# 上海第七次Spark meetup会议资料分享

上海Spark meetup第七次聚会将于2016年1月23日(周六)在上海市长宁区金钟路968号凌空SOHO 8号楼进行。此次聚会由Intel联合携程举办。

## 大会主题

1、开场/Opening Keynote: 张翼，携程大数据平台的负责人
个人介绍：本科和研究生都是浙江大学；2015年加入携程，推动携程大数据平台的演进；对大数据底层框架Hadoop，HIVE，Spark等技术以及数据仓库建设相关的技术略有研究。

2、演讲者1: 李健，携程攻略社区，开发总监
个人介绍：毕业于上海大学本科，交通大学计算机系研究生学历。自2013年9月进入携程网以来，先后任攻略社区后台开发主管、数据组主管等职，在自然语言处理、图像识别和处理方面有所研究。
分享主题：Spark 自然语言处理在携程攻略的实践
概要： NLP技术简介，情感分析，特征标签抽取，特征词抽取，游记自动文摘，知识图谱关系抽取，目前研究方向。

3、演讲者2: 陶刚, Splunk上海架构师
个人介绍：2000年毕业于北京邮电大学，目前在Splunk上海担任架构师,负责数据采集和云平台产品的技术架构。 在加入Splunk之前，曾就职于SAP，EMC等企业，负责ERP，BI和存储管理产品的开发。陶刚拥有丰富的企业级产品的开发经验，对数据科学，数据可视化和机器学习等领域有着浓厚的兴趣。
分享主题：Splunk & Spark
概要： Splunk是业内领先的机器数据的平台，Splunk公司也是业内第一家在纳斯达克上市的大数据公司。本次演讲介绍了Splunk的基本功能和技术架构，并就如何在Splunk平台上利用和集成Spark进行数据处理展开讨论。 http://www.splunk.com/zh-hans_cn

4、演讲者3: Lex Yu, 英特尔高级软件工程师
个人介绍：Lex Yu, 英特尔高级软件工程师，Spark contributor，致力于大数据技术的性能优化及开发。
分享主题：Drive Best Spark Performance on IA
概要：调优是Spark开发中非常重要的一个环节，好的调优可以提升10倍以上的性能。这里，我们分享了Intel长期调优经验的总结，并展示了一种使用最新的PCIE SSD技术来加速程序的方案。最后，我们通过IO Profile揭示出SSD能够加速的根本原因。

## 资料下载

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（过往记忆）所有，未经许可不得转载。
本文链接: 【】（）