

## 【Café速递】杜堂武：互联网行业见闻及工作经验分享

**核心提示：**在互联网行业工作是什么样子？不同公司之间的工作氛围、职级薪资差距如何？数据科学从业者需要掌握哪些技能？在 SEA 与大陆的互联网公司有什么不同？我们要如何准备，才能收获到自己心仪的 Offer？让我们一起来看一看本期杜堂武师兄关于互联网行业见闻与工作经验的分享。

主持：王妍 文字：王浩成

### >>>人物名片

**杜堂武**，测绘遥感信息工程国家重点实验室 2012 届硕士，师从朱欣焰教授。主要研究兴趣是机器学习在互联网行业的应用。Singapore ETH Center 工作期间三维重建 NUS 校园；百度工作期间完成“See you again 加德满都”重建工作，获得艾菲全场大奖；从零到一搭建高德高精地图作业平台；构建 Shopee User Profile 系统与 AutoBanner 系统。

### >>>嘉宾小语

- ◇ 技术是基础，但只能决定你的下限。
- ◇ 永远 Stay hungry, stay foolish, 学会拥抱变化。
- ◇ 职场发展要有跳板思维，需要不断拓展自己的边界。

### >>>报告现场

11 月 14 日晚上 7 点，武汉大学校友、现就职于新加坡 Shopee 的杜堂武师兄做客 GeoScience Café 第 279 期。杜堂武师兄结合自己在大厂工作的见闻和经历，介绍了在中国以及东南亚互联网工作的所见所闻、总结市场上需要的技术和技能，为同学们带来了很多信息与启发。

在讲座之初，杜堂武师兄叮嘱道，他希望自己的经历和见闻能够给大家带来一定的借鉴意义，但同时也希望同学们能结合自己的认知和实际情况，在思考之后再做出属于自己的人生选择。

杜堂武师兄针对以下的五个问题为我们展开了分享：

- 第一部分：个人工作经历及职业发展脉络；
- 第二部分：与就业选择相关的思考；
- 第三部分：中国和东南亚的互联网市场中的见闻；
- 第四部分：互联网市场中对技术技能的要求；
- 第五部分：一些思考与总结；

## 个人工作经历及就业选择思考

在讲座伊始，杜师兄首先展示了自己的领英 Profile 页，为我们简要梳理了他的个人工作经历。通过师兄的分享，我们感受到了他在选择职业道路的过程中不断追寻社会价值的内在动力。

杜师兄的职业生涯始于 Singapore ETH Center，他主要负责利用无人机回传数据构建校园三维模型。在这段经历中，杜师兄通过解决密集匹配、精细建模等问题，感受到如何运用在本科和研究生阶段习得的知识重建三维模型，并运用在实际场景中解决问题。离开新加坡后，杜师兄加入了百度。在百度完成“See you again 加德满都”重建工作项目，获得了艾菲全场大奖。

但由于百度的项目偏向运营及展示方向，离实际运用、创造社会价值仍有一定的距离，杜师兄选择在 2017 年加入阿里巴巴高德事业部，开始进行与高精地图和无人驾驶相关的项目。

什么是高精地图？杜师兄介绍道，高精地图有别于传统地图之处在于它的道路的属性非常丰富，我们甚至可以看到一路面标牌和电线杆这样的面状和杆状结构。为了提升高精地图的精度，研发人员需要依次提取街景中的面状、线状、点状结构。由于地图精度很高，研发人员常常需要人工标记许多数据，比如严丝合缝地把路面标识标记出来，标记被遮挡或磨损的边缘线等等。

回顾完 Singapore ETH Center、百度、阿里的三段经历后，杜师兄为我们总结道，很多时候我们的所学的东西只是一块门砖；有了敲门砖，我们才能够进入一个行业，“其实说我就是把三维重建技术当作敲门砖，进入到了互联网行业。”随后，杜师兄进一步补充道，当我们进入到这个行业了，就可以去做一些触类旁通的事情，去探索更多的职业发展的可能性。

也正是因为希望接触更多线上的业务、看到自己创造的价值、探索职业发展的可能性，杜师兄离开了高德，加入了东南亚第一大本地化电商互联网企业——Shopee。

杜师兄为我们介绍道，Shopee 的团队结构包括 Image、Supply Chain、Translation、Knowledge Graph、Security 等，杜师兄所在的团队为 Marketing Science。

作为数据科学家，杜师兄主要支持两个业务：

第一部分是 User Profile 的数据建设。它分成两部分，第一个是 demographic，即用户的一些属性，包括用户的收入、性别、子女状况、种族；第二部分是对用户感兴趣的类目做标记，以预测用户的流失概率、偿还能力等。这两部分的数据主要的应用场景是推荐和搜索策略。

第二部分业务是 Auto Banner。Auto Banner 是一个由 Template、主题、商品三部分构成的促销条幅。为了制作这样的横幅，需要将商品主体分离出来，并

通过用户分群和商品分群的办法，将商品和用户进行匹配。

介绍完自己的经历，杜师兄简单介绍了同学们的毕业去向，拓宽了同学们就业的思路。

## 同学的就业情况

### 1. 互联网公司：

阿里巴巴(3)，百度(1)，美团(1)，滴滴出行(2)...

### 2. 大学研究所：

国内 -- 中山大学(3)，吉林大学，天津大学(1)，南京邮电(1)，中国公安大学，中科院

...

国外 -- 哈佛，密歇根，路易斯安那州立，俄亥俄州立

### 3. 创业：

极验验证，大势智慧科技，珈和科技

### 4. 事业单位

中石油，规划设计院，地理信息中心

### 5. 自由职业者

从事炒股或者经商等

图 2 其他校友们的就业去向

## 中国&东南亚的互联网行业概况

杜堂武师兄将互联网行业分成国内和国外两个部分，分别从工作体验、人才流动、薪酬体系等方面概述了互联网大厂的基本情况。

针对国内互联网，杜师兄介绍了百度和阿里的工作情况。

从工作体验来看，百度重视技术投入，鼓励大胆和创新；而阿里巴巴重视价值观，重视项目的落地回报，361 制度给员工带来了较大的竞争压力。

## 国内互联网之我见:百度阿里的的工作体验

### 百度:

1. 重视技术投入
2. 鼓励技术创新和大胆探索
3. 不一定以赚钱为导向
4. 技术人才流失严重

### 阿里:

1. 重视价值观, 由于价值观导致阿里会对收购来的公司进行改造, 这在一定程度上会带来文化的冲突和人事的动荡, 阿里当前最成功的收购可能是高德地图。
2. 看重业务落地;
3. 由于有众多BU, 而年底KPI评比差异较大, 导致年终奖及股票分配差异很大;同时由于361的设置, 组内成员之间的竞争压力也较大。

图3 百度与阿里的工作体验对比

从人才流失角度看, 百度的技术人才大部分流向了阿里和字节跳动; 阿里的技术人才则流向腾讯、字节、快手, 内部转岗的去向为蚂蚁金服和阿里云。

## 国内互联网之我见:人才流动

### 百度:

流向阿里较多, 其次是造车新势力(理想汽车), 出来创业的比较多。(仅限于我认识的群体)

### 阿里:

流向腾讯、字节跳动, 我所在的组有很大一批人去了快手, 内部转岗去蚂蚁金服, 阿里云的人较多。出来创业的人比较少, 喜欢吸收有过创业经验的人。(仅限于我认识的群体)

有部分群体, 因比特币而财富自由。



图4 国内互联网公司间的人才流动规律

这部分的最后, 杜师兄为我们分享了国内互联网大厂的薪水职级分布图片。同时, 杜师兄介绍道, 阿里的职级体系更受到市场认可。

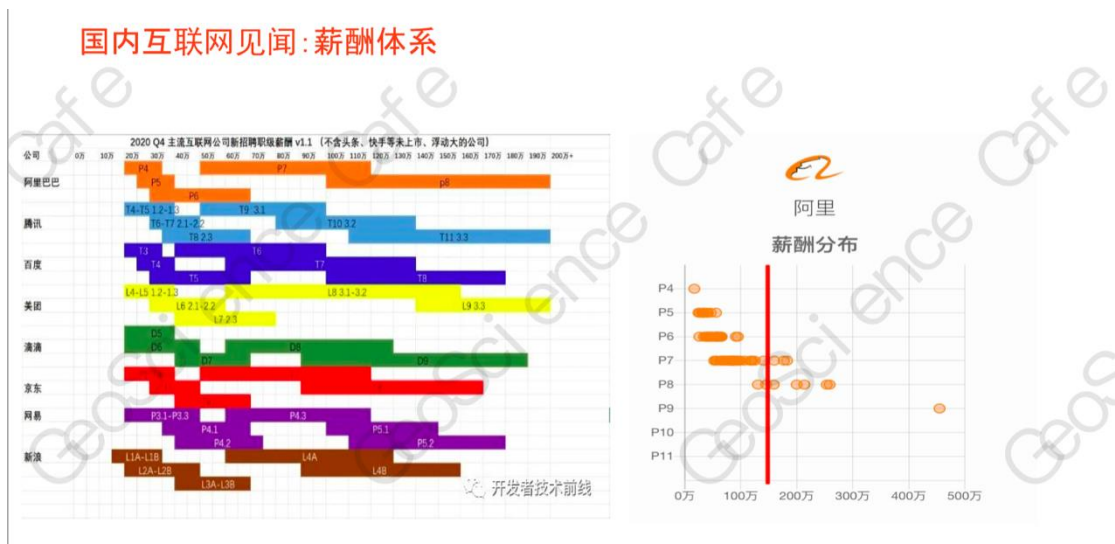


图 5 国内互联网大厂的职级与薪酬体系

谈及东南亚互联网科技公司时，杜师兄依据电商、出行、支付、互金、物流、游戏等类别，列举了在东南亚比较大型的互联网公司。

**全球大型互联网科技公司：**

阿里云，蚂蚁金服，腾讯（主要游戏相关），字节跳动（Tiktok, Lark），Apple, Google, Facebook, Microsoft, Twitter...

**电商类：**Shopee, Lazada, Tokopedia, bukalapak...

**出行类：**Grab, Gojek

**支付类：**基本上大型互联网公司和银行都有自己的支付，各种支付方式纷繁复杂

**互金类：**Advance AI

**物流类：**LalaMove, Ninja Van

**游戏类：**Garena, YOOZOO, Riot Games（英雄联盟）

**其他：**虎牙, bigo, 商汤, 依图

图 6 东南亚地区的大型互联网公司一览

本部分的最后，杜师兄将中国和新加坡的互联网工作进行了对比。中国是全球最大的单一化市场，但互联网市场较为饱和，而东南亚正成为互联网公司扩张的首要阵地。同时，中国和新加坡在工作理念和生活成本上都有着较大差异。在国内，师兄每晚 9 点多到家，而在新加坡，公司会更注重私人时间和合理的休息；在生活上，新加坡买房的压力没有那么大，生活的条件也会稍微好一点，但相对来说缺少一些房产增值的机会。

**就业市场所需技能**

针对就业问题，杜师兄全面地介绍了数据科学家所需的技能，强调了研发与产品的配合、技术的选择和思考、行业转变的趋势，并介绍了提升技术水平

的若干方式。

杜师兄介绍道，企业越来越发现纯数据科学家是是很难有产出的，为了解决这个问题，公司逐渐认识到需要有一些 data engineer 的加入，帮助企业进行一些数据搜集和管理的工作。譬如说 MySQL，或者 Hadoop 这样的来做数据上的一些管理，然后尤其这种 Data Warehouse 的一些事情，同时有一些 software engineer，他们会帮我们去 deliver 我们的产品到前端、后端。

随着数据科学团队的发展，行业发生了许多转变。目前的趋势是：许多企业引入了 PM 来驱动项目，PM 最主要的角色是跟外部去协调资源，然后判断内统筹项目能否实现、业务要求是否合理；

另外一个趋势是，企业把数据科学团队给打散，并入到各种业务线中，服务于业务的发展。

接下来，杜师兄展示了数据科学家的技术全栈图。在实际工作中，最主要使用的软件包括 Database、MySQL、Hadoop 和 Spark，同时要熟练的掌握 SQL 语言。

## 数据科学

1. 纯数据科学团队很难有产出。As a minimum complement, SRE and Data Engineers are needed.  
背后的原因是因为产品落地能力
2. 数据科学团队发展趋势：  
需要引入PM来驱动项目的进展，协调资源，同时当前数据科学团队更多的并入业务线，服务于业务的发展

算法与工程的关系：

Netflix Never Used Its \$1 Million Algorithm Due To Engineering Costs

<https://www.wired.com/2012/04/netflix-prize-costs/>

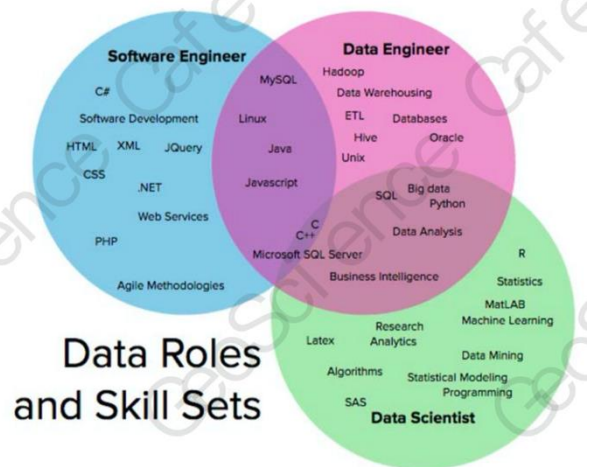


图 6 数据科学的不同角色与所需技能

第二部分需要用 Python 做一些 prototype 的工作，最后他们会用到一些深度学习的算法，也就是一些 framework，他们用的最多的是 PyTorch 来做一些人工智能的一些算法。然后第三部分是一个工程化的任务，需要最终能够去部署到线上。

同时，杜师兄非常鼓励同学们针对实际目标来补充自己的技能树：“假如说你未来想要做一些高精地图和无人驾驶，就需要去了解图像检测、图像识别、图像分割，以及一些典型的数据处理，并且还需要掌握一些图形学上的知识。”

为了增加就业技能，杜师兄给出了 3 条建议：

首先，通过刷题提升自己对基础的算法的理解。比如我们去做一些 Leetcode 的题。对于应届生而言，题目可以检验基本功，大家在刷题时应该尽量关注中等难度的题目；

其次，参加一些竞赛，比如 Kaggle、阿里巴巴的天池大数据竞赛；以 Kaggle 为例，比赛包含了数据科学的方方面面，有各种各样的题材和题目，在参加比赛时，可以学习到许多其他的经验。

关于求职，杜师兄给出了自己的建议：学习与推荐搜索相关的算法。毕竟推荐搜索广告是很多互联网公司最核心的业务。

### 总结与思考

最后，杜师兄为同学们介绍了自己的职场经验，与后辈们展望了未来的可能性。

第一，我们的基础是块敲门砖，但这只能决定我们的发展下限，也就是能否进入到一个互联网公司；

其次，stay foolish, stay hungry, 互联网公司需要员工触类旁通地去学更多的知识，掌握更多的技能。

第三，不管在读书还是在工作中，都要学会主动反馈。“当我们的老板或导师给我们布置了一个任务时，我们必须在完成一定阶段后主动向他们反馈。即使他们有时由于事务繁忙，暂时忘记了这项任务，但当他们想起来的时候，他们对我们的印象就会由于我们没有主动反馈而打折扣，甚至觉得我们缺乏主动性，进而影响自己的职业发展。”

第四，学会拥抱变化。一个公司的组织架构可能会每半年或者每一年就发生一个比较大的变化，因此同学们要能适应这样的变化。

最后，不断提升个人认知，这要求同学们能够从与别人的交流中、从网络资料中学习经验，从全局来考虑职业发展。

对于未来，杜师兄分享了两点自己的看法：一，未来是信息化的时代，云是一个基础设施，这也就是说我们现在小作坊式的 exe 的形式将不复存在，未来会是一个提供服务的模式；二，中国企业需要通过出海来扩大自己的影响力，因此也需要拥有国际化视野的人才，如果同学们想朝着国外的公司发展，建议去提升自己的语言水平，通过申请国际学校、参加国外招聘等方式，建立更广阔的视野。

### >>>互动交流

**提问人一：**您在工作之余如何提升自己的知识体系和技能？

**杜堂武：**善于利用网络给自己充电，比如 Google、Youtube 和 B 站最主要



的是针对要解决的问题，我们要去想该怎么做，然后通过搜索来解决问题，并且获得知识的积累。另外，在 Github 上面，我们也可以把别人做的一些东西拷贝过来，然后仔细地学习。

**提问人二：**高精地图这个领域目前很火，想知道师兄当初为什么决定转行？

**杜堂武：**高精地图现在虽然说是很火，但是距离应用还较远，这也是我刚刚一直强调的。我总觉得我们在高德的时候跟车厂联系非常少，没有看到车厂对我们高清地图使用的一些反馈，所以让我觉得离实际使用太远，这是我决定转行的原因。

另外一个原因就是刚刚提到的，在电商行业，触类旁通的机会比较多，在 Shopee，我也可以用图像技术进行图像的分割，跟高精地图做的技术没有什么太大的差别，而我们用同样的技术解决了其他方面的问题。

**提问人三：**第一个问题：您的职业前期从事的是图像的相关工作，后来您转到了数据业务，可能是用户画像，这样一个业务的转移在招聘过程中的核心竞争力是怎样体现的？第二个问题：新加坡的永久居留权如何申请？

**杜堂武：**虽然我的工作有很大一部分与用户画像相关，但也有 Auto Banner 的部分，这一部分就用到了图像分割技术。而且我的工作范围内也有许多与图像相关的部分。因此我们部门需要我去做这些事情，我自己也很感兴趣，所以很顺畅地就转来了。其实在招聘时，公司想让我做图像相关，但进来之后，电商公司鼓励员工做新的尝试，所以我们在公司内部可以比较自由地选择转去其他业务。

新加坡的永久居留权我也还在申请，第一次被拒之后，这一次的申請还在判定之中，暂时没有成功的经验可以分享。PR 申请像一个黑盒子，一些单身同事或是年纪稍大的朋友，他们第一次申请就过了。所以，我还没能总结出一些肯定能成功的条件供你参考。

**提问人四：**GIS 专业在互联网地图部门是不是没有计算机科班出身有优势？

**杜堂武：**从就业来说，我觉得应该是计算机科班出身有优势，但是他们不一定选数据科学这个方向，但他们可能去做一些其他事情，比如说前端后端。因为现在所有算法都已经被打包了，可能我们去调个包就可以解决很大一部分问题。而计算机科班出身的人可以直接去调这个包，实现从 0~1 的这样一个过程。同时，他知道如何去把它去部署上线，如何把它去规模化，这就是他们的优势。这也就是我刚刚强调的。做算法的人一定要具备工程上的一些能力，去学习一些工程上的经验，这样你才可能有更大的竞争优势。

**提问人五：**如何获得国外公司的工作机会，他们更注重的是什么方面？



**杜堂武：**我没有参加过校招，所以说我不是很清楚，我觉得国外公司工作机会肯定与国内的互联网大厂是一样的，而且在新加坡这边，需要很多技术人员，他们中的很多都来自国内，即使英语能力只能进行基本沟通也问题不大，最重要的还是你的基础能力，就像前面说的，建议大家去 LeetCode 刷题、参加 Kaggle 之类的比赛，面试时可能会问参加 Kaggle 的一些经验，包括我刚刚提到的如何对缺失数据进行处理等。

GeoScience Café 以“谈笑间成就梦想”为目标，于每周五晚 7:00 在实验室四楼休闲厅，邀请 1-4 位嘉宾，为大家带来学术报告或经验分享。报告内容包括摄影测量与遥感、地理信息系统、导航与定位服务等研究方向，听众可在报告结束后向嘉宾提问、与嘉宾交流探讨，同时每学期还会举办 2 期人文类讲座和 2 场导师信息分享会。每期报告会根据嘉宾意愿在 B 站开设直播，使不能来到现场的听众同步参与。报告 PPT 和视频会在征得嘉宾同意的情况下在 qq 群和 B 站上发布。

更多精彩内容（讲座预告、讲座回顾、报告 PPT、报告视频）敬请通过以下方式获取：



QQ群



微信公众号



B站直播