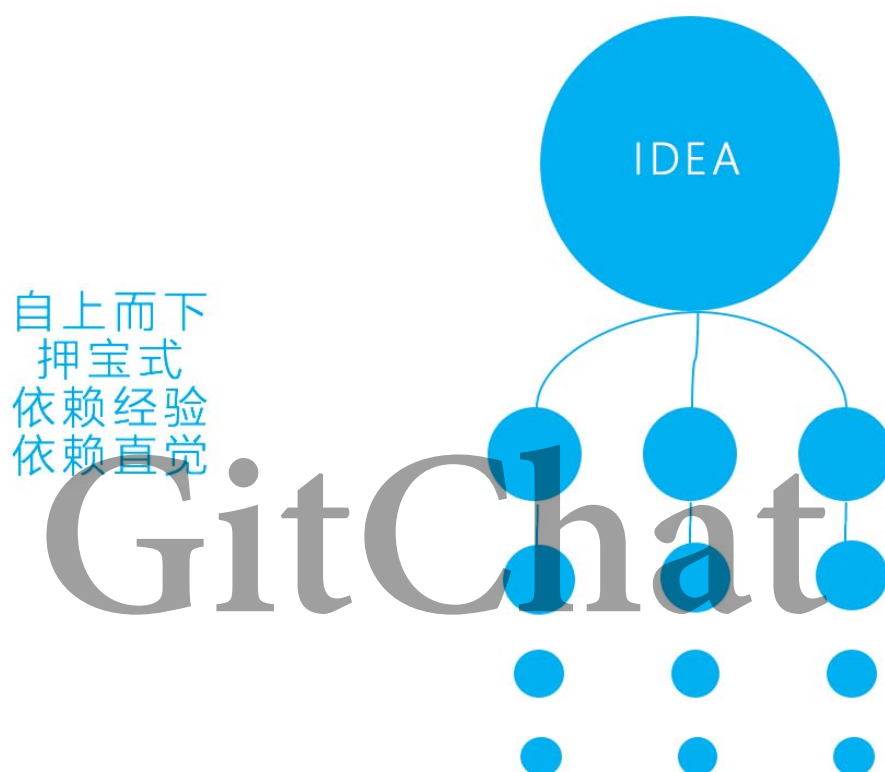


大数据和 AI 驱动创新实例分享

传统企业创新流程的特点

传统企业的管理模式决定了，他们的创新流程具有以下特点：



自上而下

传统的创新模式大部分都是自上而下的，很多的时候领导层不仅把方向指出来了，而且对于创新具体执行的方法也会给出具体的指导，并且有的领导倾向于将某个创新拆解成一个个的任务，然后分给不同的部门去完成，最后他们认为，只要每个部门完成了各自的部分，那么这个创新也就成功了。

抛开创新想法是否正确之外，这样的问题，在于：

1. 严重扼杀了中低层员工的主观能动性，他们成为一个任务的执行者，并没有全貌的理解这个任务背后的意图和目标，而往往经过多层的传递，信息基本上严重失真，最后很难保证任务完成与设想的一致性。
2. 反馈非常的缓慢，层级之下的创新工作一层层的下达，执行过程中的结果和反馈又一层层的向上传递，导致反馈非常缓慢，往往会失去最佳的判断机会。

押宝式的

自上而下的创新带来的就是领导层独家一言堂指明了创新的具体行动任务，而导致其他人都很难提出不同的观点。这样的创新项目就变成了下级人员对于高层意识的执行力的考核，大家都会全力以赴的去执行，导致这个创新的投入会非常的大，而不是精益的小步快跑的方式，变成了一个押宝式创新，这样的优点是集中了绝对优势的资源，但是风险就是这些创新的想法是否经过了验证，都不一定。

基于经验，依赖直觉

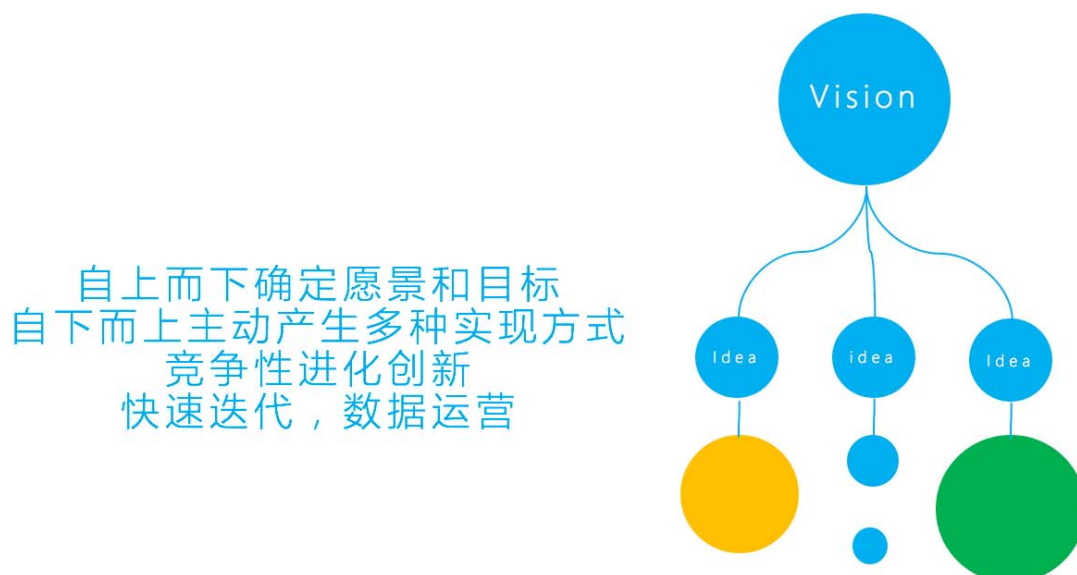
传统的创新更多的是基于经验和直觉的，而在现在的时代，数据基础已经有了很大的提高，我们在经验和直觉的基础上要加入更多的对于数据的探索和总结的内容，才能更加客观理性的去分析每一个创新的想法。

小结

自主创新的经济组织主要有两种形式：自上而下的创新和自下而上的创新，两种组织形式的一个重要差别在于投资的来源以及创新的动力来自何处。罗森伯格和伯泽尔指出，要使技术变革有效而持久，当局必须放弃它们对革新进程的直接支配权，并使之分散。这实际是为技术进步的发生创造了一个重要条件：成功的创新者要有富足的机会和决策权。根据经济发达国家的经验，技术进步很少是自上而下的，通常也不需要得到权威部门的许可。

自上而下往往是少数人的选择或决策决定了多数人的选择，在这种组织形式下，下边的人为了迎合上边人的偏好或需要，往往放弃了自己的选择。而且这种组织形式会导致重复研究或建设，同构化严重。自下而上的创新是草根式创新，有利于发挥每个人、每个经济主体的创造性，所产生的创新能量远大于自上而下的创新。

所以，我们所提倡的创新应该是：



自上而下的给出方向和愿景，自下而上的产生创新的想法和行动

高层不具体指出创新的行动，而是给出创新的愿景和目标。具体行动则自下而上的去产生，每一个创新想法都围绕这个愿景和目标。将全部的信息都赋予所有的团队，激发团队的主观能动性。而来自不同部门，不同领域的员工对于愿景目标实现的方法会有不同的见解，这样能够全面的产生更多的想法。

小规模试点，多样化启动，快速的迭代

不要在想法被验证之前，一开始就投入太多资源。加速创新的迭代周期。

建立观察和数据反馈体系

在创新的过程中，实时的采集和分析每一个创新产品产生的数据，从而进化竞争出最优的产品。

每个企业都需要一个数据驱动的创新农场

每一个企业都需要的数据驱动的创新农场：



这个创新农场由6个组件构成：

- 土壤：快速交付平台

创新的想法需要在土壤中快速生长，从而得到反馈，所以如何能够快速的将想法实现出来，变成可以运行的产品，将产品投入市场是非常重要的。

这里的快速交付平台，主要包括以下内容：

- 弹性基础设施，即交付团队使用底层云计算平台的方式，既包括各种虚拟机和镜像的管理，也包括生产环境的水平伸缩能力。
- 持续交付流水线，交付团队编写的代码需要通过这条流水线最终变成可以上线运行的软件。部署运行时，软件在开发、测试、试运行、用户验收、培训、生产等各种环境需要部署的环境。快速交付平台可以让开发团队，低成本，快速将一个想法实现出来，并且投入市场验证。
- 肥料：DevOps。为什么把DevOps比喻成肥料，是因为DevOps也可以加速软件产品的迭代。
- 种子：创新想法。创新想法就是这个农场中的种子，将一个想法，不论是痛点驱动的还是愿景驱动的想法，都可以在创新农场中生长发芽。
- 水：数据湖。这么多年的业务发展，每一个行业的成熟业务基本上都已经固化。光靠经验，很难找到有效和有价值被验证的创新点。但是跨领域的数据，跨流程的数据的碰撞和融合，是可能发现价值的，这个数据湖中，包含着很多的创新种子。
- 农业无人机：数字化运营。像一个真实的农场一样，要想让农场的作物茂盛生长，一样需要运营。而运营最重要的是需要实时掌握作物的生长情况。这就是产品的运营，每一个企业都需要一个Control Tower。在这个控制塔里，企业管理者能够随时掌控所有运营数据，从而调配各类资源，更好的支持生产。

如何将大数据和AI技术利用到业务创新当中

凯哥在给企业做数字化转型的时候，经常会碰到这样的问题：

“我也知道数据很有价值，但是我不知道如何利用数据产生价值”

“我的老板很关注人工智能技术，你能不能想办法看看如何让我们也用上人工智能？”

“人工智能看上去很虚，到底他能帮我们解决什么呢？”

不论大家对于人工智能是相信还是不信，但是大部分的企业都希望能够利用这个趋势做点什么，但是技术成熟度如何，到底要花多少的资源能做到什么程度，产生什么样的价值，很多企业都不清楚。

如何将AI技术切实的利用到业务创新中，如何让AI和业务场景结合紧密并且让业务人员接受？

在过去的一年中，凯哥和团队一起在做这方面的实践，结合项目实践，总结出了一套方法论，Data Discovery，分四个阶段来探索大数据和AI在企业的创新实践。

1. 现有业务架构梳理 2. 相关行业/业务研究 3. 用户调研访谈	1. 业务场景设计 2. 业务流程定义 3. 利益相关者定义	1. 目标定义 2. 采样数据获取 3. 数据历程图设计	1. 目标和范围 2. 交付物和成果 3. 架构设计
--	--------------------------------------	------------------------------------	----------------------------------

业务探索

第一个阶段，先进行全面的业务探索，从而从痛点和创新点两个方面来驱动创新。

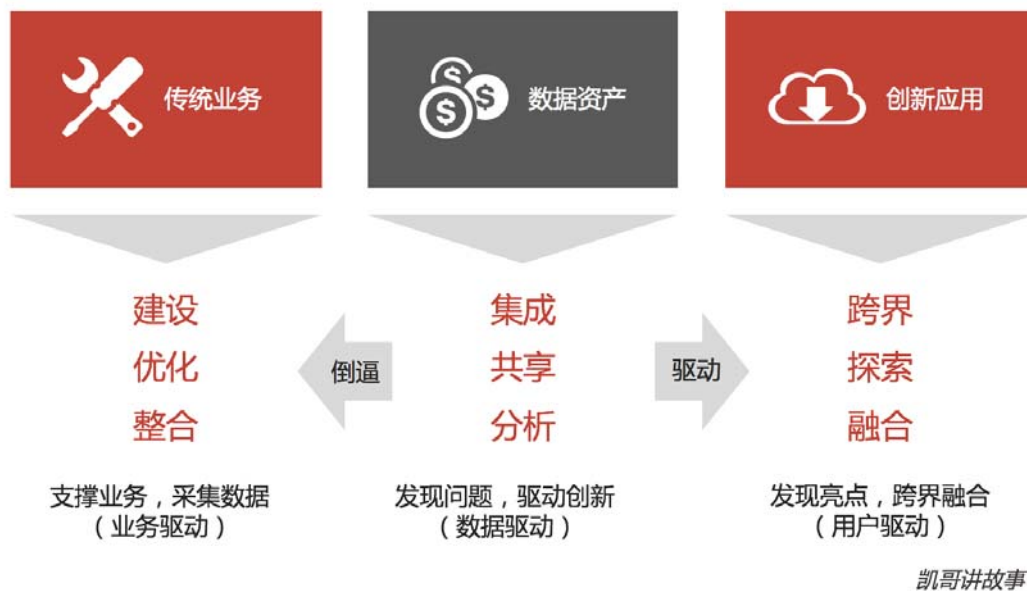
这个阶段一般可以分为6个步骤：

- 现有业务架构梳理
- 相关行业/业务研究
- 用户调研
- 创新业务价值探索
- 数据蓝图浏览
- 创新价值蓝图

通过对行业的相关最佳实践的理解，结合现有业务架构，通过用户访谈，创新实验室这样的活动，挖掘出用户的痛点和创新想法。

同时对于每一个创新的想法和问题，进行分析，包括相关的数据探索，也就是说，如果要实现这样的想法，或者解决这样的问题，我们需要什么数据，需要什么样的算法和技术。

在我们有的客户中，我们发现，这样做的结果，同时不仅驱动了应用的创新，也倒逼了传统业务的改造和优化。比如有的业务系统在过去的历程中，经过多次变更和迭代，业务错综复杂，已经很难从业务层面去优化和梳理，但是当我们梳理数据的时候发现，原来有些流程或者步骤是可以被优化的。



最后，要制定出这个企业的数据创新业务蓝图，也就是将调研得到的所有的创新点和痛点都梳理出来，一览无余。

产品定义

得到创新业务蓝图后，我们要逐个梳理出每一个业务的高阶场景，要逐个对这些产品进行定义，包括以下工作：

- 定义核心业务场景。
- 发现该产品的业务价值。
- 梳理业务流程及对应的利益相关者。
- 对这个产品需要的资源（数据、算法、技术、平台、人员）进行梳理。
- 进行价值成熟度评估。

最后，通过每一个产品的定义和梳理，可以排序出优先级最高（成熟度最高，价值最高）的产品，并且制定出后续这些产品的实施规划及路线。

试验验证

对于高优先级的产品，我们就要进行试验和验证，从而判断可行性和投入产出比：

- 定义试验的方式。
- 定义和获取采样数据。
- 进行数据里程图的梳理和设计。
- 进行技术验证（POC）。

- 对POC的结果进行总结和分析，并且尽可能形成可度量的数据，供客户决策方参考。

当那些高优先级的产品验证通过后，就进入了计划和执行阶段，那就是每一个敏捷项目交付的MVP阶段了。

计划启动

- 确定项目计划和范围。
- 制定资源和时间进度计划。
- 设计交付物成果的架构。
- 梳理数据流。
- 然后就进入到具体的开发实施阶段。

数据和AI驱动创新实例分享

下面的几个真实的案例，就是基于上面的Data Discovery的方法论从众多创新想法中筛选出来，最后有了实际落地的应用。

深度学习构建移动导盲犬

背景

盲人过马路是非常困难的事情。对于盲人来讲，走盲道没问题，有拐杖作为探测仪。但是过马路则很危险，因为马路上没有盲道，盲人也看不到红绿灯，更看不到对面走来的行人和旁边的机动车。最重要的是，盲人一旦偏离了直线，就很难根据感觉在走向直线。而在世界上，导盲犬是非常稀有的，培训成本非常高，并不是每一个盲人都能够享受得到的。

愿景

ThoughtWorks和澳洲一个公益组织一起，希望通过技术的手段来解决盲人过马路的问题。

创新实践

尝试了多种可能性，在拐杖上安装传感器陀螺仪来识别方向，当出现偏差的时候给出提示。但是发现效果很不好，方向定位很不准确。中国区大数据和人工智能团队参与了其中的一个尝试，利用深度学习来训练模型，利用手机陀螺仪来检测方向，从而纠正和指引盲人的行动。

想法的起源是自动驾驶技术，我们期望通过图片识别和深度学习，建立模型，让它学习正常人过马路的模式，然后给盲人以指引，然后用陀螺仪来确定和侦测方向。

所用的是CNN模型，用这个模型来学习何时需要改变方向，从而来预测何时需要提醒盲人改变方向以及向那边拐。



Hold camera still in hand and walk across different streets.

训练数据来自于每天上下班的手机视频分帧图片，然后同步将手机的传感器的状态同步的进行训练，从而让模型能够识别其中的规律，做到在出现类似的模式的时候，进行类似的行动预测。

遇到的困难及解决办法

在训练的模型完成后，第一准确度需要提高和优化，第二真实的场景是，人走路的时候很难保证不晃动，所以APP就会频繁的提醒往左拐，往右拐，让人很容易就晕了。后续的解决办法是将视线前方分成均等的区域，只有在手机检测到偏离超出了区域范围再去提醒用户。

机器学习助力精准营销

背景

这是一个银行客户，他每天的潜在分期付款用户有20万左右，但是他的呼叫中心呼出能力每天只有2万，所以如何在20万的潜在用户中筛选出最有可能的那些潜在用户，成为了他们的痛点。

愿景

如何利用数据分析更加精准的筛选出潜在用户。

创新实践

最早我们利用人工建模的方式做的，采样的特征值也不够多，只是采用了一些常用能想到的特征值，比如消费总额，最大金额，平均消费额等，但是测试的时候发现准确度不高。然后我们加入了更多的数据，利用神经网络进行训练，加入了不仅是信用卡，而且这个消费者的储蓄卡的数据以及用户的基本信息和相关信息。监督学习的方式，将准确度提高了5倍以上。

数据分析优化物流配送

背景

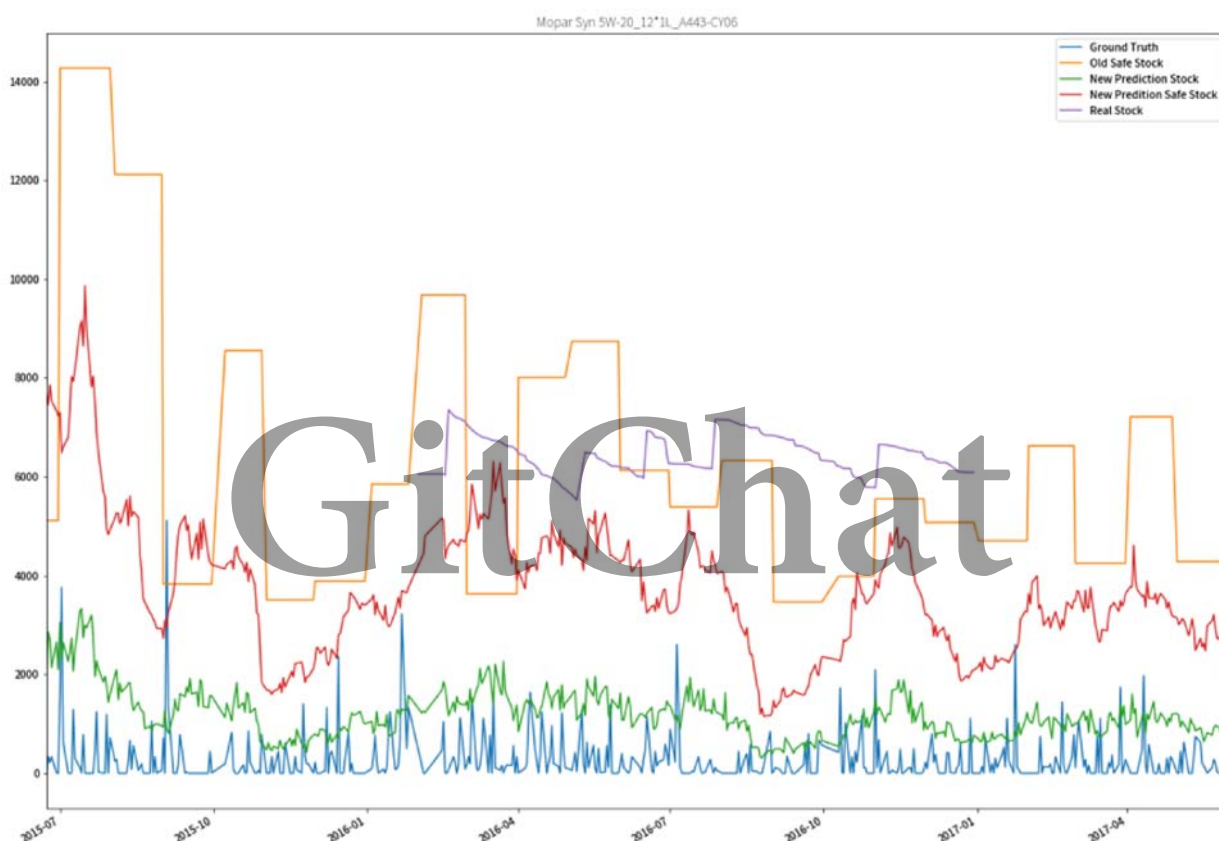
这个用户是一个零售客户，他原来在全国有4个大仓库，从这四个大仓库往全国的3000家客户发货，但是随着市场竞争的加剧，他们希望通过数据分析来优化仓储物流的设置，从而将配送时间从7天缩短到2天。

愿景

核心是要解决两个问题，一个是在哪里租仓库仓储成本和运输成本最低，第二个是每个仓库的安全库存如何设置。

创新实践

通过对这个用户的历史订单数据的建模分析，我们发现以下创新点：



安全库存预测能够大幅降低客户的成本，这个给客户以非常强的信心去推动这个想法的深入执行：

- 预测发生库存XXXXX kg·天 能够节省库存 82.5 %，而无法兑付的比例是30%多。
- 预测发生库存XXXXX kg·天 节省 30.3 % 出现无法兑付天数1天。



同时，发现一些产品的SKU之间有潜在的替代和关联销售的关系。

总结

数据基础是非常重要的，所以这一类的项目，数据质量是成功的关键，在每一个项目做分析之前，数据的清理，转换，映射的处理工作非常大。

要有非常发散的业务想象力，因为非监督学习目前还是不够成熟。所以还是要先有一些预设，然后基于这些预设，去制定特征值，在进行建模和训练。（涉及到一些用户的信息，所以很多数据和技术细节无法展示，如有需求，可以线下讨论）。