# 智能化软件开发大脑 – UDev

Table of Contents

[智能化软件开发大脑 – UDev 1](#_Toc529807200)

[定义 1](#_Toc529807201)

[核心技术 1](#_Toc529807202)

[评价标准 1](#_Toc529807203)

[案例分析 – 物体识别系统 2](#_Toc529807204)

[痛点 2](#_Toc529807205)

[设计原则 3](#_Toc529807206)

[介绍视频需求 3](#_Toc529807207)

[举例 3](#_Toc529807208)

[难点 3](#_Toc529807209)

[优势 4](#_Toc529807210)

[劣势 4](#_Toc529807211)

[未来 4](#_Toc529807212)

[TODOs 4](#_Toc529807213)

[附加 4](#_Toc529807214)

[参考 4](#_Toc529807215)

## 定义

* 一款基于人工智能的软件研发和流程优化平台，为项目需求方，交互设计师及算法工程师等团队成员提供研发流程优化，代码自动生成等技术。基于敏捷软件开发原则，对从业务初始需求分析到最终成品交付流程进行深度优化。

## 核心技术

* 以深度学习作为工具，提供草图到代码生成功能（如pix2code ）
* 以深度学习作为工具，提供机器人和工程师结对编程，以加速软件开发过程
* 预设定技术架构
* 预集成业界先进的库

## 评价标准

* 更短的软件研发周期
* 更低的项目风险
* 更低的项目成本
* 更高的需求方，设计师，算法工程师满意度

## 案例分析 – 舆情分析与控制

1. the system should able to check simultaneously the different web sites following a key word
2. the global matching and localization of the sites following their IP address
3. able to reduce the access or to restrict the access to these web sites
4. following the activities of different pages in the different social media
5. ability to restrict the access to the pages (shouting) the activities using URL or other tool
6. checking the activities through VPN or other cyber soft and limit or restrict the flux
7. identify the users and location of the persons behind the different accounts
8. to secure the activities (virus and other massive attacks)
9. the global trend and the different connections of the cyber activists
10. to reach all the information based on the influence and the followers of the different pages in the social network and social media
11. All other items who could be integrated, and related to the cyber activities, cyber crimes and counter insurgency

(Globally, we need the equipment, tools and software and training)

## 案例分析 – 物体识别系统

需求方 & 研发团队

步骤：

* 给出 问题和数据集。在这里，Happy Whale DB中的2,5000图片，做Whale Identification
* 给出 问题评价标准。在这里，每幅图片对应5个标签。这幅图片里面出现了哪条鲸（每条鲸有一个ID编号）
* 定义 目标：构建单条鲸鱼物体识别引擎
* 定义 关键成果：整体准确度需达到80%上（和现有准确度作交叉对比）
* 尽早发现 潜在挑战。对于每一只鲸，只有有限的几张图片

原则

* 数据驱动
* 强调目标与关键成果

人员构成

* 1产品经理
* 1全栈工程师
* 2算法工程师
* 1个领域专家

引擎部件与架构体系：

* 数据源与爬虫系统
* 预处理系统
* 深度学习物体识别引擎
  + 注：当准确度（或其他模型衡量指标）在某阀值下时被触发再训练。通常为离线训练，在线预测。
* UI前端交互
  + 预测数据展示，用户审核，AI正负反馈闭环，打标签数据。前端交互是人工智能系统必不可少的组成部分。

## 痛点

* 需求频繁变更
  + 敏捷和迭代式开发
* 项目维护漫长且困难
  + 云端，SaaS
* 风险，成本重要指标预测不准
  + 预测大约范围，而不是准确值
* 人员流失造成供给方缺陷
  + 舒适环境，提高工资待遇
* 流程的繁杂，需求说明书，信息传递的损失
  + 删减掉冗余中间环节，让初始需求方直接参与到软件研发全流程中

## 设计原则

* UDev不遵循瀑布式软件开发模式，立足于敏捷软件开发原则，为用户提供更为高级的智能化软件研发新模式
* 小步快跑，快速迭代
* 大量使用脚本进行集群优化
* 人工智能流水线和可切分的小步骤（如数据预处理，模型训练和参数调节等）
* 我们最优先要做的是通过尽早的、持续的交付有价值的软件来使客户满意
* 大量使用预训练模型，并在其基础上作10%左右的个性化训练，大幅度节约时间
* 即使到了开发的后期，也欢迎改变需求
* 经常性地交付可以工作的软件，交付的间隔可以从几周到几个月，交付的时间间隔越短越好
* 在团队内部，最具有效果并且富有效率的传递信息的方法，就是面对面的交谈
* 敏捷过程提倡可持续的开发速度

## 举例

* 计算机自动设计一通讯录，背后数据为(name, value) pair。由数据触发出多种表现形式，加入用户个性化需求。

## 难点

* 数据集及其转换操作
* 如何描述用户的交互场景
* 需求的明确，把信息转化为用户场景，子场景的过程
* 定义何种UI 交互操作
* 部件复用
* 前端技术落后，需要漂亮界面，可重用性
* 数据多样性
* 高效交互

## 优势

* 个体和交互胜过过程和工具
* 可以工作的软件胜过面面俱到的文档
* 客户合作胜过合同谈判
* 响应变化胜过遵循计划

## 劣势

* 研发难度大

## 介绍视频需求

* 按照时间顺序
* 场景化
* 对大多数客户的统计分析
* 对客户的个性化需求分析
* 体现适应能力强且快的特点
* 有可以演示的视频

## 未来

* 生态圈建立 = 平台 + 云润内部研发 + 外部开发者

## TODOs

* 在垂直搜索引擎业务线上开始收集用户潜意识行为和动作
* 支持更多数据源和增强数据兼容性
* 增强UI前端模组化能力

## 附加

* 用户操作即新增，搜索，删除一个data object
* 云润整体：垂直搜索引擎与舆情分析，to B
* 人才储备：缺高级前端开发

## 参考

* aiXcoder
* pix2code
* Neural Programmer Interpreters
* DeepCoder: Learning to Write Programs