



# Weaving 作品展示

By JerryMouse



# What's Weaving?

- 编织一张Web人际网
- 帮助认识新朋友

# More Details

- 人际关系搜索引擎
- Go to <http://weaving-gae.appspot.com/>



# Content

- 产品
  - 产品背景
  - 功能
  - 竞品分析
- 技术
  - 工作机理
  - 具体工作
  - 子项目

# 产品背景

A decorative floral pattern in a lighter shade of blue, featuring a central flower with a long stem and several leaves, and a large, stylized leaf or petal shape at the bottom right.

# 产品背景

- 随着社交网站的逐渐增多
  - 更多的潜在社交需求被满足
  - 需要与更多不同的平台上和朋友进行互动
  - 需要一个工具来帮助他们拓宽自己的社交圈，获取新信息、新知识

# 产品目的

- 帮助人们认识自己想要进入的社交圈的人，从而拓宽自己的社交圈
- 语义网的方式向各企业提供互联网上开放的个人及其社交信息

# 用户需求——一个人用户

帮助人们认识自己想要进入的社交圈的人，  
从而拓宽自己的社交圈。

- 对特定的某个人很感兴趣，想要了解他的信息
- 对某个人有兴趣，顺带想要了解他认识和关心的人
- 想要了解某件事情，需要找到和这件事有关联的人



功能



# 对于个人信息的搜索

名称	描述
1 搜索此人的博客	
2 搜索此人的SNS账号	搜索结果呈现此人注册过的SNS网站的连接方式
3 搜索此人的邮箱	
4 搜索此人的产出信息	即此人在使用互联网过程中有意产生的信息
4-1 此人的博文	
4-2 此人的状态更新	此人在使用社交网站所产生的状态更新
4-3 此人的阅读分享	如 Google Reader， Delicious等网站的分享
4-4 此人的知识贡献	如在百度知道， 爱问等网站的知识贡献
5 搜索关于此人的资讯	与此人相关的互联网资讯
6 搜索关于此人的图片	
7 形成关于此人的信息聚合	通过订阅此feed， 可以全面跟踪此人信息更新

# 对于个人关系信息的搜索

名称	描述
1 搜索此人认识的人	博客的友情链接，双向follow等线索都可推断出两人是否相识
2 搜索此人感兴趣的人	根据他所follow的人来推断出他所感兴趣的人
3 搜索对此人感兴趣的人	Follow此人的人

# 小众知识搜索

名称	描述
1 搜索某知识领域的专家	通过具有代表性的搜索关键字，向用户显示与此知识领域有关的人物。
2 搜索能够解答自己疑问的人	人们常常在学习工作中产生
3 搜索和自己有相同兴趣的人	

通过搜索与某知识  
相关联的人物



与此人物互动的过  
程中创建新知识

与其说是知识搜索，不如说是知识创造。

# 竞品分析

A decorative floral pattern in a lighter shade of blue, featuring a central flower with a long stem and several leaves, and a smaller flower to the right, all set against a dark blue background.

# Weaving和普通搜索引擎对比

	普通搜索引擎	Weaving
搜索目的	搜索现成的资讯；学习现有的知识	认识人；获得个性化的知识
搜索流程	单层的关键字搜索	多层的推理式搜索
搜索结果	以与关键字相关的页面为单位	以人为单位的搜索结果，定位到个人社交网络

# 类似竞品分析——产品思想

产品名称	产品思想
Google Social Search (Google)	从你的朋友和社交圈获得的内容网网比从陌生人获得的内容对你具有更大的价值
Follow Finder (Google)	假设两种社会联系：拥有类似follow列表的用户，你可能想要follow的其他用户
Aardvark (Google)	通过用户提供的个人信息以及技能信息来帮助找到适合回答问题的人以及他能回答的问题
SOSO华尔兹（腾讯）	能够把与你要搜索的人物相关的其他人物都连成一个关系网，按照热门程度来进行相关联，可以发现此人与其相关联人物之间的关联
人立方关系搜索（微软）	当用户给定任意搜索关键词，它能够找出与关键词最可能相关的人名、地名和机构名，并且根据它们与关键词之间的相关度排序
百度知道	和搜索引擎完美结合。让用户所拥有的隐性知识转化为显性知识，用户既是内容使用者，又是内容创造者
猫扑人肉	通过他人来搜索自己搜索不到的东西

# 类似竞品分析——搜索范围

产品名称	范围
Google Social Search (Google)	以Google Profile为基础在用户已有社交圈进行搜索
Follow Finder (Google)	以Twitter为基础，根据所提供的账号圈定搜索范围
Aardvark (Google)	根据用户提供的信息在所有用户中找到最适合的问题解答人
SOSO华尔兹 (腾讯)	以已经建立的人物数据库搜索名人
人立方关系搜索 (微软)	通过对网页信息的抽取计算在整个互联网搜索
百度知道	在已注册用户中进行手动式知识搜索
猫扑人肉	在已注册用户中进行手动式搜索



# 类似竞品分析——数据来源

产品名称	博客	邮箱	SNS账号	标签	人	二级联系人	其它
Google Social Search	√	√	√	×	×	√	①
Follow Finder	×	×	√	×	×	×	②
Aardvark	×	√	√	√	√	×	
SOSO华尔兹	×	×	×	×	×	×	③
人立方关系搜索	×	×	×	×	×	×	④
百度知道	×	×	×	×	√	×	
猫扑人肉	×	×	×	×	√	×	

# 类似竞品分析——获得信息

产品名称	博客	SNS 账号	网站	图片	关系	新闻	博文	其他
Google Social Search	√	×	√	√	×	√	√	①
Follow Finder	×	√	×	×	√	×	×	②
Aardvark	×	×	×	×	×	×	×	③
SOSO华尔兹	×	×	×	√	√	√	√	④
人立方关系搜索	√	√	√	√	√	√	√	
百度知道	×	×	×	×	×	×	×	
猫扑人肉	√	√	√	√	√	√	√	

# 竞品分析总结

- 精确的人际关系搜索必须建立在详实的用户信息的基础之上，即深入用户的社交网络。
- 目前的人际关系搜索多着眼于名人搜索以及固定社交圈的内容搜索，对个人社交圈的拓展以及个人知识贡献的关注度仍然较低。

# Weaving定位

## 产品思想

人们社交的最终目的就是获取帮助。不同的人需要不同的帮助，因此不同目的、背景的人就会形成不同的社交圈。而Weaving的目的就应该是帮助人们认识自己想要进入的社交圈的人，从而拓宽自己的社交圈。

## 范围

通过对页面信息的分析、提取、计算对整个互联网进行搜索

## 信息源

博客	邮箱	SNS账号	标签	人	二级联系人	其它
√	√	√	×	×	√	

## 获得信息

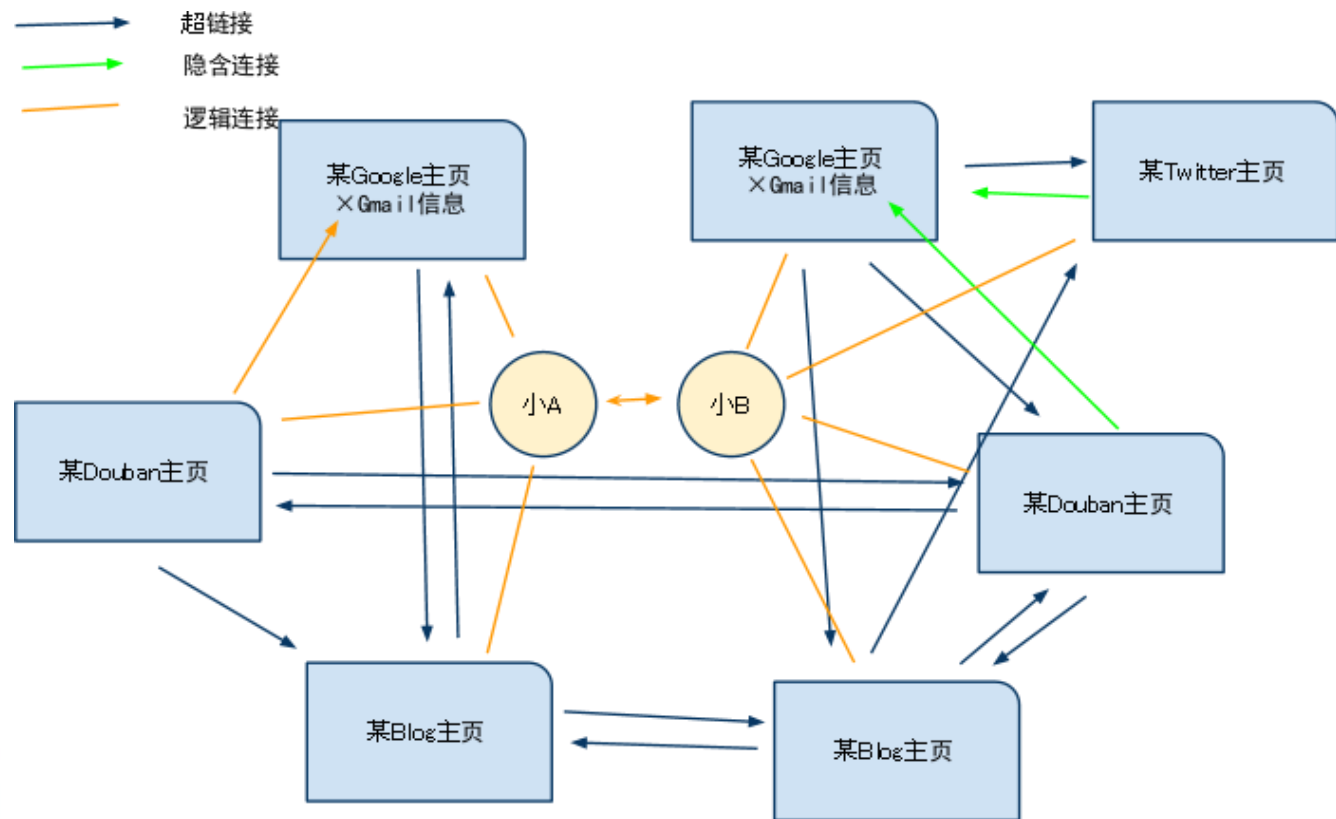
博客	SNS账号	网站	图片	关系	新闻	博文	其他
√	√	√	√	√	√	√	①

# 工作机理



# 网络现状

- 杂乱无章的链接关系
- 隐含有语义信息



# 加以整理→

## Web抽取

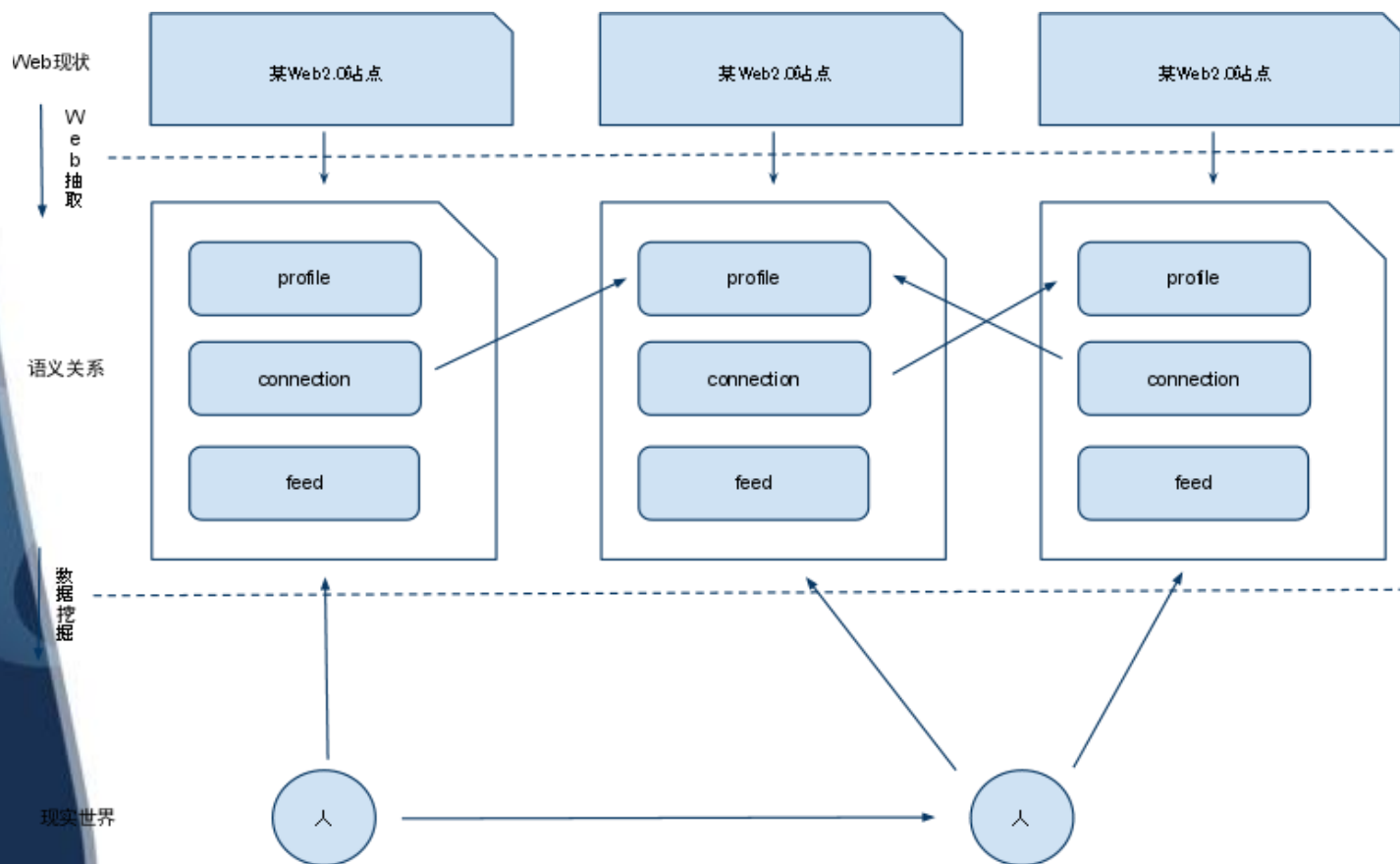
将互联网上的信息抽取为语义网

## 数据挖掘

从得到的语义网挖掘出“人”的信息

# 整理后

- 语义网
- 包含“人”的关系



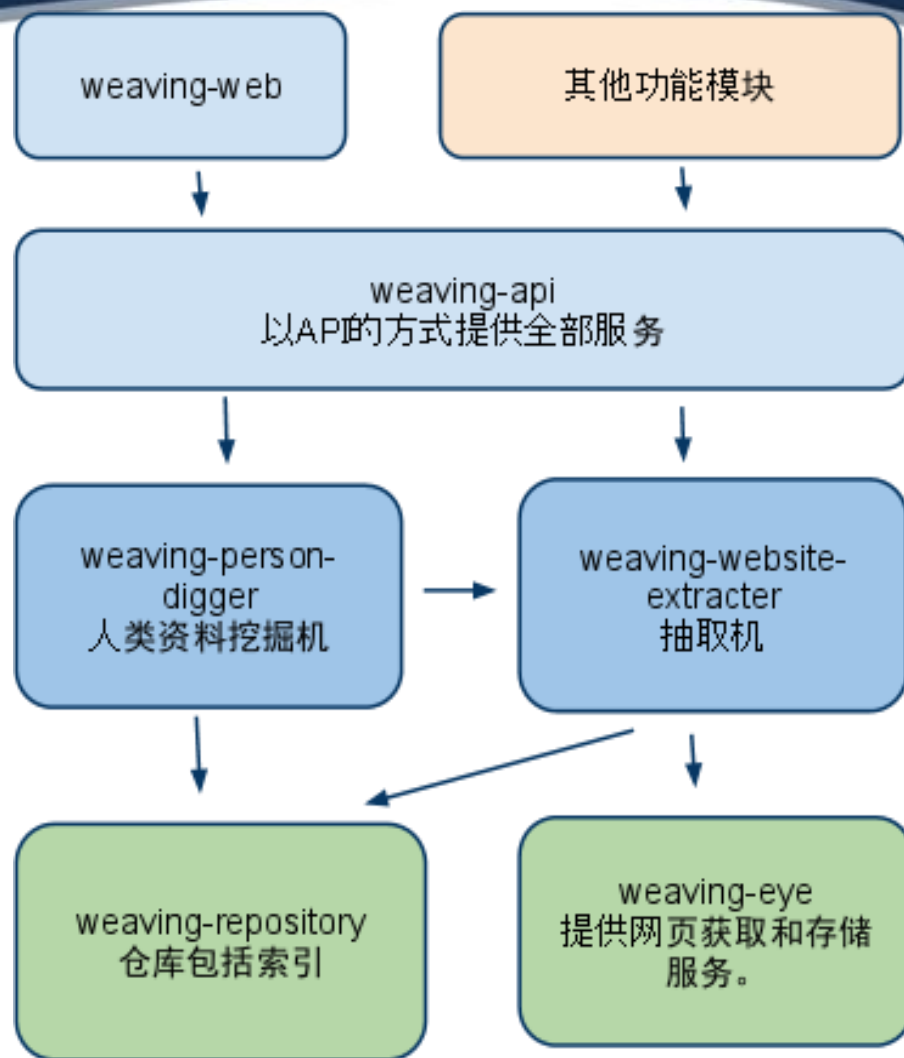


# 具体工作

A decorative floral pattern in a lighter shade of blue, featuring a central flower with a long stem, leaves, and a large, swirling vine-like element.

# 系统结构

- Weaving-API
  - dig(url,email or id):人
  - extract(url):语义网
- Weaving-person-digger
  - 从语义网挖掘出人
- Weaving-website-extractor
  - 从互联网抽取出语义网
- Weaving-repository
  - 信息存储索引
- Weaving-eye
  - 网页获取,存储。可翻墙



# 语义网抽取大法

困难

- 各大网站各不相同。页面千奇百怪。
- 独立博客
- 访问权限
- 墙

# 语义网抽取大法

## 解决办法

- 社交网站结构归纳
  - 个人资料
  - 关系
  - 资料
- 脚本引擎。三行搞定一类网站。 [示例代码](#)
- 动态策略选择
- 使用现有API(如Facebook Graph API)
- 翻墙模块

# 挖人神功

## 困难

- 语义网本身不准确。有假信息和信息不足
- 语义网太大，分析困难

## 解决办法

- 信息分级
- 从语义网提取出局部进行运算
- Agent(滚雪球的判别机器人)

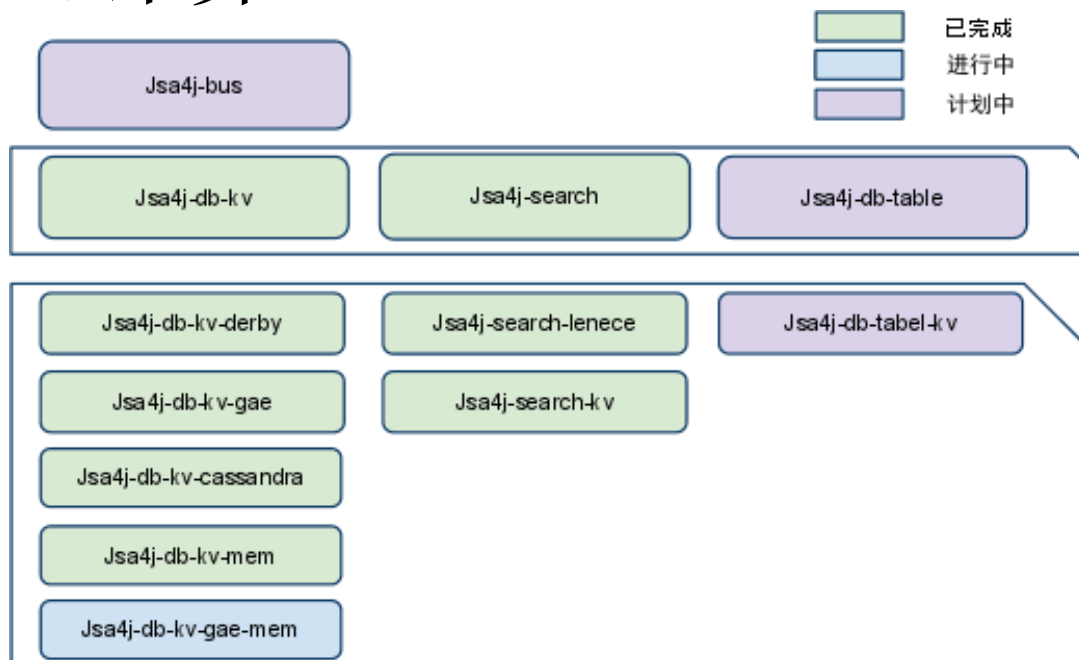
# 两个有趣的子项目

A decorative floral pattern in a lighter shade of blue, featuring stylized leaves, a small flower, and swirling lines, positioned on the right side of the slide.

# Jsa4j [网址](#)

通用数据底层

- 简洁
- 高效
- 支持云计算



# getFavicon [网址](#)

- 通过URL获取网站图标
- 图标编辑小工具

The screenshot shows the 'Favicon Getter' website. On the left is a navigation menu with links: 'home', 'Get Icon', 'Edit Icon', 'Doc', and 'contact'. Below the menu are two red buttons: 'Follow me on' (with a Twitter bird icon) and 'Send me an e-mail' (with an envelope icon). At the bottom left, it says 'powered by Google App Engine' and 'By Jerrymouse'. The main content area is titled 'Get Icon' in red. It contains four input fields, each with an icon and a label: a '@' icon for 'Domain', a person icon for 'Format', a ruler icon for 'Width', and a ruler icon for 'Height'. Below these fields is a grey button labeled '提交查询' (Submit Query).



Thanks

