# Id1

## what-在花旗让信息的传播和获取更智能

## why

S:让我们想象一下，作为协作平台的一员，你现在在协作平台上创建/分享了自己的一个文本作品（或者带有文本的富媒体作品），

你应该非常期望会有下面的结果：

1.你的该作品能被真正感兴趣的同事或者协作者看到或者关注，并期待有所反应。（我称之为依据内容来定向传播）

2. 如果你分享完该作品之后，能自动地接收到一些与该作品同主题的其它作者的作品，

这也许也是你希望发生的。

而目前，我们的协作平台只能依据个人（关注个体）来实现推送share。

如果能够实现以上期望， 根据内容来自动识别感兴趣的群体，从而实现主动地推送，和定向内容的反馈会更智能一些。

## How

Q1:那么，哪些同事会对我的作品真正感兴趣，如何智能地识别感兴趣的同事？

识别的依据：

作品内容的相似性，即 对基于词库的特征向量进行余弦定理运算，计算文本作品内容的相似性）

实现过程概述：

section1：生成词集和词库

stp1.将所有用户的所有文本作品作为我们的工作对象，

逐一对每个文本作品依据分词法（简单起见-依据空白符进行拆词）进行分词，

去重后，再去掉无语义的词（eg. the, a, an）得到第i个作品词集Wi(i=1,2,3,...,n).

step2.对所有Wi词集整合再去重后，再按照既定的排序方法（eg.红黑树排序法java TreeMap实现）排序得到排好序的目标词库WR。

step3. ste2中对词集去重之前，对词库每个词运用统计的方法计算得到每个词的权重。基于逆向-词频法（词出现的频率高，权重越小）得到每个词的权重。所以可以想象我们有个词库表，每个词WRi有它自己的权重WTi。

step4. 对step1中的每个词集Wi基于词库WR构建基于权重的特征向量，即得到每个词集的特征向量FWi。

step5. 针对当前提交的作品，用其特征向量FC和其它所有作品的特征向量进行余弦定理运算。

cos<FC,Fi>= (FCxFi)(FCxFi)/(|FC|X|Fi|)

step6：如果该值越接近于1，FC与Fi越相似。0.5以上的值我们暂定认为是相似的，或者筛选出值降序排序后的前若干个作品，得到匹配到的相似作品集。

我们暂定大于0.5的值是我们所需的保留值，其Fi对应作品的作者找到（用户表找到作者的邮箱或者协作平台账号）

Q2.我的该作品如何主动地到达感兴趣的同事？

step5: 我们只需将你当前提交的作品地址和标题推送给 在section1部分得到的用户。发送email或者推送到他们的c账号下的推荐栏（如果没有可web端可设计一小块区域）

Q3:当你在协作平台上分享完了自己的一个作品后，如何能自动地接收到自己想要的其它类似作品呢？

step6:我们只需将

---------------------------------------------------

图解说明：

用例图：

包图：

需要的资源：

----------------------------------------------------

Sumarry:

地图上有字

描述已自动生成手机屏幕截图

描述已自动生成

# id2

## what-让我们的项目更坚固，让客户的信息更安全

# id3

## what-让花旗的员工工作更有活力

坚持做广播体操-第八套广播体操，让你重温青春活力

饭后百步走-

避免久坐-每天偶尔站立办公：

每周或每两周一次长途Hiking或者骑行

# id4

## what-在花旗让信息获取更便捷

手机屏幕截图

描述已自动生成

# Id5

## what-让代码更易读懂和维护-养成给代码添加注释的习惯

# id6

## what-SQL注入的危害与避免

# id7

## 让花旗的国际化环境，培养花旗员工的独特的多边交流优势

坚持英文书写

抓住机会敢于表达自己，融入多元文化中。你中有我我中有你。