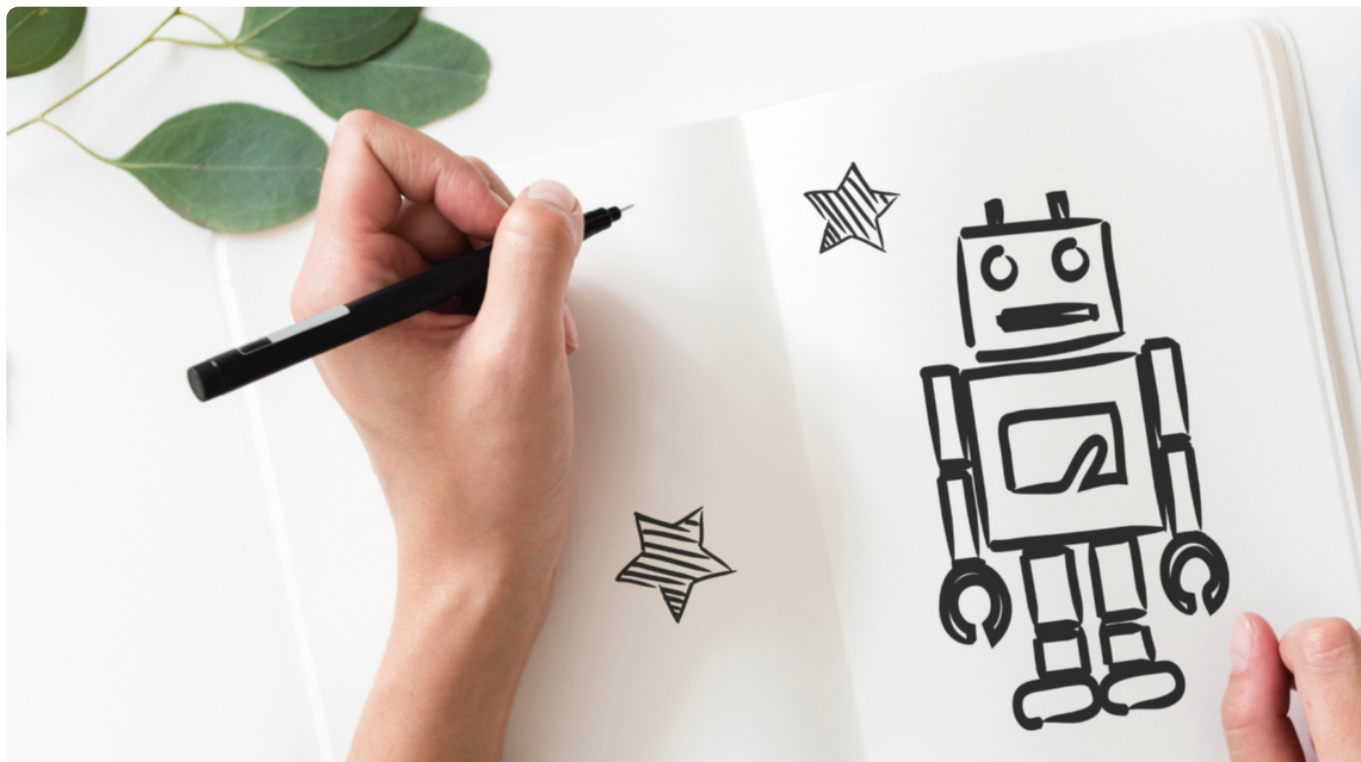


## 开篇词 | 为什么说每个软件工程师都应该懂大数据技术？

2018-10-29 李智慧

从0开始学大数据

[进入课程 >](#)



讲述：李智慧

时长 06:09 大小 2.82M



2012 年的时候，我从阿里跳槽到 Intel 做大数据开发。当时很多人不理解，我为什么会从如日中天的互联网公司跳槽到“传统”的 IT 公司。

我是这样考虑的，软件编程技术出现已经半个多世纪了，核心价值就是**把现实世界的业务操作搬到计算机上，通过计算机软件和网络进行业务和数据处理**。我们常见的软件系统，不管是电子商务还是库存管理，不管是搜索引擎还是收银终端，都是如此。这一点价值巨大，可以成百上千倍地提高我们的生活和工作效率。

但是时至今日，能用计算机软件提高效率的地方，几乎已经被全部发掘过了，计算机软件成为人们日常生活的必备品，人们已经习惯了计算机软件的存在。在这种情况下，如果想让软件再成百上千倍地提高我们的生活和工作效率，使用以前的那套“分析用户需求和业务场景，进行软件设计和开发”的做法显然是不可能的了。

那如何走出这个困局呢？我觉得，要想让计算机软件包括互联网应用，能够继续提高我们的生活工作效率，那就**必须能够发掘出用户自己都没有发现的需求，必须洞悉用户自己都不了解的自己**。

计算机软件不能再像以前那样，等用户输入操作，然后根据编写好的逻辑执行用户的操作，而是应该能够预测用户的期望，在你还没想好要做什么的情况下，主动提供操作建议和选项，提醒你应该做什么。

这听起来很科幻，但实际上已经出现了，那就是大数据技术和机器学习技术，也就是后来我们都耳熟能详的人工智能 AI 技术。

现在回过头来看，我当时的选择是正确的。就在我加入 Intel 从事 Hadoop 开源软件开发的第二年，也就是 2013 年，大数据技术开始火热起来，从 BAT 到传统的商业公司，都纷纷在自己的软件系统中大规模使用大数据技术，有的公司甚至称自己为大数据公司，而 2013 年这一年，后来也被称为“大数据元年”。

又过了 3 年，也就是 2016 年，Google 的 AlphaGo 横空出世，让我们见识到了“**大数据 + 机器学习**”的巨大威力。

所以，我同意这样一种说法：在未来，软件开发将是“**面向 AI 编程**”，软件的核心业务逻辑和价值将围绕机器学习的结果也就是 AI 展开，软件工程师的工作就是考虑如何将机器学习的结果更好地呈现出来，如何更好地实现人和 AI 的交互。

我曾经跟一个同学讨论这个观点，他认同面向 AI 编程，但是他认为：这并不意味着我一定要懂 AI，也不一定要懂大数据和机器学习，我只要懂业务，理解机器学习算出的结果就可以了。

真的是这样吗？只需要懂业务就能在“面向 AI 编程”的时代胜任软件开发的工作吗？

在阿西莫夫的科幻经典巨作《基地》中，描述了这样一个场景：在银河系，随着战争的蔓延，很多星球的科技在逐渐退步，到后来，他们虽然还有核电站等高科技产品，但是已经不知道是如何运作的了。而在银河系的边缘，有一颗小星球，在大战爆发前从银河系各处转移来了大量的科技文献，这颗小星球也没有加入战争并将科学技术一直传承了下去。

后来，当其他星球的科技产品出现问题的时候，就会向这个小星球求援，小星球会派工程师前去维修。但是他们并不管工程师叫工程师，而是叫“僧侣”；也不管核电站叫核电站，而是叫“圣殿”；维修也不是叫维修，而是叫“祈祷”。

他们的说法是：因为这个星球上的人做了不该做的事，比如发动战争、破坏环境等，触怒了神，所以神怪罪下来，让他们失去能源，如果想恢复能源，就必须纠正自己的错误行为并向神祈祷赎罪。所以，当工程师进入核电站维修的时候，整个星球的人都跪下祈祷，当电力恢复的时候，大家纷纷称颂神的伟大。

你看，科学和宗教并不是互斥的，科学也可以成为宗教，当人们面对自己不懂的东西的时候，会倾向于用宗教的原理去解释。

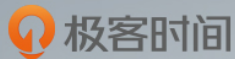
**如果未来是面向 AI 编程的，希望软件工程师不要把 AI 当作什么万能的东西。当机器学习结果出现问题的时候，我们既不要陷入某种不可知的“玄学”之中，也不要无谓的抱怨什么“人工智障”，而是应该积极参与到问题的讨论、分析和解决中去。这也是我的观点，即使自己不做大数据与机器学习相关的开发，每个程序员也应该懂大数据和机器学习。**

将来，数据会越来越成为公司的核心资产和主要竞争力，公司的业务展开和产品进化也越来越朝着如何利用好数据价值的方向发展。不懂大数据和机器学习，可能连最基本的产品逻辑和商业意图都搞不清楚。**如果只懂编程，工程师的生存空间会越来越窄，发展也会处处受限。**

如果说，大数据技术和应用是一个技术的殿堂，希望我的专栏不仅可以带你找到进入大数据“殿堂”的钥匙，也能透视“殿堂”里的结构、装饰、家具，告诉你为什么用这些元素可以构建恢弘的“殿堂”，以及如何更好地利用这个殿堂的空间与设施，而不是你进入“殿堂”看到一张床就舒服地躺下，错失了更美的风景。

**学习大数据最好的时间是十年前，其次就是现在！**

欢迎你进入我的专栏，我们一起一探究竟。



# 从0开始学大数据

智能时代你的大数据第一课

李智慧

同程艺龙交通首席架构师  
前 Intel 大数据架构师



新版升级：点击「 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

下一篇 预习 01 | 大数据技术发展史：大数据的前世今生

## 精选留言 (55)

写留言



newbee 置顶

2018-10-30

38

读作作者的《大型网站技术架构》，深为作者的见解和对后来者的引导所折服，没想到今天能开始追作者的专栏。对于程序员必须要懂数据和机器学习的观点十分赞同，特别是互联网公司，产品一旦做大了之后，必然要拥抱数据处理。有些产品甚至生来就基于数据及机器学习。



芒果

2018-10-29

12

期待老师的专栏！

展开

**熊出没**

2018-10-29

👍 8

买了，听大佬讲讲和大数据、机器学习、人工智能、云计算有关的内容，因为太火了，很多人都讲，但是听谁的？我选择听极客时间的~(￣▽￣)~

**linazi**

2018-11-05

👍 7

我是一个女程序猿 从15年开始接触大数据 一直到现在 但还是不能够很好的理解大数据的整体生态体系架构 别人问我 只能说出个大概 脑袋中没有一个完整的体系和学习思路 希望听了智慧老师的课 能够有所启发

展开 ▾

**F pure...**

2018-10-29

👍 6

我设想，如果可以把自己大脑中的一些东西价值理念输入到计算机里面，通过数据分析。找到这些信息之间的相互关系，那么就可以创造更好的自己。比如说，如果你会弹吉他 并且 你的数学学的 特别好。你可以通过数学知识创造出关于吉他的相关内容。再抽象点儿说，就是 你可以 找到 一种 两者之间更美的，你更喜欢的东西！这种东西，肯定是有价值的。

展开 ▾

**AI**

2018-10-29

👍 6

顶起

展开 ▾

**sunlight00...**

2018-10-29

👍 5

看了大型网站技术架构，很期待老师的专栏！

展开 ▾

**MJ**

2018-10-30

👍 4

请教老师，关于大数据时代，学习一门编程语言，是java还是python好呢？系统的话，

## Ubuntu还是centos呢？求老师指点



**斯盖丸**

2018-10-29

👍 4

老师，您也会像左老师那样在专栏里列个书单让我们按图索骥一步步“练级”吗？

作者回复: 大数据涉及的知识技能领域比较多，你需要选择一个具体领域，然后开始练级，这个专栏就是帮你全面了解大数据的知识体系，然后选择合适自己发展方向的领域。



**龙华强**

2018-10-29

👍 4

第一次买专栏，因为这是李智慧的。看了课程目录，确实干货满满，开栏第一章看了好几遍，有种理应如此的感觉。作为大数据开发工程师，不仅仅是为了写代码挣钱，更应有一种对大数据的清晰认识，后面的路也不会迷茫疲倦

展开 ∨



**小老鼠**

2018-10-29

👍 4

现在有种说法，AI又遇到了困境，无法向前，你如何看？

展开 ∨



**A洒脱的风**

2018-10-29

👍 3

比心

展开 ∨



**K**

2018-10-29

👍 3

支持智慧哥错不了

展开 ∨





杰之7  
2019-01-14



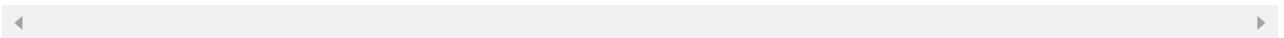
回过头来重新阅读学习，过去的软件逻辑是通过人的需求编写程序开发软件产品，使得人们生活的效率提高了多倍。

大数据技术和机器学习技术要做的事是挖掘人自身都可能不清楚喜欢的事和物，因为需求的再次通过数据挖掘可以发觉未知的东西，这样在原先的基础之上更有效率的提高了我...

展开 ∨

作者回复: 其实已经没有什么传统金融了，金融行业因为资金充裕，对大数据的投入和使用其实非常大的，不亚于互联网，未来的前景也非常看好。

PS: 我认为科技才是国运，金融是润滑油，科技才是发动机。



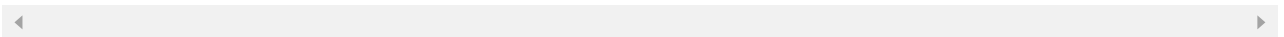
Mr.zhang

2018-11-22



看了作者的《大型网站技术架构》非常棒，一看作者开专栏了，就果断买了，希望跟着作者能够开拓眼界。

作者回复: 谢谢



林锋

2018-11-20



李老师就是我的架构启蒙老师，支持一下

展开 ∨



夏一Sunny

2018-11-09



选择大数据是为了帮助用户挖掘连用户自己都不知道的需求，这是在传统编程的基础上更进一步，创造更上一层的价值。很认同老师的选择，我认为这是一种进步，更是一种趋势。希望通过学习老师的课程，能让自己有一个更系统的思维去认识大数据，去学习大数据开发。



**落叶飞逝的...**

学习大数据最好的时间是十年前，其次就是现在！

展开 ▾

---

**My dream**

2018-10-31

老师，希望你在课程中讲一下具体的操作步骤哈，不要讲理论讲的太多了，实战现理论想结合，我们这些零基础的才学的会

**行云**

2018-10-30

小白用户，因为感兴趣定了专栏

展开 ▾