

今日头条的评论都在说什么？
话多又有趣的人都在这里。

就你话多

——今日头条用户评论数据报告（上）



今日头条算数中心

01

PART ONE

请开始表演吧



这里，是我们捕获的话痨

——什么样的人最爱评论？

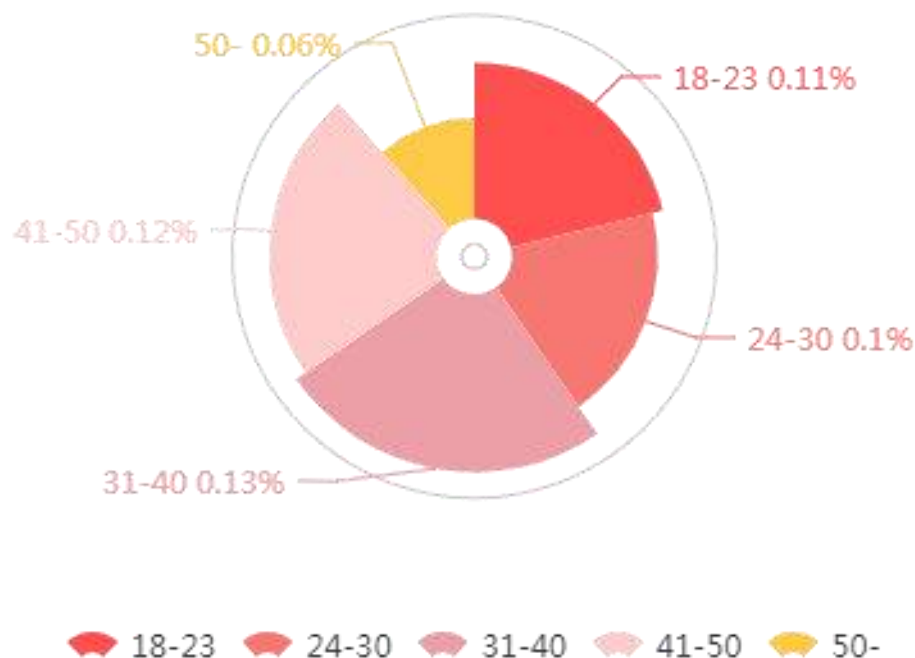
我们以“评论率”衡量话痨，
评论率即评论数/阅读数。

什么样的人评论率最高？

在过去一年中（2016.7.15-2017.7.15）

爱说话的人原来都长这个样子。

$$\text{♂} : \text{♀} = 1.7 : 1$$



数据告诉我们，

男人 才是话更多的那个。

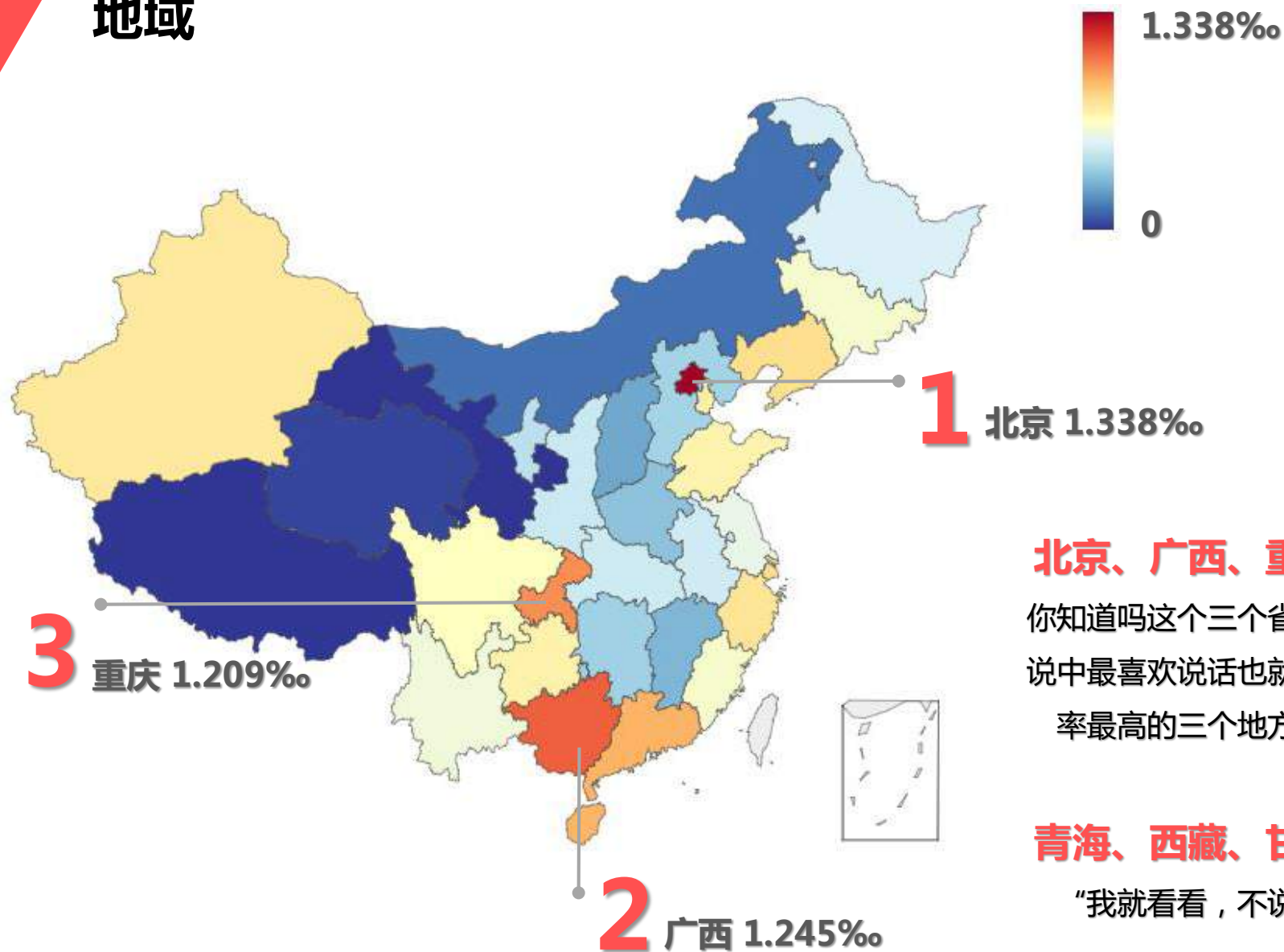
31-50岁 的青壮年，
比年轻人更有表达欲。

数据来源：头条指数

注：本章展示了某个年龄段/性别的用户的评论率。

© 2017.07 今日头条算数中心

地域



北京、广西、重庆，
你知道吗这三个省就是传说中最喜欢说话也就是评论率最高的三个地方呢。

青海、西藏、甘肃，
“我就看看，不说话”

数据来源：头条指数

注：本章展示的是某个省份的用户评论率，经统计数值分布在0-1.338%之间。

© 2017.07 今日头条算数中心

从 **超一线城市** 到五线城市，
评论率呈现下降的趋势。

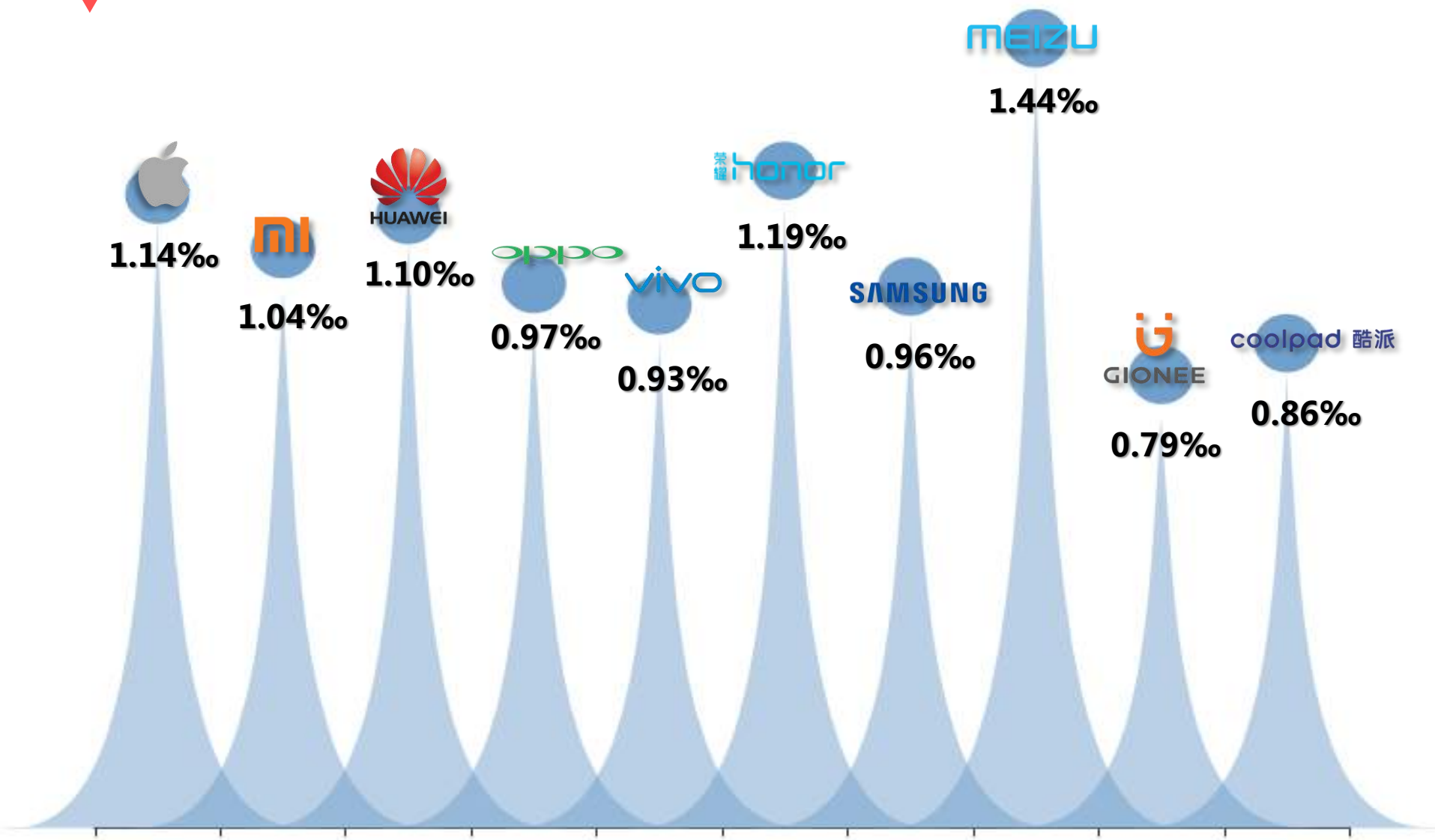


数据来源：头条指数

注：本章展示了不同级别城市的用户的评论率变化状况。

© 2017.07 今日头条算数中心

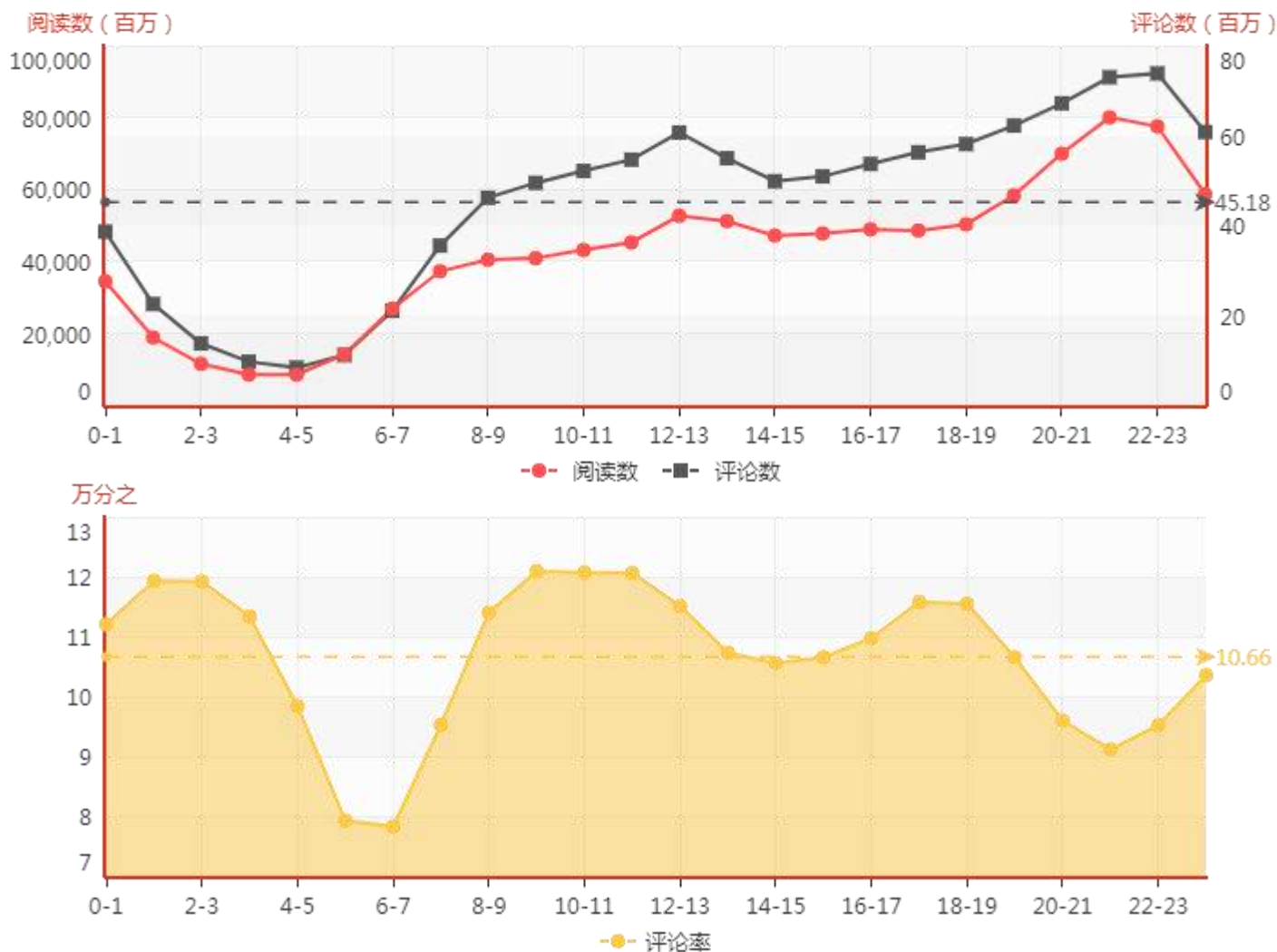
手机机型



数据来源：头条指数

注：本章展示了使用某个机型的用户的评论率，本图仅展示使用人数前十的机型。

© 2017.07 今日头条算数中心



数据来源：头条指数

注：本章展示了一天24小时内用户总体阅读数、评论数和评论率变化状况。

© 2017.07 今日头条算数中心

02

PART TWO

有自己的一套原则
我行我素



话多的人一向爱憎分明

——爱评论的人在表达什么？

我们挑选出过去一年中的三组热点事件，
通过对600条超热门评论进行词频分析，
向大家展示我们的用户在评论中表达了什么样的观点
以及对待不同的人 and 事的不同态度。

你们真的很酷哦 (*'v`)~♥



一个英雄开黑大乱斗，一个氪金跪求SSR。

哪里有评论区，哪里就是**技术讨论区**，

当然，**未成年和防沉迷**，也是每个研发者都需要思考的问题。



注：词云图展示了与某个关键词相关的前100条最热门评论的词频分布，图中某个字词所占面积越大，则该字词出现得频率越高。下同。

在过去一年中，
两位因为出轨事件处在舆论风口浪尖的女星，
与两者相关的新闻评论中，
大部分是对出轨者和第三者的**批判**、以及对原配的**心疼**。

白百何 VS 马蓉



对白百何事件来说，

大家会将它与众多明星夫妻出轨的**八卦**联系在一起，

而对于马蓉事件，由于牵扯到财产欺诈，

因而大家会有更多的**愤怒和谴责**。





是评论里提及最频繁的话题。

司机用户似乎是表达观点最鲜明的那群人。

其中还有司机用户大笔一挥地算了一笔账 → 往后翻



三太几 10月前

专职网约车能赚钱不？能赚多少钱？你算过吗？没有！好！现在我给你算一算！先算收入 60万公里报废，这60万公里中你能挣钱的最多50万公里，因为你去接客人是不记费的。50万乘上每公里1.8元，你收入90万，50万公里收费里程按照你慢点开，2分钟一公里，你跑100万分钟，每分钟收费0.35元，你收入35万，合计总收入为125万元。下面算支出部分 125万，网络平台抽成最少23%，平台抽走28.75万，按照每公里0.4元油钱计算，油费共计24万，5000公里保养一次，共120次，每次按照250计算，要花费3万元，购买车辆按照10万计算，保险按照8年，每年4000，要花费3.2万，轮胎刹车片等维修费用8年最少4万，罚款一定避免不了，每年按1000多计算，8年，按1万计算，到此已经73.95万，8年的审车和交管部门要求的手续费用以及要求安装的定位系统行车记录系统等按照1万多点，合计总支出就75万了，你8年挣了多少钱？ $125-75=50$ 万 按照8年，你每年收入6.25万 每个月5000多点 然后我们算一下我们的工作量 60万公里，8年，每年12个月，每个月工作25天，每天行程就是250公里。每天要挣多少钱才行啊？125万除8年，除12个月，再除每月出车25天，平均每天必须收入520元。你知道收入520元你要工作多长时间吗？写到这里我只能对平台说：呵呵！

回复 · 481条回复

2764

“共享单车是国人素质的照妖镜。”

这是围绕着共享单车，
大家在评论中发出最多的感慨。

共享单车为日常出行提供了**便利**，
也引发了一场素质大讨论。

同时也存在着对**哪家更好、未成年安全隐患**
等问题的关注。





老杨32263891 2月前

我是共享单车公司的一员、现在我每天的工作是开着皮卡带着队员去清理掉大河、挂树上、丢垃圾房、被喷漆的单车、还有四分五裂的单车，我们不图赚钱只为方便大家。心里很嘈杂！我的城市在创文明城市、我只想说我们都是这个城市的一份子，希望大家都爱护呵护我们的家园。谢谢你们了

4199

回复 · 298条回复



auv8630 2月前

小黄车最老实，被人骑，被人打，还被没素质的人不给钱，真怕有一天小黄车不在世了。只剩下怀念了。

1631

回复 · 67条回复

03

PART THREE

话多的人一向很智慧

——爱评论的人都知道些什么？



我们发现不仅热门事件的评论很热闹
许多冷门内容的评论也藏龙卧虎。
来看看这些领域里的“扫地僧们”。

听说你想聊聊影视剧《人民的名义》？



别人家的男喷友 2月前

赵德汉——贪，丁义珍——奸，育良书记——阴，祁同伟——坏，程度——横，高小琴——精，欧阳菁——辣，沙瑞金——威，侯亮平——睿，赵东来——实，陈岩石——忠，季昌明——庸，陆亦可——真，李达康——难，陈海——躺，郑胜利——烦。

回复

2820



网络首席执行官 3月前

英语达人陈清泉，心怀宇宙孙连城，婚姻楷模高玉良，一分文没动赵得汉，扫地神僧汤姆丁，世界臣服祁同伟，撩妹高手赵东来，爱送饺子吴大妈，我要换台郑胜利，一脸嫌弃陆亦可，朝阳群众陈岩石，三味真火王文革，一动不动是陈海，我是好人蔡成功，育子无方郑西坡，无限背锅李达康，京剧名家高小琴，黑白通吃赵公子，大闹天宫侯亮平，笑看风云沙瑞金。

回复

5941

评论惊现“补同伟”本人！！↓↓↓



补同伟 2月前

非常感谢高总当初看得起我 找我演李亚军这个角色 不然我补同伟不会有今天啊

回复 · 184条回复 ▾

3226

听说你种的韭菜活不了？



从心开始151882368 3月前

错了，韭菜在第一次栽苗之后，2到3年需要刨了，掰开根把腐烂的根去掉，根须子剪短，每簇6，7棵根，间距6寸左右，重新栽。这样才能吃到又肥又大的韭菜。你说的吃一辈子是不可能的，那样韭菜会越来越小，最后变成韭菜毛。割韭菜不能太深，每过2，3年刨一次。7年没刨，证明你家老人太懒。不喜勿喷，实话实说。

回复 · 收起回复 ^

270



我是大灰狼你是小白兔吗 3月前

说的不错！

回复

4



润物细无声143664374 3月前

不是懒而是不懂耕作

回复

3



从心开始151882368 3月前

看图片应该在农村，农村人是不可能不懂的怎样栽韭菜，我从小在农村长大，18岁到城里，18岁之前没种过菜，只是有些印象罢了。现在在市里有个小院子，平时吃的辣椒，茄子什么都会种了。

回复

4

请问这道JAVA题怎么解？



头条里有你有我还有他 10月前

我是这么理解的：栈和常量池中的东西，都是值传递，放在堆里东西都是引用传递，然而堆里的东西的引用都在栈里，故而没有引用传递一说了。所以java里传递的都是值传递，没有引用传递。也不知道这么理解有没有错，大家指点一下。

回复 · 收起回复 ^



用户26940067 10月前

并不是都是值传递。比如说你传一个list到下一层的方法，改变了list里面的值，上一层也是变的。

回复



头条里有你有我还有他 10月前

这时候传递的其实是这个list的引用，也就是它的引用的值，这个值放在栈里面的。值传递，我说的是这个意思。因为没有方法会把这个list对象本身传递进去，比如传个可序列化的对象（以数据的形式表示）给某个方法，这才是传递那个对象本身嘛。

回复



lessismore32017636 10月前

传地址和传值，还是有引用传递的，Person b = new Person, Person a 等于b，这不是传的地址么，地址也就是引用。。ab指向了同一个堆里的Person

回复



头条里有你有我还有他 10月前

嗯嗯，是这意思，只不过我理解的是：a=b，其实是拷贝了b的引用（这个引用就是个值），我是这意思。

回复



胶片和数码相机你更喜欢哪个？



没睡醒的懒猫 3月前

区别在这里:你拿个数码相机再带个增强电池手柄,只想着突突突扫片,反正最后再差也能从一百张照片里选出来几张好看的;你拿个胶片相机,一卷胶卷就三四十张,每一次按下快门之前,都会很用心构图,很用心调光圈快门。

回复 · 收起回复

87



弓老爷 3月前

弓老爷 3月前

80年代中期,我用的理光9,拍摄了天津三年煤气化工程,数千张珍贵资料片,保持在原公用局,不知失落于何处?后来自己买了东方S3,一直精心保存着。如果胶卷价钱合理,我还会用胶卷拍摄。那是与数码相机完全不同的体验。每张照片都需要精心设置,光圈,快门,构图等。很有意思。

回复 · 10条回复

33



清风亭外 3月前

我手里也有一堆胶片机,看着它们,就想起过去的日子:斤斤计较着每张照片的所有数据,哪怕是有一点儿的不足也不愿意按下快门!所以,那时照片的成功率几乎是满卷!

回复

3



清风亭外 3月前

我也是这么做的。在装胶卷之前,先把相机的快门拨开,胶卷只从胶卷盒里拉出来短短的一段,够卡到转轴上用的就行。然后钻到被子里,再把胶卷拉开卡到拨轴端,盖好后盖。选好拍摄对象后,只要一按快门就可以拍一张,这样就能增加一、两张照片。一股可以拍三十八张以上。

回复

4

哦，原来你还看得懂书法啊？



千年兰亭 20天前

草书发展可分早期草书，章草和今草。今草始于汉末，张芝是。代表书法家，有终年帖，冠军帖等传世。诸葛亮这件作品属今草，故仅从书体判断不能断定其为伪作。当然，这件作品本身确实存疑

2 1

回复



暗淡了岁月 21天前

这字体应该不是三国时代的

1 1

回复



春天温暖在你心 21天前

书法界的老手

0 1

回复



老游59 20天前

楷书和行书最早见于晋代的钟繇，卫夫人和敦煌石窟里晋唐的经书！汉朝主要是隶书（汉隶）！该碑帖存疑，不敢下结论，

回复 · 收起回复 ^



千年兰亭 20天前

草书发展可分早期草书，章草和今草。今草始于汉末，张芝是。代表书法家，有终年帖，冠军帖等传世。诸葛亮这件作品属今草，故仅从书体判断不能断定其为伪作。当然，这件作品本身确实存疑

2 1

回复

1 1

什么！你连AI都懂？！

头条

渗蓝12小时前

人类会降低消耗，放弃追求更高的目标，而机器完全没有这方面的限制，机器一定会比人类强大很多倍。如果机器的智能释放出来，产生的科学成果超出人类的理解范畴。爱因斯坦的理论经过一百年才逐渐被证实，如果机器产生的理论超过人类认知的一万年一百万年？人类根本不知道怎么驾驭机器智能的

回复

1 1

原生心理11小时前

不能，设计的反击，和受到伤害的身体与精神的双重创伤，完全不同的概念。所以人类才会出现反抗的过激行为，例如有意识的防卫过度

回复

0 1

再线12小时前

假设智能机器人受到伤害（
输入的开始吗？

回复 · 1条回复



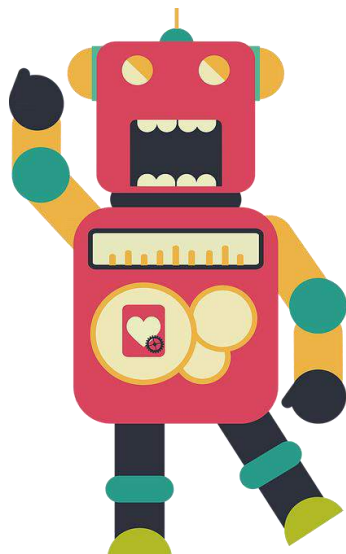
机器人智造师周雨浩 5月前

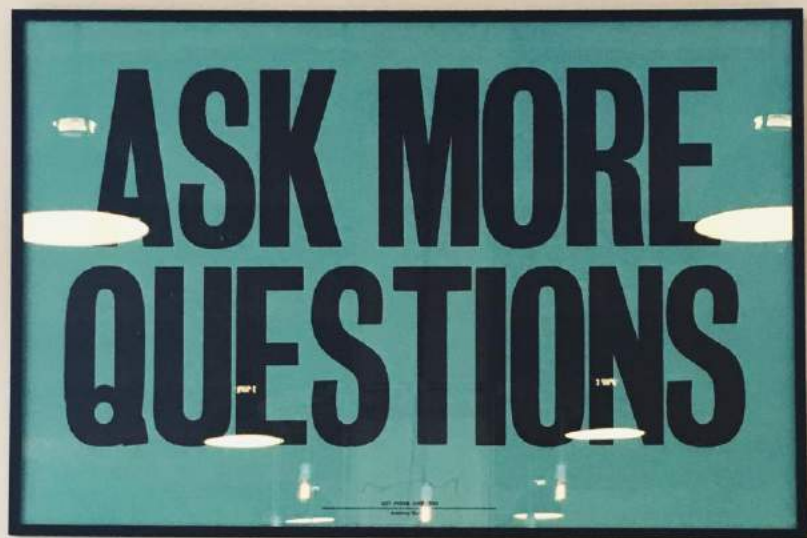
科大讯飞的语音语义智能系统在国内算是比较领先的，但是在实际环境应用中还是有一定的问题。第一：实验室环境得到的测试数据和实际环境中测试的数据相差太大啦，在实际商业环境中，同时5个人说话，距离多远，科大讯飞的语音系统几乎无法识别，即使是5麦克风环绕阵列都没有用，不要拿实验室数据做判断数据。第二：和现在语音语义在全球都是难以克服的技术难题，最多只能理解到第三层(不信大家可以试一下苹果的siri和微软的小冰)，汉语就更加无法语义解析了，我说白话，你说四川话，他说闽南语，机器人就傻了！语音语义解析这条路太漫长了，我们的机器人就盼着这一天早点到来。希望科大讯飞继续努力啊！

回复 · 2条回复

0 1

11 1





就你话多

——今日头条用户评论数据报告（上）



今日头条算数中心



重要声明

Declaration

本文《就你话多——今日头条用户评论数据报告（下）》（以下简称“本报告”）由今日头条算数中心（以下简称“算数中心”）制作。报告中文字、数据等受中国知识产权法律法规保护。除注明的引用第三方数据及公开信息，本报告所有权归今日头条算数中心所有。

算数中心致力于更及时、更全面、更深度的阅读数据挖掘和数据应用创新。未经允许，不得对本报告进行加工或改造。如有转载或引用，需及时与我们联系并注明出处“**今日头条算数中心**”，同时不得删减或改写报告内容。

特此声明。



关于我们 About Us

今日头条算数中心依靠今日头条海量用户每天数以百亿次级的阅读行为产生的数据，测量每一则资讯的传播力，描绘每一个用户更精准的画像，洞悉用户阅读行为下更深层的动机，致力于更广更深的阅读数据挖掘，用数据服务各个行业。

下载更多原版报告，请访问**今日头条媒体实验室** (mlab.toutiao.com) 报告下载区。

联系方式

电邮：datadance@bytedance.com

地址：北京海淀区知春路63号中国卫星通信大厦F17

研究人员：刘志毅&吴珏&裴俊良



报告下载区
二维码



公众号二维码