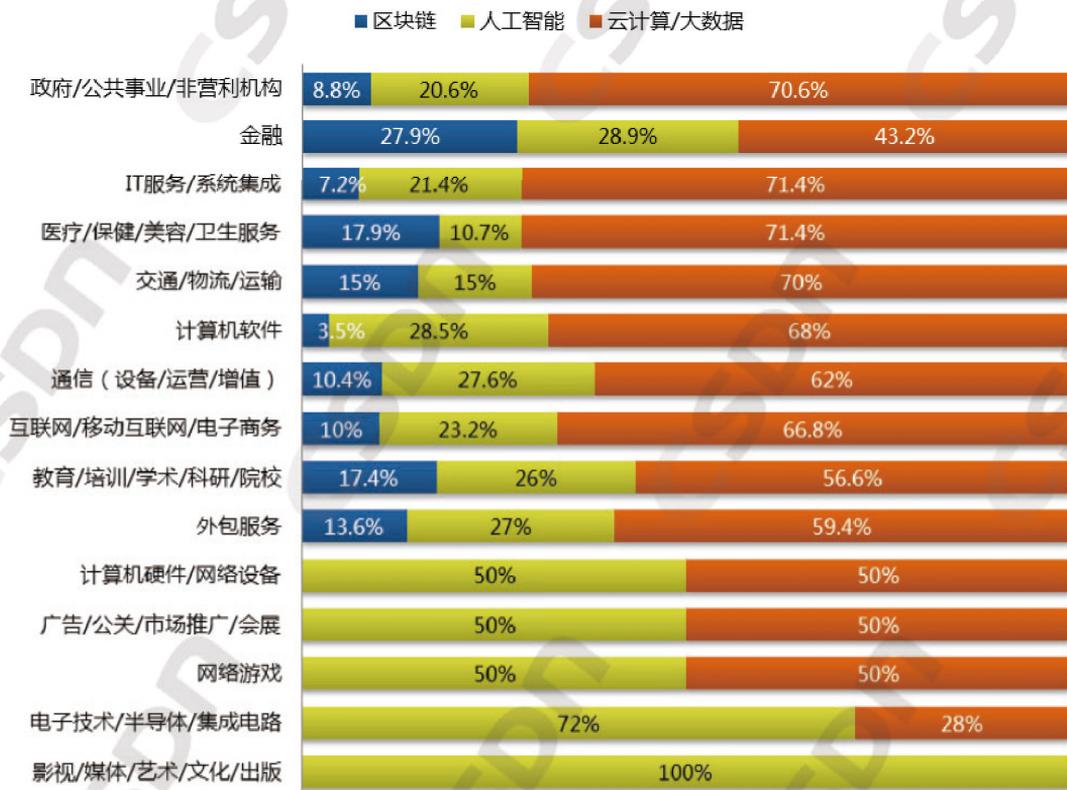


2017年度·中国软件开发者白皮书

图表 3：不同岗位软件开发者的性别分布

## 不同行业对新技术关注程度不同

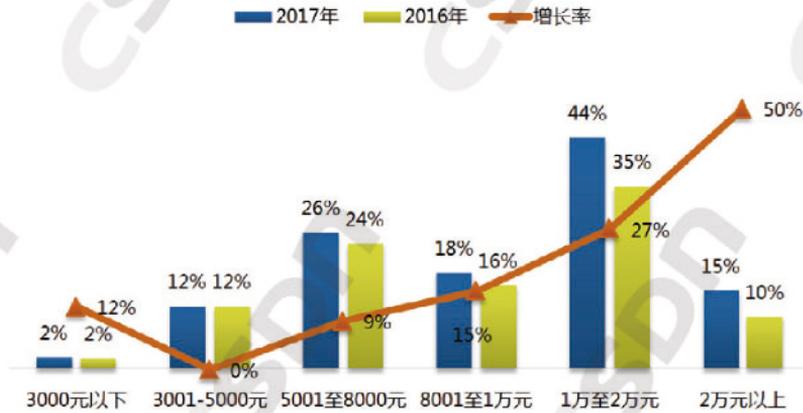
从不同行业对云计算/大数据、人工智能、区块链等新技术的关注来看，金融行业、政府/公共事业/非营利机构对这些新技术的关注更明显，相关从业人员占比也最多。



### 3.2 软件开发者薪资状况分析

全国有4成以上开发者月均收入过万，其中超5成来自一线城市

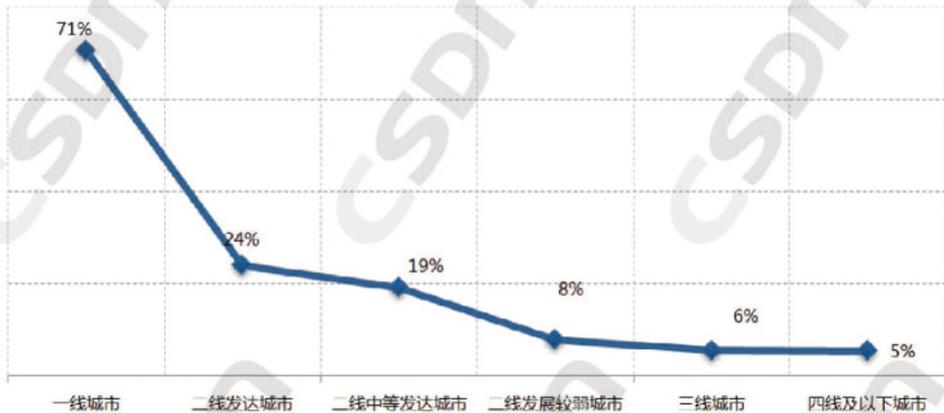
月薪在1万元以上的开发者占总体数量的59%，较去年又上升了14个百分点。



2017 年度·中国软件开发者白皮书

图表 5：软件开发者月薪分布

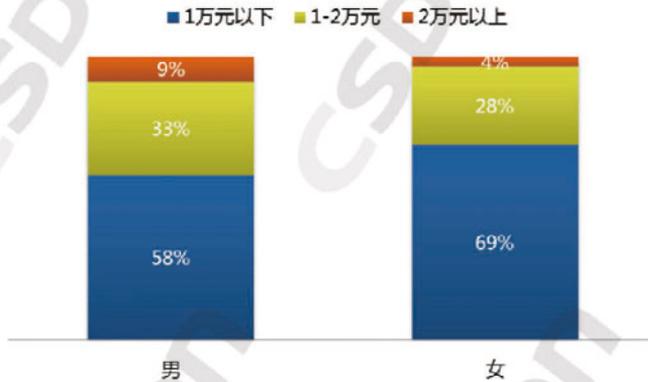
在一线城市（北京、上海、广州、深圳、天津）中，月薪过万的软件开发者的占比为71%，该比例远高于国内其它城市，但其他城市月入过万的开发者也出现了较大幅度的增长。



2017 年度·中国软件开发者白皮书

图表 6：月薪过万开发者所在城市分布

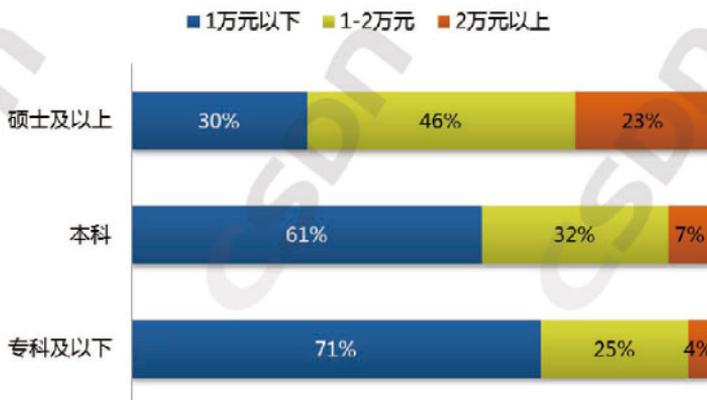
数据显示，男女软件开发者薪资水平有一定的差异。在男性软件开发者中，月收入超过万元的占42%，而女性月入过万的为32%。



2017 年度 · 中国软件开发者白皮书

图表 7：不同性别开发者月薪对比

通过结合受教育程度和薪资水平的数据特点来看，学历越高的人群中，月薪2万元以上的高收入比例越高。

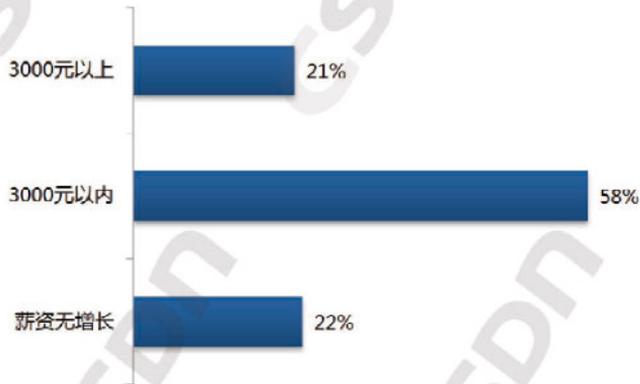


2017 年度 · 中国软件开发者白皮书

图表 8：不同教育程度的软件开发者月薪分布

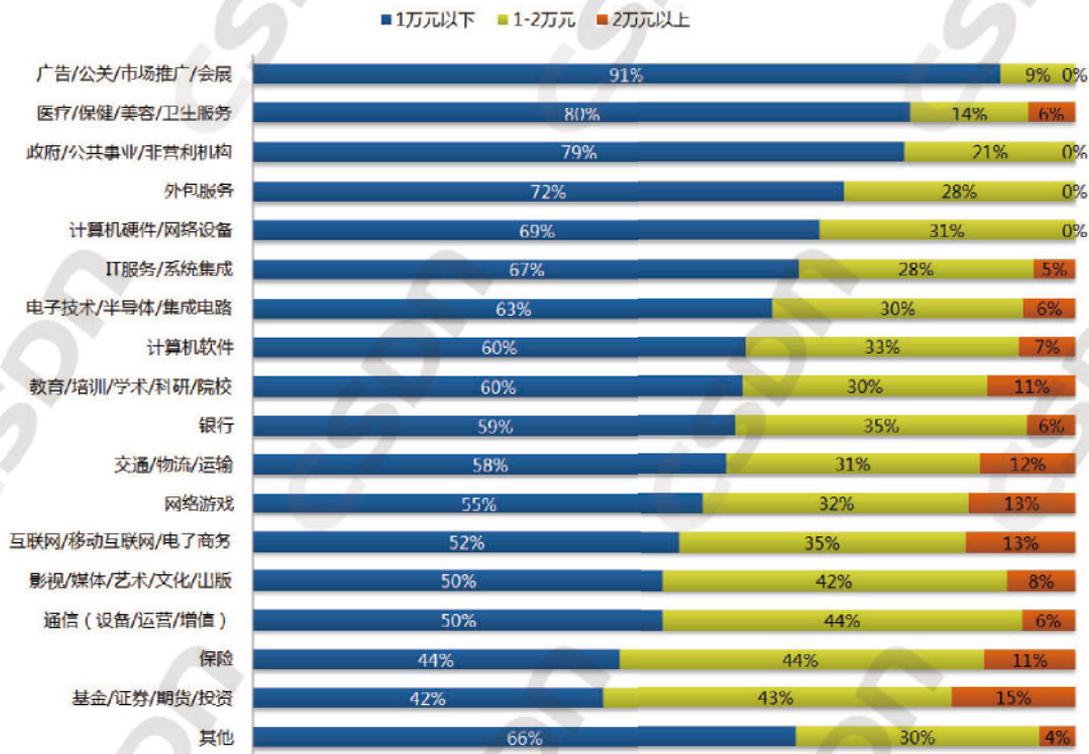
### 近六成开发者本年度月薪增加幅度在3000元以内

本次调研数据显示，在本年近8成软件开发者的月薪不同程度地得到了增长。近6成的开发者月薪增长数字在3000元以内。超2成的开发者工资增长幅度更大，达到3000元及以上。



## 基金/证券/期货/投资行业的软件开发者高收入人群比例最高

基金/证券/期货/投资行业的软件开发者高收入人群占比最高，58%的人群月薪超过1万元。

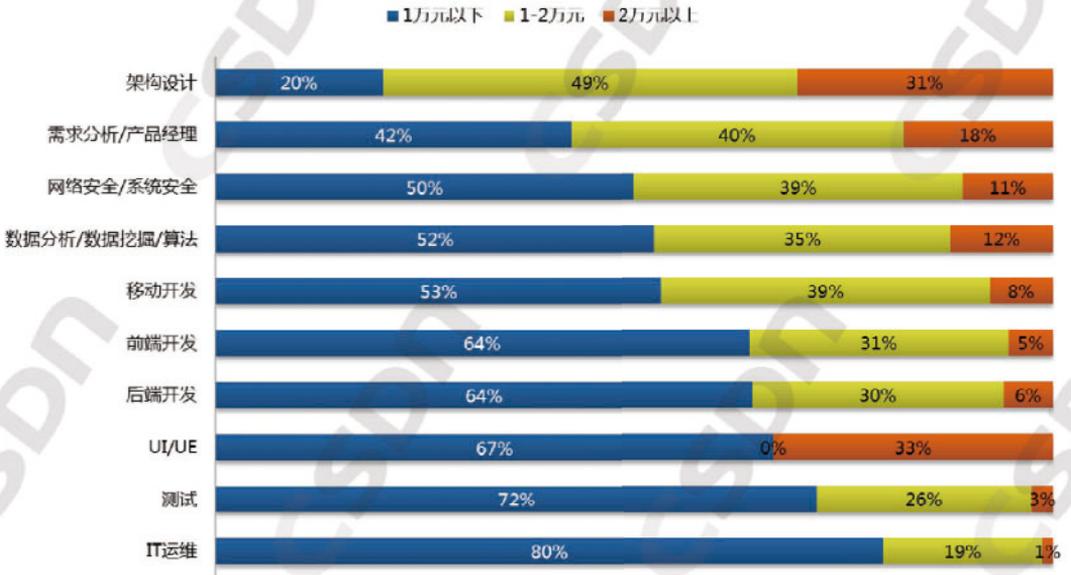


2017 年度·中国软件开发者白皮书

图表 10：软件开发者不同行业的薪资范围分布

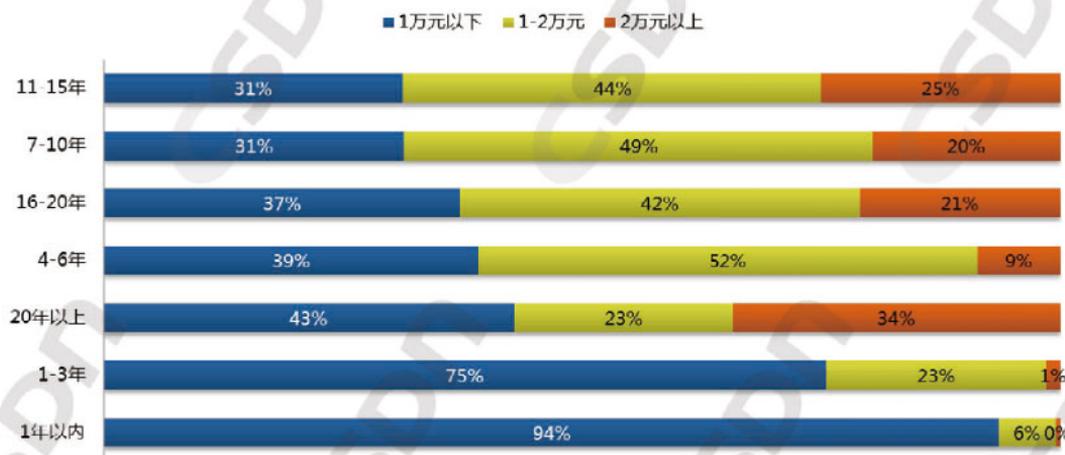
## 从事软件架构设计岗位的开发者8成月薪过万

通过数据交叉对比分析发现，8成架构设计师月薪过万，超3成的架构师平均月薪超过2万元。



## 从业年限在4-15年的人群超6成月薪过万

超7成的从业年限3年以内，普遍薪资月薪低于1万；但从业年限4-20年的人群超过6成月薪过万。

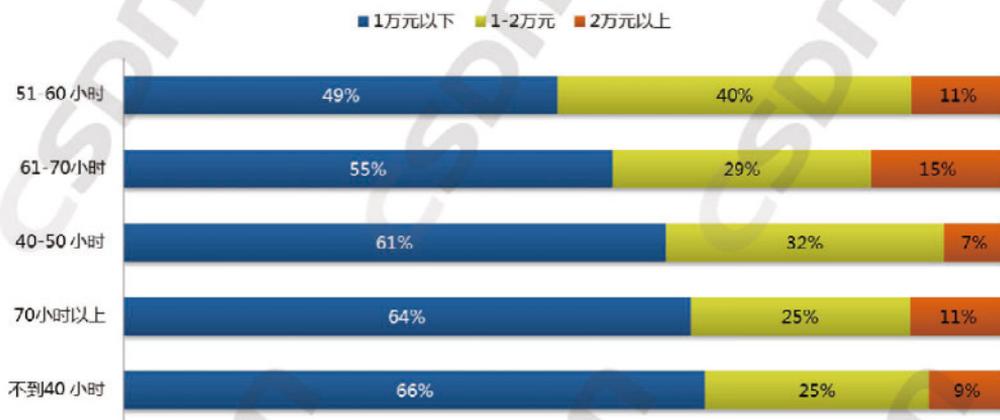


2017 年度·中国软件开发者白皮书

图表 12：软件开发者不同从业年限薪资范围分布

## 周工作时长在51-60小时的人群中，5成收入过万

周工作时长在51-60小时的人群中，5成收入过万。但工作时长并不与收入正相关。周工作70小时以上的人群，与周工作不足40小时的人群收入比例基本一致。



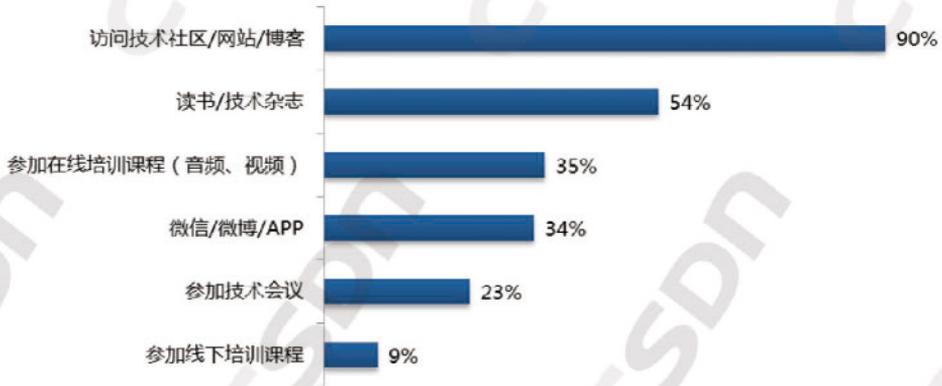
2017 年度·中国软件开发者白皮书

图表 13：软件开发者不同周工作时长薪资范围分布

### 3.3 软件开发者学习特点分析

#### 信息技术类网站是开发者了解行业动态和学习知识的主要渠道

在本次调研中，90%的开发者会通过技术媒体、技术社区、技术论坛等网站了解行业最新动态、获取新知识，阅读技术类书籍也是开发者进行系统性学习的主要方式，比例占到54%。

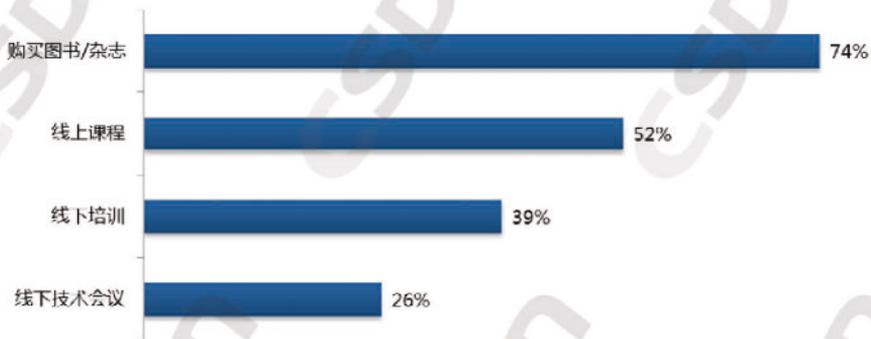


2017 年度·中国软件开发者白皮书

图表 14：软件开发者获取知识来源分布（调查项为多选）

#### 超7成人员愿意为学习购买图书和杂志，超五成人员愿意购买线上课程

74%的人群愿意通过购买图书杂志方式进行学习。超5成人员愿意为线上课程付费。愿为线下培训付费的仅占39%。

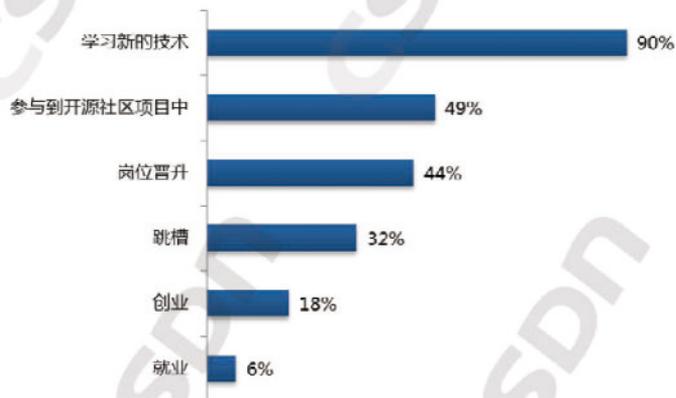


2017 年度·中国软件开发者白皮书

图表 15：软件开发者愿意付费的学习方式分布（调查项为多选）

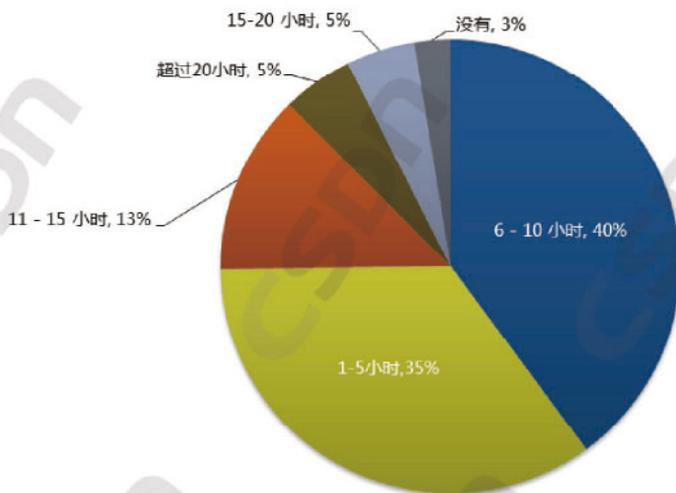
超9成开发者计划通过学习新技术来提升自己的事业，6成开发者每周学习6小时以上

超9成开发者计划通过学习新技术来提升自己的事业。近5成开发者愿意参与到开源社区项目中。超过63%的开发者每周学习6小时以上。



2017 年度 · 中国软件开发者白皮书

图表 16: 软件开发者计划提升事业的方式分布 (调查项为多选)

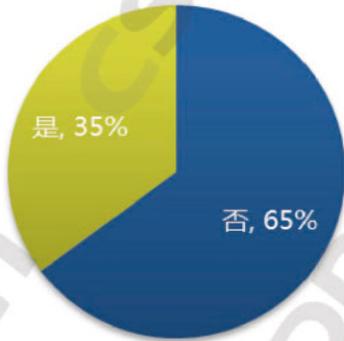


2017 年度 · 中国软件开发者白皮书

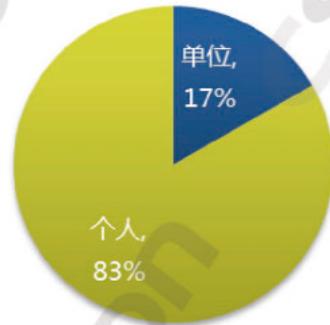
图表 17: 软件开发者每周学习时长

### 35%的开发者未来半年有培训计划。超过8成的开发者参加培训的预算来自个人

35%的开发者未来半年有培训计划。超过8成的开发者参加培训的预算来自个人，来自单位的仅占17%。



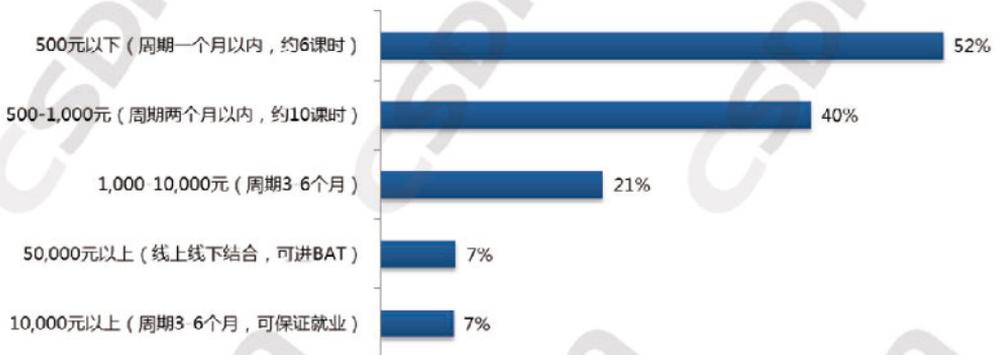
2017年度·中国软件开发者白皮书  
图表 18: 软件开发者未来6个月是否有培训计划



2017年度·中国软件开发者白皮书  
图表 19: 开发者参加培训的预算来源

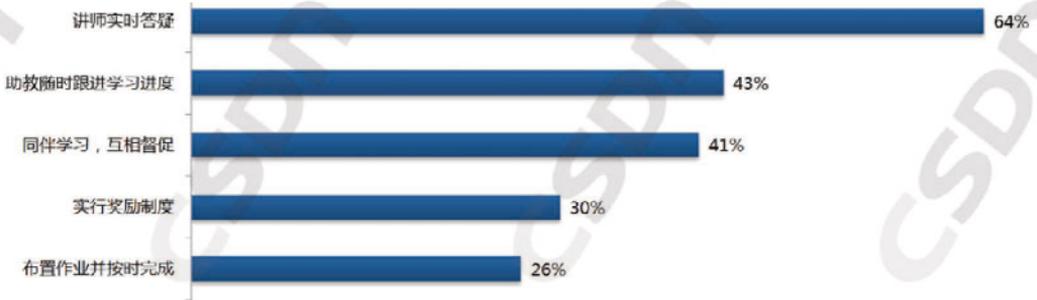
### 时长越短的线上培训越受欢迎，较受欢迎的线上培训服务为讲师实时答疑

时长越短的线上培训越受欢迎，500元以下且周期1个月以内的线上培训接受率超过5成。64%的开发者倾向于线上培训的讲师实时答疑服务。



2017年度·中国软件开发者白皮书

图表 20: 开发者可接受的线上培训类型

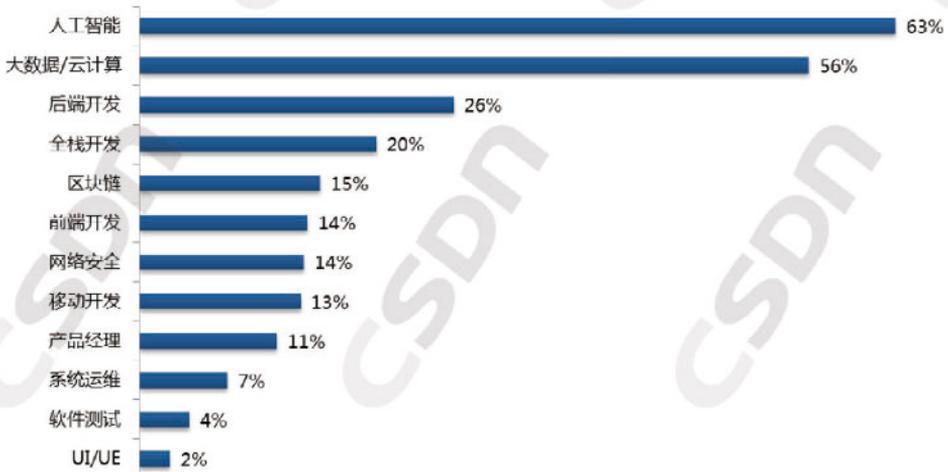


2017 年度 · 中国软件开发者白皮书

图表 21：吸引开发者线上培训服务

### 开发者最感兴趣的培训方向是人工智能、大数据 / 云计算

63%的开发者对人工智能培训方向感兴趣，56%对大数据 / 云计算感兴趣，区块链培训人数仅占15%。



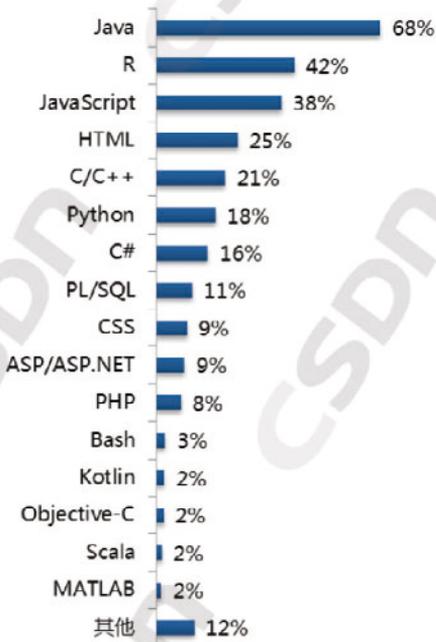
2017 年度 · 中国软件开发者白皮书

图表 22：开发者最感兴趣的培训方向（调查项为多选）

### 3.4 软件开发工具使用特点分析

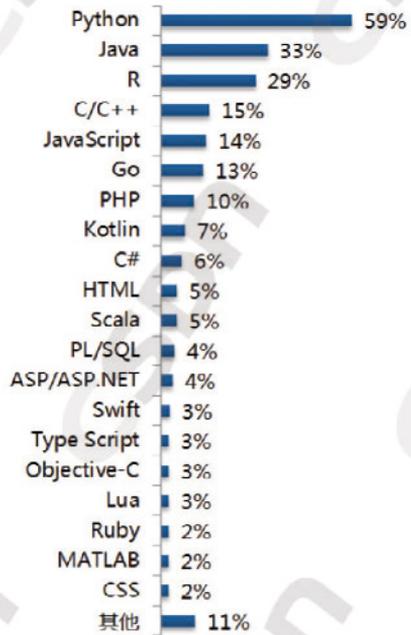
#### 近7成开发者在使用Java语言，近6成开发者近期最想学Python语言

Java、R和JavaScript成为开发者最多使用的编程语言。Python是开发者近期最想学的开发语言，其次是Java和R。



2017 年度·中国软件开发者白皮书

图表 23：开发者最常用的开发语言（调查项为多选）



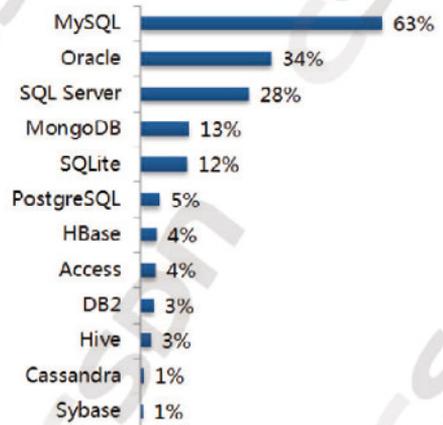
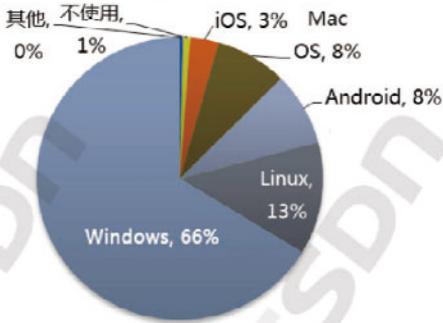
2017 年度·中国软件开发者白皮书

图表 24：开发者想学的开发语言（调查项为多选）

**【孙玄点评】**Java语言虽然是老牌语言，依然非常有生命力。第一是因为它的持续更新升级优化，Java 9在2017年已经推出；第二，它的生态很完善，项目需要的功能都能够找到对应的实现Jar；第三，它的使用场景很广泛，从电商领域、到信息流分发流量，再到大数据、云计算、AI领域，都可以如鱼得水的使用它。因此Java成为使用最多的语言也就不足为奇了。人工智能是技术发展的下一个风口，必将改变目前行业发展的格局，拥抱人工智能，将拥有未来。Python作为解释型的语言，它提供了大量的计算包，可以帮忙学习人工智能的同学很快上手，并且Python语法简单，表达的语义丰富，入门门槛相对较低。人工智能发展的火热，成为促进Python语言应用的催化剂。

## 6成以上开发者在使用Windows操作系统，62%的开发者在使用MySQL数据库

6成以上开发者在使用Windows操作系统，13%在使用Linux操作系统。63%的开发者在使用MySQL数据库，34%的在使用Oracle数据库。



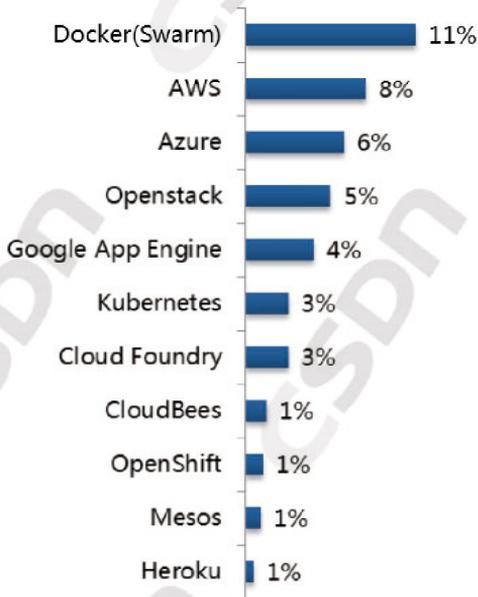
2017 年度·中国软件开发者白皮书  
图表 25: 开发者正在使用的操作系统

2017 年度·中国软件开发者白皮书  
图表 26: 开发者正在使用的数据库 (调查项为多选)

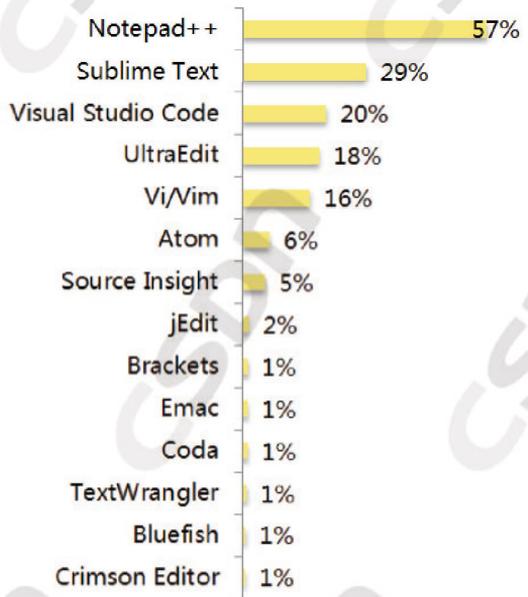
**【孙玄点评】**Windows操作系统成为开发者使用最多的系统，MySQL关系数据库依然很有生命力，传统的X86机器相对Mac机器，最大的优势就是价格低。因此在公司开发普及层面会占据一定的优势。Linux在GUI、软件生态等方面相对Windows有不少的短板，因此Windows会成为开发者的首选操作系统。数据库领域经过不断的迭代发展，从RDBMS发展到NoSQL，再到NewSQL。每个阶段的发展都是工业界需求驱动的结果。MySQL作为单机版的RDBMS，在满足公司结构化关系型的数据存储需求方面，依然很有生命力。但是MySQL在可扩展性方面有一定局限，NewSQL的出现可以很好的解决扩展性的问题，个人看好NewSQL的发展。

## 开发者的云 / 容器使用率仅11%，近6成开发者在使用Notepad++ 文本编辑器

开发者的云 / 容器使用率不高。使用最普遍的 Docker ( Swarm ) 普及率也仅11%。在文本编辑器方面，Notepad++ 较为流行，57%的开发者使用了Notepad++，其次是Sublime Text，29%的开发者在使用。



2017 年度·中国软件开发者白皮书  
图表 27：开发者正使用的云 / 容器（调查项为多选）

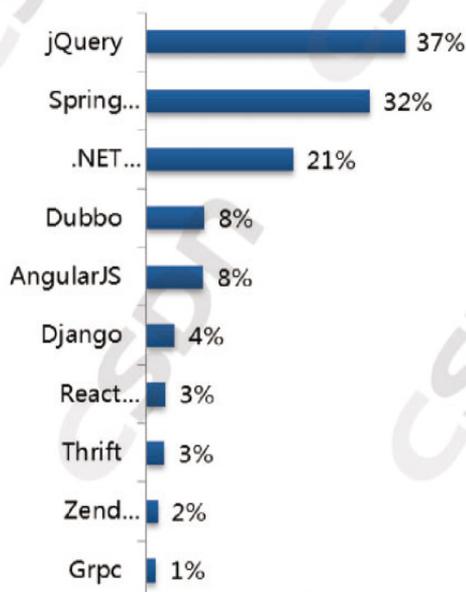
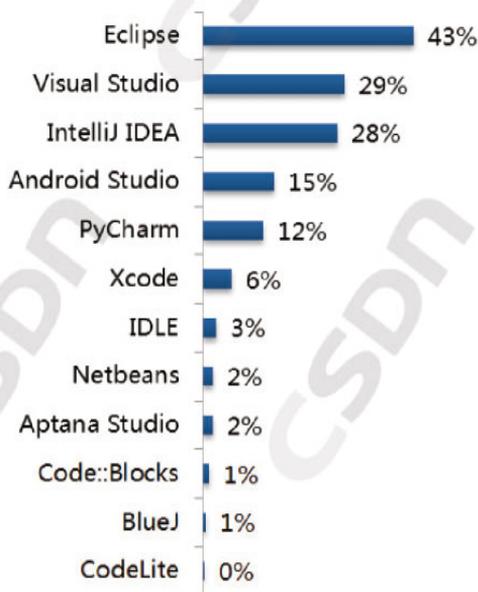


2017 年度·中国软件开发者白皮书  
图表 28：开发者使用的文本编辑器（调查项为多选）

**【孙玄点评】**随着微服务如火如荼地发展，微服务架构模式能够让项目快速迭代、持续交付，能够帮助公司快速试错，更好的发展。因此互联网公司、传统公司都在积极拥抱微服务。采用了微服务架构后，服务颗粒较小，服务数量多，如何运维管理这些微服务成为各个公司不得不面对的新问题，Docker (Swarm) 容器的出现大大降低了微服务管理的成本。任何事物的出现到普及都需要一定的周期，Docker 容器化普及也不例外，2018年Docker容器发展普及会越来越广泛。文本编辑器领域，产品众多，竞争非常激烈，如何能够占据大量市场份额，靠的是简单易用，Notepad++做到了。

Eclipse是最受欢迎的IDEs，jQuery、Spring Framework是较受欢迎的Frameworks

Eclipse是最受欢迎的IDEs，有43%的开发者在使用。Frameworks方面，jQuery、Spring Framework较受欢迎，3成以上开发者在使用这两者。



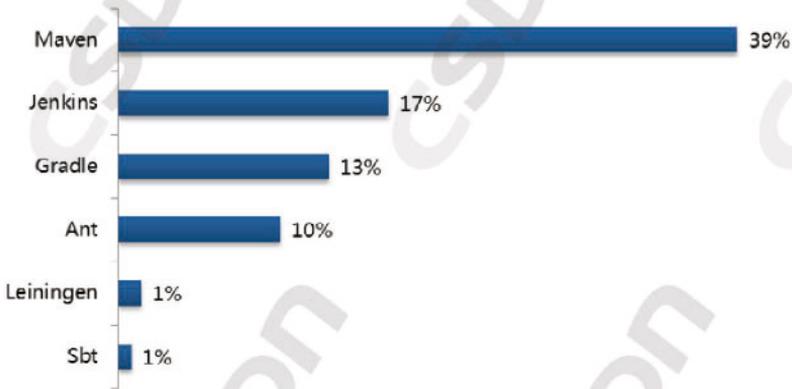
2017 年度·中国软件开发者白皮书  
图表 29：开发者正在使用的 IDEs

2017 年度·中国软件开发者白皮书  
图表 30：开发者正在使用的 Frameworks

**【孙玄点评】**Eclipse成为最受欢迎的IDE，和Java语言的市场占用率是分不开的，在Java的世界使用Eclipse作为IDE开发工具非常的自然，同时Eclipse支持C/CPP/Python等语言开发。总之Eclipse占用份额最高，是意料之中。随着微服务架构的普及，Spring Framework使用越来越多，在开发微服务架构中，使用Java作为开发语言，那么开发框架使用Spring Framework往往是首选。在前端的世界里，虽然不断涌现出大量的前端开发框架，jQuery由于简单、易用，在项目过程中往往能够较好满足需求，使得其一直很坚挺，可谓廉颇虽老，尚能饭！

## 近4成开发者在使用Maven

在自动构建工具方面，Maven较受欢迎，有近4成开发者在使用。



2017年度·中国软件开发者白皮书

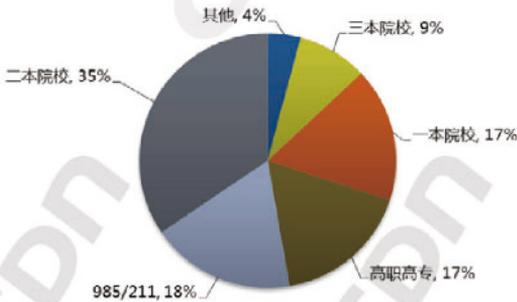
图表 31：开发者正在使用的自动构建工具

**【孙玄点评】**Maven项目对象模型(POM)，除了以程序构建能力为特色之外，还提供高级项目管理工具。由于Maven的缺省构建规则有较高的可重用性，所以常常用两、三行Maven构建脚本就可以构建简单的项目。由于Maven的面向项目的方法，许多Apache Jakarta项目发文时使用Maven，而且公司项目采用Maven的比例在持续增长。互联网公司以及传统的公司也都大量使用Maven，使得Maven在2017年成为最受欢迎的项目管理工具。

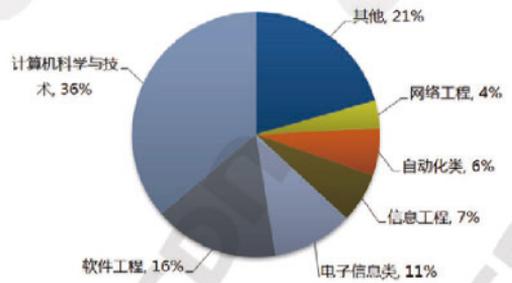
### 3.5 软件开发者大学毕业就业特点分析

#### 超7成软件开发者主要来自二本以上院校，超五成毕业于计算机与软件专业

来自二本院校的开发人员占35%，来自985/211、一本院校的分别占18%、17%。从专业来看，来自计算机科学与技术专业的最多，占36%，其次是软件工程专业，占16%。



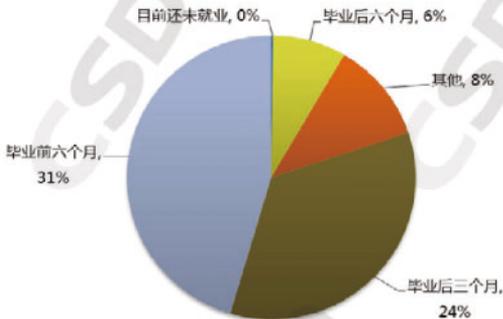
2017 年度·中国软件开发者白皮书  
图表 32: 开发者毕业的院校分布



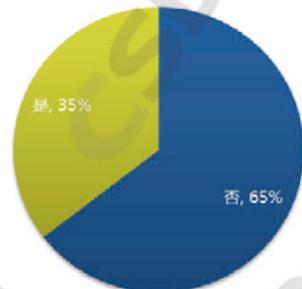
2017 年度·中国软件开发者白皮书  
图表 33: 开发者毕业的专业分布

#### 超6成开发者毕业前就找到工作，超过6成开发者未曾参加过针对大学生的就业培训

毕业前六个月找到工作的占31%，而毕业后3个月找到工作的占24%。在大学生就业培训方面，目前普及率还较低，超过6成开发者没有参加过针对大学生的就业培训。



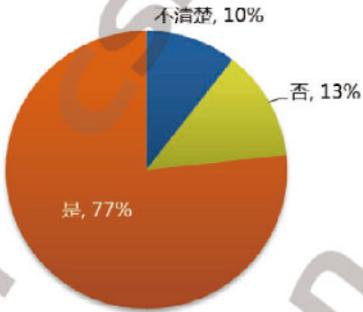
2017 年度·中国软件开发者白皮书  
图表 34: 开发者毕业后找到工作的时间分布



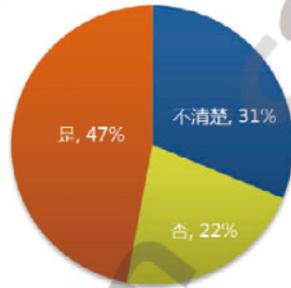
2017 年度·中国软件开发者白皮书  
图表 35: 开发者是否参加过针对大学生的就业培训

### 大学生就业情况良好，培训机构的学生就业也较乐观

从就业情况来看，77%的软件开发者所在公司招聘大学生，47%的开发者所在公司招聘培训机构的学生。



2017 年度·中国软件开发者白皮书  
图表 36：开发者所在公司是否招聘大学生



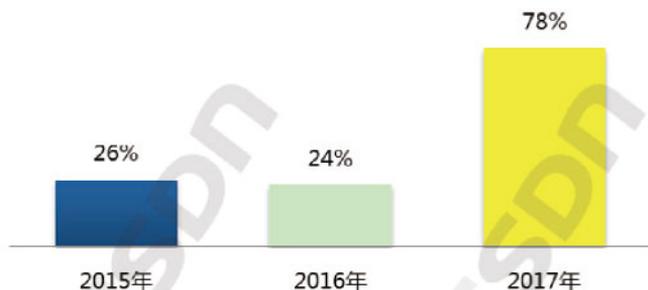
2017 年度·中国软件开发者白皮书  
图表 37：开发者所在公司是否招聘培训机构的学生

## 第四章 大数据技术应用现状分析

### 4.1 企业大数据平台构建及应用情况

#### 78%的企业在运用大数据技术进行应用开发

本次调研数据显示，78%的企业在进行大数据相关的开发和应用，这个比例与2015年、2016年相比，有非常大的提升。

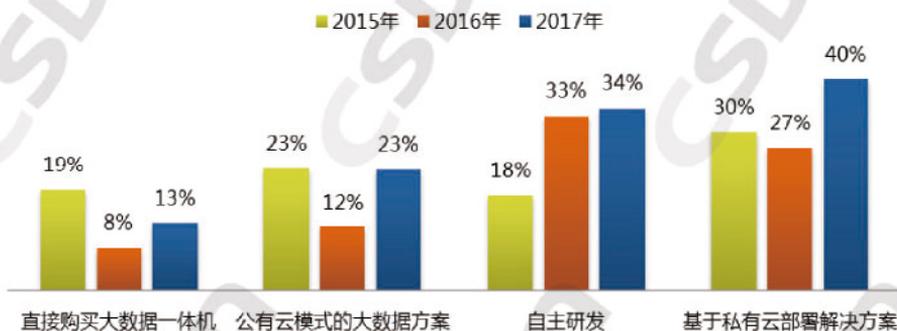


2017 年度·中国软件开发者白皮书

图表 38：从事大数据开发和企业的占比

#### 私有云部署解决方案是企业构建大数据平台的主要方式

本次调研数据显示，随着云计算技术的成熟及广泛应用和出于对数据安全性的考虑，2017年有4成企业选择私有云解决方案来部署大数据应用。这一比例相对2015、2016年有较大幅度增长。34%的企业选择自主研发，这一比例与2016年基本持平。

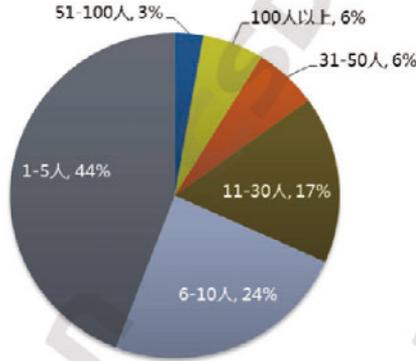


2017 年度·中国软件开发者白皮书

图表 39：企业大数据平台构建方式分布

## 超八成企业大数据团队规模在30人以下，其中近7成的数据团队小于10人

企业大数据团队规模集中在30人以下，占比为85%，其中10人以下规模的团队占比接近7成，100人以上占6%。

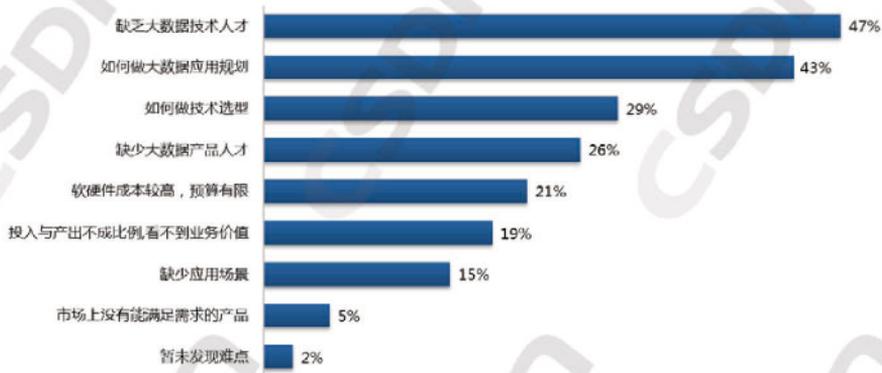


2017年度·中国软件开发者白皮书

图表 40：企业大数据团队规模

## 专业人才的缺乏是影响企业大数据应用落地的最大障碍

大数据时代，大数据所蕴含的价值已经毋庸置疑，然而许多企业却不知道如何获取大数据的价值，如何让大数据应用落地，企业内部也没有可以胜任这方面工作的人才，本次调研数据也说明，大数据人才的缺乏是企业构建大数据应用时遇到的主要问题。如何做大数据应用规划以及技术选型也是困扰大数据技术发展的现实问题。

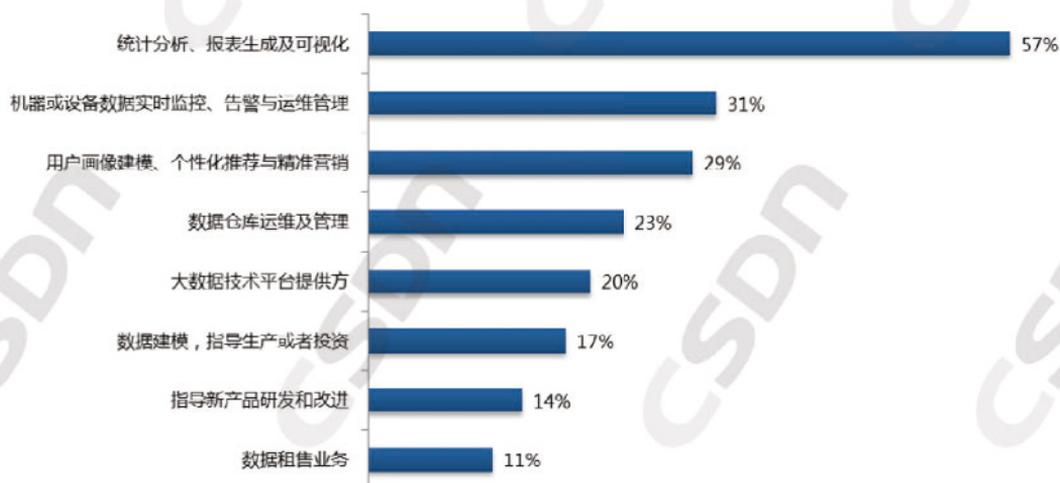


2017年度·中国软件开发者白皮书

图表 41：企业构建大数据平台面临的问题（调查项为多选）

## 大数据技术应用场景逐渐多元化，以数据分析及可视化最为普遍

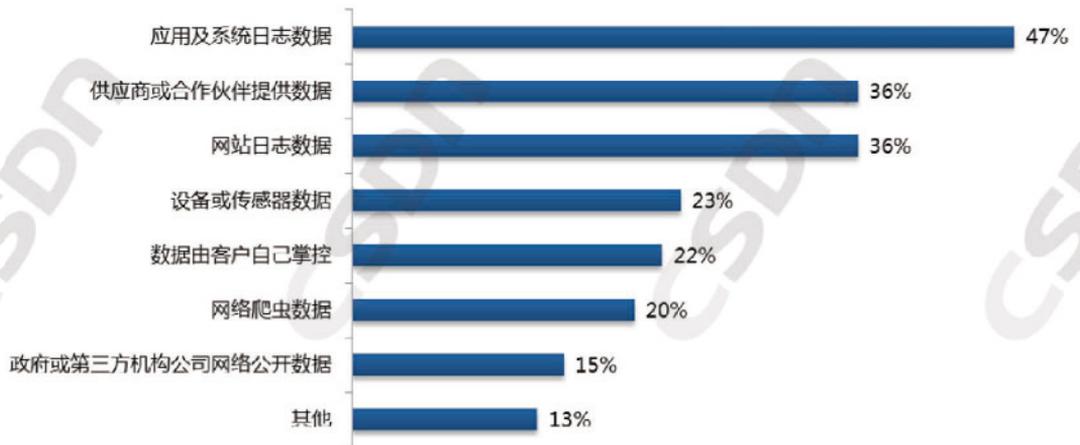
本次调查发现企业大数据应用场景相对单一，目前大部分企业对大数据的应用更多的体现在统计分析、报表及数据可视化上，占比为57%。机器或设备数据实时监控、告警与运维管理以及基于用户画像的数字化营销方案的次之，分别占31%、29%。



## 4.2 企业处理数据及平台集群规模情况

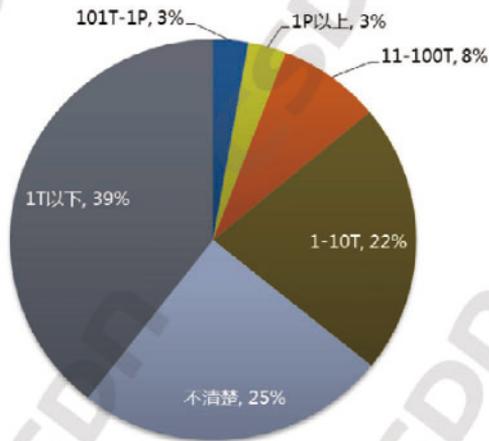
### 日志数据是企业大数据平台处理的主要数据类型

企业数据主要来源于企业内部的日志数据，主要包括系统日志和用户行为日志。在本次调研中，企业数据来自内部应用及业务系统日志数据的占47%，来自供应商或合作伙伴提供数据及网站访问日志数据的均占36%。



## 超6成的企业日均处理数据规模在10TB以下

调研数据显示, 39%的企业日均处理数据规模在1TB以下, 在1-10TB的企业仅占22%, 在10TB以上的仅占14%。

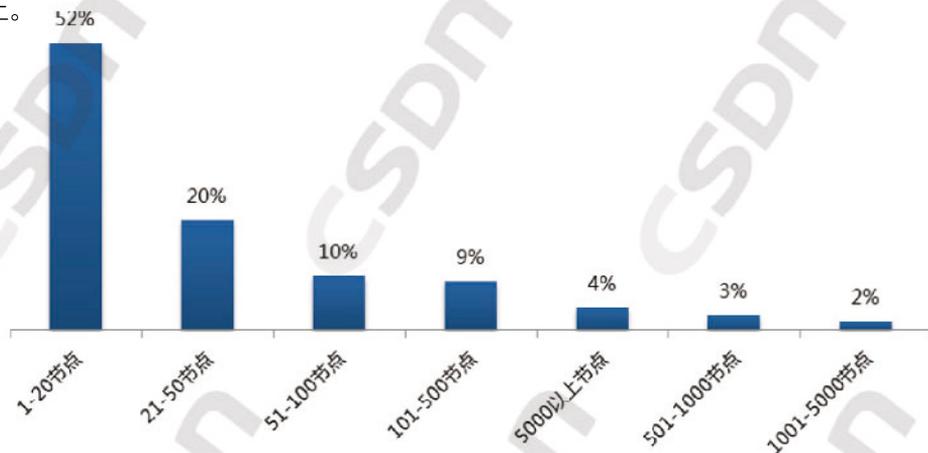


2017 年度 · 中国软件开发者白皮书

图表 44: 企业大数据平台日均处理数据量

## 48%企业大数据平台的集群规模在20个节点以上

本次调研中48%的企业大数据平台集群规模拥有20个节点以上, 其中2%的企业集群规模在5000个节点以上。



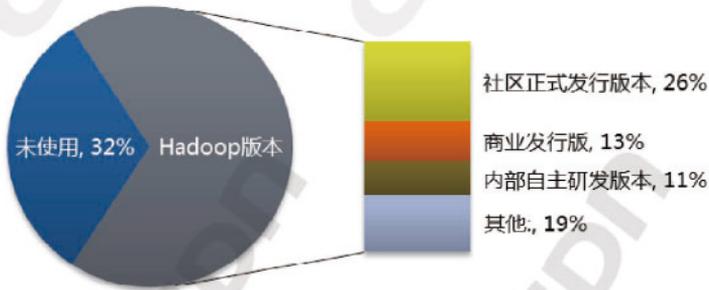
2017 年度 · 中国软件开发者白皮书

图表 45: 企业大数据平台集群规模

### 4.3 企业构建大数据平台所用技术情况

#### 仅13%企业使用商业发行版Hadoop版本搭建数据平台

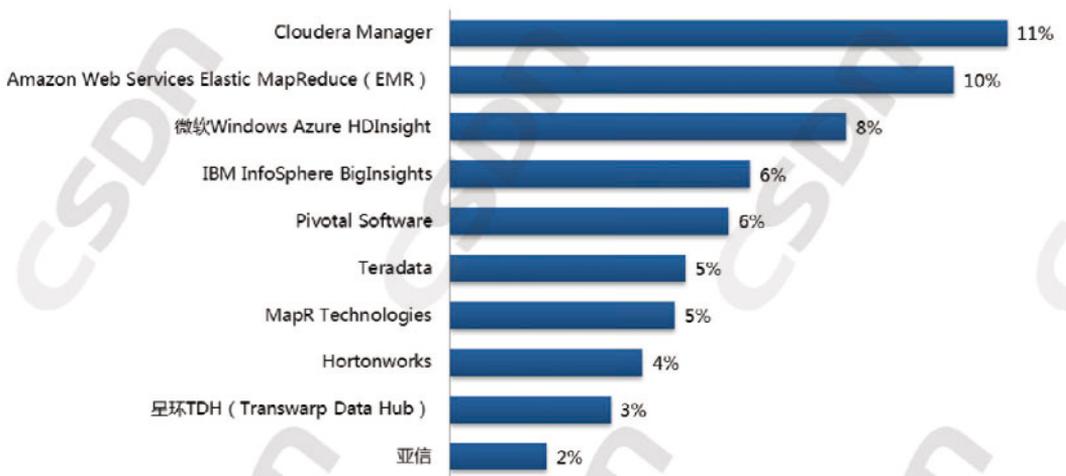
32%的企业未使用Hadoop搭建数据平台。使用社区正式发行版的占26%，而使用商业发行版的仅占13%。内部自主研发版本的占11%。



2017 年度·中国软件开发者白皮书

图表 46：企业使用 Hadoop 搭建大数据平台情况

在Hadoop商业版本中，Cloudera Manager使用率最高，但也仅有11%的受访者在用；其次是 Amazon Web Services Elastic MapReduce (EMR)，10%的受访者在用。



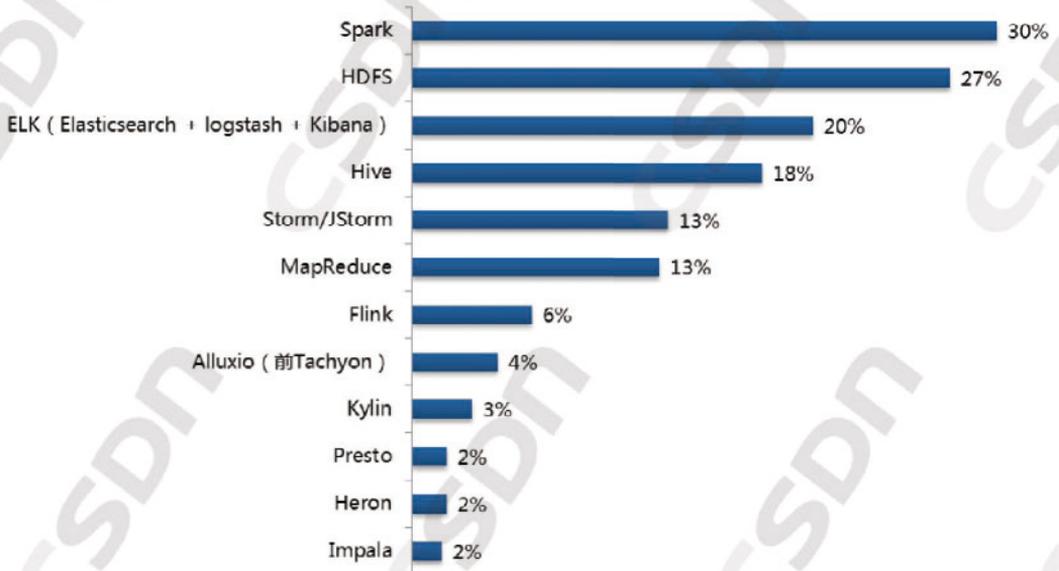
2017 年度·中国软件开发者白皮书

图表 47：Hadoop 商业版使用情况

## Spark是企业大数据平台最普遍的组件

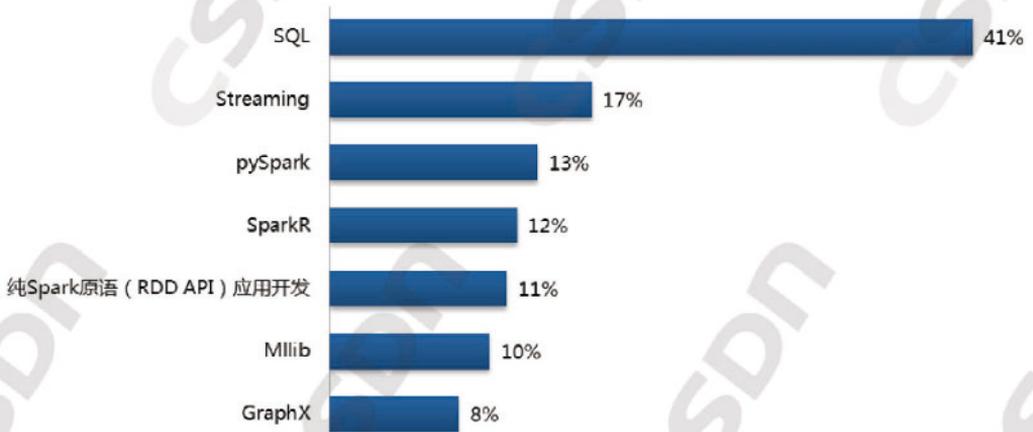
Apache Spark是一个处理大规模数据的快速通用引擎，它可以独立运行，也可以在Hadoop、Mesos、云端运行，它可以访问各种数据源，包括HDFS、Cassandra、HBase和S3，可以提升Hadoop集群中的应用在内存和磁盘上的运行速度。Spark生态系统中除了核心API之外，还包括其他附加库，可以为大数据分析和机器学习领域提供更多的能力。本次调研中，Spark是使用最普遍的大数据平台组件，使用率达到30%。而MapReduce使用率仅为13%。

分布式文件系统HDFS作为核心组件之一，使用率也达到了27%。企业对大数据平台应用最多的场景是统计分析、报表生成及数据可视化，20%企业使用ELK (ElasticSearch + Logstash + Kibana) 实时日志分析平台。



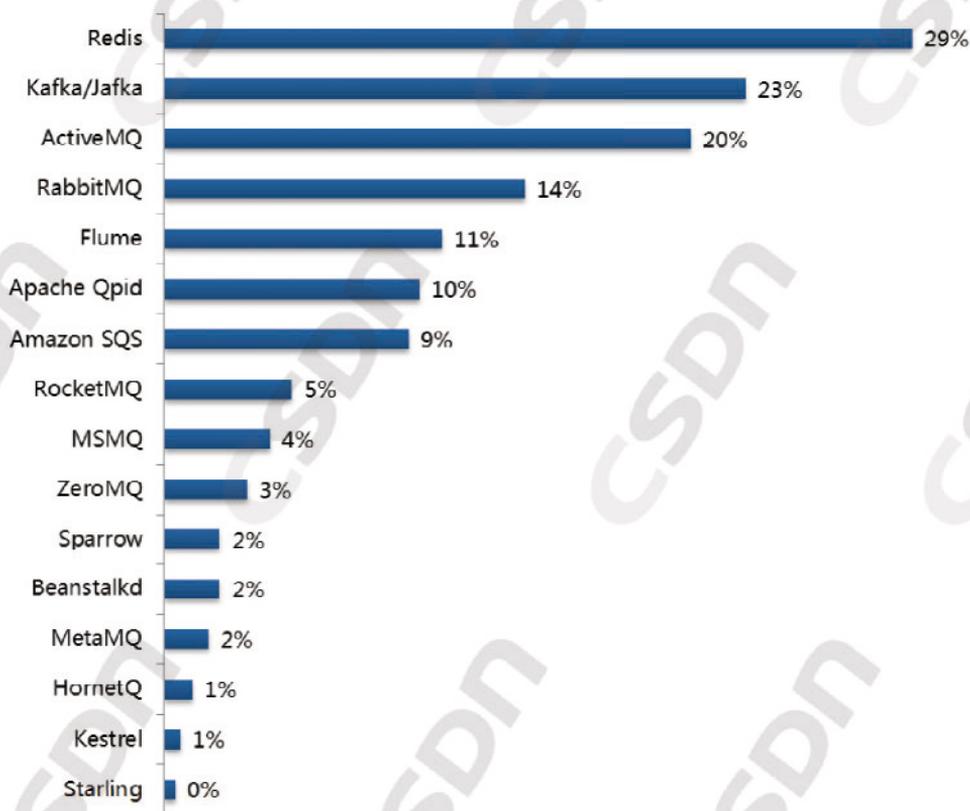
### Spark组件中使用SparkSQL和Streaming最多

Spark组建中，SparkSQL处理速度快并且能够和Hive完全兼容，以41%的使用率排在使用Spark组件首位。Streaming使用率17%。



## Redis和Kafka是最普遍使用的消息队列和数据采集技术组件

消息队列中间件是分布式系统中重要的组件，主要解决应用解耦、异步处理、流量削峰、消息通讯等问题。Redis使用最为普遍，占29%。Kafka占比为23%。



**【毕洪宇点评】**从报告中我们可以看出大数据技术的应用在2017年有个突增的发展，2017年是AI元年，感兴趣的读者可以在google trends检索几个关键词“big data”“machine learning”“IoT”的趋势，不难发现这个突增是和IoT及人工智能有极大关系的；

首先，在这样的背景下相应的大数据人才供给自然是跟不上市场需求的，报告中“企业大数据应用落地最大障碍”也佐证了这点；有7成的团队规模是不到10人团队的，而这个比例在2016年是49%，同时8成以上的企业集群规模在100个节点以下，在供需的矛盾下2017年通过私有云的方式来部署大数据平台成为了主要方式，而2016年是自主研发的方式为主；

其次，在应用场景上“统计分析、报表及数据可视化”依然是最多，因为这个场景是数据处理的基本工作，高频，随着更多企业对大数据的投入，这个场景自然就从统计上是最高的；从服务提供者的角度来说，基础工作的服务化是个很好的切入点，也为企业减少了重复的基础设施投入，更专注于用数据创造价值；从企业的角度来看，是持续重视对数据的洞察，从“big data”到“big insight”；

最后，在技术栈上有这么几点值得注意的地方：

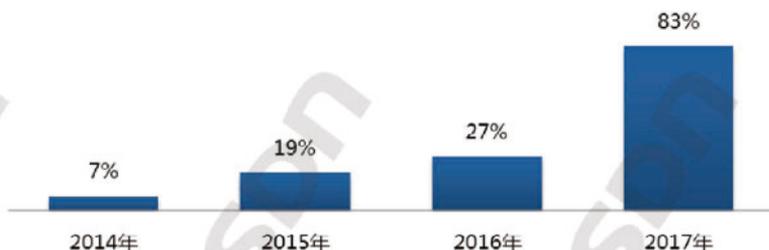
1. 企业内部的日志数据(主要包括系统日志和用户行为日志)是主要的数据源，schema-free的数据存储和传输技术需求会非常大，Kafka/Redis的统计比例也可以支撑这点；
2. HDFS和Spark已是大数据技术存储和计算的标配；
3. SQL接口依然持续是数据处理的主流形式，大多数计算引擎以及存储引擎的发展都趋向于此。

## 第五章 企业使用云服务现状分析

### 5.1 企业使用云服务方式分析

#### 超8成企业使用了云服务

调研数据显示，有83%的企业正在使用云服务，相较2016年有非常大的提升。

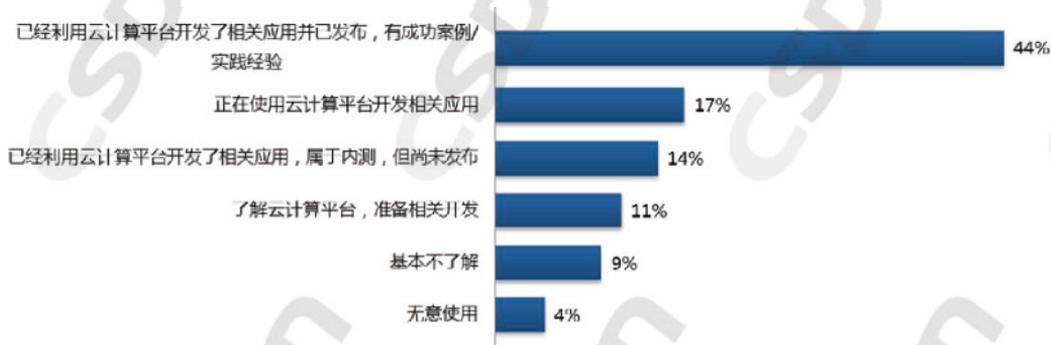


2017 年度·中国软件开发者白皮书

图表 51：使用云服务的企业占比

#### 超7成企业利用云平台已经开发应用或正在开发应用

2017年的数据显示，云计算已经相当普及，仅有不到1成左右企业对云计算平台基本不了解或者无意使用。75%的企业利用云平台已经开发应用或正在开发应用，11%的企业了解云计算平台，正在准备相关开发。

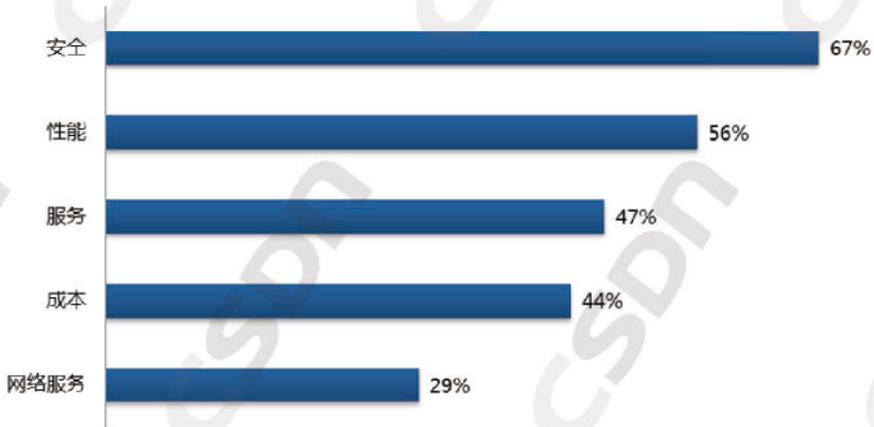


2017 年度·中国软件开发者白皮书

图表 52：企业对云计算平台的使用程度

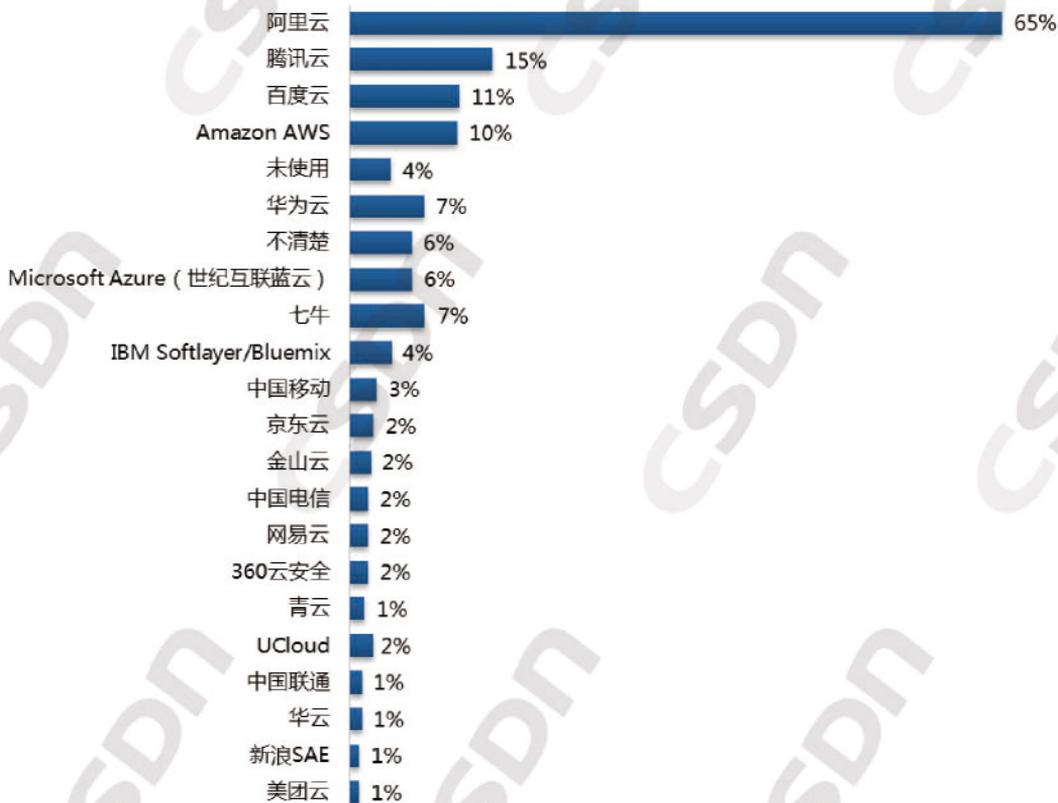
## 安全是企业选择公有云服务时最大的顾虑

云服务在带来大数据量存储、计算优势服务的同时，如何确保云服务环境的安全性及个人隐私性是企业最担忧的问题。本次调研中，云服务的安全性问题，成为企业选择公有云服务时最大的顾虑，67%的企业担心云服务的安全问题，56%的企业对云服务的性能表示担心。



## 阿里云领跑国内公有云平台市场

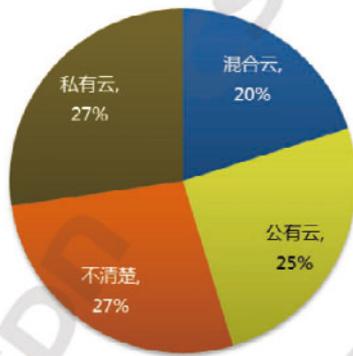
本次调研数据显示，阿里云服务以64%的使用率遥遥领先于其他云服务厂商，排在首位。其次是腾讯云服务，使用率为14%。



## 5.2 企业云平台部署特点

企业未来一年在私有云、混合云、公有云方面均有部署

私有云、混合云、公有云均有发展，未来一年部署的比例分别为27%、20%、25%。

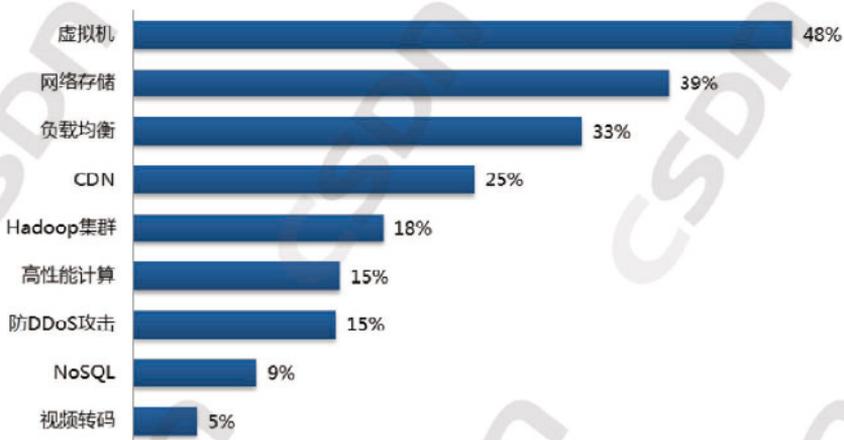


2017年度·中国软件开发者白皮书

图表 55: 企业下一年云服务部署模式分布

虚拟机、网络存储、负载均衡是云计算的三大使用场景

企业在虚拟机、网络存储、负载均衡方面应用云计算较为普遍，使用率分别是48%、39%、33%。

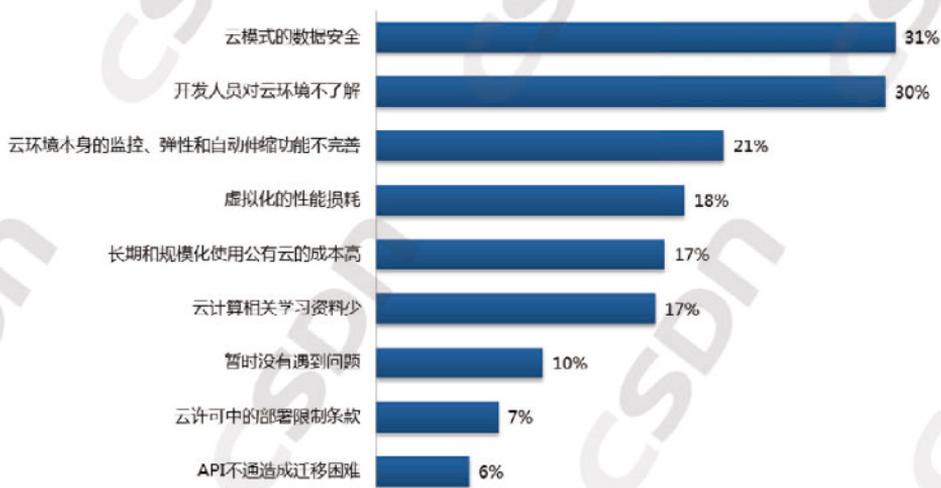


2017年度·中国软件开发者白皮书

图表 56: 企业云计算的使用场景分布 (调查项为多选)

## 数据安全是企业云技术开发中所面临的最主要问题

企业在云技术开发中所面临的最主要问题，依然是云模式的数据安全问题，本次调研中，占比31%。其次，开发人员对云环境不了解，也是企业相对普遍的问题，30%的企业对此表示担忧。

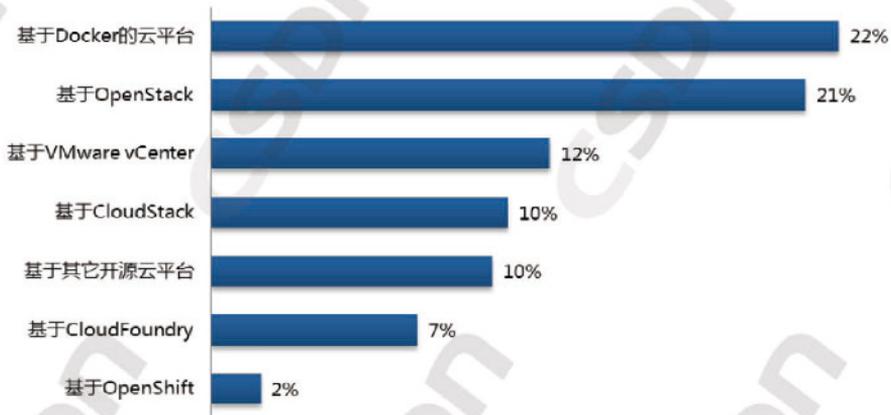


### 5.3 企业构建云服务平台技术选型分析

#### 基于Docker或OpenStack是当前云平台部署的两种主流框架

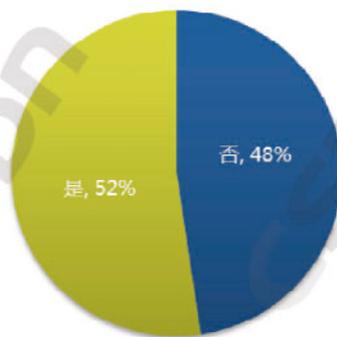
相对于VM, Docker在其轻量、配置复杂度以及资源利用率方面有着明显的优势。随着Docker技术的不断成熟,越来越多的企业开始考虑通过Docker来改进自己的IT系统。在本次调研中,基于Docker搭建的云平台占比为22%,位列第一位。

OpenStack是侧重于IaaS(基础设施即服务)的云平台,开放性是其最大的优势,由于参与社区的厂商众多,用户在基于OpenStack构建云环境时,计算、存储和网络方案的选型有很大的空间。在本次调研中,基于OpenStack的云平台占21%。



## 52%的企业在生产环境中使用了Docker技术

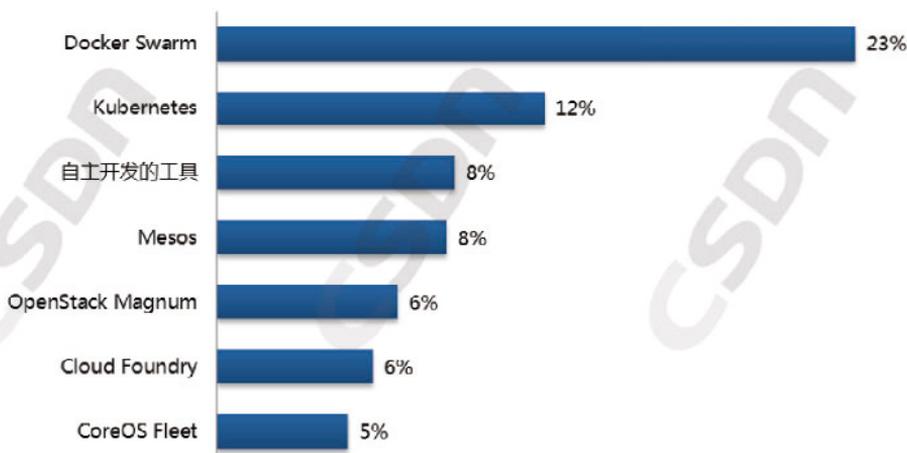
传统虚拟化技术是对硬件资源的虚拟，容器技术则是对进程的虚拟。云计算借助容器虚拟化技术，通过资源调度的方式实现虚拟机的动态迁移、监控等多种功能。通过调研数据发现，52%的企业在生产环境中使用了Docker技术。



2017 年度·中国软件开发者白皮书  
图表 59：企业生产环境中是否使用了容器技术

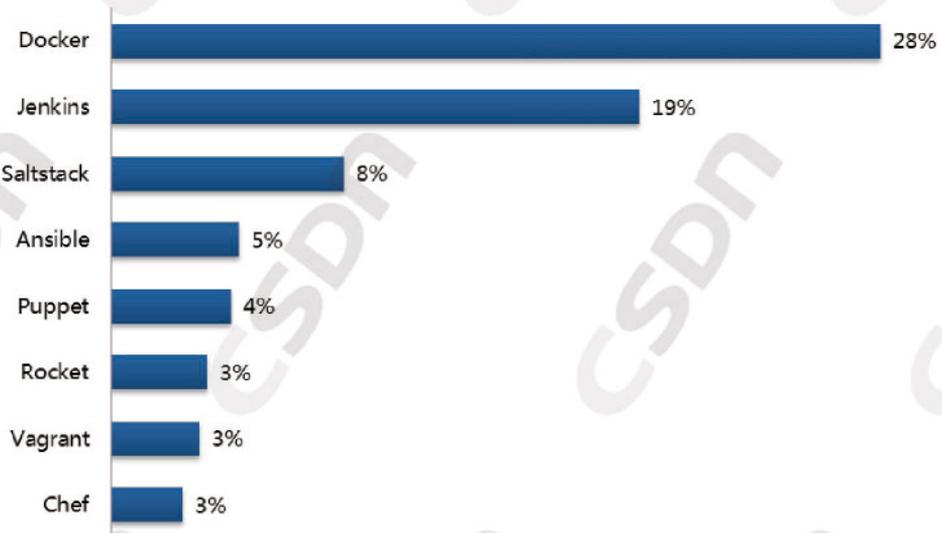
## Docker Swarm是企业使用最多的服务编排工具

在本次调研中，Docker Swarm是Docker的原生集群工具，被23%的企业使用。



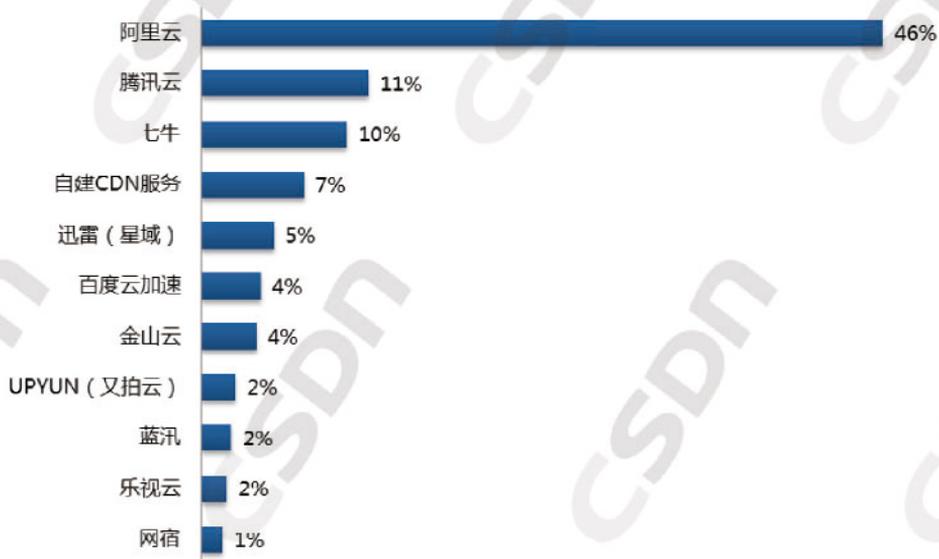
### DevOps工具中Docker和Jenkins的使用率较高

DevOps是以自动化工具与流程为基础，更快速、高效和稳定地实现环境搭建、代码部署、交付、测试以及发布的协同机制，可以促进开发与运维的紧密合作与交流。随着企业对DevOps理念认知度的提高，我们看到越来越多企业开始尝试在开发运维环节中使用DevOps工具。在本次调研中，Docker和Jenkins作为实现DevOps的重要技术手段被使用最多，占比分别为28%和19%。



## 阿里云是企业使用最为普遍的CDN服务

本次调研中，46%的企业在使用阿里云CDN服务，使用腾讯云的占11%。



**【李明宇点评】**就云计算部分来说,经过多年发展,云终于成为了企业建设IT系统的首选方案。在过去的一年中,超过80%的企业在使用云服务,超过70%的企业已经或正在利用云平台开发应用,相对于2016年只有27%的企业在使用云服务有大幅的提升,这跟我们从事云计算业务的各个公司在2017年从市场获得反馈是一致的。而对于未来一年,在各企业计划的部署中,私有云以27%的占比略高于公有云的25%,另外还有20%的混合云,这说明在未来市场预期中基本上是私有云和公有云平分秋色。

白皮书对调研数据已经进行了比较详细的分析,相信每位读者看完都会有自己的收获,就我个人而言,有两点体会比较深刻:

综合云计算与大数据相关调研数据,我们会发现,以往因为性能、数据规模等原因,许多人认为不太适合将大数据应用部署在云上,现在这个局面已经得到改变,2017年有超过半数的企业在部署大数据平台时选择了基于云的方案,而对于云上的应用,Hadoop和高性能计算加在一起占到了三分之一,这得益于云计算社区和大数据社区的共同努力,例如,从2008年开始,Hadoop社区经过近十年迭代开发,支持S3对象存储接口的模块已经相当完善,性能和稳定性俱佳,这些都为企业在云上部署大数据应用提供了便捷。

我们也能够从白皮书中了解到一些存在的问题,调研结果多处显示安全性是用户关心/担心的首要问题。确实,由于成熟的传统安全方案并不能直接套用到云环境中,企业在部署私有云平台时,如何设计和实施安全方案,企业在使用公有云时,如何确信服务商对数据做了足够的安全保护,都是云厂商和服务商经常被问到的问题。随着云技术得到越来越广泛的应用,人们必须回答这些问题。

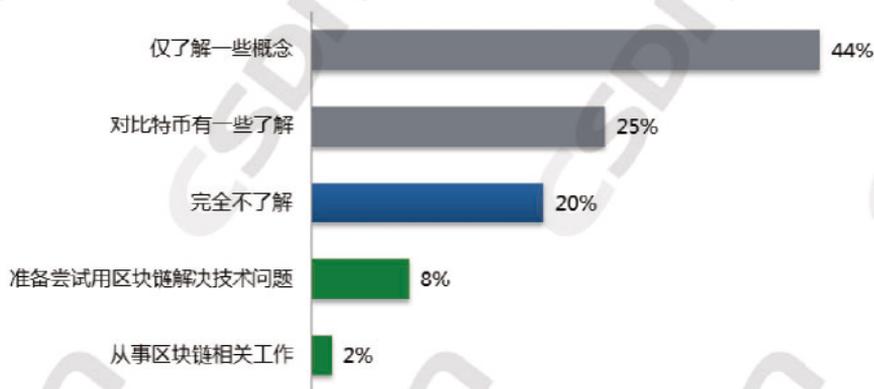
“春江水暖鸭先知”,应用的开发者正是云平台和云服务的使用者,CSDN《中国软件开发者白皮书》云计算部分的调研数据显示云技术正在得到大规模的应用。整个云计算产业方兴未艾,处于上升期,期待通过我们2018年的努力,在明年的白皮书中能够看到更好的表现。

## 第六章 区块链技术应用现状分析

### 6.1 区块链技术应用现状分析

#### 区块链技术当前应用仅一成

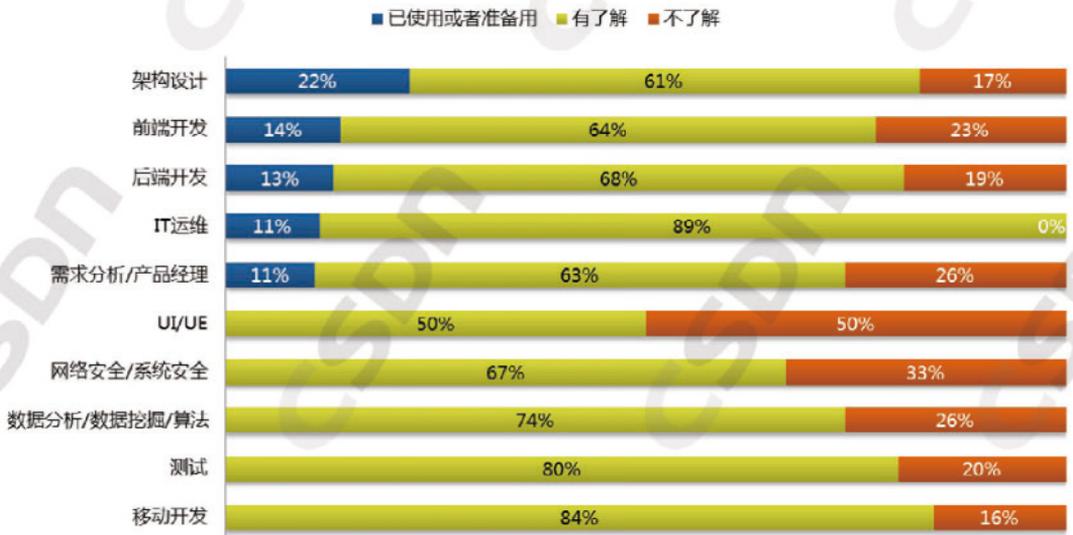
区块链作为近年来兴起的技术，当前应用相对较少。本次调研中，正在用或者准备用区块链技术解决技术问题的人群仅占受访者的10%。44%的受访者仅对区块链了解一些概念，25%的受访者对比特币有一些了解。有20%的人对区块链完全不了解。



**【王伟点评】**区块链的概念是2014年左右才被很多咨询公司和TMT领域媒体所提出并追捧，并且由于一直以“分布式账本”的概念出现，一开始基本上都是金融和相关行业领域专业人士所关注，即使在2015—2016年产生了大量POC项目，也很难辐射到所有IT从业人员。因此，绝大部分人应该还是从各类媒体和比特币的角度才了解到区块链。

## 架构设计岗位人员对区块链技术更为敏感

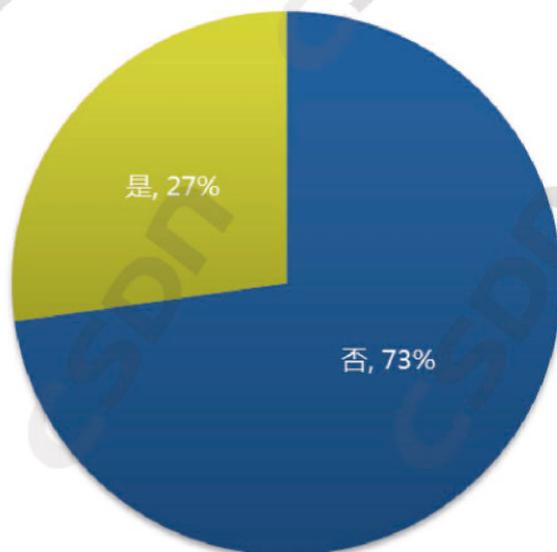
架构设计岗位人员有22%已经使用或者准备使用区块链技术解决技术问题。该岗位61%的开发者对区块链或比特币有部分了解，完全不了解的仅占17%。前端开发、后端开发人员在区块链技术的应用上也有所尝试，已使用或准备用的分别占14%、13%。



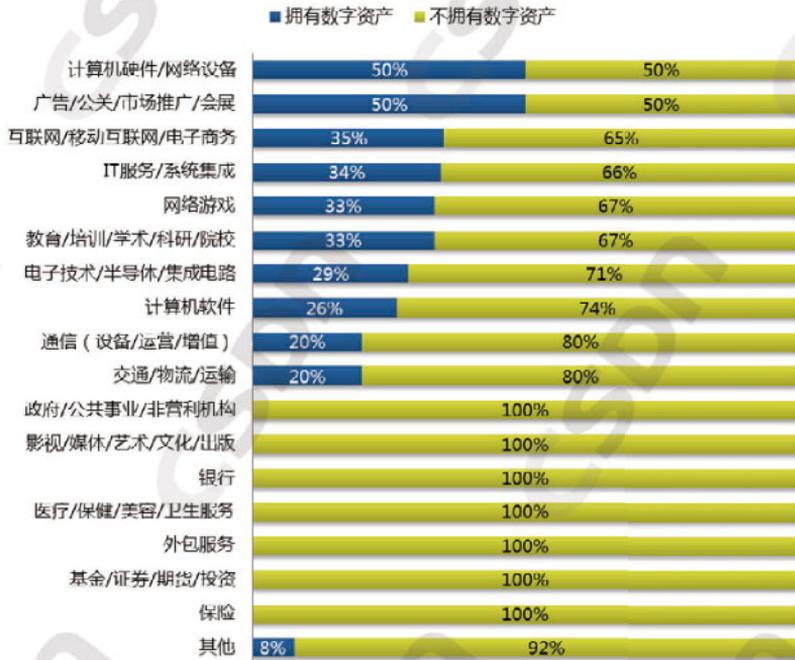
**【王伟点评】**现阶段，区块链从技术上讲还是属于后端架构类，因此迄今为止从事过与区块链技术相关的开发工作的人员，也都是这个领域的从业者。其他方向，如：测试、UI/UE、数据分析等，则由于系统形态和行业渗透度的原因，最多只做到听说/了解过这项技术的程度，在日常工作中都很难直接参与。未来随着领域覆盖面和影响力的扩大，会有更多的垂直领域产生与区块链的关联。

### 近3成开发者拥有数字资产，计算机硬件/网络设备行业的数字资产普及率较高

此次调研数据显示，27%的开发者拥有数字资产。



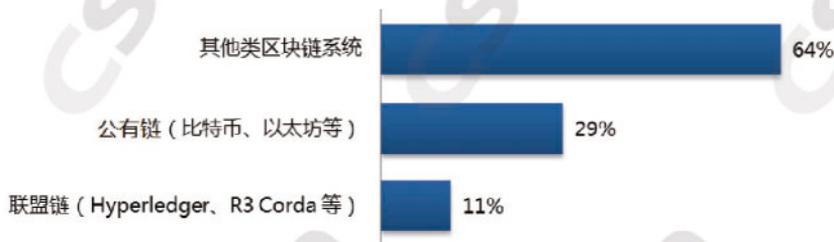
计算机硬件/网络设备、广告/公关/市场推广/会展行业的数字资产普及率较高，此次调研中50%的受访者拥有数字资产。互联网/移动互联网/电子商务、IT服务/系统集成、网络游戏、教育/培训/学术/科研/院校等行业拥有数字资产的开发者也达到或超过33%。



**【王玮点评】**从今天往回看，拥有数字资产的人普遍都是所谓“先知先觉”者，这一方面是因为他们所处环境能够更早接触到相关事物（如：计算机行业），或者是对新生事物比较感兴趣或者与国外信息接触比较多（如：广告、公关、市场推广等）。实际上，如果进一步分析计算机/网络领域的人员，应该会发现大量从事安全领域开发或研究，以及本身具有黑客属性的人员，拥有量是最多的，因为他们最早接触比特币等数字货币领域。

## 6.2 区块链技术开发特点

本次调研中，29%的受访者在从事公有链（比特币、以太坊等）的开发，11%在从事联盟链（Hyperledger、R3 Corda 等）的开发。

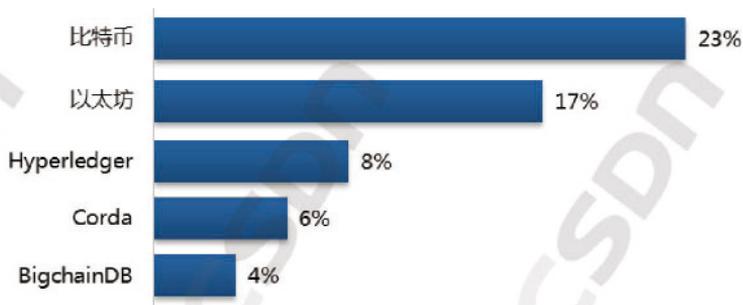


2017 年度 · 中国软件开发者白皮书

图表 67：从事区块链开发的形式（调查项为多选）

### 比特币和以太坊是当前两种主流的区块链开发平台

区块链开发平台发展迅速，截止本次调查时间（2017年11月），比特币和以太坊是当前两种主流的区块链开发平台，本次调研中，分别占23%、17%。



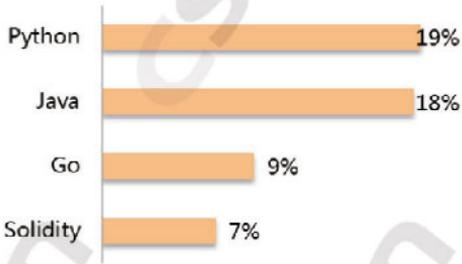
2017 年度 · 中国软件开发者白皮书

图表 68：开发者正在使用的区块链开发平台（调查项为多选）

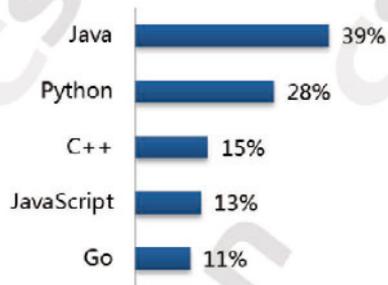
**【王玮点评】**从开发者的角度可以看出，人员分布基本服从于时间先后和领域影响力两个维度，这与任何一个技术门类在初始阶段的规律是相同的。相信未来随着公有链和联盟链各自的发展，开发者的分布会逐渐向影响力的方向倾斜，最终会产生几个主要的产品和技术体系。

## Python和Java是区块链的主流开发语言

智能合约的开发语言中，Python和Java应用较多，分别占19%、18%。在核心应用的开发中，Java使用更为普遍，占39%，Python其次，占28%。



2017年度·中国软件开发者白皮书  
图表 69: 智能合约开发语言 (调查项为多选)

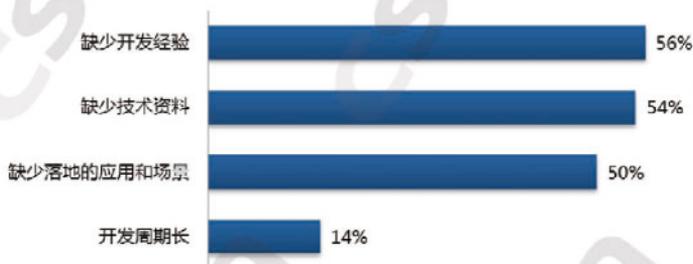


2017年度·中国软件开发者白皮书  
图表 70: 核心应用开发语言 (调查项为多选)

**【王玮点评】**从智能合约的角度看，Python、Java类开发者居多，应该反映出真正参与智能合约的开发者，普遍还是在联盟链领域，说明联盟链的开发者更多地倾向于将系统功能放在智能合约上，也是联盟链的主要能力和诉求之一。从核心开发者的角度看，应用开发仍然是Java为主，因此即使开发区块链相关应用，仍然会有很大一部分在传统技术栈上，采用与区块链对接的方式来运行，这应该说是一种合理的选择。

## 缺少开发经验、缺少技术资料、缺少落地的应用和场景是当前区块链开发的主要挑战

作为新兴技术，缺少开发经验、缺少技术资料、缺少落地的应用和场景是当前区块链开发的主要挑战，此次调研中分别占56%、54%、50%。



2017 年度·中国软件开发者白皮书

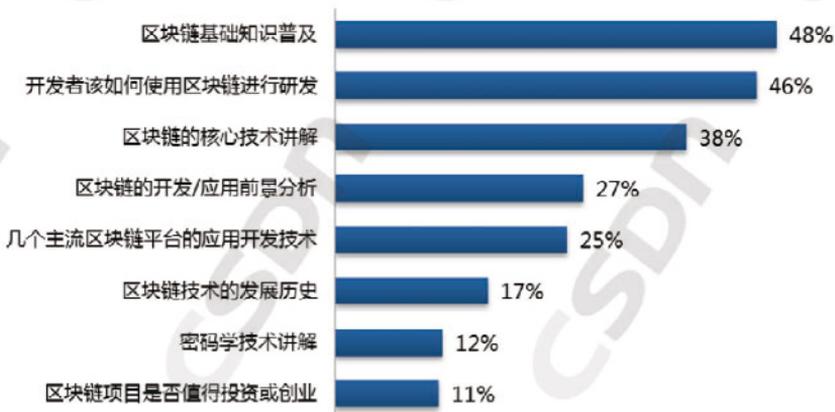
图表 71：区块链开发中面临的最大挑战（调查项为多选）

**【王伟点评】**缺少开发经验是任何一个领域初始阶段必然的现象，这个应该没有太多可担心的，正如网上的段子所说，“招聘十年以上区块链开发经验的程序员”只能是玩笑，这也正是区块链相关技术培训业务的机会。缺少技术资料应该也是同样的问题，不会成为技术和应用发展的阻碍。缺少落地的应用和场景，则某种程度上体现了区块链领域的隐忧，关于这一点，大家可以参考“通证派”有关的观点和文章。

### 6.3 区块链应用前景展望

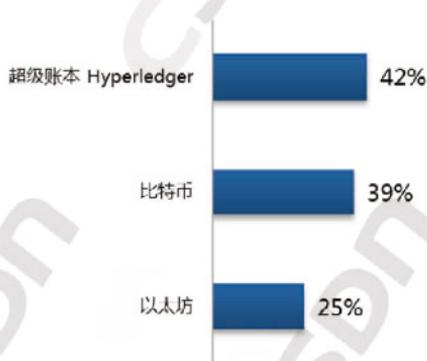
#### 超四成开发者希望了解区块链基础知识普及以及如何使用区块链进行研发

此次调研中，开发者最希望了解区块链基础知识普及和开发者该如何使用区块链进行研发，分别占48%、46%。可见开发者对新技术充满了兴趣，学习热情高涨。

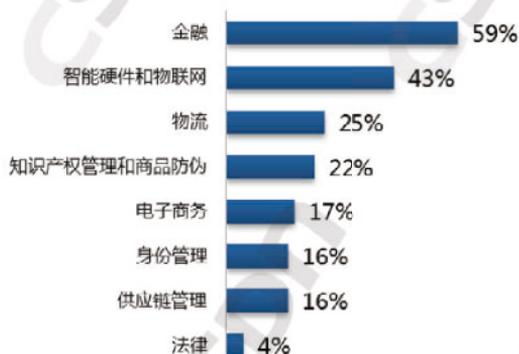


## 超级账本和比特币是普遍认为的技术方向，金融是普遍认为的行业应用方向

此次调研中，超级账本Hyperledger是普遍认为的技术方向，占42%，金融行业是普遍认为的行业应用方向，占59%。其次，智能硬件和物联网也被认为是主流应用方向，占43%。



2017 年度·中国软件开发者白皮书  
图表 73：区块链技术应用方向（调查项为多选）

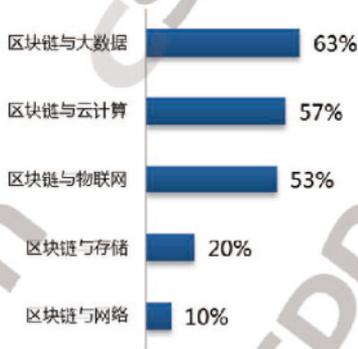


2017 年度·中国软件开发者白皮书  
图表 74：区块链行业应用方向（调查项为多选）

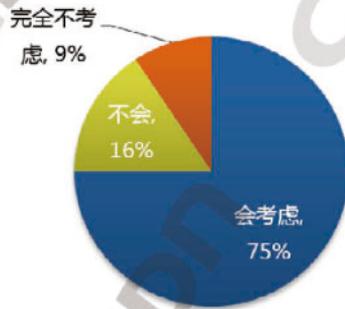
**【王伟点评】**超级账本的得票率高，应该跟程序员所处的行业关系比较密切。区块链做为后台技术，还是以行业应用的姿态出现的比较多，C端用户能接触到的很少。因此比特币/以太坊等公有链在2C行业中的影响力远不如超级账本等系统在2B行业中的影响力。金融行业作为区块链最早应用的领域，关注度一直遥遥领先，这与区块链技术的特性是密不可分的，即使在其他领域中应用，也是以“账本”的形态出现，很多与实物相关的应用，推进难度比较大，也有其内在原因。

## 超6成开发者看好区块链与大数据结合，超7成考虑1年内进入区块链行业

区块链与大数据、云计算、物联网行业的结合普遍被看好，此次调研中分别占63%、57%、53%。超7成开发者表示会考虑1年内会进入区块链行业。



2017年度·中国软件开发者白皮书  
图表 75: 被看好的区块链技术结合 (调查项为多选)



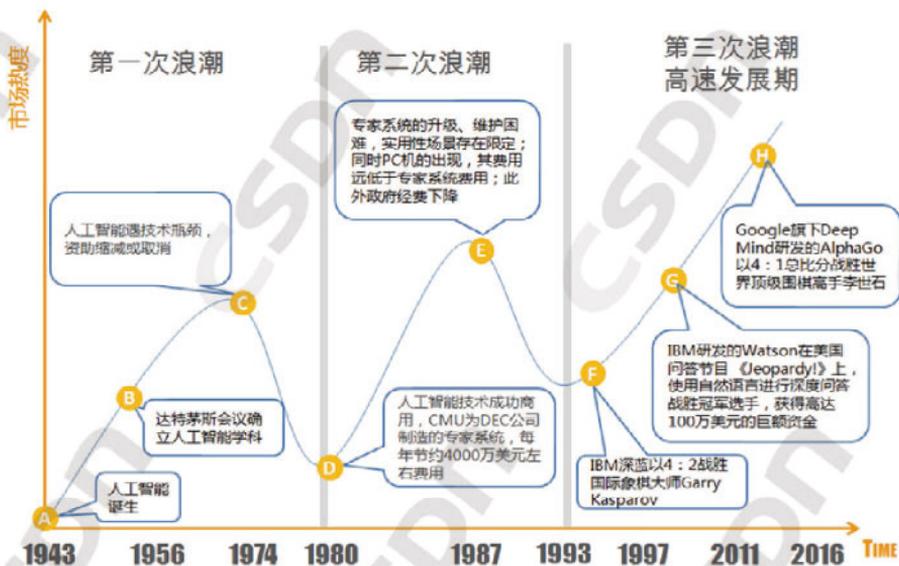
2017年度·中国软件开发者白皮书  
图表 76: 未来一年是否会考虑进入区块链行业

**【王伟点评】**因为区块链是个比较新的技术，大家难免容易产生与原有技术结合的视角，但事实上这些看法都存在未定之数，原则上说区块链还是相当特殊的一种技术，很难说未来发展与哪个领域会走得比较近。

从本次调查的总体情况看，区块链目前主要作为一种新兴的“后端”技术被开发者社群所认知，这与其技术特性和媒体宣传的形象基本匹配。同时，由于这种技术比较早期，又有相当的特殊性，因此从行业、应用场景采纳的程度以及开发者学习的速度来看，又都明显低于某种具体开发技术或者框架，大致与大数据、云计算、物联网等宏观层面的计算体系处于一个水平。考虑到区块链以及数字资产在技术、经济和社会领域的影响力，以及其所覆盖的应用场景的多样性，未来发展应该也与上述几个领域类似，会成为计算模型方面的一次新突破的尝试，对整个IT界及其他行业产生重大的影响。这些影响中，联盟链与公有链的影响又可能是截然不同的，这也是区块链技术的独特之处，时刻显现着其独有的魅力。

## 第七章 人工智能技术应用现状分析

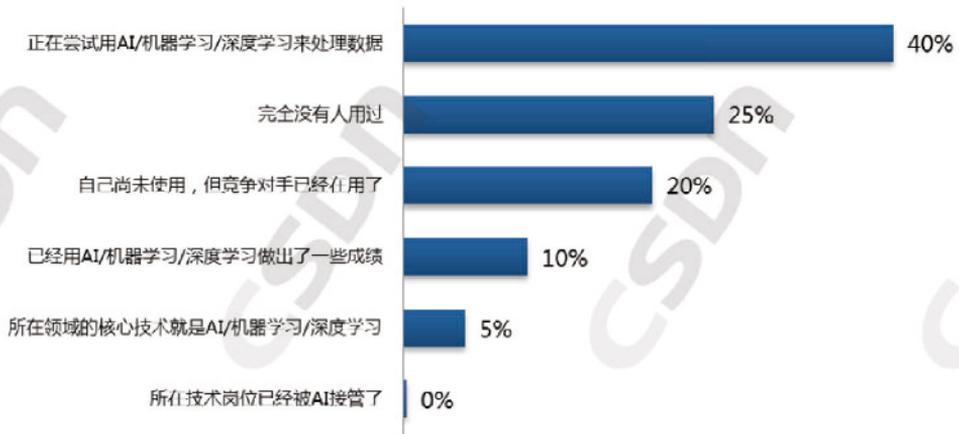
从1956年人工智能成为一门学科至今，人工智能发展经历几起几落。随着谷歌AlphaGo与世界围棋冠军开展对弈并获得胜利以来，人工智能技术应用再次火遍全球并获得资本和市场的热捧。人工智能技术在美国、韩国和日本飞速发展，其技术也开始应用到不同的领域，比如机器人、基于自然语言处理的智能客服、基于计算机视觉技术的智能安防等。



## 7.1 企业人工智能应用状况及面临挑战

### 4成开发者表示其技术领域正在尝试使用AI/机器学习/深度学习

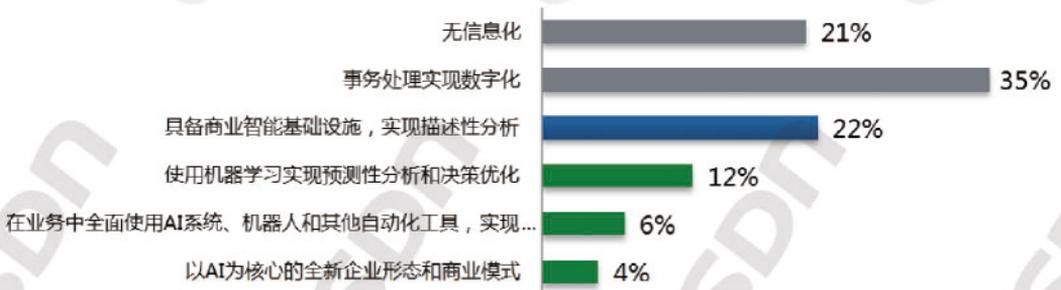
当前AI/机器学习/深度学习的普及率还偏低，但发展潜力很大。本次调研数据显示，已经使用AI/机器学习/深度学习技术的开发者仅占15%。40%表示正在尝试用AI/机器学习/深度学习来处理数据。20%表示自己尚未使用，但竞争对手已经在用了。只有25%的开发者表示完全没有人用过。



## 76%的企业尚未实现智能化

此次调研数据显示，21%的企业尚无信息化基础。35%的企业实现了事务处理数字化，20%的企业具备商业智能基础设施，实现描述性分析。

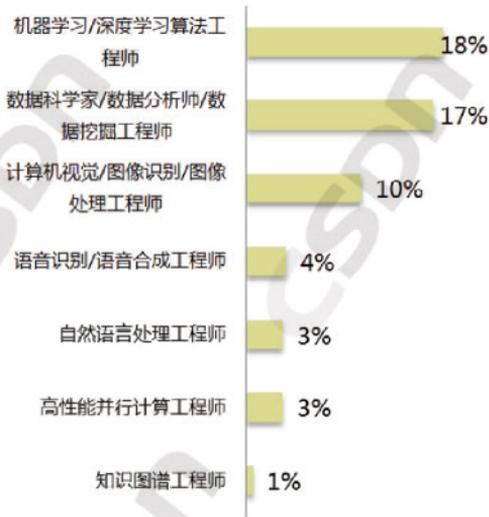
使用机器学习实现预测性分析和决策优化的企业占12%，而在业务中全面使用AI系统、机器人和其他自动化工具，实现创新和效率的提升或以AI为核心的全新企业形态和商业模式的企业仅占10%。



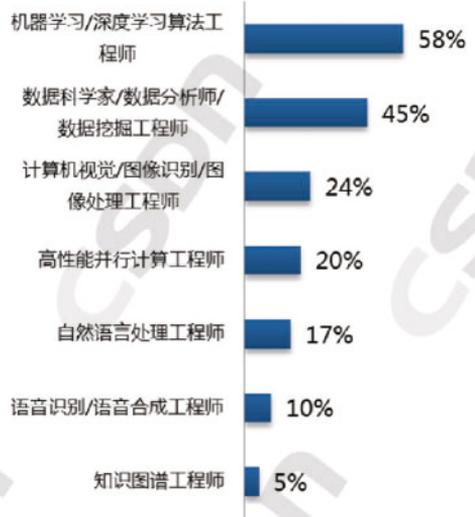
## 人工智能对机器学习/深度学习算法、数据分析的岗位需求更大

此次调研中机器学习/深度学习算法工程师、数据科学家/数据分析师/数据挖掘工程师岗位从业人员更多，分别占18%、17%。

当前最急缺的岗位也是机器学习/深度学习算法工程师、数据科学家/数据分析师/数据挖掘工程师岗位。58%的开发者表示其团队急缺机器学习/深度学习算法工程师，45%表示急缺数据科学家/数据分析师/数据挖掘工程师岗位。



2017 年度·中国软件开发者白皮书  
图表 80：开发者在 AI 领域所从事的具体工作岗位



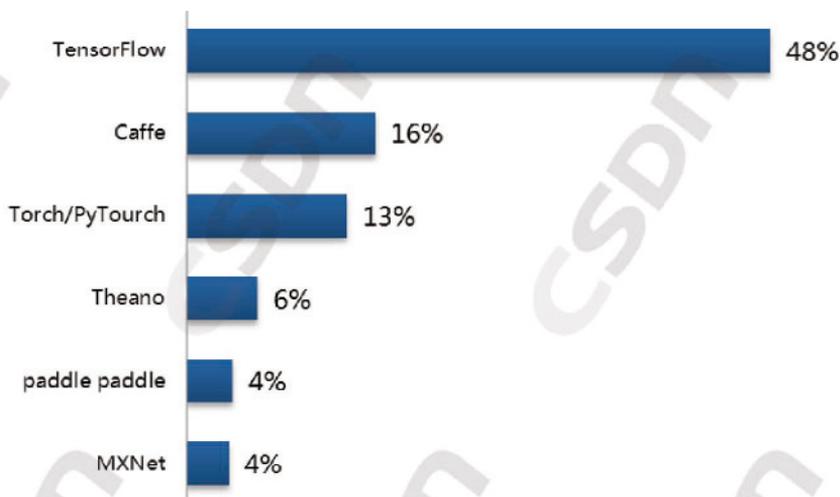
2017 年度·中国软件开发者白皮书  
图表 81：最急缺的 AI 工程师（调查项为多选）

## 7.2 人工智能技术应用现状

### TensorFlow是人工智能领域主流深度学习框架

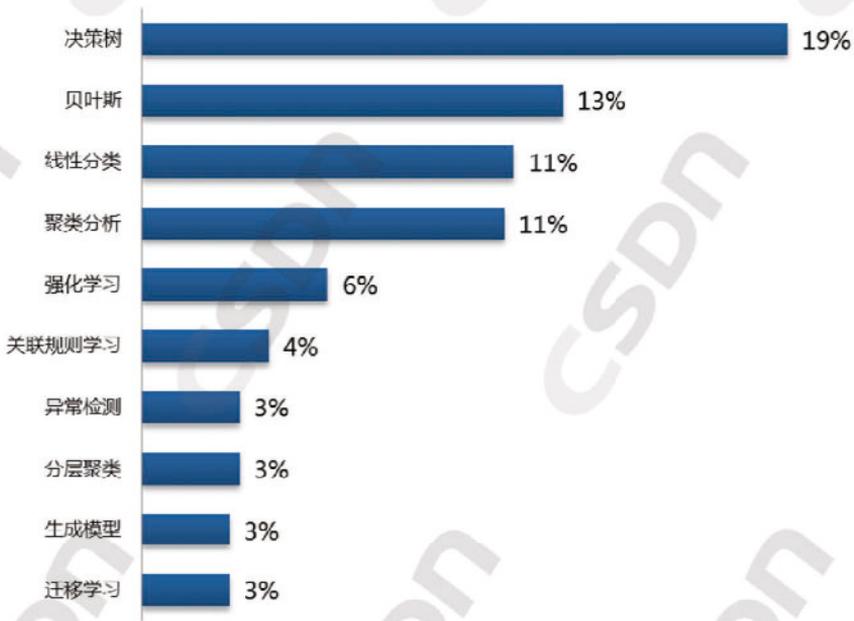
TensorFlow是谷歌基于DistBelief进行研发的第二代人工智能学习系统，是将复杂的数据结构传输至人工智能神经网络中进行分析和处理过程的系统，TensorFlow可被用于语音识别或图像识别等多项机器学习和深度学习领域，它可在小到一部智能手机、大到数千台数据中心服务器的各种设备上运行。

TensorFlow将完全开源，任何人都可以用。此次调研中，TensorFlow试用普及率达到48%。



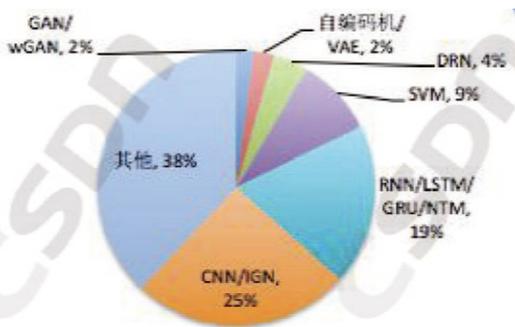
### 传统机器学习方法使用量接近2/3

此次调研中，传统机器学习方法在开发者中的使用量接近2/3。其中，有监督的线性分类、决策树占比30%，无监督的分层聚类、聚类分析、关联规则学习、异常检测占比21%，还有占比为13%的贝叶斯方法。而最新的生成模型、强化学习、迁移学习方法使用量也有12%，表现出一定的热度。决策树是开发者使用最多的机器学习类型，占19%，其次是贝叶斯，占13%。



### CNN类、RNN类深度学习模型最为火热

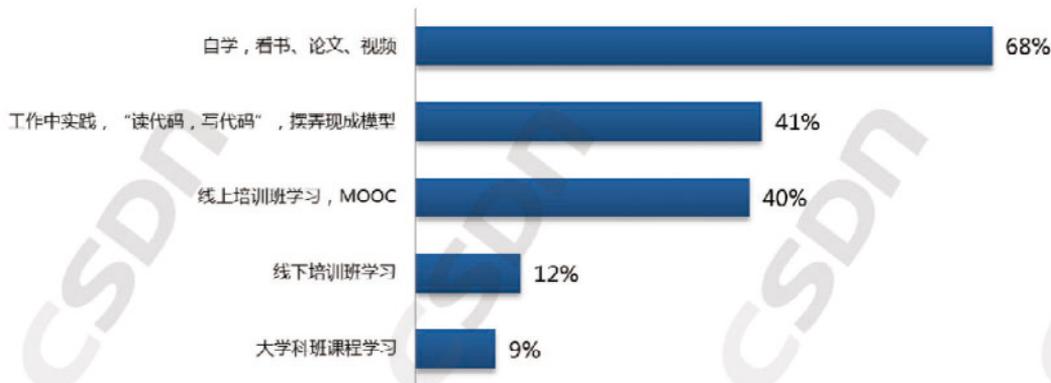
此次调研中，CNN类神经网络是使用最多的模型，占25%；RNN类其次，为19%；这两类深度学习模型占比达到44%。然后，还有38%是未做进一步细分的传统及最新模型，这说明具体的AI工作几无定式。



2017 年度·中国软件开发者白皮书  
图表 84：开发者使用最多的神经网络模型

### 自学是当前开发者学习人工智能技术的普遍方式

此次调研中，68%的开发者通过自学方式学习AI/机器学习/深度学习技术。但有四成开发者通过工作中实践、线上培训课程来学习AI技术。



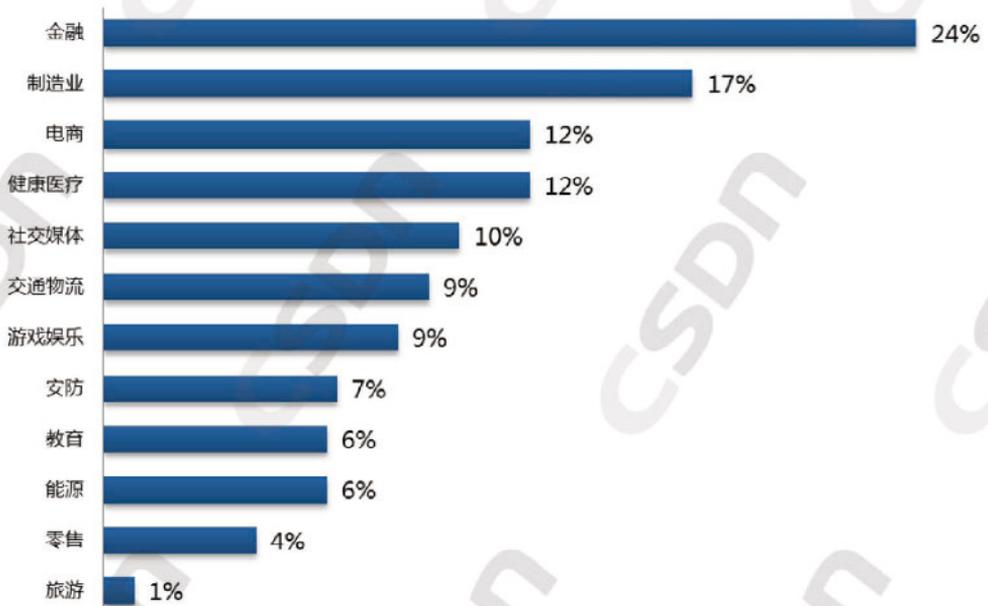
2017 年度·中国软件开发者白皮书

图表 85：开发者学习 AI/机器学习/深度学习技术的方式（调查项为多选）

### 7.3 人工智能行业应用及展望

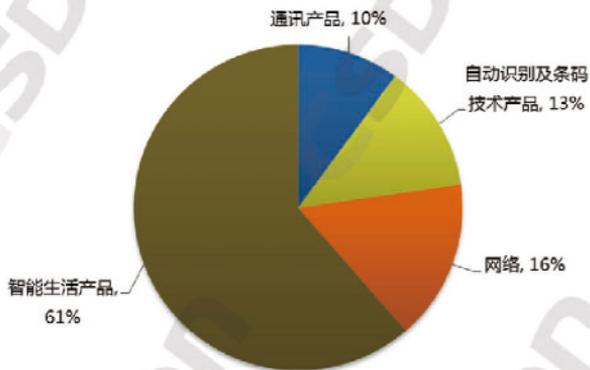
#### 金融行业是AI技术结合最多的行业

此次调研中，24%的开发者表示其所在企业AI技术正在结合金融行业业务。其次是制造业，占比17%



## 智能生活产品是最被期待的展会产品

智能生活产品以61%的比例高居AI业界技术会议的榜首。

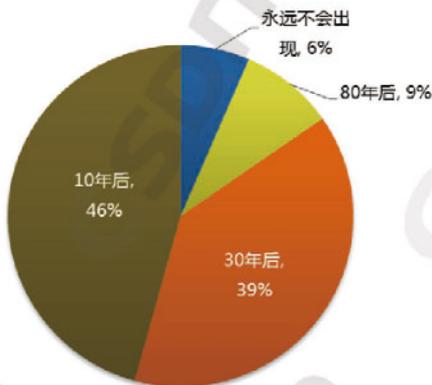


2017 年度 · 中国软件开发者白皮书

图表 87: AI 业界技术会议最希望看到的产品

## 超4成开发者预测10年后会出现人类级别的AI

46%的开发者表示10年后会出现人类级别的AI，但39%的开发者表示需要30年以后，9%的表示需要80年以后，甚至6%的悲观主义者表示永远不会出现。



2017 年度 · 中国软件开发者白皮书

图表 88: 预测人类级别的 AI 将在什么时候出现

**【智亮点评】**这里我们谈论的AI并非特指最近两年爆红的深度学习，而是另一个更大的范畴，机器学习，或者可以称为统计学习。它的诞生到现在其实也有了几十年的历史，随着运算能力的增强，它的作用范围也越来越广。但是它擅长于处理结构化的数据，更多被应用到后台系统中，通过对数据的分析和学习，在用户感受不到的地方提供着大量的支持。到了深度学习出现，终于从幕后走向了前台，开始发挥出改变这个世界的力量。从上个世纪70年代便已经有所发展的机器学习，目前在各行各业其实都在发挥着重要的作用。例如预测，分析，广告，检索。而这些算法，也一直在切切实实地为这些行业创造着价值和利润。这些并不是什么新鲜事儿，这些领域以前就在使用机器学习的技术，以后也会更多地使用。

而深度学习的意义在于，将特征选取和抽提的触角从结构化数据，探入了更接近人类的领域。它将传统机器学习算法的适用范围扩展到了视觉、听觉甚至未来的其他感官。就如同在移动互联网这个大环境下，共享单车和移动支付都是因为移动这个新的属性，而形成了新的业务场景。这一个简单的场景变化，形成了如此巨大的社会变革和商业机会。而机器学习也需要面对各种场景，解决各种问题，将能力和需求结合起来，才会焕发神奇的生命力。从这个角度来讲，AI会是一个产业，但是就像“互联网+”一样，会是“AI+”的表现形式。

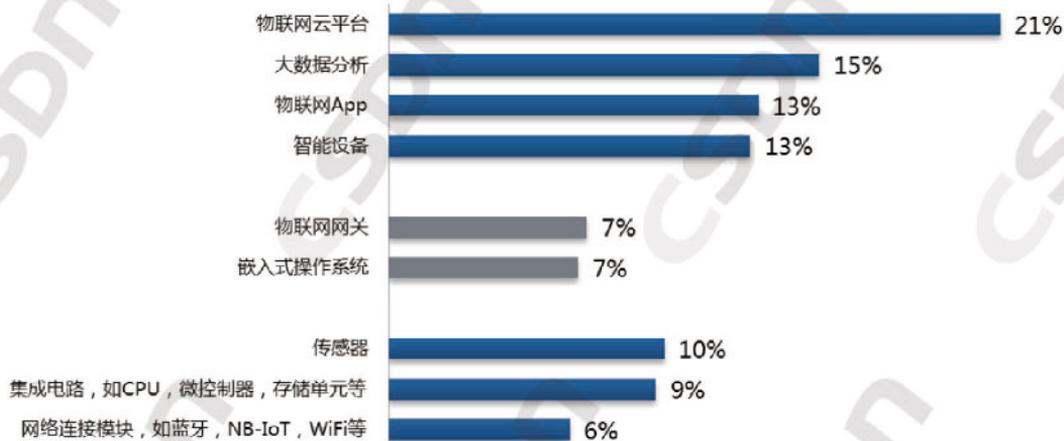
就如同调查的结果中显示的，在各个行业，都将会尝试着用机器学习/深度学习来解决自身行业的需求。而在这个过程中，最为稀缺的会是两类人才，第一类是需求定义人员。他应该熟悉行业自身，同时能够明确地判断存在哪些需求，其中哪些是可以用机器学习实现得很好，哪些则很难做到，是了解行业需求，同时也了解人工智能的复合人才，可以看做是机器学习时代的产品经理。第二类是需求实现人员，他在得到了明确的需求之后，能够综合运用各种技术和算法，提供一个能够满足需求的算法/模型方案，他们可以看做是机器学习时代的系统架构师。这两类人才的紧缺目前可以说是产业界应用机器学习的瓶颈。

## 第八章 物联网技术应用现状分析

### 8.1 物联网技术应用状况

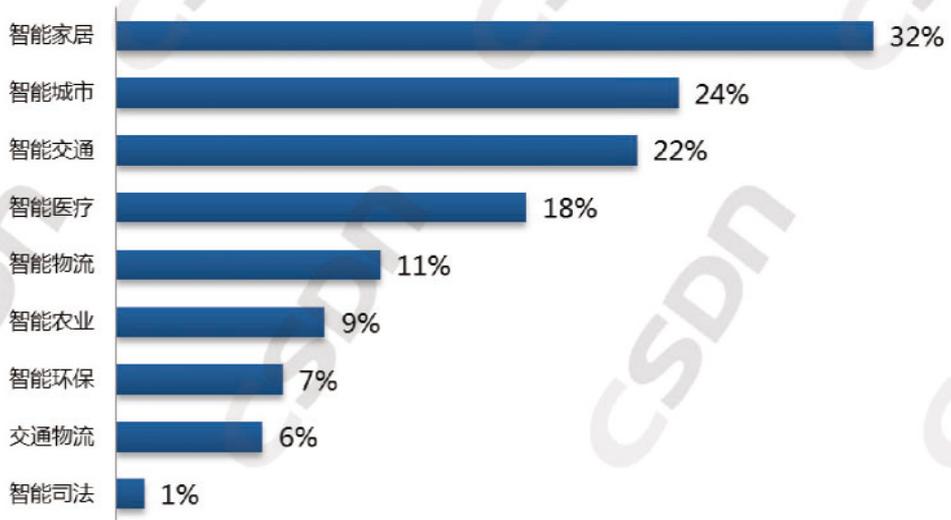
#### 物联网产业链上企业最多的在云平台环节

物联网云平台、大数据分析、物联网APP、智能设备是物联网产业链上企业较多的环节，分别占21%、15%、13%、13%。



## 智能家居是物联网最大的应用领域

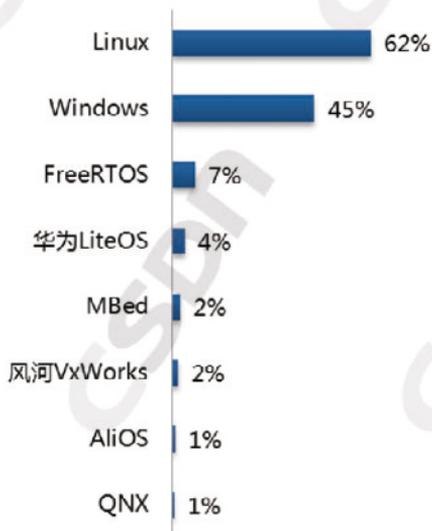
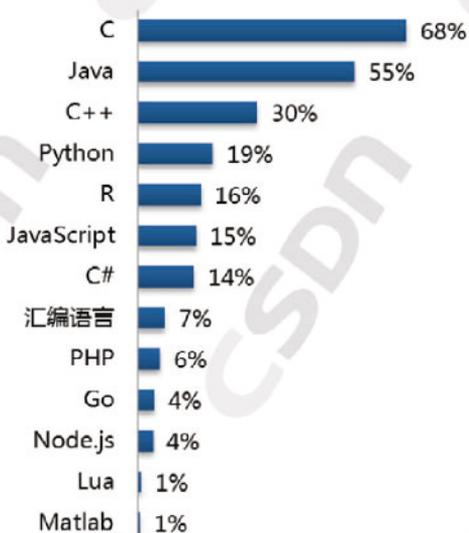
从此次调研结果来看，智能家居是物联网最大的应用领域，32%的开发者公司在用IoT技术结合智能家居行业业务。其次是智能城市、智能交通，分别占24%、22%。



## 8.2 物联网技术开发现状

C和Java是物联网领域较普遍的开发语言。Linux和Windows是较普遍的操作系统

C和Java是物联网领域较普遍的开发语言，此次调研中，使用率分别为68%、55%。Linux和Windows是较普遍的操作系统，使用率分别为62%、45%。



2017 年度·中国软件开发者白皮书

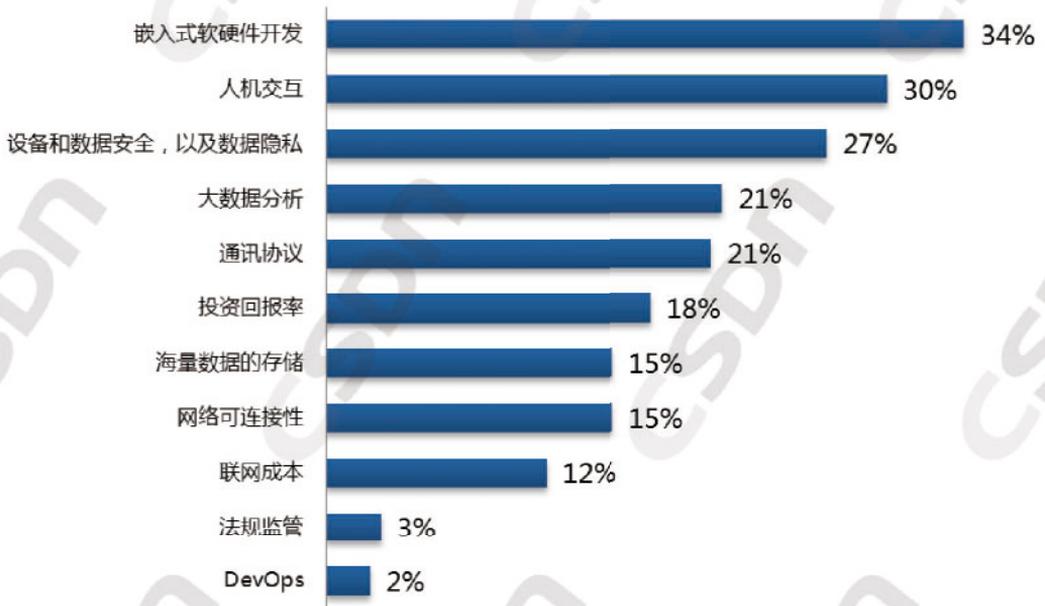
图表 91：开发者使用的开发语言（调查项为多选）

2017 年度·中国软件开发者白皮书

图表 92：开发者使用的物联网操作系统（调查项为多选）

### 开发者面临的最大挑战是嵌入式硬件开发

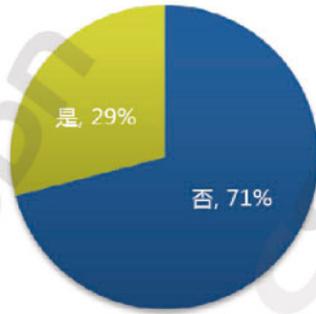
此次调研结果显示，34%的开发者面临嵌入式硬件开发困难，30%的开发者面临人机交互挑战。27%的开发者对设备和数据安全以及数据隐私表示担忧。



### 8.3 业务涉及工业物联网的企业调查

#### 近三成开发者所在企业业务涉及工业物联网

此次物联网调研中，29%的开发者所在企业业务涉及工业物联网。

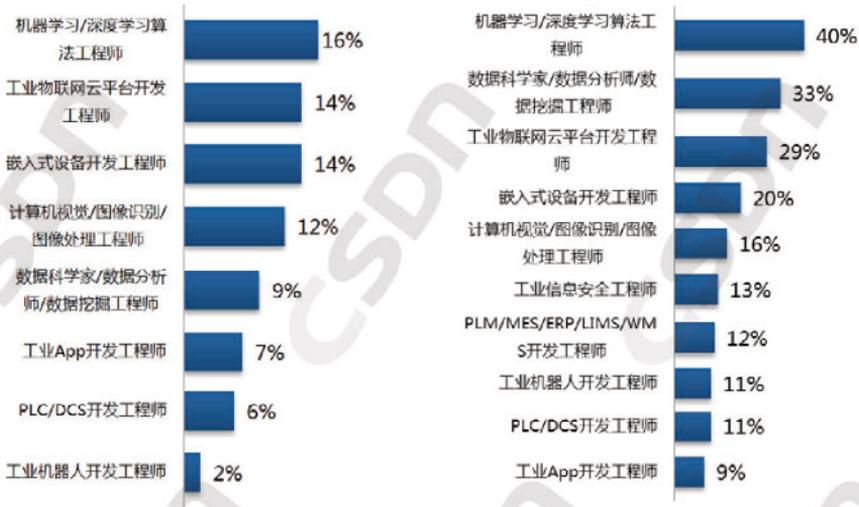


2017 年度·中国软件开发者白皮书  
图表 94：开发者所在公司是否涉及工业物联网

#### 机器学习/深度学习算法工程师是工业物联网领域需求最大也最急缺的岗位

此次工业物联网调研中，开发者最多的岗位是机器学习/深度学习算法工程师，占16%，其次是工业物联网云平台开发工程师和嵌入式设备开发工程师，分别占14%、14%。

当前开发者所在团队，40%缺乏机器学习/深度学习算法工程师，33%缺乏数据科学家/数据分析师/数据挖掘工程师。

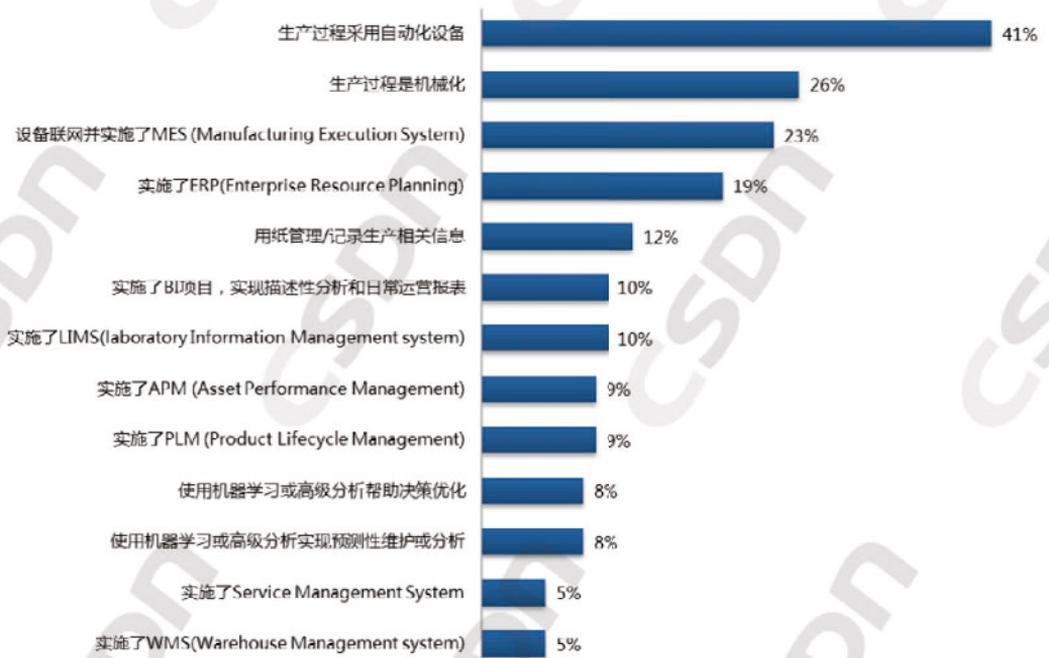


2017 年度·中国软件开发者白皮书  
图表 95：开发者在 IIOT 领域的岗位

2017 年度·中国软件开发者白皮书  
图表 96：目前最急缺的 IIOT 工程师（调查项为多选）

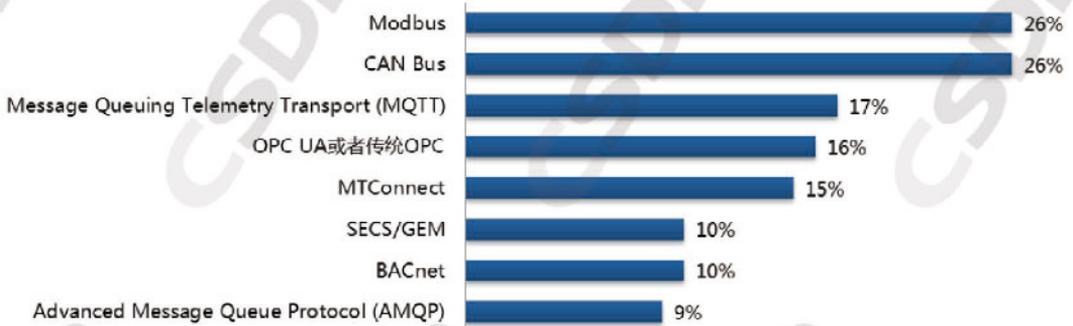
## 工业物联网开发者所处的公司或客户的主营业务目前的工业智能化水平参差不齐

从工业物联网开发者所处的公司或客户的主营业务目前的工业智能化水平来看，41%的工业物联网开发者所处的公司或客户生产过程中采用了自动化设备，但有26%的企业生产过程仍然是机械化。有23%的受访者表示设备联网并实施了MES，19%表示实现了ERP，但是也有12%表示当前用纸管理/记录生产相关信息。



## Modbus和CAN bus是两种应用最多的工业物联网协议

26%的受访者表示在应用Modbus或CAN bus协议。



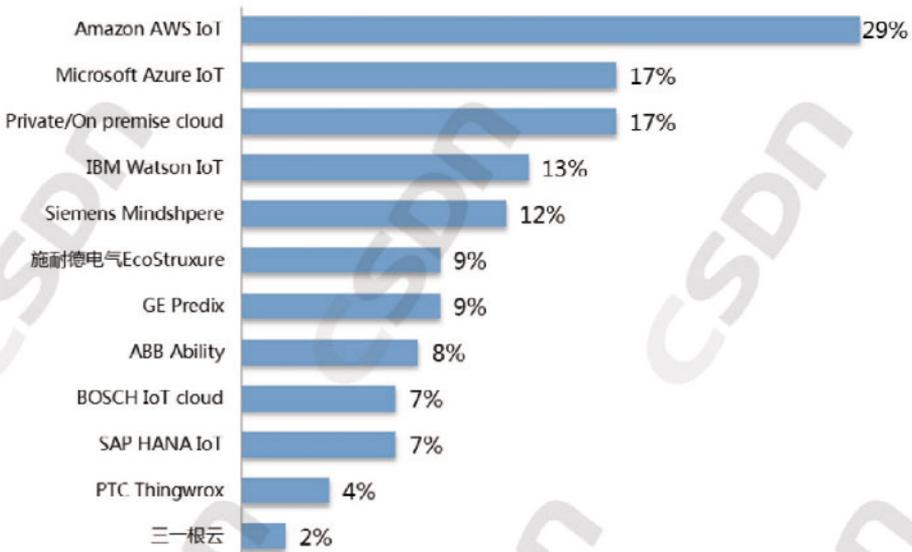
2017 年度·中国软件开发者白皮书

图表 98：开发者所在团队正在使用的 IIoT 协议

## 近3成工业物联网开发者在使用Amazon AWS IoT云平台

Amazon AWS IoT是当前应用最普遍的工业物联网云平台，有29%的工业物联网开发者在应用。

Microsoft Azure IoT 与Private/On premise cloud排其次，均有17%的工业物联网开发者在应用。



2017 年度·中国软件开发者白皮书

图表 99：开发者正在使用或了解的工业云平台

**【唐翎国点评】**刚刚过去的2017年，国内针对工业领域的物联网市场是百花争鸣，精彩纷呈。继GE Predix，西门子MindSphere，博世Bosch IoT Suite之后，国内外各路豪杰也纷纷推出了自己的平台或数字化解决方案。另一方面，如果我们对比Gartner在2016，2017两年的新兴技术周期曲线，就会发现IoT Platform刚刚从萌芽区进入炒作区。如果我们再回顾互联网行业在过去十几年的发展历程就可以看到，大规模的互联网应用是在网络连接的电脑和用户达到一定数量之后才开始爆发式增长的。因此，我是认同Gartner的分析的，也就是说，尽管工业物联网的未来市场巨大，但现在只是这场漫长马拉松的开始阶段，要想摘到胜利的果实，所有选手都必须在后面持续不断地投入。

同时，也必须现实地意识到，只有在连接的工业设备数量达到一定的数量后，真正意义上的大数据分析和各种工业应用才会有夯实的数据源。这又牵涉到若干挑战：

1.海量存量工业设备的数字化连接，也就是如何使没有传感器或者联网能力的老旧设备能够联网的问题。工业设备的特点是生命周期长，并且多数受到各种行业法规的约束，不是说企业想联网就可以联网，想改造就可以改造的；还要考虑到升级换代带来的资金问题以及对正常生产经营的影响。可喜的是，不少企业已经注意到了这个问题，推出了无损的“外挂”模块来采集老旧设备的部分关键影响参数，比如温度、振动等。

2.如何建立安全、可靠的网络连接，特别是最近爆发的英特尔、ARM等公司芯片级的安全问题，使得必须通过各种边缘网关才能连接到企业私有数据中心或者公有云的大量工业设备，都面临着被非法访问甚至控制的风险。这需要和英特尔、ARM以及主流操作系统供应商沟通协调来解决，并要通过严格测试甚至重新认证来确保各个厂家的工业设备可以被安全、可靠的接入网络。

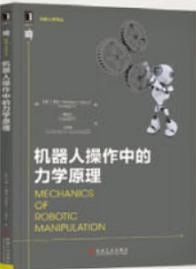
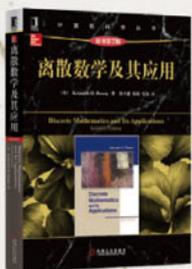
因此，我始终认为，只有脚踏实地的花3-5年甚至更长时间，安全、可靠地帮助工业用户连接更多的设备到物联网平台的企业，才是负责任的和有希望的企业。

## 附录：中国城市线具体划分

城市线	城市名称
一线城市	北京、上海、广州、深圳、天津
二线发达城市	杭州、南京、济南、重庆、青岛、大连、宁波、厦门
二线中等发达城市	成都、武汉、哈尔滨、沈阳、西安、长春、长沙、福州、郑州、石家庄、苏州、佛山、东莞、无锡、烟台、太原
二线发展较弱城市	合肥、南昌、南宁、昆明、温州、淄博、唐山
三线城市	乌鲁木齐、贵阳、海口、兰州、银川、西宁、呼和浩特、泉州、包头、南通、大庆、徐州、潍坊、常州、鄂尔多斯、绍兴、济宁、盐城、邯郸、临沂、洛阳、东营、扬州、台州、嘉兴、沧州、榆林、泰州、镇江、昆山、江阴、张家港、义乌、金华、保定、吉林、鞍山、泰安、宜昌、襄阳、中山、惠州、南阳、威海、德州、岳阳、聊城、常德、漳州、滨州、茂名、淮安、江门、芜湖、湛江、廊坊、菏泽、柳州、宝鸡、珠海、绵阳、三亚
四线及以下城市	除以上城市外的其余城市

# 计算机科学丛书 20周年

20本跨世经典 | 夯筑科技基石



机械工业出版社  
China Machine Press



扫描二维码了解  
更多经典著作

# 博文视点 IT出版旗舰品牌

## 2018热门好书优选



博文视点Broadview

### 人工智能



### 网站架构



### 编程语言



### 新书推荐



# GitChat

---

聚合业界顶级技术专家，是技术菁英知识付费分享社区与学习平台，打造“阅读+交流=知识+社交”新理念，满足IT从业者日新月异的知识学习需求，改变知识学习方式。

## 平台优势：

- 不仅是知识付费平台，更是人人参与分享与交流的高品质社区
- 语音文字互动，文章、课程、书籍，为用户提供最合理的一站式学习内容
- 依托微信，发挥社交的力量，为分享者觅得知音
- 付费会员入驻，积累技术口碑与品牌的最佳渠道



## CSDN新媒体



**CSDN公众号**是CSDN旗下最具影响力的开发者新媒体。目前在新榜科技类自媒体排名位列前二十，CSDN资讯关注于技术人所必须要知道的业界热点和大事，以技术人的独特视角，描述技术人关心的焦点事件。截止2018年3月，CSDN微信公众号矩阵总关注人数288万。



**程序人生**传承CSDN论坛最活跃板块“程序人生”，是一个以程序员日常工作和生活紧密相关且垂直服务于程序员群体的新媒体。内容涵盖程序员关心的方方面面。新榜平均排名前70。拥有“404音频电台”、“代码传奇”“程序员有话说”等明星栏目。



**AI科技大本营**短短一年中，成长为AI行业最有影响力的大号，为中国AI行业全面赋能，取得了强大而自传播的用户口碑。一时间直击行业热点，解剖背后逻辑，提供扎实干货，并通过线上线下互动、培训、读者体系、作者体系等一系列建设，激活AI产业的所有人员，以四两拨千斤的杠杆挑动行业神经，完成传播、产品、线下再到线上的完美闭环。



**区块链大本营**以技术开发者为核心，为技术开发者提供资讯、解读、观点、访谈等知识服务，并配合以线下沙龙、线上培训、深度孵化、区块链大会为技术开发者提供交流和培训服务。区块链大本营的核心宗旨是，全心全意为技术开发者的入门、转型、升级提供立体式服务。



**人工智能头条**是CSDN纯技术类AI垂直公众号，关注AI技术的发展，坚持秉承原创的原则，为AI领域从业者提供优质国外、国内专家文献。



**CSDN技术头条**网络互连网前沿技术、一线科技公司的热门架构和应用案例。网络IT领域技术专家，向IT从业人员提供纯干货的技术内容。

**CSDN**

网址：[www.csdn.net](http://www.csdn.net)

邮箱：[business@csdn.net](mailto:business@csdn.net)

电话：010-64376055

地址：北京市朝阳区酒仙桥路10号院恒通商务园B8座二层（100015）



CSDN