



REUTERS/Pawel Kopczynski

Thomson Data Analyzer (TDA) 的高效利用

张丹丹

产品与解决方案部

汤森路透



THOMSON REUTERS

TDA简要流程概览

1.数据导入

- | |
|----------------|
| 1.1 结构化数据都可导入 |
| 1.2 添加/融合字段 |
| 1.3 数据集/记录融合 |
| 1.4 重复项（去重/合并） |

2.数据清理

- | |
|----------|
| 2.1 建立词表 |
| 2.2 修改词表 |
| 2.3 使用词表 |

3.数据分析

- | |
|---------------------------------|
| 3.1 各种脚本，比如线图、饼图、柱状、aduna、聚类圆群等 |
| 3.2 技术报告、公司报告、公司比较等脚本 |
| 3.3 矩阵 |
| 3.4 图谱 |
| 3.5 词频逆文档(TF-IDF) |

4.数据导出

- | |
|-------------|
| 4.1 导出字段化记录 |
| 4.2 导出原始记录 |

TDA简要流程概览

1.数据导入

- | |
|----------------|
| 1.1 多种格式数据导入 |
| 1.2 添加/融合字段 |
| 1.3 数据集/记录融合 |
| 1.4 重复项（去重/合并） |

2.数据清理

- | |
|----------|
| 2.1 建立词表 |
| 2.2 修改词表 |
| 2.3 使用词表 |

3.数据分析

- | |
|--------------------------------|
| 3.1各种脚本，比如线图、饼图、柱状、aduna、聚类圆群等 |
| 3.2 技术报告、公司报告、公司比较等脚本 |
| 3.3矩阵 |
| 3.4 图谱 |
| 3.5 词频逆文档(TF-IDF) |

4.数据导出

- | |
|-------------|
| 4.1 导出字段化记录 |
| 4.2 导出原始记录 |



TDA简要流程概览

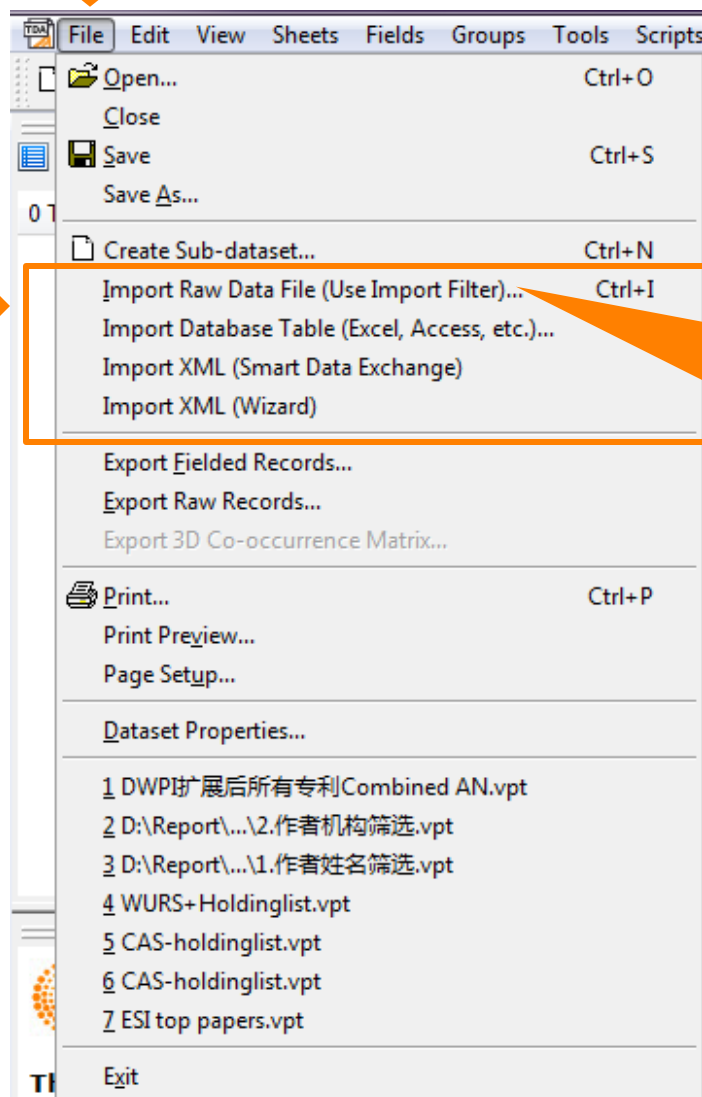
1.数据导入

1.1 结构化数据均可导入

1.2 添加/融合字段

1.3 数据集/记录融合

1.4 重复项（去重/合并）



Raw data需要过滤器，目前TDA已有的过滤器是WOS/TI/Inspec，如果想导入其他平台的数据到TDA中，需要利用Tools->Import Filter Editor

TDA简要流程概览

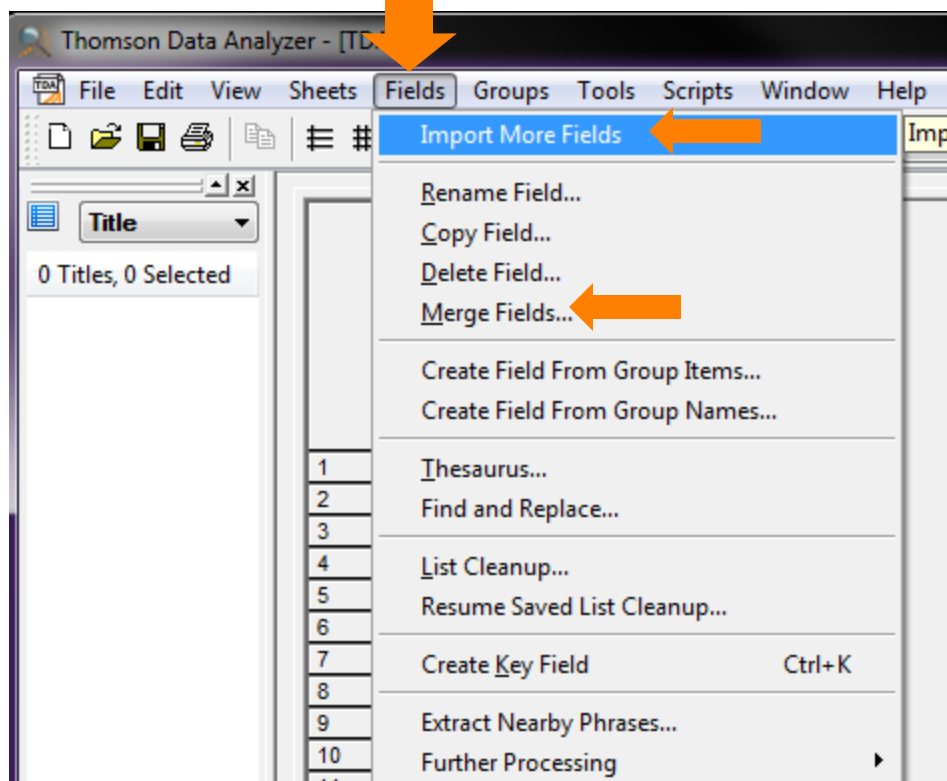
1.数据导入

1.1 结构化数据均可导入

1.2 添加/融合字段

1.3 数据集/记录融合

1.4 重复项（去重/合并）



THOMSON REUTERS

TDA简要流程概览

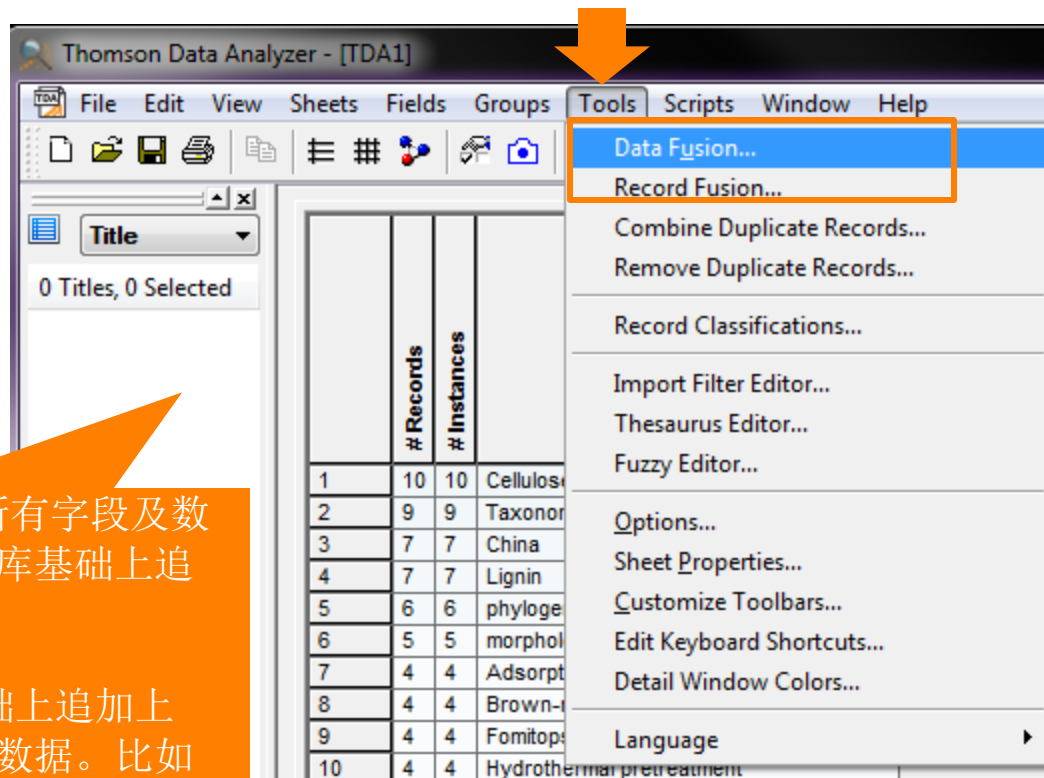
1.数据导入

1.1 结构化数据均可导入

1.2 添加/融合字段

1.3 数据集/记录融合

1.4 重复项（去重/合并）



Data Fusion: 是两个数据集中所有字段及数据的融合，比如在历史机构知识库基础上追加上新年度的数据；

Record Fusion: 是在主集合基础上追加另外一个文档中的某几个字段及数据。比如在SCI机构知识库中加入期刊影响因子信息

TDA简要流程概览

1.数据导入

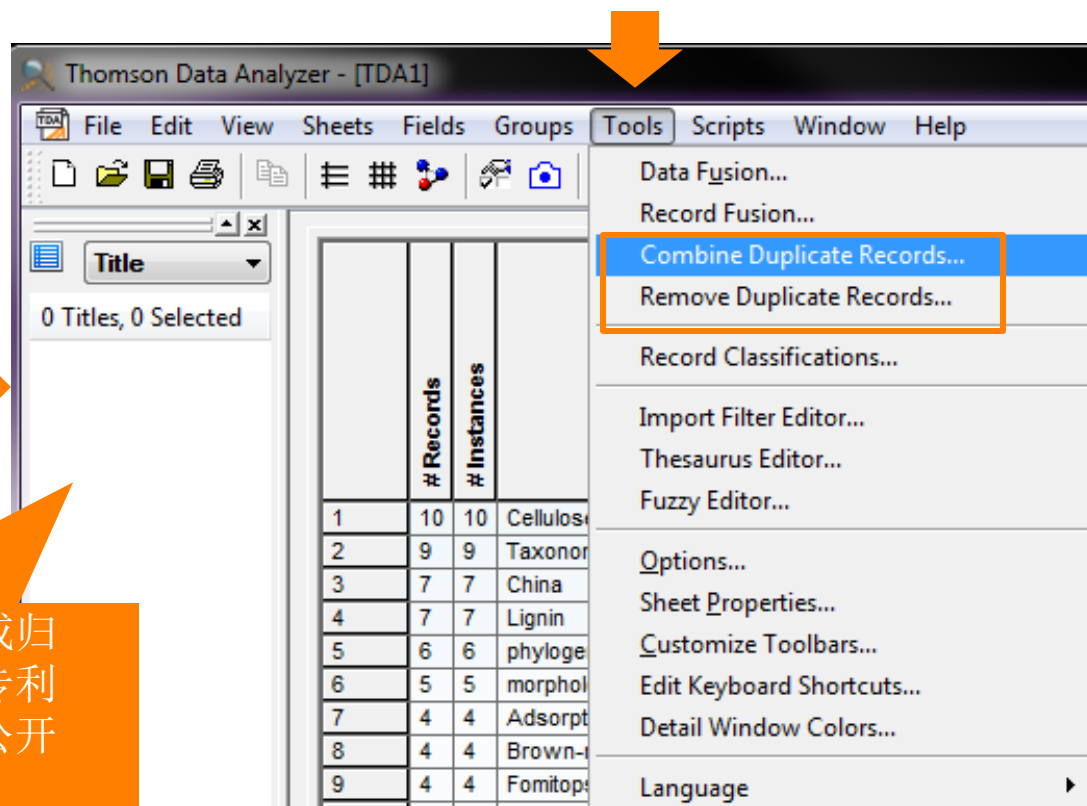
1.1 结构化数据均可导入

1.2 添加/融合字段

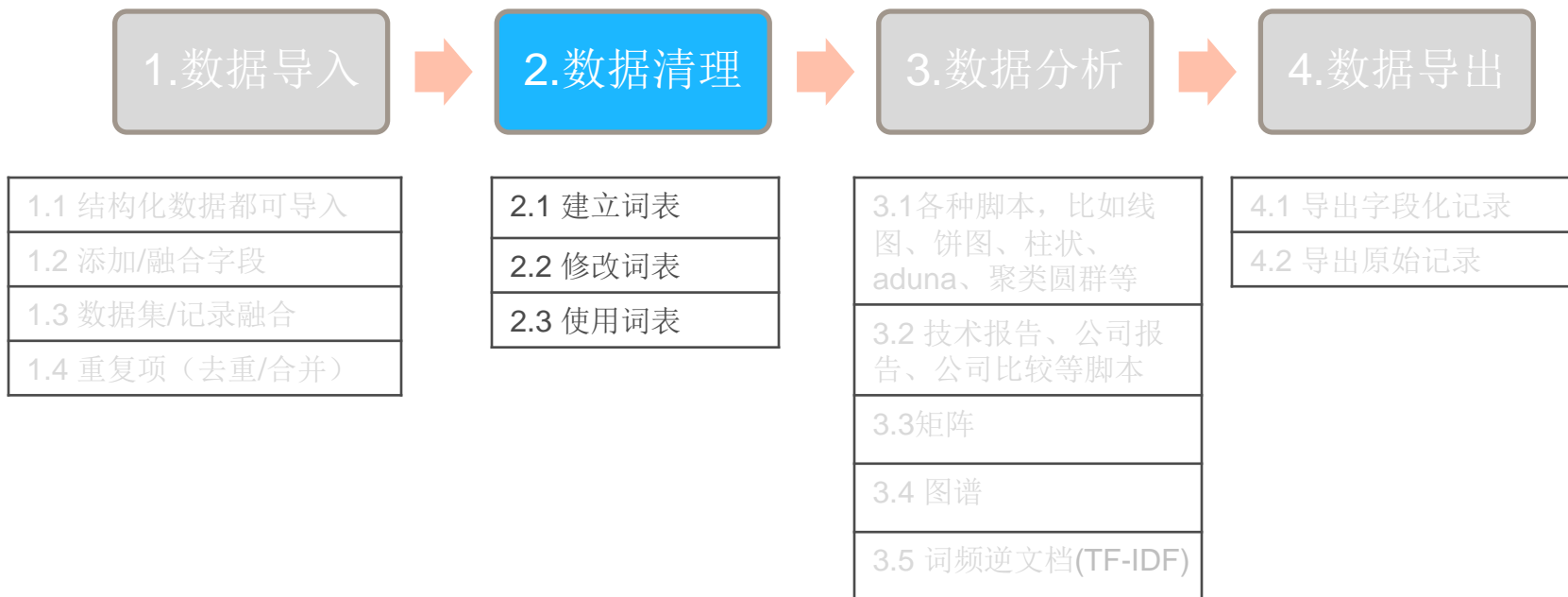
1.3 数据集/记录融合

1.4 重复项（去重/合并）

涉及到归并去重时可利用去重或归并功能。比如专利数据，一件专利会对应多个文本：申请文本、公开文本、授权文本等。



TDA简要流程概览



TDA简要流程概览

2. 数据清理

2.1 建立词表

2.2 修改词表

2.3 使用词表

两种方法:

其一，自动清洗: **Fields->list cleanup**，比如发明人、专利权人、文本字段

其二，查找分组进行标引:

比如，如何提炼成立本校各学院**SCI**知识库，以北京林业大学为例。大致思路:

通常会利用**author affiliation (full)** 字段，在其中搜索各学院然后添加到不同的组，但因为**author affiliation (full)** 中还有其他机构，这给搜索学院带来噪声。所以，最好首先挑出北京林业大学的项，单独添加到组后，再单独将这些项成立一条新字段

(**Fields->Create field From Group Items**)。之后把各学院变体放到不同组后，可利用组成立叙词表 (**Group->Create Thesaurus Using Groups**)

动画研究领域主要国家/地区的研究主要方向

| Reset | | Countries (Group Names) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | | |
|-----------------------------------------|-----------|--------------------------------------------------------------------------------------------------------------------------------------|-----------|-----------------|------------|---------------|-------|-------------------|-------------|--------|---------------------|--------------------|-------------------------|-----|-------------------|--------|----------------|--------------|----------------|-------|------------|----------------|-------|---------------------|-----------|---------|-----------|----|---|
| | | # Records | 31 | 37 | 13 | 11 | 11 | 11 | 32 | 20 | 7 | 7 | 66 | 14 | 8 | 52 | 10 | 13 | 93 | 8 | 26 | 23 | 10 | 15 | 10 | 35 | 20 | | |
| Keywords (author's) + Keywords Plus (1) | # Records | <div><div>▼▲</div><div>Show Values >= 1 and <= 118</div><div>Cooccurrence # of Records</div><div><div>▼▲</div></div></div> | animation | virtual reality | simulation | 3D Technology | model | COMPUTER GRAPHICS | deformation | fluids | galaxies: evolution | DIGITAL SKY SURVEY | Japan animation & manga | GPU | VISUAL SIMULATION | motion | Computer games | optimization | multitieracies | SMOKE | algorithms | motion capture | Water | collision detection | real-time | reality | knowledge | | |
| | | 1 | 1031 | USA | 10 | 11 | 44 | 34 | 39 | 41 | 15 | 5 | 4 | 5 | 11 | 5 | 2 | 19 | 34 | 3 | 36 | 1 | 11 | 6 | 4 | 5 | 2 | 13 | 8 |
| | | 2 | 381 | UK | 36 | 35 | 9 | 12 | 13 | 8 | 2 | 1 | 4 | 5 | 9 | 1 | | 7 | 20 | 1 | 3 | | 6 | 3 | 2 | | | 1 | 4 |
| | | 3 | 174 | CANADA | 20 | 24 | 5 | 2 | 9 | 9 | 3 | 1 | 3 | 4 | 4 | 2 | | 3 | 5 | | 15 | 1 | | | | 1 | 2 | 2 | 3 |
| | | 4 | 172 | GERMANY | 18 | 28 | 8 | 13 | 10 | 4 | 3 | 3 | 4 | 4 | 1 | 1 | | 3 | 6 | | 2 | | | 1 | | 2 | 1 | 1 | |
| | | 5 | 160 | AUSTRALIA | 11 | 13 | 5 | 2 | 7 | 2 | | | 3 | 3 | 8 | 1 | 2 | | 5 | | 18 | | | | | | | | 2 |
| | | 6 | 152 | JAPAN | 19 | 12 | 2 | 6 | 5 | 8 | 1 | 2 | 4 | 5 | 10 | 2 | 1 | 4 | 2 | 3 | 2 | 1 | | 2 | | | | | 1 |
| | | 7 | 144 | CHINA | 33 | 22 | 15 | 14 | 14 | 8 | 6 | 6 | 6 | 6 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 3 |



动画研究领域主要国家/地区的研究主要方向

| USA | UK | CANADA | GERMANY | AUSTRALIA | JAPAN | CHINA |
|---------|-------|---------|---------|-----------|--------|--------|
| 虚拟现实 | 动画 | 虚拟现实 | 虚拟现实 | 多元化读写能力 | 动画 | 动画 |
| 动画 | 虚拟现实 | 动画 | 动画 | 虚拟现实 | 虚拟现实 | 虚拟现实 |
| 模拟 | 电脑游戏 | 多元化读写能力 | 3D技术 | 动画 | 日本动画 | 模拟 |
| 计算机图形学 | 观众 | 模型 | 模型 | 教育 | “恐怖谷” | 3D技术 |
| 模型 | 性能/表现 | 教育 | 增强现实 | 读写能力 | 计算机图形学 | 模型 |
| 多元化读写能力 | 模型 | 计算机图形学 | “恐怖谷” | 日本动画 | 日本 | 计算机图形学 |
| 3D技术 | 环境 | 读写能力 | 模拟 | 模型 | 系统 | 变形 |
| 电脑游戏 | 3D技术 | 多峰性 | 系统 | 教育学 | 3D技术 | 流体 |
| 性能/表现 | 感知 | 观众 | 虚拟环境 | 媒体 | 感知 | 数字化 |
| 环境 | “恐怖谷” | 性能/表现 | 面部 | 电脑游戏 | 模型 | 日本动画 |

TDA简要流程概览

2.数据清理

2.1 建立词表

2.2 修改词表

2.3 使用词表

对已保存的词表进行修改，有三种方法：

其一，在词表编辑器中打开词表进行修改（Tools->Thesaurus Editor）；

其二，直接在TDA系统文件中，在thesaurus文件夹中直接打开之前保存的词表进行编辑。

第一种方法适合少量修改；第二种方法适合批量修改

其三，对现有的excel表加工编写成词表（Scripts->Utility-Make Thesaurus）。

TDA简要流程概览

2. 数据清理

2.1 建立词表

2.2 修改词表

2.3 使用词表

如何利用之前建好的叙词表？

两种方式：

其一，直接替换字段下各项（Fields->Thesaurus）；

其二，以组的形式呈现（Groups->Group Using Thesaurus）

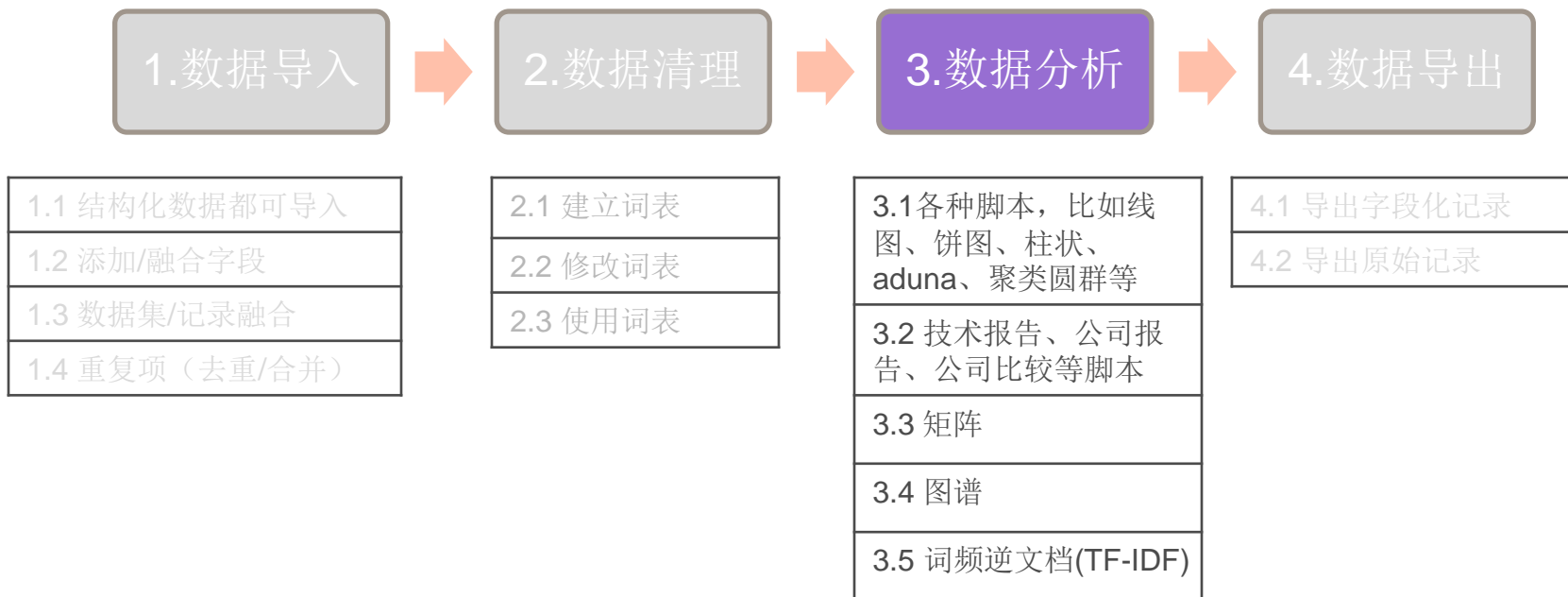
需要注意：

一，词表在用之前一定要确保匹配规则（精确匹配、前方一致、后方一致还是包含）；

二，待匹配的字段必须要滤除噪声。比如如果拿北京林业大学只包含各学院的变体词表去清理author affiliation（full）极可能命中的物理学院有其他学校的物理学院。



TDA简要流程概览



TDA简要流程概览

3.数据分析

3.1 各种脚本，比如线图、饼图、柱状、aduna、聚类圆群等

3.2 技术报告、公司报告、公司比较等脚本

3.3 矩阵

3.4 图谱

3.5 词频逆文档(TF-IDF)

可以解决哪些问题：

- ①与excel兼容，直接生成柱状图、线图、饼图、线图、泡泡图等；
- ②看合作（个人/机构等）
- ③直接生成技术报告、公司报告、公司比较报告；
- ④专利健康报告：patent vital signs

矩阵和图谱有共现、自相关、互相关

- ① 共现矩阵：比如看各学院历年的发文量，他们在最近几年的活跃度；各专利权人专利的全球布局
- ② 自相关：看已有专利权人之间的合作；
- ③ 互相关：看哪些人正在做着相似的事；

TDA简要流程概览

3.数据分析

3.1 各种脚本，比如线图、饼图、柱状、aduna、聚类圆群等

3.2 技术报告、公司报告、公司比较等脚本

3.3 矩阵

3.4 图谱

3.5 词频逆文档(TF-IDF)



词频逆文档 TF-IDF

TF: Term Frequency;

IDF: Inverse Document Frequency

用统计学语言表达，就是在词频的基础上，要对每个词分配一个"重要性"权重。最常见的词（"的"、"是"、"在"）给予最小的权重，较常见的词（"中国"）给予较小的权重，较少见的词（"蜜蜂"、"养殖"）给予较大的权重。这个权重叫做"逆文档频率"（Inverse Document Frequency，缩写为IDF），它的大小与一个词的常见程度成反比。



TDA简要流程概览

3.数据分析

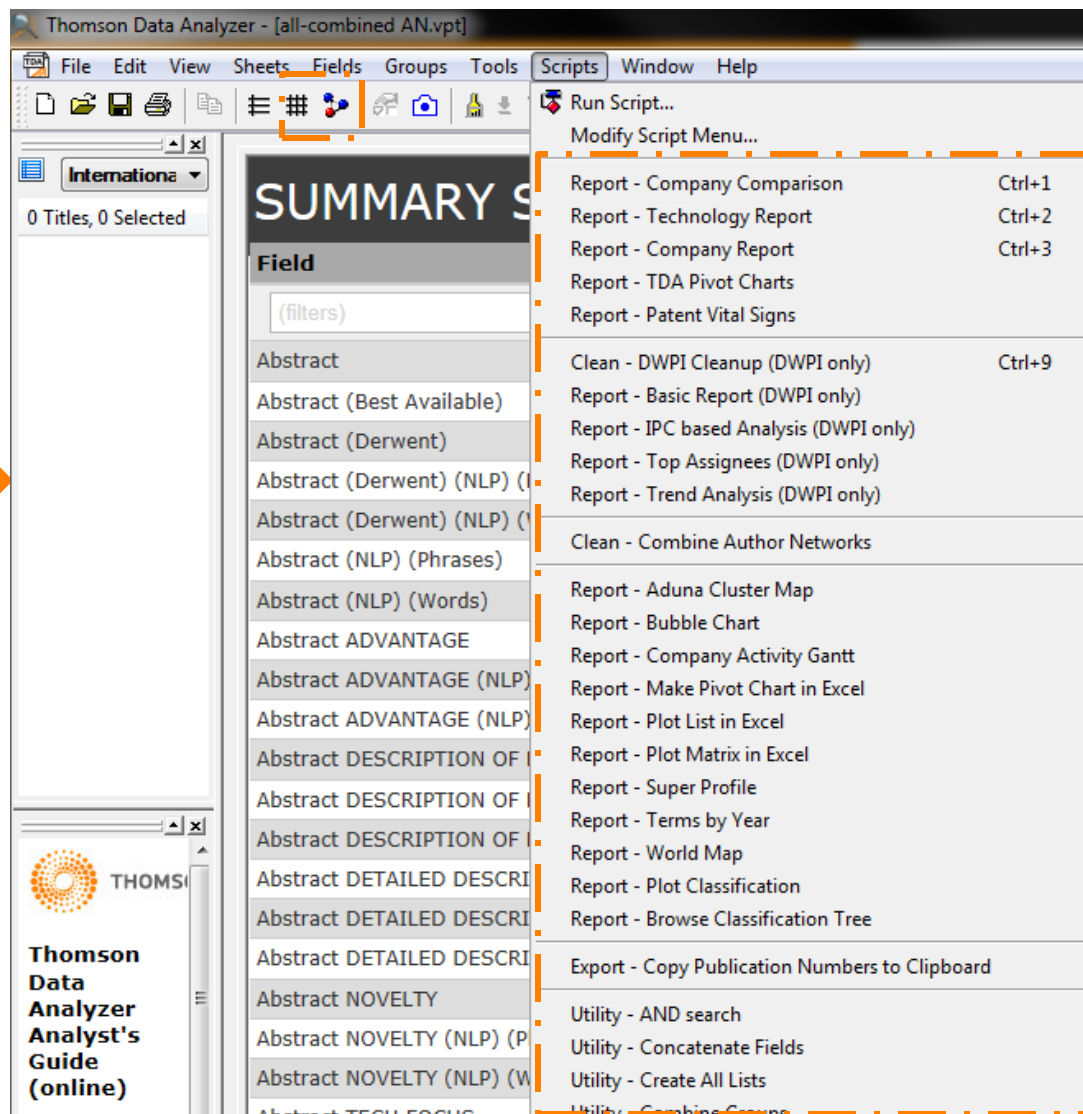
3.1 各种脚本，比如线图、饼图、柱状、aduna、聚类圆群等

3.2 技术报告、公司报告、公司比较等脚本

3.3 矩阵

3.4 图谱

3.5 词频逆文档



Technology Report:

TDA Tech Report.xls

| | A | B | C | D | E | F |
|----|---|-----------------------------|---|---|---|---|
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | Summary Information | | | | |
| 4 | | | | | | |
| 5 | | Year Chart | | | | |
| 6 | | | | | | |
| 7 | | Organization Chart | | | | |
| 8 | | | | | | |
| 9 | | Country Chart | | | | |
| 10 | | | | | | |
| 11 | | Technology Chart | | | | |
| 12 | | | | | | |
| 13 | | Technology v Year Chart | | | | |
| 14 | | | | | | |
| 15 | | Technology v Country Chart | | | | |
| 16 | | | | | | |
| 17 | | Country v Year Chart | | | | |
| 18 | | | | | | |
| 19 | | Organization Holdings Chart | | | | |
| 20 | | | | | | |
| 21 | | Organization Profile | | | | |
| 22 | | | | | | |
| 23 | | Technology Profile | | | | |
| 24 | | | | | | |
| 25 | | People Profile | | | | |
| 26 | | | | | | |
| 27 | | Country Profile | | | | |
| 28 | | | | | | |
| 29 | | Time Profile | | | | |
| 30 | | | | | | |
| 31 | | | | | | |
| 32 | | | | | | |
| 33 | | | | | | |
| 34 | | | | | | |
| 35 | | | | | | |
| 36 | | | | | | |
| 37 | | | | | | |
| 38 | | | | | | |

THOMSON REUTERS

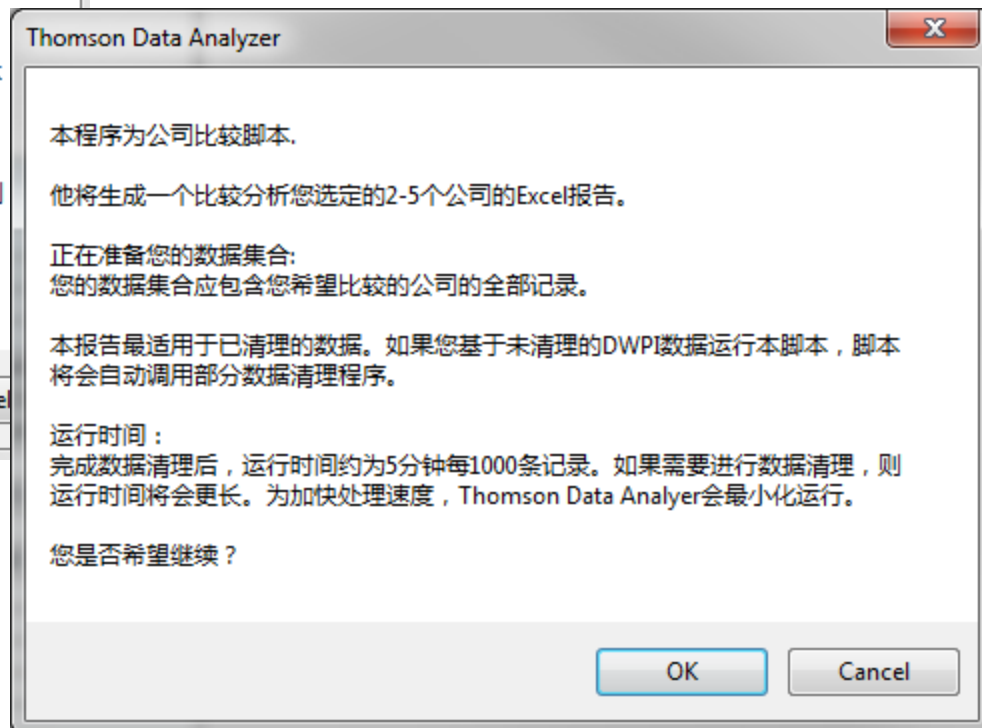
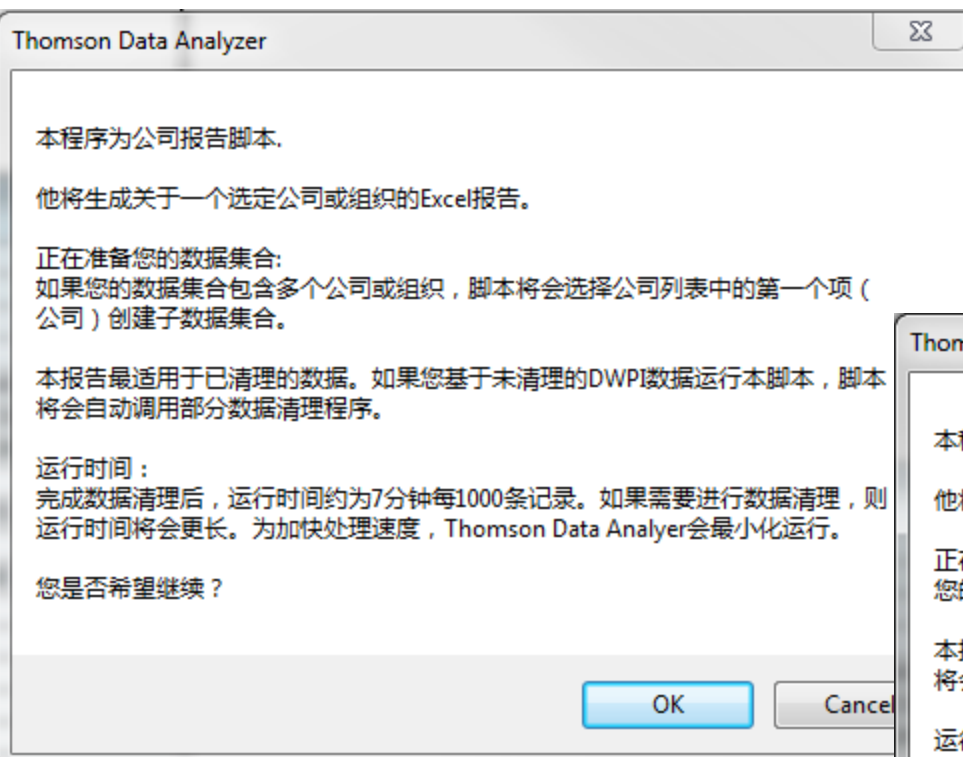
THOMSON DATA ANALYZER
Technology Report

Total Number of Records: 139
Years Range From: 2010 - 2015
Peak Year: 2013 [78 Records]
Date of report creation: 11/30/2015
Technology: Manual Codes

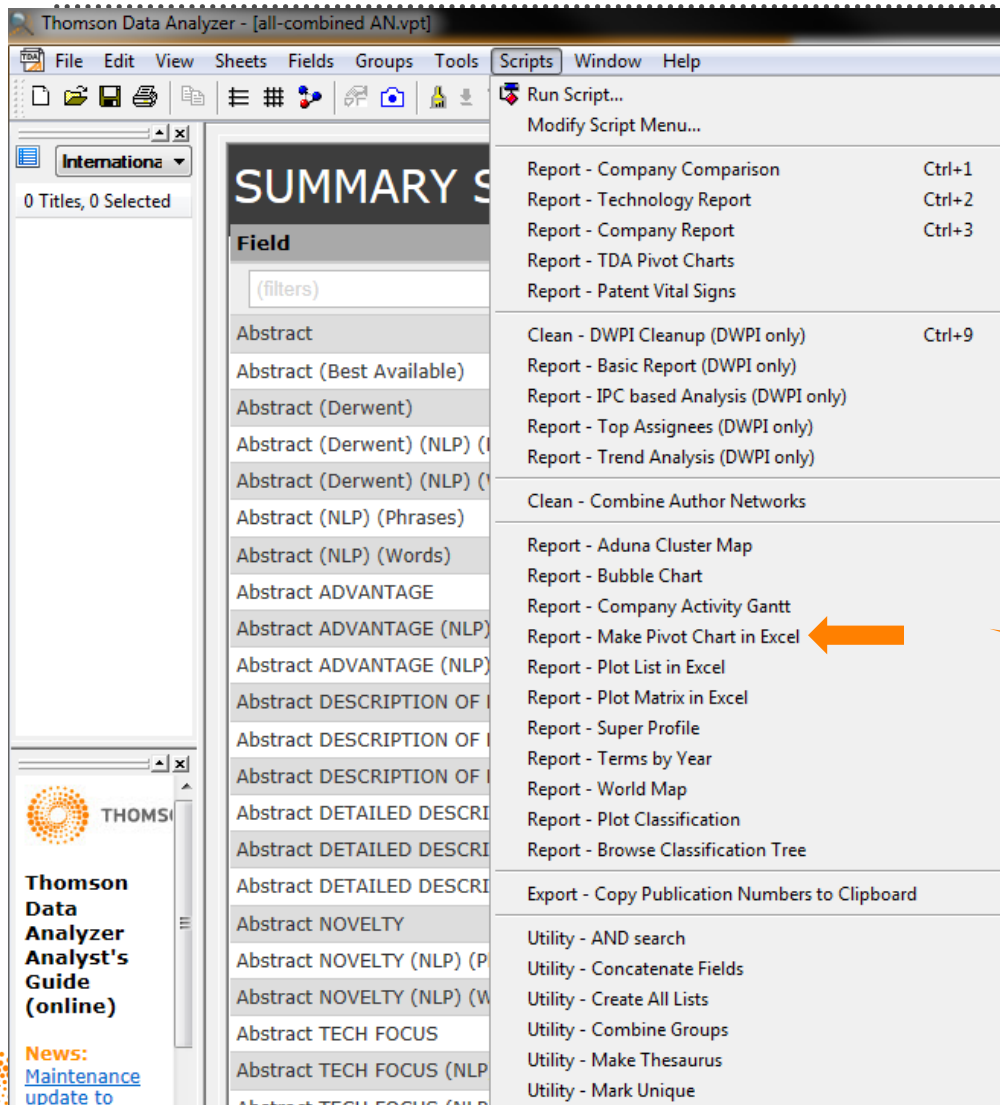
Introduction | Summary Info | Year Chart | Organization Chart | Country Chart | Technology Chart | Country v Year Chart | Technology v Year Chart | Technology v Country Chart



Company Report / Company Comparison



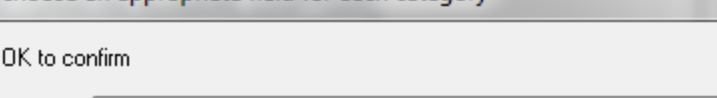
Make Pivot Chart in Excel



比如查看各国家市场每年中的主要专利权人

Make Pivot Chart in Excel

比如查看各国家市场每年中的主要专利权人



Please choose an appropriate field for each category

Click OK to confirm

Row Field: Application Countries

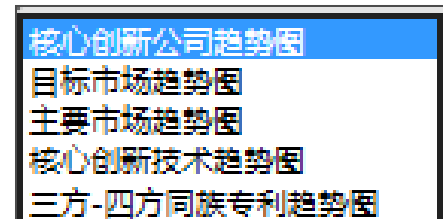
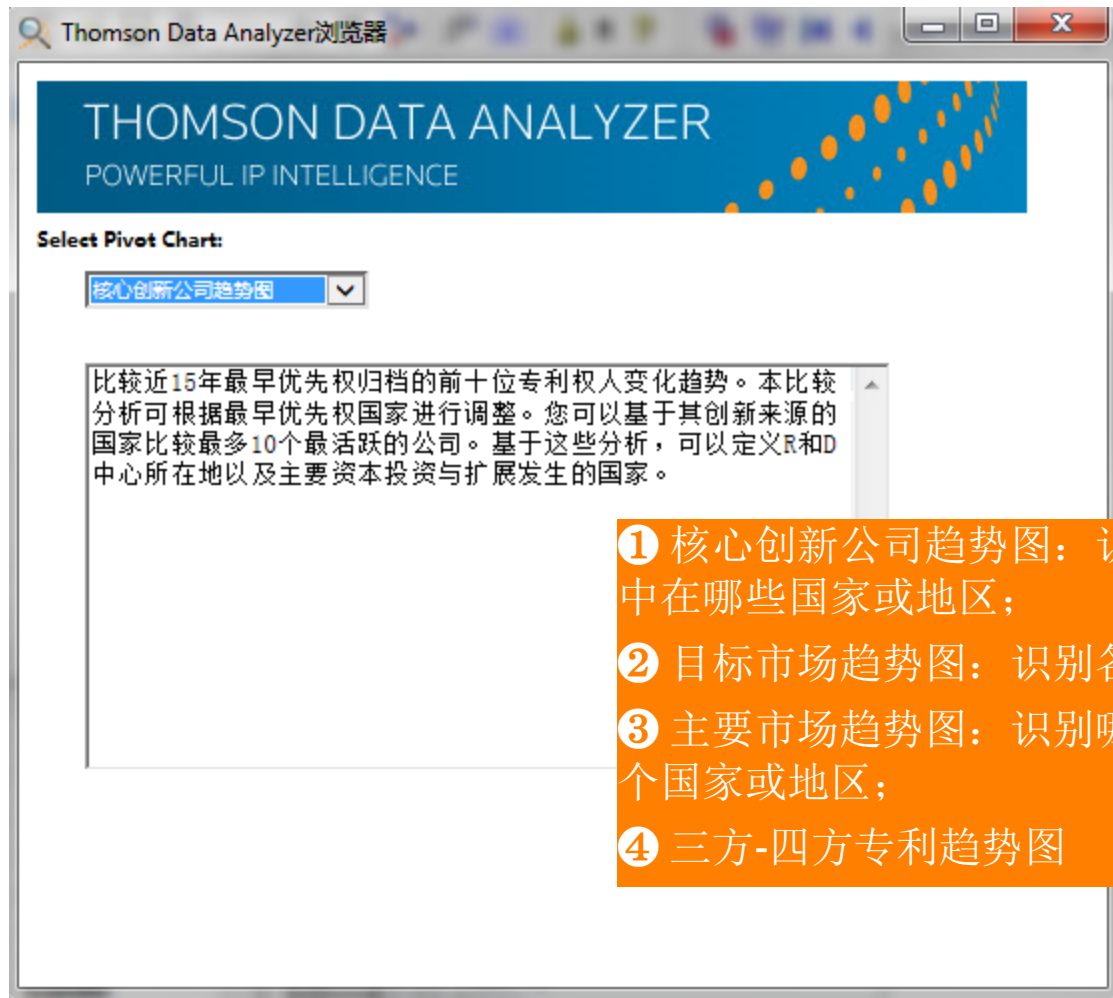
Column Field: Publication Year

Pivot Field: Patent Assignees (Cleaned)

Restore Defaults OK Cancel

| Application Countries | | Publication Year | Patent Assignees (Cleaned) | Number of Records |
|-----------------------|------|------------------------------------------|----------------------------|-------------------|
| CN | 2014 | AHN J Y | 1 | |
| CN | 2014 | ALCATEL LUCENT | 2 | |
| CN | 2014 | BEIJING INST TEC | 1 | |
| CN | 2014 | BROADCOM CORP | 1 | |
| CN | 2014 | ELECTRONICS & TELECOMMUNICATIONS | 1 | |
| CN | 2014 | NAM J Y | 1 | |
| CN | 2014 | NEC LAB AMERICA | 1 | |
| CN | 2014 | SAMSUNG ELECTRONICS CO LTD | 2 | |
| CN | 2014 | SHANGHAI LANG | 1 | |
| CN | 2014 | UNIV BEIJING POSTS & TELECOMMUNICATIONS | 2 | |
| CN | 2014 | UNIV BEIJING TECH | 1 | |
| CN | 2014 | UNIV BEIJING UNIV OF AERONAUTICS & SPACE | 2 | |
| CN | 2014 | UNIV CHONGQING COMMUNICAT | 1 | |

Pivot Charts



- ① 核心创新公司趋势图：识别各公司的研发中心及投资集中在哪些国家或地区；
- ② 目标市场趋势图：识别各公司的主要生产和销售区域；
- ③ 主要市场趋势图：识别哪些公司主要定位在中日美欧四个国家或地区；
- ④ 三方-四方专利趋势图



Super Profile

| A | B | C | D |
|--------------------------------------------|---------------------------------------------------------------|-----------------------------------------------------------|--------------------------------------|
| Patent Assignees (Best Available) TOP10 | Priority Years (earliest) Top 5 Items | Priority Years (earliest) Top Terms (minimum 5 record) | Application Years Year Line Chart |
| FERRO CORP[35] | 1992 [12]; 2012 [7]; 2002 [6]; 2000 [6]; 1993 [3] | 1992 [12]; 2012 [7]; 2002 [6]; 2000 [6] | |
| NGK SPARK PLUG CO LTD[29] | 2000 [16]; 1998 [7]; 2003 [4]; 1983 [1]; 1999 [1] | 2000 [16]; 1998 [7] | |

Patent vital signs

| | | Patent Vital Signs | | | | | APPLICATIONS | | | | | | | | | |
|----|---|--------------------|-------------|----------------|--------------------------------|----------------------------------------|-------------------------------------------------|------------------------------------------------|-----------------------------------------------|------------------------------------|-----------------------------------------------|-------------------------------------|---------------------------------------------|----------------------------------|---------------------------------------------|-----------------------------------|
| | | # Records | # Instances | Number | Legal Status (most recent +/-) | Legal Status (year of most recent +/-) | Years Remaining (Earliest Priority Year + 20 -) | Years Remaining (Publication Year + 10 - 2015) | GRANTS | | | | | | | |
| | | | | | | | | | Grant - probably active with positive (+) LLS | Grant - probably active no +/- LLS | Grant - probably active with negative (-) LLS | Grant - probably expired (20 years) | App - probably active with positive (+) LLS | App - probably active no +/- LLS | App - probably active with negative (-) LLS | App - probably expired (10 years) |
| 1 | 1 | 1 | 1 | CN104601297A | | | | 10 | | | | | | | | |
| 2 | 1 | 1 | 1 | KR2015073126A | | | | 10 | | | | | | | | |
| 3 | 1 | 1 | 1 | CN104753628A | | | | 10 | | | | | | | | |
| 4 | 1 | 1 | 1 | CN104717730A | | | | 10 | | | | | | | | |
| 5 | 1 | 1 | 1 | CN104703189A | | | | 10 | | | | | | | | |
| 6 | 1 | 1 | 1 | CN104702324A | | | | 10 | | | | | | | | |
| 7 | 1 | 1 | 1 | CN104683074A | | | | 10 | | | | | | | | |
| 8 | 1 | 1 | 1 | CN104678350A | | | | 10 | | | | | | | | |
| 9 | 1 | 1 | 1 | CN104644164A | | | | 10 | | | | | | | | |
| 10 | 1 | 1 | 1 | CN104618921A | | | | 10 | | | | | | | | |
| 11 | 1 | 1 | 1 | CN104618080A | | | | 10 | | | | | | | | |
| 12 | 1 | 1 | 1 | CN104617996A | | | | 10 | | | | | | | | |
| 13 | 1 | 1 | 1 | CN104601312A | | | | 10 | | | | | | | | |
| 14 | 1 | 1 | 1 | CN104579439A | + | 2015 | | 10 | | | | | | | | |
| 15 | 1 | 1 | 1 | CN104579443A | + | 2015 | | 10 | | | | | | | | |
| 16 | 1 | 1 | 1 | CN104539335A | | | | 10 | | | | | | | | |
| 17 | 1 | 1 | 1 | WO2015096423A1 | | | | 10 | | | | | | | | |
| 18 | 1 | 1 | 1 | WO2015095844A1 | | | | 10 | | | | | | | | |

List::DPCI Citing Patent Numbers

List::DPCI Cited Patent Numbers

Export-Copy Publication Numbers to clipboard

Thomson Data Analyzer - [all-combined AN.vpt]

File Edit View Sheets Fields Groups Tools Scripts Window Help

Application

17 Titles, 0 Sel...

CN
CN
EP
EP
JP
KR
KR
TW
TW
US

| | # Records | # Instances | Inventors (Cleaned) | Derwent | Potential Experts | test |
|----|-----------|-------------|---------------------|---------|-------------------|------|
| 1 | 20 | 25 | Marzetta Thomas L. | | | |
| 2 | 18 | 18 | Kim Youn Sun | | | |
| 3 | 18 | 19 | Lee Ju Ho | | | |
| 4 | 17 | 20 | Ashikhmin Alexei | | | |
| 5 | 17 | 17 | Lee Hyo Ji | | | |
| 6 | 14 | 14 | CHO Young | | | |
| 7 | 13 | 13 | L Louay | | | |
| 8 | 13 | | | | | |
| 9 | | | SHER Amin | | | |
| | | | K Sam | | | |
| 16 | 8 | 9 | Ahn Jae Young | | | |
| 17 | 8 | 8 | LEE Kilbom | | | |
| 18 | 8 | 9 | Nam Jun Young | | | |
| 19 | 8 | 8 | OKUMURA Yukihiro | | | |
| 20 | 8 | 8 | SUYAMA Satoshi | | | |
| 21 | 8 | 10 | Yang Hong | | | |

Thomson Data Analyzer Analyst's Guide (online)

News

①选中需要导出publication no.的项(比如 Ashikhmin Alexei), ②然后, Scripts-> Export-Copy Publication Numbers to Clipboard

Scripts Window Help

Run Script...
Modify Script Menu...

Report - Company Comparison Ctrl+1
Report - Technology Report Ctrl+2
Report - Company Report Ctrl+3
Report - TDA Pivot Charts
Report - Patent Vital Signs

Clean - DWPI Cleanup (DWPI only) Ctrl+9
Report - Basic Report (DWPI only)
Report - IPC based Analysis (DWPI only)
Report - Top Assignees (DWPI only)
Report - Trend Analysis (DWPI only)

Clean - Combine Author Networks

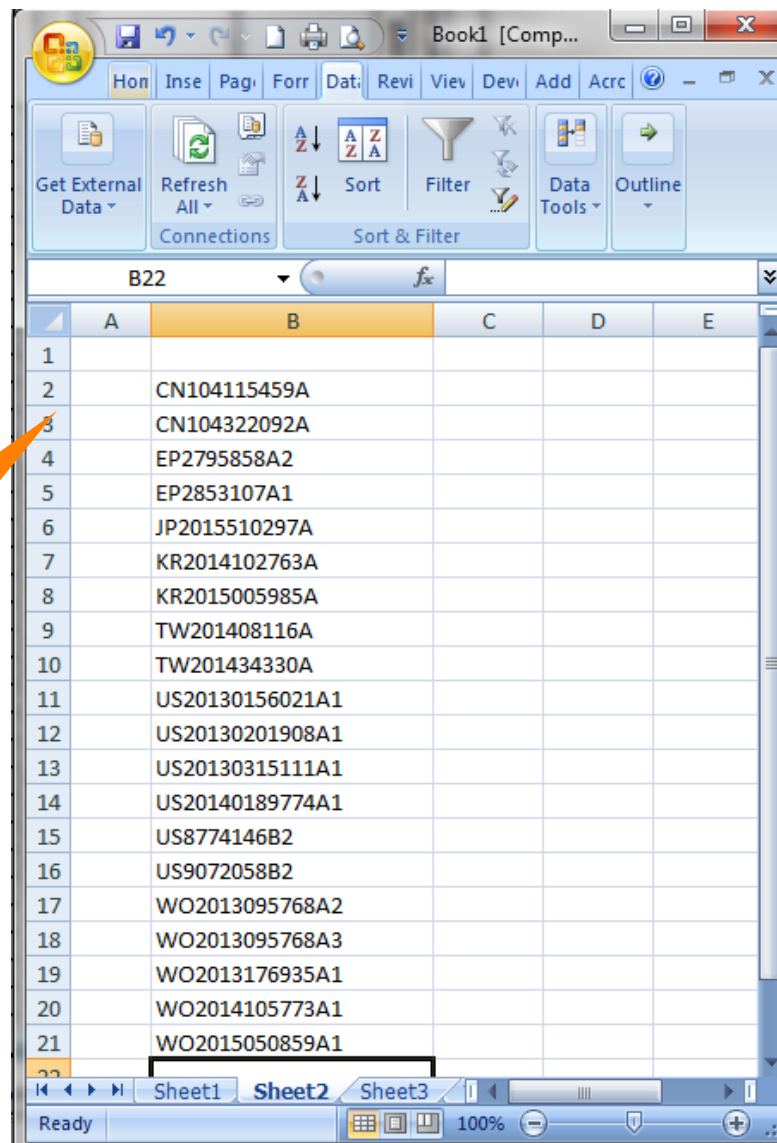
Report - Aduna Cluster Map
Report - Bubble Chart
Report - Company Activity Gantt
Report - Make Pivot Chart in Excel
Report - Plot List in Excel
Report - Plot Matrix in Excel
Report - Super Profile
Report - Terms by Year
Report - World Map
Report - Plot Classification
Report - Browse Classification Tree

Export - Copy Publication Numbers to Clipboard

Utility - AND search
Utility - Concatenate Fields
Utility - Create All Lists
Utility - Combine Groups
Utility - Make Thesaurus

Export-Copy Publication Numbers to clipboard

③ 复制到excel中



Book1 [Comp...]

Home Insert Page Layout Formulas Data Review View Developer Add-Ins Account

Get External Data Refresh All Connections Sort & Filter Sort Filter Data Tools Outline

B22

| | A | B | C | D | E |
|----|---|-----------------|---|---|---|
| 1 | | | | | |
| 2 | | CN104115459A | | | |
| 3 | | CN104322092A | | | |
| 4 | | EP2795858A2 | | | |
| 5 | | EP2853107A1 | | | |
| 6 | | JP2015510297A | | | |
| 7 | | KR2014102763A | | | |
| 8 | | KR2015005985A | | | |
| 9 | | TW201408116A | | | |
| 10 | | TW201434330A | | | |
| 11 | | US20130156021A1 | | | |
| 12 | | US20130201908A1 | | | |
| 13 | | US20130315111A1 | | | |
| 14 | | US20140189774A1 | | | |
| 15 | | US8774146B2 | | | |
| 16 | | US9072058B2 | | | |
| 17 | | WO2013095768A2 | | | |
| 18 | | WO2013095768A3 | | | |
| 19 | | WO2013176935A1 | | | |
| 20 | | WO2014105773A1 | | | |
| 21 | | WO2015050859A1 | | | |
| 22 | | | | | |

Sheet1 Sheet2 Sheet3

Ready 100%

Utility-combine Groups

The screenshot shows the Thomson Data Analyzer interface. The main window displays a list of records with columns for #Records, #Instances, Title (Best Available), test 1, test B, and combine group test. The third record is highlighted, showing a title 'Method for mitigating interference due to p' and checkboxes for test 1 and test B. An orange callout box points to the 'combine group test' column, indicating that all groups from a specific field are added to a new group using Utility-Combine Groups. The third group is noted as a fusion of test1 and testB.

| | #Records | #Instances | Title (Best Available) | test 1 | test B | combine group test |
|----|----------|------------|-----------------------------------------------|-------------------------------------|--------------------------|--------------------------|
| 1 | 6 | 8 | Method for mitigating interference due to p | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2 | 6 | 6 | Method for serving cell wireless network | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 3 | 6 | 6 | Method for transmitting channel state infor | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4 | 5 | 5 | Multiple-input-multiple-output transmission | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5 | 4 | 4 | Method for beamforming transmitter in lon | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 6 | 4 | 4 | Method for feeding back channel state inf | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 7 | 3 | 4 | Method for detecting downlink physical ch | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 8 | 3 | 3 | Method for performing fractional beamfor | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 9 | 3 | 3 | Method for supporting service-antenna in | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 10 | 3 | 3 | Method for transmitting pilot reference sig | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 11 | 3 | 4 | Multi-input multi-output (MIMO) transmissio | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 12 | 3 | 3 | Transmitter for use in multi-antenna wirele | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 13 | 4 | 4 | Method for implementing hierarchical chan | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 14 | 3 | 3 | Base station method for communicating wi | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 15 | 3 | 3 | Channel transmission method for base sta | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 16 | 3 | 4 | Dual-tier wireless communication system, | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 17 | 2 | 2 | Channel state information reporting metho | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 18 | 2 | 2 | CSI feedback method, involves determinin | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 19 | 2 | 2 | Method for allocating uplink pilot in base st | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 20 | 2 | 2 | Method for beam coordination between ba | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

将某个字段中的所有组全部添加到新组中，利用Utility-Combine Groups。比如图中第三组是对test1和testB的融合

Utility-concatenate Fields

- Report - World Map
- Report - Plot Classification
- Report - Browse Classification Tree
- Export - Copy Publication Numbers to Clipboard
- Utility - AND search
- Utility - Concatenate Fields**
- Utility - Create All Lists
- Utility - Combine Groups
- Utility - Make Thesaurus
- Utility - Mark Unique

Concatenate two fields
The second field can

| | # Records | # Instances | International Classifications 8 |
|---|-----------|-------------|---------------------------------|
| 1 | 198 | 410 | H04 |
| 2 | 3 | 3 | G01 |
| 3 | 3 | 5 | H01 |
| 4 | 1 | 5 | A61 |

| | # Records | # Instances | International Classifications 8 |
|---|-----------|-------------|---------------------------------|
| 1 | 199 | 415 | H |
| 2 | 3 | 3 | G |
| 3 | 1 | 5 | A |

| | # Records | # Instances | International Classification |
|---|-----------|-------------|------------------------------|
| 1 | 195 | 195 | H04testH |
| 2 | 3 | 3 | H01testH |
| 3 | 2 | 2 | G01testG |
| 4 | 1 | 1 | A61testA |
| 5 | 1 | 1 | A61testG |
| 6 | 1 | 1 | A61testH |
| 7 | 1 | 1 | NONEtestNONE |

将两个字段项融合，利用
Utility-concatenate Fields

矩阵

生成矩阵

行

- Abstract
- Abstract (Best Available)
- Abstract (Derwent)
- Abstract (Derwent) (NLP) (Phrases)
- Abstract (Derwent) (NLP) (Words)
- Abstract (NLP) (Phrases)
- Abstract (NLP) (Words)
- Abstract ADVANTAGE
- Abstract ADVANTAGE (NLP) (Phrases)
- Abstract ADVANTAGE (NLP) (Words)
- Abstract DESCRIPTION OF DRAWINGS
- Abstract DESCRIPTION OF DRAWINGS (NLP) (Phrases)
- Abstract DESCRIPTION OF DRAWINGS (NLP) (Words)
- Abstract DETAILED DESCRIPTION

列

- Abstract
- Abstract (Best Available)
- Abstract (Derwent)
- Abstract (Derwent) (NLP) (Phrases)
- Abstract (Derwent) (NLP) (Words)
- Abstract (NLP) (Phrases)
- Abstract (NLP) (Words)
- Abstract ADVANTAGE
- Abstract ADVANTAGE (NLP) (Phrases)
- Abstract ADVANTAGE (NLP) (Words)
- Abstract DESCRIPTION OF DRAWINGS
- Abstract DESCRIPTION OF DRAWINGS (NLP) (Phrases)
- Abstract DESCRIPTION OF DRAWINGS (NLP) (Words)
- Abstract DETAILED DESCRIPTION

矩阵：共现（基于记录数量）
190 行：Abstract --全部项
0 列：无字段选定

矩阵类型

- ☒ 共现
- ☐ 自相关
- ☐ 交互相关

基于

- ☒ 记录数目
- ☐ 出现频次

范围

- ☒ 记录
- ☐ 源项

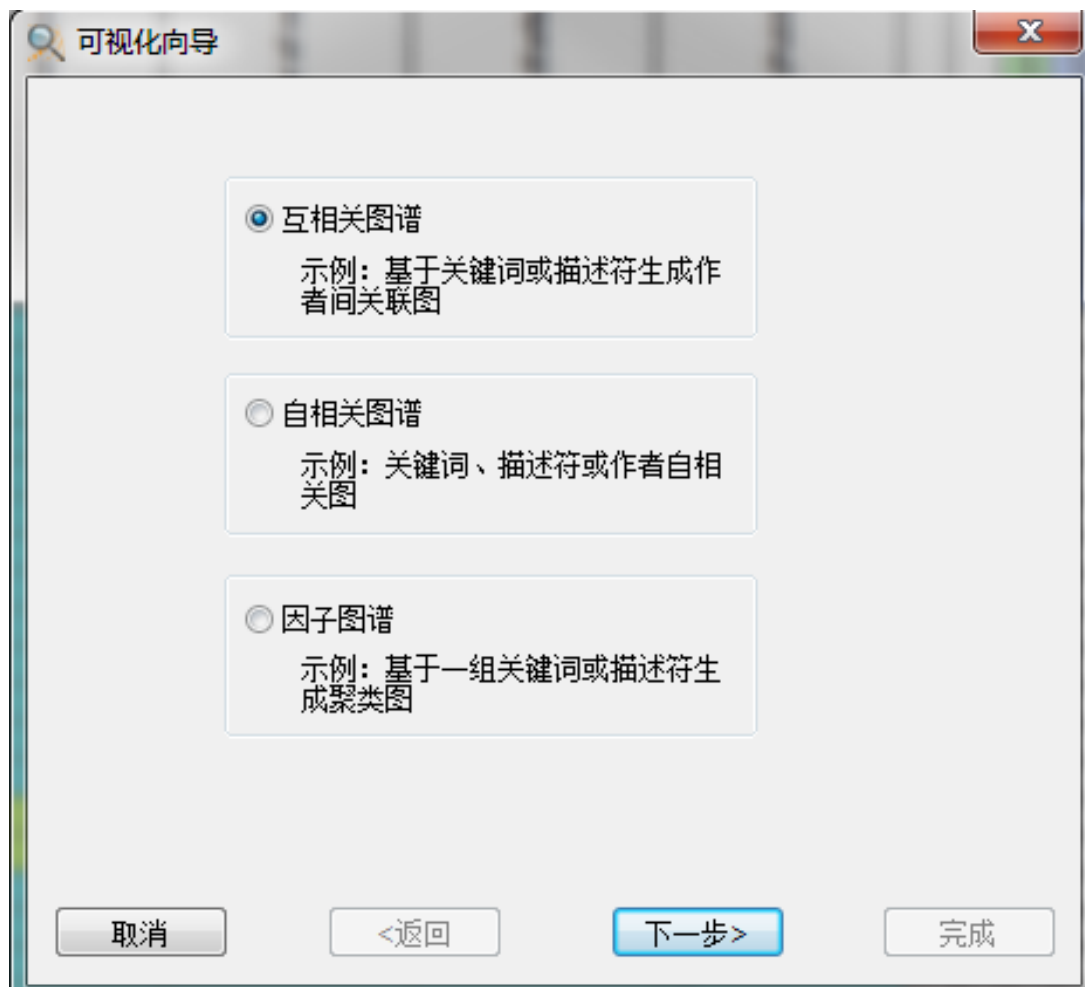
相关函数

- ☒ 皮尔逊系数
- ☐ 余弦
- ☐ 最大比例

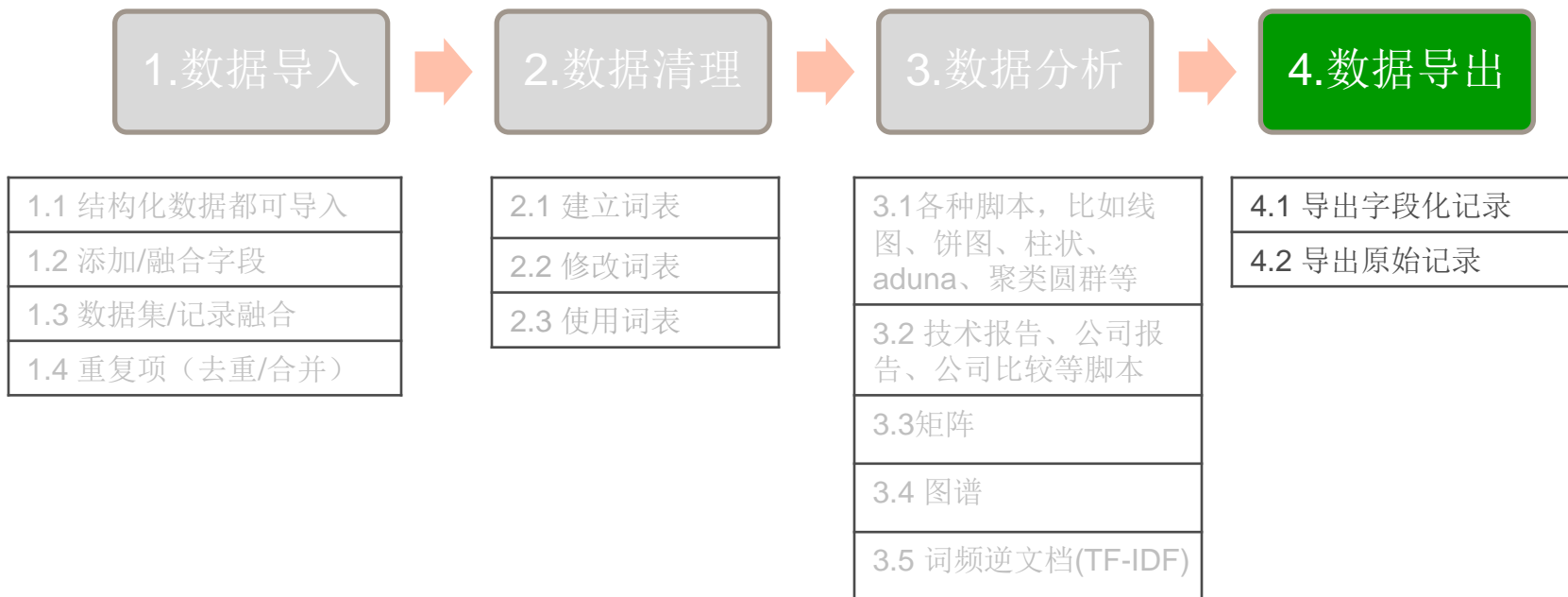
帮助 确定 取消



图谱:



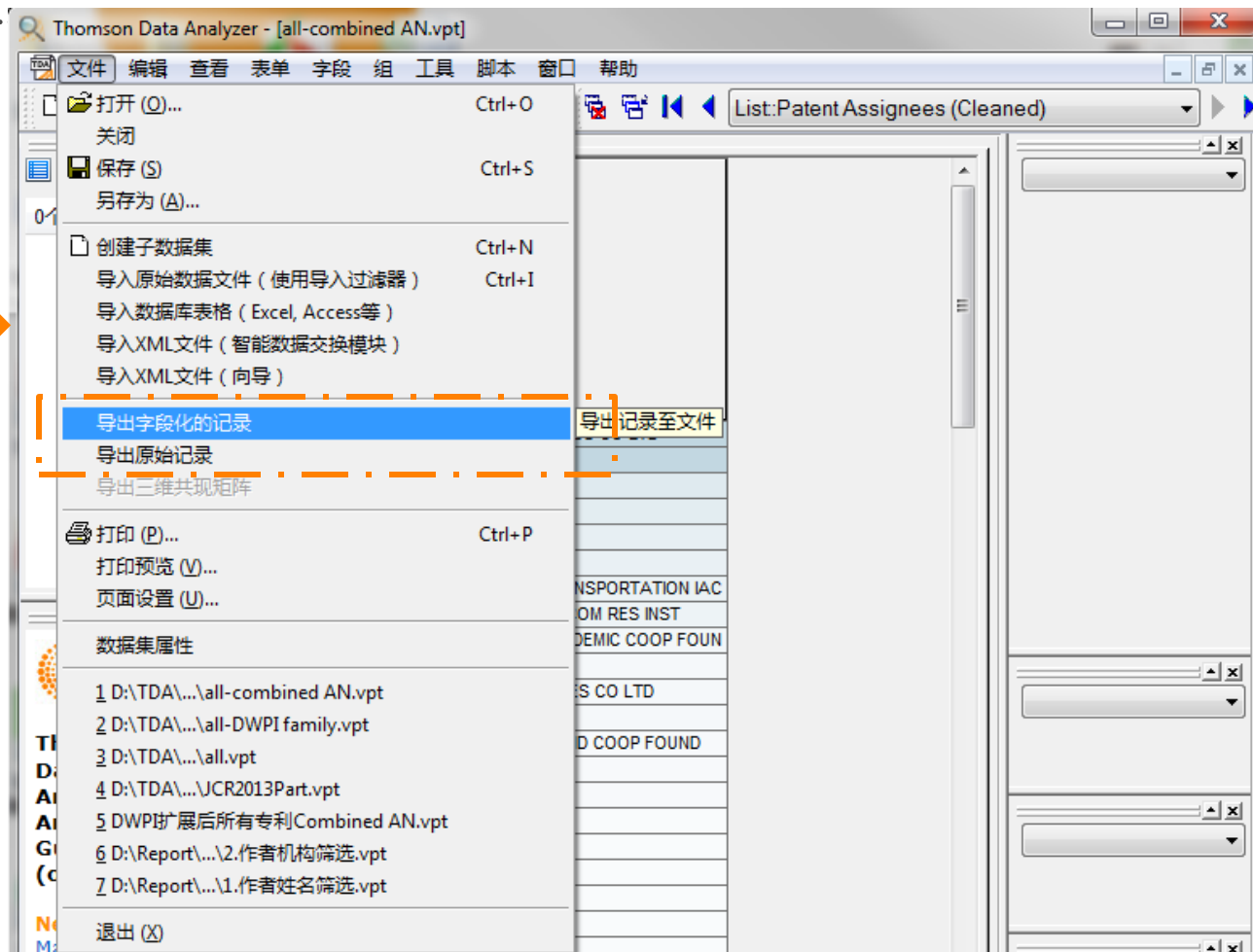
TDA简要流程概览



4.数据导出

4.1 导出字段化记录

4.2 导出原始记录



THOMSON REUTERS

TDA简要流程概览

1.数据导入

- | |
|----------------|
| 1.1 结构化数据都可导入 |
| 1.2 添加/融合字段 |
| 1.3 数据集/记录融合 |
| 1.4 重复项（去重/合并） |

2.数据清理

- | |
|----------|
| 2.1 建立词表 |
| 2.2 修改词表 |
| 2.3 使用词表 |

3.数据分析

- | |
|--------------------------------|
| 3.1各种脚本，比如线图、饼图、柱状、aduna、聚类圆群等 |
| 3.2 技术报告、公司报告、公司比较等脚本 |
| 3.3矩阵 |
| 3.4 图谱 |
| 3.5 词频逆文档(TF-IDF) |

4.数据导出

- | |
|-------------|
| 4.1 导出字段化记录 |
| 4.2 导出原始记录 |





REUTERS/Pawel Kopczynski

联系我们！

技术支持: ts.support.china@thomsonreuters.com

Tel: 4008 822 031(工作时间：周一至周五, 9:00—17:00)

Fax: 010-82862088

北京市海淀区科学院南路2号 融科资讯中心C座北楼610室
汤森路透知识产权与科技集团



THOMSON REUTERS