

OASIS需求文档

项目名称：OASIS

版本号：V0.1

文档修订记录

版本编号	日期	参与人员	修改内容	备注
V0.1	2020.2.18	濮宗悦、丁玲燕、王雪君	构建草稿	无
V0.2	2020.2.29	陆放明	补充接口	
v0.3	2020.3.7	丁玲燕	增加数据初始化需求	
v0.4	2020.3.14	濮宗悦、丁玲燕、王雪君	增加迭代二用例	
v0.5	2020.3.16	陆放明、濮宗悦、丁玲燕、王雪君	细化迭代二用例	

1 引言

1.1 目的

说明OASIS系统主要功能、规则细节，指导开发、测试，并作为产品基础文档存档，便于理解回顾，减小产品迭代或系统交接风险成本。

1.2 预期读者

开发、测试、最终用户

1.3 参考/引用文档

暂无

1.4 术语与缩写解释

OASIS: (Online grAph System for acdemIcS) 学术关系图谱系统

1.5 遵循的标准或规范

通用标准。

2 系统概述

2.1 系统介绍

通过将不同数据源的学术数据集成在数据库，完成实体和实体关系的抽取，从而构建一个学术关系图谱系统（OASIS系统），为广大用户提供高效的学术关系查询、学术同行评价、专家推荐系统、学术机构评价等

2.2 应用范围

前端客户、后端管理人员

2.3 系统角色

系统中各种角色及其职责，各种角色的具体行为将在功能性需求中描述

角色	操作
用户	进行相应查询

3 业务处理流程

3.1 流程1-查询流程

- 刺激：用户在主搜索入口输入要搜索内容
- 响应：系统返回搜索结果和按规则排序的前十条作者、关键词信息以及前五条发布组织信息
- 刺激：用户进行二次搜索：（可叠加）
 - 用户输入二次查询需要的信息
 - 用户改变指定的年份区间
 - 用户选择具体的作者
 - 用户选择具体的关键词
 - 用户选择具体的发布组织
- 响应：系统返回二次搜索的结果

3.2 流程2-查看Ranking

- 刺激：用户在主页中点击Ranking button
- 响应：系统跳转至Ranking页面，显示
 - 论文总数折线图
 - 被引用论文数最多和论文数最多作者TOP5的堆叠柱状图（包含作者被引用数最多的前5篇论文的名字、引用数和PDF link）
 - 被引用数最多的论文TOP10的环状图（包含其标题、作者、引用数、PDF link）

- 年度热门方向（词及热度）的词云
- 刺激：用户将鼠标移动至图表的某一部分
- 响应：系统显示该部分对应的细节数据
 - 柱状图显示该论文对应的标题、发表年份
 - 环状图显示论文的标题、发表年份、作者
 - 词云显示该词热度
- 刺激：用户点击柱状图和环状图的某一部分
- 响应：系统跳转至相应论文

3.3 流程3-查看排序

3.3.1 查看学者排序

- 刺激：用户点击首页学者排序图或者"Authors" button
- 响应：系统跳转至学者排序界面，包括：
 - 学者的研究领域筛选工具栏
 - 按照活跃度排序的学者列表
 - 搜索框
- 刺激：用户选中工具栏中的某个或多个研究领域
- 响应：系统返回该领域下按照活跃度排序的学者列表
- 刺激：用户选择按照论文总数 / 总被引用数 / 热度排序
- 响应：系统返回按照相应规则排序的学者列表
- 刺激：用户使用搜索框搜索特定学者
- 响应：系统返回名字和搜索内容匹配的学者
- 刺激：用户点击某个学者项
- 响应：系统返回该学者的详细画像页面

3.3.2 查看会议 / 机构 / 领域排序

- 刺激：用户点击首页会议 / 机构 / 领域排序图或者"Conference" / "Affiliation" / "Fields" button
- 响应：系统跳转至相应排序界面，包括：
 - 按照活跃度排序的会议 / 机构 / 领域列表
 - 搜索框
- 刺激：用户选择按照论文总数 / 总被引用数 / 热度 / 拥有学者数排序
- 响应：系统返回按照相应规则排序的会议 / 机构 / 领域列表
- 刺激：用户使用搜索框搜索特定会议 / 机构 / 领域
- 响应：系统返回名字和搜索内容匹配的会议 / 机构 / 领域
- 刺激：用户点击某个会议 / 机构 / 领域项
- 响应：系统返回该会议 / 机构 / 领域的详细画像页面

4 功能需求

4.1 功能清单

- 1. 数据初始化
- 2. 单一关键词搜索
- 3. 复合关键词搜索
- 4. 二次搜索
- 5. 报表展示

4.2 模块1-初始化

模块功能描述：

在使用网站功能前先进行数据的初始化，即将论文数据导入数据库。

4.2.1 功能1-数据初始化

用例编号	UC1-1
名称	数据初始化
创建者	丁玲燕
创建日期	2020-03-07
最后一次更新者	丁玲燕
更新日期	2020-03-07
参与者	管理员
触发条件	进入管理员界面
前置条件	无
后置条件	无
优先级	高
正常流程	1. 管理员进入管理员界面 2. 管理员选择初始化数据 3. 系统进行数据的初始化
扩展流程	无
特殊需求	无

4.3 模块2-查询

模块功能描述：

提供查询入口，使用户可以通过查询关键字或其组合查找目标论文。

4.3.1 功能1-单一关键词搜索

用例编号	UC2-1
名称	单一关键词搜索
创建者	濮宗悦
创建日期	2020-02-18
最后一次更新者	濮宗悦
更新日期	2020-02-18
参与者	普通用户
触发条件	用户选择了特定类型的搜索。用户输入关键词，点击搜索按钮。
前置条件	无
后置条件	无
优先级	高
正常流程	用户输入关键词，点击搜索按钮。系统展示用户的论文搜索结果
扩展流程	无
特殊需求	无

4.3.2 功能2-复合关键词搜索

用例编号	UC2-2
用例名称	组合关键词搜索
创建者	王雪君
创建日期	2020-02-18
最后一次更新者	濮宗悦
更新日期	2020-02-26
参与者	普通用户
触发条件	不修改搜索类型/选定搜索类型为ALL
前置条件	无
后置条件	无
优先级	高
正常流程	用户输入一个或多个关键词，用空格隔开。点击搜索按钮。系统展示搜索结果的论文及附加信息
扩展流程	1. 返回的搜索结果要附带产生结果数（即若有n篇论文的作者包含A，A的结果数就是n）最多的前十位作者（可以少于十位，但不能超过十位） 2. 返回的搜索结果要附带产生结果数最多的五个机构/会议（细节定义同上） 3. 返回的搜索结果要附带产生结果数最多的十个关键词（细节定义同上）
特殊需求	无

4.3.3 功能3-二次搜索

用例编号	UC2-3
用例	

名称	二次搜索
创建者	濮宗悦
创建日期	2020-02-26
最后一次更新者	濮宗悦
更新日期	2020-02-29
参与者	普通用户
触发条件	在结果页面的侧边栏继续进行二次搜索
前置条件	已经得到一次搜索结果
后置条件	无
优先级	高
正常	用户已经通过查询关键词A得到一个结果，在侧边栏操作得到搜索附加关键词B1,B2...Bn,C1,C2....Cn,....（一个字母代表一个维度，字母相同下标不同代表一个维度下的不

流程	同内容。具体获取方法见扩展流程），确认搜索后实际结果等价于一次搜索中使用关键词A AND （B1 or B2 or ... or Bn） AND (C1 or C2 or ... or Cn) AND...
扩展流程	1.可以在侧边栏里的搜索框输入得到一个维度的关键词 2.可以通过在侧边栏指定时间区间得到一个维度的关键词 3.可以通过进一步指定作者/机构/关键词（参见UC1-2的扩展流程）得到一个维度的关键词
特殊需求	二次搜索后附加信息不会改变。若再次进行二次搜索，是基于第一次搜索的基础上。

4.4 模块3-展示

模块功能描述：

用统计图、报表等方式展示用户可能感兴趣的关于学者、机构、论文等的统计结果

4.4.1 功能1-报表展示

用例编号	UC3-1
用例名称	报表展示
创建者	丁玲燕
创建日期	2020-02-18
最后一次更新者	王雪君
更新日期	2020-02-28
参与者	普通用户
触发条件	用户点击报表展示按钮
前置条件	无
后置条件	无
优先级	高
正常流程	<p>用户点击“报表展示”按钮系统展示预设好的各类统计数据图</p> <ol style="list-style-type: none"> 1. 论文总数折线图 2. 被引用论文数最多和论文数最多作者TOP5的堆叠柱状图（包含作者被引用数最多的前5篇论文的名字、引用数和PDF link） 3. 被引用数最多的论文TOP10的环状图（包含其标题、作者、引用数、PDF link , year） 4. 年度热门方向（词及热度）的词云
扩展流程	<p>2a. 用户点击某项统计数据的具体部分</p> <ol style="list-style-type: none"> 1. 系统展示具体数据
特殊需求	无

4.5 模块4-画像展示

4.5.1 功能1-领域画像

用例编号	UC4-1
名称	领域画像
创建者	濮宗悦
创建日期	2020-03-13
最后一次更新者	濮宗悦
更新日期	2020-03-23
参与者	普通用户
触发条件	用户点击某个领域词条。
前置条件	无
后置条件	无
优先级	高
正常流程	<p>系统跳转至该领域词条页面，展示：</p> <ol style="list-style-type: none"> 1. 领域的基本信息：名称、活跃度、发表论文数、引用数、热度、H-index 2. 历年论文数、被引用数、H-index、活跃度、热度（总被引用数/发表论文总数）指数统计图。 3. 与此领域相关的领域（词云形式展示）。 <ol style="list-style-type: none"> 3.1 包含词与相应的活跃度。 3.2 词云中元素的大小为活跃度。 4. top作者、top论文和top机构。 <ol style="list-style-type: none"> 4.1 作者需包含其姓名、活跃度、论文数和被引数目。按照活跃度排序 4.2 论文需包含其标题、被引用数。按照被引用数排序 4.3 机构需包含其名称、活跃度、论文数和被引用数。按照活跃度排序。
扩展流程	<ol style="list-style-type: none"> 1.1 活跃度：paper数的和（以3年为周期，每向前3年paper数除以Math.ceiling（今年-当年年份/3））
特殊需求	无

4.5.2 功能2-学者画像

用例编号	UC4-2
名称	学者画像
创建者	王雪君
创建日期	2020-03-13
最后一次更新者	濮宗悦
更新日期	2020-03-23
参与者	普通用户
触发条件	用户选择了某个学者
前置条件	无
后置条件	无
优先级	高
正常流程	<p>系统跳转至该学者主页，展示：</p> <ol style="list-style-type: none"> 1. 学者基本信息：姓名、所属机构，活跃度、发表论文数、引用数、热度、H-index，领域。 2. 历年论文数、被引用数、H-index、活跃度、热度（总被引用数/发表论文总数）指数统计图。 3. 学者之间合作关系 <ol style="list-style-type: none"> 3.1 关系图，按照活跃度表示点的大小。 3.2 图中节点只包含和当前学者的一度关系 4. top论文 <ol style="list-style-type: none"> 4.1 按照引用数排序。 5. 主要学术领域饼图 <ol style="list-style-type: none"> 5.1 展示该会议主要涉及领域。
扩展流程	见 4.5.1扩展功能
特殊需求	无

4.5.3 功能3-机构画像

用例编号	UC4-3
名称	机构画像
创建者	丁玲燕
创建日期	2020-03-07
最后一次更新者	濮宗悦
更新日期	2020-03-23
参与者	普通用户
触发条件	用户选择了某个研究机构
前置条件	无
后置条件	无
优先级	高
正常流程	<p>系统跳转至该研究机构主页，展示：</p> <ol style="list-style-type: none"> 1. 研究机构基本信息：名称、现有研究人员数、活跃度、发表论文数、引用数、热度、H-index 2. 历年论文数、被引用数、H-index、活跃度、热度（总被引用数/发表论文总数）指数统计图 3. 主要学术领域饼图 <ol style="list-style-type: none"> 3.1 展示该机构主要涉及领域。 4. 研究机构关系图，合作关系密切的研究机构。 <ol style="list-style-type: none"> 4.1 关系图，按照活跃度表示点的大小。 4.2 图中节点只包含和当前机构的一度关系 5. top作者和top论文 <ol style="list-style-type: none"> 5.1 作者需包含其姓名、活跃度、论文数和被引数目。按照活跃度排序 5.2 论文需包含其标题、被引用数。按照被引用数排序
扩展流程	见 4.5.1扩展功能
特殊需求	1. 作者名称和研究方向可以点击进入详情画像

4.5.4 功能4-会议画像

用例编号	UC4-4
名称	会议画像
创建者	丁玲燕
创建日期	2020-03-07
最后一次更新者	丁玲燕
更新日期	2020-03-13
参与者	普通用户
触发条件	用户选择了某个会议
前置条件	无
后置条件	无
优先级	高
正常流程	<p>系统跳转至该会议主页，展示：</p> <ol style="list-style-type: none"> 1. 会议基本信息：名称、活跃度、发表论文数、引用数、热度、H-index 2. 历年论文数、被引用数、H-index、活跃度、热度（总被引用数/发表论文总数）指数统计图 3. 主要学术领域饼图 <ol style="list-style-type: none"> 3.1 展示该会议主要涉及领域。 4. 会议中发表的论文列表。 <ol style="list-style-type: none"> 4.2 论文需包含其标题、被引用数。按照被引用数排序
扩展流程	无
特殊需求	无

4.6 模块5-实体排序

4.6.1 功能1-领域实体排序

用例编号	UC5-1
名称	领域实体排序
创建者	濮宗悦
创建日期	2020-03-13
最后一次更新者	濮宗悦
更新日期	2020-03-13
参与者	普通用户
触发条件	用户进入领域排序界面。
前置条件	无
后置条件	无
优先级	高
正常流程	1. 领域排序默认方法为活跃度，可以改为论文总数、总被引数、研究人员数、热度。 2. 在本页可以按照名称搜索领域。
扩展流程	无
特殊需求	无

4.6.2 功能2-学者实体排序

用例编号	UC5-2
名称	学者实体排序
创建者	王雪君
创建日期	2020-03-13
最后一次更新者	王雪君
更新日期	2020-03-13
参与者	普通用户
触发条件	用户进入学者排序界面。
前置条件	无
后置条件	无
优先级	高
正常流程	<ol style="list-style-type: none"> 1. 学者排序默认方法为活跃度，可以改为paper总数、总被引数、热度。 2. 学者可以按照方向进行筛选。 3. 在本页可以按照名称搜索学者。
扩展流程	无
特殊需求	无

4.6.3 功能3-机构实体排序

用例编号	UC5-3
名称	机构实体排序
创建者	丁玲燕
创建日期	2020-03-13
最后一次更新者	丁玲燕
更新日期	2020-03-13
参与者	普通用户
触发条件	用户进入机构排序界面。
前置条件	无
后置条件	无
优先级	高
正常流程	1.机构排序默认方法为活跃度，可以改为论文总数、总被引数、研究人员数、热度 2. 在本页可以按照名称搜索领域
扩展流程	无
特殊需求	无

4.6.4 功能4-会议实体排序

用例编号	UC5-4
名称	会议实体排序
创建者	丁玲燕
创建日期	2020-03-13
最后一次更新者	丁玲燕
更新日期	2020-03-13
参与者	普通用户
触发条件	用户进入会议排序界面。
前置条件	无
后置条件	无
优先级	高
正常流程	1.会议排序默认方法为活跃度，可以改为论文总数、总被引数、研究人员数、热度 2. 在本页可以按照名称搜索会议
扩展流程	无
特殊需求	无

5 接口需求

5.1 系统接口

- 本系统依赖IEEE获取最基本的源数据

5.2 其它接口

- 系统需要能够及时响应外界的扩展变化

6 非功能性需求

6.1 系统性能需求

- 输入完网址后，进入应用系统的时间不应超过3~5秒。
- 输入搜索内容后，从提交到相应得到搜索结果的平均时间不应超过10s，平均时间应在2~3秒。
- 在用户第一次使用本系统时，在使用搜索、切换关键词和二次搜索功能时，寻找和理解的时间不应超过3秒。

- 在预计的高峰负载条件下（最大并发用户数为500），10%处理器能力和15%系统可用内存留出备用，偶然情况下处理器满负荷（100%）运行时间不超过30秒。
- 系统在更新或者重大维护时，不可用时间不超过十分钟

6.2 数据管理能力需求

1. 系统记录icse和ase近5年从IEEE官网导出的论文简要数据

6.3 软硬件环境需求

硬件环境	阿里云学生服务器： CPU：单核 内存：2G 硬盘：40GB
软件环境	后台操作系统Ubuntu（前台采用浏览器，对操作系统没有特别要求，支持用户使用Windows2003、windows7+、MacOS、Linux发行版等主流操作系统） 编程工具：java+ docker +vue全家桶+v-chart 数据库：MySQL
浏览器及分辨率	浏览器：支持IE9+、chrome50+、Firefox45+及其兼容浏览器 分辨率：1024×768,1366×768,1280×800,1280×1024，1440×900,1600×900，1920×1080

6.4 可靠性需求

- Reliability1：系统每年非计划宕机时间不能超过3小时
- Reliability2：系统要能够采用适当的方法检测并发现故障
- Reliability3：系统出现故障时，能够在30分钟内恢复并正常使用

6.5 安全性需求

- Security1：系统要确保服务器上的数据一致性
- Security2：使用加密的数据传输协议，保证传输中数据不丢失、不失真、不被窃取、不被篡改
- Security3：数据需要存储在相对安全的硬件设备上，并对存储的数据进行定期备份

6.6 可移植性需求

前端无

6.7 兼容性需求

- Compatibility1：系统能够在各大主流浏览器上使用，包括IE、Firefox、Google Chrome、Safari、Opera

6.8 扩展性需求

Scalability1: 系统可以自由进行副本数据的添加

Scalability1: 论文的添加、数据删改需要支持灵活的变更

6.9 易用性需求

Usability1: 系统布局要合理, 方便用户找到所需功能, 同时设置初次使用引导功能方便用户上手操作。Usability2: 系统显示各种论文的信息, 或者用户、账户的信息时, 行距要为1倍或以上行距, 方便看清各种信息。Usability3: 系统要对重点信息(论文题目、概要等)突出显示, 方便用户确认。Usability4: 系统应针对不同的终端屏幕尺寸展示出不同界面, 提高用户的体验度。

7 附录

1. 产品采用响应式界面设计, 保证在不同分辨率浏览器上运行
2. 主要功能采用基于WebKit的WebView技术, 也就是浏览器技术, 能够在不同配置的设备上运行, 具有比较好的可移植性