4.1 人工智能发展的现状

核心问题：典型案例、问题剖析、解决之道

首先通过典型案例让大家直观感受到当前人工智能发展的状态；其次，针对当前的典型案例所表现出来的问题进行深度剖析，最后根据这些项目问题提出相应的解决方法和路线。

1.章节目标 了解人工智能当前发展的方向，了解人工智能产品的现状，了解人工智能发展的制约因素，熟悉人工智能发展的驱动模式。

通过这四部分，第一让大家了解人工智能当前在哪些方向上有所发展，第二当前人工智能产品是什么样的状态，第三分析问题产生的原因，到底是什么制约了人工智能的发展，最后，分享一下人工智能驱动模式的一些研究。

1.1人工智能发展的现状



如果从业务视角来看人工智能，人工智能的发展可以从三个方面或者三个维度来衡量，第一个维度是交互能力，也是人工智能当前发展的一个主要方向。这个方向基本上聚焦了国内外70%-80%企业的精英。第二个是思考能力，第三个是服务能力。这也是当前人工智能发展从始到终的一个过程，当前人工智能技术研究主要停留在交互能力部分上，至于思考能力和服务能力目前来讲，零零散散有一些产品出现，但是并没有取得实质性的进展，也就是说只有在交互能力，有两个技能，一个是语音识别，就是把人的声音信号翻译成文字。另外一种技能是人脸识别，也就是图像识别。图像识别的能力成熟度远高于语音识别的能力成熟度，根源在于，图像识别这件事情完全是交互能力的一个技能，但是语音识别不是，它还涉及到一个思考能力的问题，就相当于我们把声音信号翻译成文本信号，这部分文本和信号所代表的的内容和含义是什么，它背后紧跟着的是认知能力，也就是我们说的思考能力。这三种能力的划分也就是当前人工智能研究的主要方向。这三点里面，每一点的驱动因素是不一样的。交互能力主要的驱动因素在于技术，这里包含了语音识别算法模型的成熟度，包括图像识别算法模型的成熟度，我们知道在传统的语音模型的识别方面，采用的算法模型一直制约了语音识别能力的发展，直到最近几年，当深度学习和神经网络被引入到语音识别领域来之后，语音识别的准确率才突破人的认知的上限，就是95%，也就是说目前为止，对于语音识别的准确率才刚刚超越人的识别能力，即95%，也就是市面上搞人工智能，搞语音识别的公司无论是谁，它所敢宣称的语音识别的准确率仅仅在95%上下。这是google，包括科大讯飞，都是以此来作为自己能力的一个衡量指标。第二个思考能力。对于思考能力，它发展的核心制约因素是什么呢？其实是数据。我们知道，思考即认知。认知其实是，前面你说了一句话，这句话是声音信号，我翻译成文字信号以后，这部分文字信号所代表的的知识和内容到底是什么，其实它是依赖于数据的，或者更进一步讲是数据里的知识部分。怎么讲呢？当你告诉你的智能交互的机器人说，“你帮我检查一下身体，看我有没有感冒”。这时，语音识别可以把这句话翻译成文本，然后再这里面它的思考能力在于，第一能不能够把感冒识别出来，把感冒和疾病关联起来，这种疾病在西医里面是什么，在中医里面是什么，它的临床表现是什么，它的诊断依据是什么，这个时候才叫具备了思考能力，否则不然的话，它仅仅是一种交互能力。其实，有了思考能力并不是终点，听得清、听得懂，完了以后我要对外去输出。因为，我们不需要仅仅听我们说话或是复述我们说话的机器人，我们需要的是能够帮我们解决问题的，这背后就是服务能力，服务能力的建设依赖于具体的业务场景。所以，这三者可以理解成为人工智能发展的主要方向，发展的主要驱动因素，还有一点就是商业变现。为什么这么讲呢？人工智能发展前期，无论是研究交互能力、思考能力还是服务能力，最终商业的本质是为了生意，一定是要变现的。后面的变现模式并没有真真正正成熟起来，没有成熟的模式，所以在视频中暂时没有写出。

1.2典型案例一

这是关于人工智能语音识别的一个典型例子，代表了当前语音识别发展的瓶颈和现状。

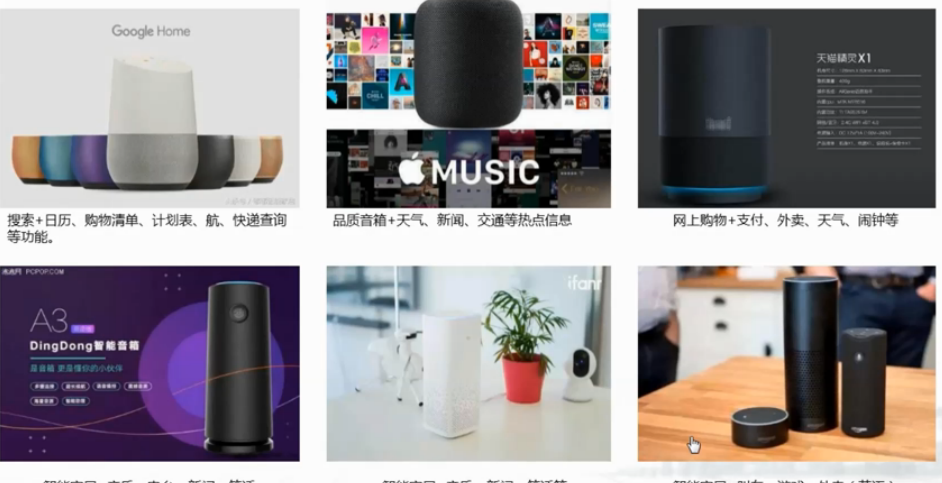
视频



从小视频里面可以总结暴露出来的问题，就是在语音识别里面所面临的几个困境。第一个困境是，同义词的识别。“联系”和“连续”分不清，“扫货”和“骚货”分不清，很尴尬。对语境的理解，仅仅是情绪的宣泄，结果连接到自己老板的信息。在这个小case中，在轻松愉悦中，可以清晰感受到人工智能当前在语音识别这一块存在的问题。

1.3典型案例二

这是当前市面上最火的一套东西的评估，就是智能音箱。我们知道，在人工智能发展的里程碑上，智能音箱可以理解为第一站，或者说是前站。智能音箱都有哪些呢？智能音箱目前为止，对外都提供了什么思考能力和服务能力呢？智能音箱本身代表的是交互能力，交互能力背后它有什么？首先，google home，google大家都知道，是搜索。Google home提供的



核心也是基于搜索，然后加了一些辅助功能，比如日历、购物清单、计划表，包括航空信息、快递信息等内容。另一个是苹果的音箱，对于苹果音箱来讲，反而走了另外一个捷径，因为我们知道人工智能音箱，无论你卖199、99还是1699、2999，其实它都是一个摆设，至少它现在是一个摆设。它不是一个功能性的产品，不是刚需。这也是为什么很多的企业把这个AI音箱定价在99、299的原因，但是苹果反倒走了一个捷径，它是怎么做的呢？它首先是个音箱，它把音质做得非常好，它定位成品质音箱再加上常用功能，这种常用功能是天气、新闻、交通等等。再有就是天猫音箱。天猫音箱，其实也发布有一段时间了，之前价位在双十一期间也出现过一轮价格调整，叫做99元来买天猫音箱，不排除为清货所做的处理。天猫音箱的定位就在于网络购物。围绕网络购物来做，与此相关的包括支付、天气、外卖、闹钟等其他的一些功能。涉及到网上购物的复杂性，可以理解成，音箱所设想的业务场景基本上没有实现。因为，我们可以看一组数据，具体多少人通过天猫音箱进行了网络购物，这个数字可以说是惨不忍睹。接下来，我们看一下叮咚。叮咚音箱是京东出的一款音箱，它能做什么呢？它是围绕智能家居、音乐、电台、新闻效果，在这里面你会发现一个问题，做智能家居，京东哪有什么智能家居的产品真正来占领市场呢？有一些东西市场份额蛮低的，你可以理解成为基本上就是一个摆设。再一个就是小米的智能音箱。其实，从个人感觉来讲，还是蛮喜欢小米的这种模式的，而且在中国的这种人工智能发展的趋势下，唯有小米是真正的走在前列的这样一家公司。它有智能家居的应用场景和硬件，所以，它这款音箱定价299，它定位的这个点就是首先把它应用到智能家居这一块，然后再加上辅助的一些其他功能。另外还有一家就是我们最熟悉的亚马逊。这是智能音箱的鼻祖，亚马逊智能音箱的定位在智能家居，再加上一些辅助的功能，叫车、外卖、游戏等，它局限在于，老外做的一套东西，局限在英语，在国内市场上并没有售卖。我们盘点一下，前面我们讲的这六款音箱，首先有没有主打的应用场景，这里面google、苹果、亚马逊、小米是有的，京东和天猫，天猫还好一点，至少支付是它擅长的领域。但京东去主打智能家居，有卡位的诉求，但竞争力是最弱的，因为你没有核心的竞争力。另外，其他的一些功能都是辅助类的，可有可无，有他不多，没他不少。通过这些，我们可以看出，目前而言，人工智能这些产品在服务这一块还不是有太多创新性的东西。

接下来，来看另外一个小例子。

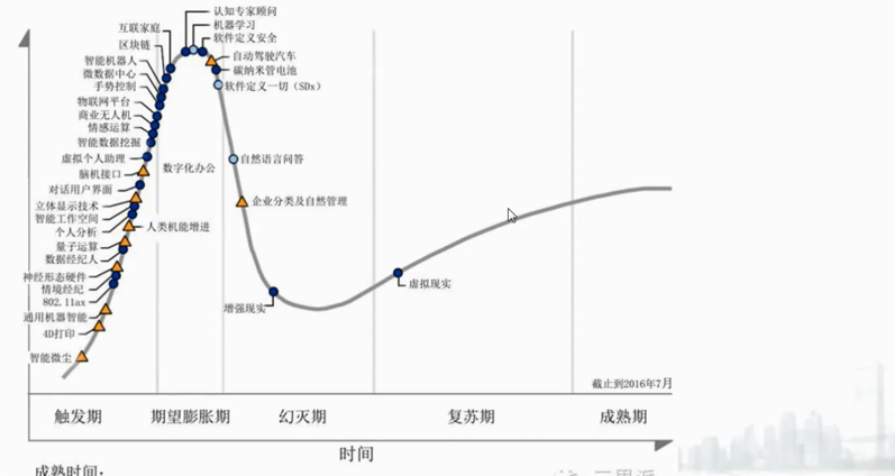
刚刚那段小视频，JARVIS为我们展示了它所预期的智能家居的场景。在这个里面，我们可以看到语音对话、智能播报、智能提醒、智能家居的控制，还有其他的一些种种能力。在这样一个水平，可以理解成为智能家居的未来发展方向。当然它也存在很多的问题。

1.4 归纳和总结：[www.gartner.com](http://www.gartner.com)

这张图是在2016年gartner（互联网信息研究公司）发布的技术成熟度的预测模型。在这个模型里面，它把一项技术成熟分为五个阶段：触发期、期望膨胀期、幻灭期、复苏期和成熟期。另外一个，又针对不同的技术当前它所在的状态做了不同元素的标注，在这个里面重点关注两个东西。第一，智能机器人部分，它处于期望膨胀期，基本处于触发期的顶峰；第二个自然语音回答，居然处于幻灭期，很诧异。在里面，我们仅是举了两个简单例子。可以在这个网站上获取自己所关注的技术。

思考：人工智能时代真的会到来吗，还需要等待多久才能到来？

前面，我们看到了智能语音存在的问题，也看到了智能家居前景展望，这个时候，我们是否可以坚信人工智能就会真正到来，另外一个问题就是如果真的会到来，还需要等待多久。



4.2 人工智能发展的阶段

核心问题：划分阶段、发展阶段、典型特征

章节目标：

了解人工智能发展的划分标准、

了解基于核心驱动力的阶段划分、

了解基于主要技术的阶段划分、

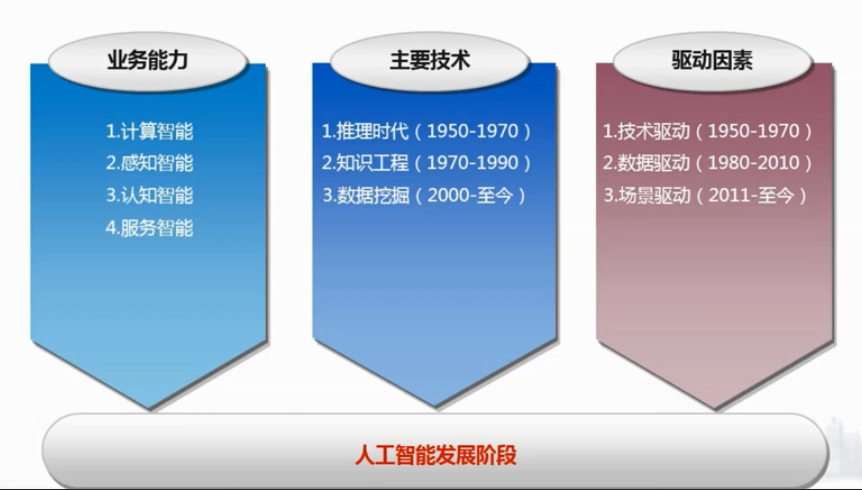
了解基于业务能力的阶段划分。

4.2.1 人工智能发展的三次潮起和两次低谷

1955年，达特茅斯会议标志AI诞生；1957年，罗森布拉特发明第一款神经网络，进入第一个高峰；1970年，受计算能力所限，AI进入第一个低谷；1982年，霍普菲尔特神经网络提出；1986年，BP算法实现了神经网络训练的突破，进行第二个高峰；1990年，人工智能计算机DAPPA失败，进入第二个低谷；2006年，深度学习神经网络被提出；2013年，深度学习算法在语音和视觉识别率获得突破性进展，进入第三个高峰。

重点关注几个时间点：第一个是1957年，第一款神经网络被发明，这就进入人工智能的第一个发展高峰；1970年，受计算能力所限，AI进入第一个低谷，在这里面我们回顾在第一部分讲到的人工智能发展简史讲到的，刻意把分布式计算技术的提出作为一个重要的时间节点放到人工智能的发展历程里面，就是基于这样的原因，因为人工智能是计算能力、数据能力和技术能力三者有机的融合。第二个高峰期是在BP神经网络训练的突破，现在讲的BP神经网络相对深度神经网络更靠前一些，它属于机器学习的范畴。1990年，人工智能计算机DAPPA失败，进入第二个低谷；关于人工智能计算机的发展，大家可以去详细了解相关的体系。再一个时间点是，2006年，深度学习神经网络模型被提出，2013年，人工智能里面的两个关键技术智能语音和视觉识别获得突破性进展，在这里面补充个信息，评判一个技术是突破还是提升的依据是什么，我们以语音识别为例，人的语音识别的上限大约在95%，这是有科学依据的，前面有人专门有这方面的科学说明。当机器对语音识别达到95%或者超越95%时，我们称之为突破。如果低于人类的水平，我们称之为提升。比如从原来的80%到90%，称之为提升；如果从90%提升到97%，就是突破。

4.2.2 人工智能发展阶段划分



人工智能发展阶段的标准是什么？这里列了3种，一种是业务能力、一种是主要技术、一种是驱动因素。人工智能的业务能力包括交互能力、思考能力和服务能力。这种划分方式是一家咨询机构，爱微资讯总结出来的一种人工智能发展阶段的描述。在业务能力里面，分为三个阶段，第一是计算智能、第二是感知智能、第三是认知智能，追加了一种服务智能。在这个里面，计算是机器学习的计算能力，包括数据处理能力；可以把感知对应到语音识别和计算机视觉，把认知对应到思考能力；把服务对应到服务能力。这四个阶段的划分，没有严格的时间限制，并没有刻意标明时间点。因为这几者之间是一直相互融合，并行向前进的，只是计算出现的更早、感知出现的更早。每个阶段相对于后一个阶段出现的更早一些而已。它和下一个阶段有个一定的延续周期。

基于主要技术。从上个世纪50年代人工智能发展以来，根据人工智能发展的主要技术，把人工智能分为三个时段：一是推理时代，就是上世纪50年代到70年代；二是知识工程，70年代到90年代，这是以专家系统为代表性的；再有一个就是数据挖掘，2000年至今。

基于驱动因素。我们讲人工智能技术发展的驱动因素，包括技术驱动、数据驱动和场景驱动。其对应的时间点如图。接下来，看一下基于每一种分类标准划分的阶段的阶段特征。

2.3业务能力

第一个发展阶段：计算智能，是在计算这个环节，它使得机器能够像人类一样进行计算，诸如神经网络和遗传算法的出现，使得机器能够更高效、快速处理海量的数据。计算智能也和前面讲得70年代出现计算能力限制了人工智能的发展，在后面一个时间点，分布式计算出现给人工智能发展带来了契机可以关联。计算机智能以数据的计算自动化完成为标志。

第二个发展阶段：感知智能，让机器能听懂我们的语言、看懂世界万物。语音和视觉识别就属于这一范畴，这些技术能够更好的辅助人类高效完成任务。听懂，会听是感知的层面，听懂是认知的层面。在感知阶段，语音和视觉识别属于这个范畴，而且是这个阶段的关键技术点。

第三个发展阶段：认知智能，在这一阶段，机器将能够主动思考并采取行动，比如无人驾驶汽车，比如说人机交互的智能问答机器人，实现全面辅助甚至替代人类工作。这里面没有刻意列入服务智能，这是因为服务智能到目前为止并没有出现成熟的产品形态。

2.4主要技术

第一个发展阶段：推理时代（这个时候研究的主要是推理机）。这一时期，一般认为只要机器被赋予逻辑推理能力就可以实现人工智能。不过此后人们发现，只是具备了逻辑推理能力，机器还远远达不到智能化的水平。

第二个发展阶段：知识工程。这一时期，人们认为要让机器变得有智能，就应该设法让机器学习知识，于是专家系统得到了大量的开发。后来人们发现，知识的归纳需要耗费大量的人力成本，把知识总结出来再灌输给计算机相当困难。举个例子来说，想要开发一个疾病诊断的人工智能系统，首先要找好多有经验的医生总结出疾病的规律知识，随后让机器进行学习，但是在知识总结的阶段已经花费了大量的人工成本，机器只不过是一台执行知识库的自动化工具而已，无法达到真正意义上的智能水平进而取代人力工作。

感知能力的发展反向推动认知智能的需求，认知智能依赖知识工程。在知识工程这部分，在一定的时间段之内，它有了偃旗息鼓，步入低谷的表现。但是到今天，感知智能发展到一定阶段之后，知识工程反倒会有重新崛起的可能和机会。因为有些知识适合通过机器挖掘的方式来学习。这个阶段的主要技术是对主题词表的构建，对本体理论的研究，对语音网络的研究。

第三个发展阶段：数据挖掘，这一时期，随着各种机器学习算法的提出和应用，特别是深度学习技术的发展，人们希望机器能够通过大量数据分析，从而自动学习出知识并实现智能化水平。这一时期，随着计算机硬件水平的提升，大数据分析技术的发展，机器采集、存储、处理数据的水平有了大幅提高。特别是深度学习技术对知识的理解比之前浅层学习有了很大的进步。

数据挖掘阶段对应的是机器学习算法，神经网络，重要的一个操作模式是通过人工去标注一些数据的训练样本，比方说语音样本、图像样本。干什么呢？对于声音信号标注对应的文字信息，然后让机器通过神经网络的模式自行学习，去发现声音信号和文字信号之间match 的关系，这个的背后是因素模型，就是这个汉字，它的发音是什么样子的，让机器学习汉字语音特征，从而实现语音到文本的翻译。这个时候，通过神经网络技术，通过大量的训练样本的就绪，慢慢出现真正人工智能技术的突破。在这个阶段，最关键的技术是深度学习、神经网络、计算机视觉相应的技术。

2.5 驱动因素

第一个发展阶段：技术驱动，是集中诞生基础理论的阶段。这个阶段奠定了人工智能发展的基本规则，并诞生了基本的开发工具，为日后人工智能的研发工具的升级开辟了先河。在这个阶段，技术的发展，尤其是算法的发展，成了推动人工智能进步的最大动力。达特茅斯会议之后，人们对于算法程序和语言开发投入了极大热情，掀起了人工智能发展的第一波高潮。

第二个发展阶段：数据驱动，是数据推动人工智能更新迭代的阶段。这个阶段，可获得和分析的数据飞速增长，不仅磨炼和提高了计算的能力，使人工智能的大规模运算成为可能，并且也反过来倒闭了数据的采集、清洗和积累，以及相应的软硬件基础设施的发展—这些都带动了大数据行业的腾飞。

这个阶段，通过就绪大量训练样本，通过对训练样本的深度学习和神经网络的训练，就绪计算机的感知能力，才出现了一些关键技术的突破，比如语音技术，比如计算机视觉技术。可以理解为，这是数据驱动，人工智能才真真正正走下了神坛，从高高在上，存在云端的东西，变成了实实在在可以感知的东西。

第三个发展阶段：情景驱动，是情境推动人工智能更深入到具体应用的阶段。随着人工智能的技术发展和数据积累，行业逐渐发现短期内通用智能和强人工智能难以实现的，数据分布的情境化特性使得人工智能在特定的情境下的垂直发展成为了可能。核心的驱动是应用场景。商业的本质在于交换，交换什么呢？到最后要输出能力。研究技术、研究数据，到最后不会有产出的，在最后还是要关键解决人的需求问题。情景就是在特定的场景下满足人的某种需求的能力。是特定的情景和人的需求推动了人工智能往更深入的更具体的应用的发展。目前已经有了一些场景，比方说基于视觉的人脸支付，比如说基于语音的各种智能音箱，虽然说这些东西目前还很难看出在服务能力方面能够提供多少具体的应用价值，但是它已经出现了，只是成熟度还有待于完善。

2.5想一想

人工智能时代，企业如何制定人工智能战略？会从什么样的出发点进行切入。