

欢迎使用 **Derwent Data Analyzer**

Derwent Data Analyzer

提供强大的知识产权情报信息。在这里，我们可以从各种丰富的信息来源（包括 **Derwent Innovation**）获取现成的数据进行集中分析。您可以查看竞争对手的技术研发趋势、相互之间的关系以及相关技术分布全景。此外，还有多种与数据准备和自动化相关的高级工具可用于快速生成优质报告。

数据分析简介

本章节将快速引导您了解 Derwent Data Analyzer 的主要功能。

1. [导入一些数据](#)
2. [检查您的原始数据](#)
3. [清理您的数据](#)
4. [分析您的数据](#)
5. [创建一份报告](#)

步骤 1：导入一些数据

Derwent Data Analyzer


是一种“离线”工具。它本身并不包含任何数据，也不检索任何数据库。因此，使用此工具的第一步是获取一些数据。

Derwent Data Analyzer 安装中包含一些您可以使用的示例数据（.vpt 格式文件）。或者，如果您检索一个受支持的数据库，则可以从导出您自己的数据，再将这些数据导入本应用程序中。

如需了解受支持数据库和导入的相关信息，请参阅帮助 (F1)。

要打开其中一个示例文件，请在“主页”标签上依次单击**打开按钮**、**浏览**（示例文件位于 C:\Program Files\Derwent Data Analyzer 文件夹中）。单击**确定**，此时将加载该文件。

步骤 2：检查您的原始数据

1. “总览”页面显示数据中 *Derwent Data Analyzer* 已经识别的所有字段以及关联项的数量
2. 右键单击任一字段，并选择**查看统计数据...**。此时将出现一个弹出窗口，其中显示了该字段的概览。您可以使用箭头按钮和**字段名**下拉菜单从总览中选择不同的字段。请查看弹出窗口中记录部分下标记有**覆盖率**的方框，其中以百分比形式显示了有多少记录含有该字段中的项。如果百分比较低，需保持警惕。单击**返回**回到总览页面
3. 双击任意字段名称将创建一个该字段中项目的排序列表。试试看！（可以是年份字段或专利权人字段）
4. 当出现列表时，请注意主工作空间的底部会出现一个标签。单击任意标签可前往查看该表单。“总览”标签将保持在您所创建的列表标签的左侧。单击此标签可回到“总览”页面。
5. 创建另一个列表并单击该列表中的任意一个项目，您将看到左侧会出现相关的标题。您可以在单击时按住 **Shift** 或 **Ctrl** 键组合选择任意项目。双击任何标题都会为您显示该记录。
6. 要关闭任何一个标签，请单击工具栏中的  按钮（注意：无法删除“总览”标签）。

不过您可能会注意到，列表没有显示出您要的内容。这可能是因为某些项是同一家公司，但拼写有误。

步骤 3：清理您的数据

1. 要开始清理字段中的数据，请在总览页面中右键单击该字段（例如“专利权人”）
2. 选择**列表清理...**
3. 此时将出现一个对话框，询问您是否选择“模糊匹配模块”：如果您的字段与公司有关，则选择**组织机构名称**；如果与人员有关，则选择**个人名称**；如果与其他任何内容有关，则选择“通用”。然后选择**使用**
4. 这可能要花费一到两分钟，且该程序可能会最小化以优化性能（如果发生这种情况，可在任务栏上单击该最小化的程序）
5. 将出现**数据清理确认**对话框，为您显示模糊匹配模块认为是相同的项。双击任意项可查看与其合并的内容。您可以从各组中拖放主题词，或通过右键单击进行重命名、移除等操作。
6. 单击项名标题可按字母顺序排序。选择**显示全部项**也将显示出不匹配的项。尝试使用拖放操作（或复制与粘贴）来创建主题词组。
7. 第一次操作可能需要一点时间。您可以中途保存进程，或将已完成的工作保存为一个叙词表，这样就无需再做一遍（都可从“清理”菜单选项中执行）
8. 叙词表的数量没有限制，您可以与同事共享叙词表、合并不同的叙词表以及向叙词表添加内容。它们特别适用于公司（例如企业重组和并购）以及区域使用，或将分类代码转换为文本
9. 无论是否已经保存，都单击“接受”
10. 现在，您已经创建了一个**新的字段**（原始信息保持不变），且排序列表会显示清理后的信息
11. 现在请在另一字段上尝试这一过程（例如“国家/地区”，您不会得到匹配项，但可使用“显示全部项”，然后拖放以创建包含洲的组，最后右键单击将这些组重命名为欧洲、亚洲等）

现在，我们的数据状况不错，接下来可以做更多的分析工作。

步骤 4：分析您的数据

1. 我们可以通过双击“总览页面”创建一个包含任意字段的列表
2. 要查看列表中任意项的更多信息，请转到“视图”/“添加细节窗口”。重复此步骤并打开第二个细节窗口。
3. 您将看到右侧出现两个窗口
4. 在每个窗口的下拉列表中选择不同的字段（例如，如果您的列表是针对专利权人，请从细节窗口中选取发明人和公开年份）
5. 单击列表中的各个项，查看其详细信息更新

6. 单击“细节窗口”中的**任意项**，此操作将在左侧标题列表中做出选择
7. 我们还可以通过共现矩阵比较任意两个字段
8. 单击功能区上**主页**标签下的**矩阵按钮**。
9. 现在您可以为行和列选择任意两个字段。请试试“**国家/地区**”和“**年份**”
10. 所得到的矩阵将显示每个交叉点处（单元）的记录数，以此来体现共现的强度
11. 您可以选择单个或多个单元，该交互式窗口依然会相应更新
12. 使用列或行标题进行排序
13. 还有其他分析工具可供使用；您可以比较任意两个列表以找到独有项或共有项，以及为任意字段中的项创建图谱
14. 很多公司的报告都使用微软 Office 产品，并经过许多分析步骤而得出。*Derwent Data Analyzer* 包含适用于多种常见场合的预配置报告，可在 Microsoft Excel 中快速生成复杂的报告。

步骤 5：为您的分析生成报告

1. 将数据导入 MS Office
产品非常简单！完全支持复制和粘贴操作。而且无论是列表还是矩阵，都可以自动捕获标题信息。
2. 使用“脚本”或“宏”可快速复制、粘贴及生成图表。
3. 创建一个列表，选择前 5 项
4. 单击“脚本”菜单，然后选择 PlotList.vpm
5. 稍作等待后图表即创建成功。相关数据同时也被复制，您现在即可如您所愿自由修改。
6. PlotMatrix 的工作原理与矩阵类似
7. 虽然这些脚本可以更加复杂，但现在我们仍使用“公司报告”脚本
8. 创建一个公司/组织（例如专利权人或机构）列表（可以根据您的意愿先对其进行清理）
9. 单击以选择主项，然后单击菜单栏中的“文件”/“创建子数据集”（或使用 Ctrl-N）
10. 单击对话框中的“确定”。此时将创建一个新的数据集，其中包含我们刚才所选定的公司记录
11. 选择菜单栏中的“脚本”/“DDA 公司报告”
12. 您将看到一条欢迎消息，单击“确定”
13. 现在请选择涉及 5 个类别的字段（可能已经选定）
14. 单击“确定”继续，此时系统会询问您是否确认，请单击“确定”继续
15. 脚本将开始运行，然后会创建一个大的 Excel 工作簿，继续处理关于公司信息组合的主要指标（趋势、主要的当事各方、组织机构关系等）。此过程将花费 1 到 15 分钟，具体取决于实际记录数

还有其他多种报告可供使用，更多详情请参见帮助文件或“**分析指导手册**”

自动化和脚本

Derwent Data Analyzer 可以运行 Visual Basic

脚本以自动执行重复性操作，从而加快工作流的速度。*Derwent Data Analyzer*

使用来自微软公司的 Visual Basic (Scripting Edition)。


您可以通过以下两种方式运行脚本：

- 从功能区中的**数据整理、分析和报告**标签中
- 使用“热键”（在编辑脚本菜单中设置）

数据整理标签包含以下选项：


图
标 描述


DWPI

 **清理工具**：清理“专利权人”和“发明人”字段，在“德温特分类”字段中添加文本信息，并为最近的同族专利年份和最早的优先权年份创建新字段。

分析标签包含以下选项：

图
标 描述


 **按年份查阅的主题词**：创建包含组的新字段，显示在哪个年份中出现了哪些主题词。


 **浏览分类树**：创建树状视图，帮助您浏览指定给选定文献的 IPC 码或 CPC 符号。


专利生命特征：创建新的列表，该列表显示数据集中专利的健康状况和有效性。这是一个比典型的专利“有效或无效”(dead or alive)更加详细的报告，基于复杂的专利法律状态数据得到。


报告标签包含以下选项：


图
标 描述

 **柱状图**：根据列表中的选定信息创建基本的垂直条形图。

 **条形图**：根据列表中的选定信息创建基本的水平条形图。

 **环状图**：根据列表中的选定信息创建基本的环状图。您可以选择饼图或开放圆环图（环状图表）。

 **线图**：根据列表中的选定信息创建基本的线图。

 **词云**：显示列表中的项在您数据集中的流程度视图。



气泡图：创建一个图表，用气泡表示所选两个字段中的数字。



甘特图：生成组织机构甘特图，显示每个公司活跃的年份范围。



矩阵查看器：在一个矩阵中显示各项之间关联强度的可视化呈现效果。



聚类图谱：创建一个动态图谱，显示列表中各个项之间的连线和重叠情况。



定制化报告表：创建一个交互式图形定制化报告表，作为 Derwent Data Analyzer 文件中的一个表单或作为一个 Excel 文件。



世界地图：最多显示两个世界地图（一个地图在另一个地图之上），以显示数据集中的专利是在哪些国家/地区公开的。



绘制分类图：创建可帮助浏览指定给选定文献的分类代码的视图。



透视表：在 Excel 中创建五个预定义透视表中的一个。



公司比较报告：创建一系列 Excel 图表和表格，其中包含对 2 到 5 家公司进行比较的指标。



技术报告：创建一系列 Excel 图表和表格，其中包含一个技术领域的各项指标。



公司报告：创建一系列 Excel 图表和表格，其中包含一个公司的各项指标。



导出公开号：将选定的公开号导出至剪贴板，用于粘贴到另一个应用程序。

主工作空间

主工作空间显示总览表以及您所创建的若干工作表视图（列表、矩阵、图谱）。

1. 您创建的每个视图都会作为一个单独的表单显示，按标签进行识别。单击标签可查看不同的表单。
2. 使用“总览”标签左侧的箭头可滚动浏览各个标签。
3. 您可以使用工作空间中的滚动条在单个表单内滚动。
4. 使用位于表单标签下方的数据集标签可在多个打开的数据集中进行选择。

收集和导入数据

如何收集数据

在收集数据进行分析之前，请先考虑以下问题。

1.您试图回答的是哪些问题？

与提问者合作，准确了解他们的要求。特别是要尝试理解以下两点：

- 他们试图支持什么样的决定？
- 哪些指标能帮助他们作出该决定？

2.应使用哪个数据库？

应使用科技文献数据库（例如，Web of Science、INSPEC），还是专利数据库（例如，搭配 DWPI 的 Derwent Innovation）？

- **科技文献数据库**可更好地阐述基础研究和早期研究。
- - 可用的数据库包括 Web of Science、INSPEC、Ei Compendex 和 MEDLINE 等，通过 Web of Science 或 Dialog 平台获取
- **专利数据库**不可避免地存在时间滞后，但因为专利权人在研发时产生了经济投入，因此会包含某些商业化数据；可能涉及应用研究或早期研发。
- - 有用的数据库包括德温特世界专利索引 (DWPI)、德温特专利引文索引 (DPCI) 和直接来自专利授予机构的数据
 - Derwent Innovation 通过便捷的支持 DDA 的导出功能，提供对所有这些数据库的轻松访问

3.如何设计一个检索以找到相关信息？

要考虑的一些因素：

- 您的问题是重点关注一家公司还是在几家公司之间进行比较？或者是针对某项特定技术？
- 如果您要分析一家公司，那么确保得到该公司及其子公司的所有记录是否是非常重要的一点？或者您是否掌握了使用各种单词/分类的技术？
- 即使包含某些干扰内容也要搜集全部记录是否比缺少重要记录相对而言更好一点？
- 考虑时间尺度 - 如果数据来自于有限的时间范围，那么趋势分析不一定准确。
- 您可以使用 Derwent Data Analyzer 合并多个检索结果集，在必要时也可以排除记录。

- Derwent Innovation
提供了大量工具来帮助您获得全面且有针对性的检索结果，例如，Smart Search
智能检索和检索结果仪表板。

从 Derwent Innovation 导出数据

Derwent Innovation 提供综合性专利信息，包括经过编辑增强的德温特世界专利索引 (DWPI) 信息、专利状态（有效或无效）的预测数据以及预计到期日期。*Derwent Innovation* 还提供来自 *Web of Science* 的科技文献数据。

从 Derwent Innovation 导出记录

1. 运行检索以查找所需的专利或文献记录
提示：您可以使用检索结果仪表盘等工具在 *Derwent Innovation* 中快速聚焦于您的结果，也可以使用 *Derwent Data Analyzer* 轻松整理您的数据并创建子数据集，以便重点分析相关内容
2. （可选）单击**获取菜单**并选择**获取 DWPI 同族专利或获取 INPADOC 同族专利**
注意：这会创建一个新的检索结果，其中包括初始检索结果中的记录的所有同族专利；新检索结果将自动按同族专利进行归并
3. 如果您打算对 DWPI 使用 *Derwent Data Analyzer* 报告生成脚本，请单击**工具**图标，选择**显示和排序选项**，将“归并方式”设置为 **DWPI 同族专利**，然后单击**确定**
提示：您可以选择在归并列表中将哪个文献显示为“首选”文献（父文献），例如，DWPI 基础记录
4. 单击**导出按钮**
提示：如果按同族专利归并记录，请单击结果上方的归并全部项按钮（加号），然后单击“**全选**”复选框，以便为每个同族专利导出一条记录（例如，仅导出 DWPI 基础记录）
5. 单击**格式下拉列表**并选择 **Derwent Data Analyzer**
6. （可选）调整要包含在导出中的字段，然后单击**创建开始导出**
7. 导出完成后，单击**当前可用**链接以下载您的导出内容
8. 从 .ZIP 文件中提取下载文件（如果需要），并打开导出的文件以将数据导入到 *Derwent Data Analyzer* 中

从 Web of Science 导出数据

要从 Web of Science 获得记录（Web of Science 核心合集、INSPEC、Derwent Innovations Index），请执行以下操作：

1. 添加选定记录到“标记结果列表”。

2. 选择所有您需要的字段。
3. 使用**保存到文件并以字段标记格式**选择输出
4. 保存生成的 .txt 文件。

Web of Science 有 500 条记录的导出限制。要创建包含多于 500 条记录的数据集，请分别对第 1-500 条、501-1000 条、1001-1500 条（以此类推）记录继续执行以上步骤。每次给文件取一个具有唯一性的名称。之后在导入时，选择所有这些文件，将它们一同导入到一个单独的 .vpt 文件中。

将数据导入 Derwent Data Analyzer

Derwent Data Analyzer 是一种离线工具，不会连接到任何外部数据库获取信息。您必须导入要在 *Derwent Data Analyzer* 中使用的所有数据。

在导入期间，*Derwent Data Analyzer* 将：

- 识别单个记录
- 将记录划分到字段
- 将字段划分到项

这项工作所需时间取决于字段的数目和类型以及数据集的规模。

创建要导入的源文件

根据不同的平台（例如，Dialog、Web of Science、Derwent Innovation）和数据库（例如，DWPI、INSPEC）而定，创建要导入的源文件的方式各不相同。请参见以下主题，了解如何创建要从主要的 Clarivate Analytics 数据库导入的文件：

- [Derwent Innovation](#)
- [Web of Science](#)

自动导入 Derwent Innovation 的导出内容

从 Derwent Innovation 导出的文件会自动打开 *Derwent Data Analyzer* 并运行导入过程。

了解更多：[自动导入 Derwent Innovation 的导出内容](#)

从其他源导入数据

要导入数据，您必须：

1. 选择要导入的文件
2. 选择源数据库和导入文件格式
3. 选择要导入哪些待分析的字段

使用[导入向导](#)可帮助您从 Derwent Innovation 之外的其他源导入数据。

自动导入 Derwent Innovation 的导出内容

从 Derwent Innovation 导出的文件会打开 Derwent Data Analyzer 并自动执行导入过程。

1. 在 Derwent Innovation 中进行检索并导出要分析的数据
2. 在您的电脑上找到已导出的文件，并且如果需要，解压缩 ZIP 文件以提取导出的文件
3. 双击从 Derwent Innovation 导出的 .PDTF 文件
4. Derwent Data Analyzer 将会打开并运行导入和自动清理功能

使用导入向导导入原始数据文件

1. 单击功能区中的**主页**标签，然后单击**文件区域中的新建分析**
2. 选择**使用导入过滤器**
3. 单击**选择文件**按钮并选择要用于导入的文件；“数据预览”窗格将更新以显示文件中的信息
4. 单击**下一步**
5. *Derwent Data Analyzer*
自动为文件选择合适的数据库，但您可以选择其他过滤器，或单击“选择新过滤器目录”以选择自定义过滤器
注意：我们建议您使用默认选择
6. 单击**下一步**
7. 选择要导入的**字段**；
注意：默认情况下会选择全部字段
8. 或者可以选择**显示所有字段**以查看默认情况下未被正常导入的字段
9. 单击**完成**以导入数据并创建数据集

较大的数据集可能需要一些时间才能导入。完成时，总览表界面会显示数据集的总览，包括总记录数、原始检索日期，以及含有每个字段中唯一项总数的字段列表。

分析您的数据

分析工具简介

Derwent Data Analyzer

提供多种可用于分析数据和文本的工具。这些工具独立于数据，因此能够以相同的方式应用于任何类型的数据。

选
项
列
表

描述

显示任意字段中每个项所对应的记录数。

基于任意两个字段之间的关系，显示每个项所对应的记录数。

- 共现：显示在任意两项的相交处对应有多少条记录（例如，多少条记录同时具备过去 10 年出现排名在前 10 位的公司这两个条件）。
- [自相关](#)：对于相同字段中的各个项，基于它们一起出现的频繁程度来显示其相关情况（例如，哪些发明人经常一起工作）。
- [互相关](#)：对于相同字段中的各个项，基于另一字段中的项来显示它们的相关情况（例如，基于 IPC 显示哪些专利权人在致力于相同的技术领域）。
- 因子：利用“主成分分析”方法，将复杂字段分解为“聚类”（即，基于词组分析来识别技术聚类）。

矩
阵

以可视化的方式展现出任意字段中各个项之间的相互关系。

- [自相关](#)：对于相同字段中的各个项，基于它们一起出现的频繁程度来显示其相关情况（例如，哪些发明人经常一起工作）。
- [互相关](#)：对于相同字段中的各个项，基于另一字段中的项来显示它们的相关情况（例如，基于 IPC 显示哪些专利权人在致力于相同的技术领域）。
- [因子](#)：利用“主成分分析”方法，将复杂字段分解为“聚类”（即，基于词组分析来识别技术聚类）。

图
谱

列
表
比
较

揭示任意两个列表之间的共同项或特有项（即使在不同的数据集中）。

分
类

可对数据集中的记录进行分类，例如先分为若干大类，然后再将这些大类进行细分。可进行自动分类，以便利用机器学习技术来加快分类进程。您可以在运行机器学习工具之前对其进行训练，以提高自动分类的准确度。

您还可以导入先前构建的[知识库](#)，将以前的训练资源应用于当前的数据集。

- [构建分类](#)：创建要应用于数据集的分类和子类别。
- [Smart Trainer](#)：手动对有代表性的记录进行分类，以提升自动分类工具的准确度。
- [自动分类设置](#)：设置自动分类的选项（如训练字段），以及启动自动分类过程。

提供各种工具，可帮助您组织整理数据集中的数据和文本。

- [VizLink](#)
[图](#)：可为各种字段同时显示三种可视化的视图。可视化选项包括：表格、柱状图、条形图、饼图、折线图或世界地图。
- [透视工具](#)：创建可编辑的透视表，其中最多可包含 10 个字段。
- [矩阵查看器](#)：在一个矩阵中显示各项之间关联强度的可视化呈现效果。
- [桑基图](#)：直观呈现从一个值到另一个值的流动情况，最多涵盖 5 个字段。
- [概念栅格](#)：通过对至少含 20 个关键主题词的字段进行主成分分析 (PCA)，将记录划分到各种主概念和子概念。这些概念以栅格形式显示，其中主概念显示在最上方的行中，并在每列所代表的主概念之下，显示其包含的子概念。

自动将计算得出的值添加到数据集中。

- [新兴趋势指标](#)：基于四个关键因子来跟踪数据集中主题词的演变过程：新颖性、持久性、社区化程度和增长。
- [日期差异](#)：自动计算两个日期之间的间隔时间，并将该信息保存到新字段中。
- [网络指标](#)：评估组织机构内的专利发明人，并跨组织机构评估科技文献作者。
- [统计数据](#)：根据数值字段计算有关字段的统计数据，例如最小值、最大值、平均值、中值、众数、和，以及标准差。

常用分析表

要解答的问题	相关工具
排名前 20 的项有哪些？	列表
排名前 5 的企业其相对增长率是多少？	共现矩阵
哪些作者/发明人一起工作？	自相关图谱
哪些企业在共同合作？	共现矩阵
哪家企业正致力于类似技术？	互相关图谱
在技术方面存在哪些“聚类”	因子图谱

？

有哪些“热点”技术领域？ 使用因子图谱来创建组，然后利用组创建共现矩阵

对于一家特定企业而言，其使用“创建子集”来分离该企业的专利组合，然后使用跨越两个数据集独特之处是什么？ 的列表比较。

文献与专利的独特之处是什么？ 使用“增量导入”以获取两个数据集中的 NLP（自然语言处理）字段（例如，词组），然后使用跨越两个数据集的列表比较。在独特主题词上使用因子图谱。

查看列表中的字段

列表显示字段中每个记录所包含的项的数量。

要打开列表，请在总览表中双击字段。

如果您不在总览表中，则可以通过以下操作打开任何字段的列表：

1. 单击功能区中的**主页**标签，然后单击**添加表单区域中的列表**
2. 选择要查看的**字段**

提示：可使用字段列表上方的文本框来检索特定字段

3. 单击**确定**可查看列表中的相应字段

标题视图

标题视图是为选定的列表项显示数据集中记录标题的窗口。选择列表视图中的项可显示数据集中包含该列表项的记录标题。当单元在共现矩阵中被选定时，标题视图窗口会显示同时包含行和列两个列表项的记录标题。

标题旁的图标表示该记录已经在[记录视图](#)中添加过注释，以及/或者该记录被标记为从新数据集中移除。

您也可以在标题显示中选择多个记录。

在标题显示中右键单击将显示具有以下选项的菜单：

选项	描述
查看记录	显示“显示记录”中选定的内容（仅限原始记录视图）。
查找相似记录	在选定字段中运行查找相似记录工具。
复制标题名称	将选定的内容复制到剪贴板，以用于粘贴到另一个应用程序中。
全选	选择标题视图中的所有标题。然后，您可以使用“复制标题名称”将所有标题复制到另一个应用程序中。

将选定记录添加到组 将选定记录添加到您选择的字段中的组。将显示建议的字段，或者您可以选择所有字段。您可以为选定字段选择一个现有组，也可以创建一个新组。

创建基于选定的 将标题视图中选定的所有记录创建为一个子数据集。

子数据集

在新数据集中忽略选定记录

标记在创建新的数据集或导出原始记录时，要忽略的记录。

在新数据集中不忽略选定记录

如果某个记录之前标记为忽略，则移除此“忽略”标签。

记录分类

允许用户对一个或多个记录设定分类方式。

专利生命特征

此报告可通过单击功能区上的**分析**标签，然后单击“添加”区域中的**专利生命特征**按钮来访问。

描述：创建新的列表，该列表显示数据集中专利的健康状况和有效性。这是一个比典型的专利“有效或无效”(dead or alive) 更加详细的报告，基于复杂的专利法律状态数据得到。

本报告检查数据集中的记录，依据 INPADOC

法律状态中的条目，计算自公开之后的时间，并向您呈现一个使用不同颜色标示的列表，该列表按照专利的“健康”情况对其进行分组，并大致估算每个专利生命周期剩余的时间。之后，文献被划分为由不同颜色标识的组，用于指示以下专利：

- 具有正面法律状态，可能有效（标识为绿色）
- 没有法律状态条目，但尚未到其过期时间（标识为蓝色）
- 尚未到其过期时间，但具有负面法律状态条目（授权专利标识为橙色，申请标识为黄色）
- 可能已过期失效（标识为红色）

为每个有效性级别的授权专利和申请创建单独的组（授权专利和申请各自均包含四个组，总共有八个可确定专利健康状况的组）。

列表同样包含多个列，分别用于最近的法律状态 (+/-

)、最近的法律状态操作年份、基于最早优先权日的预计剩余年份，以及基于公开日期的预计剩余年份。

可以通过实用工具脚本将字段作为附加列添加到“专利生命特征分析”报告中。**了解更多：**添加字段至专利生命特征

使用要求：必须打开公开级别的数据集（“专利生命特征分析”不适用于同族专利级别的数据集）。

数据备注

- 使用“专利生命特征分析”处理从 Derwent Innovation 导出的各条记录
- “专利生命特征分析”需要来自专利授予机构的数据；它不会处理 DWPI 数据
- 请不要对按同族专利（DWPI 或 INPADOC）归并的数据集使用专利生命特征分析，因为这样不能分析企业的完整专利组合

，不能为专利家族的其他成员提供有效性排序，也会导致对您数据集中最相关的授权专利或申请的分析失败

- 在 Derwent Innovation 中检索专利数据时，不要使用“仅增值专利信息 - DWPI 和 DPCI”专利集合，因为这样会使得每个同族专利仅返回一条记录

从 Derwent Innovation 导出记录的最佳做法

1. 使用 Derwent Innovation 中的获取 **INPADOC 同族专利成员**工具以包括来自所有 INPADOC 专利授予机构的专利记录
2. 按申请号归并包含所有 INPADOC 同族专利的检索结果，并使用**最新专利**作为“首选文献”
3. 单击检索结果标题行中的**全选**复选框；这会仅选择归并的文献
4. 依次单击**下载、导出和报告**，使用 Derwent Data Analyzer 格式导出选定记录

这会产生一个数据集，包含检索中的专利的所有相关 INPADOC

同族专利成员，以便您可以分析专利在其提交的每个授权机构中的健康状况。按申请号归并检索结果会从数据集中删除具有相应授权专利的申请。

分类数据集

比较矩阵中的字段

创建自相关矩阵

1. 单击功能区中的**分析**标签，然后单击**添加区域中的矩阵**
2. 单击“创建矩阵”界面中的**自相关**标签
3. 选择要用作矩阵中的**行和列的列表**（字段）；
注意：自相关矩阵始终是对称的（使用同一字段作为行和列）
4. 如果您选择含组的字段（此类字段名称旁边有一个加号），则可以选择是否要选择组；您可以通过下面两种不同的方式在矩阵中显示组：
 - **选择组/显示项：**对组内的各个列表项进行矩阵求值
 - **选择组/显示组：**将组作为一个整体进行矩阵求值
5. 如果您打算选择组，请选中要包含在矩阵中的组对应的复选框
6. （可选）单击**更多选项**以选择您要以何种方式构建矩阵（**记录数量**或**频次**）
提示：我们建议在分析中主要运用“记录数量”；对于在一条记录中可能出现一次以上给定项的字段，**频次**则可能是适当的选择（例如，NLP 单词或词组）。
7. （可选）选择您需要的相关函数：**皮尔逊系数**或**余弦**

8. 单击**确定**

请注意，在您选择字段和组时，字段列表下面的表格将会更新以提供矩阵外观的预览。

自相关矩阵

自相关矩阵显示一个列表中各个项之间的相关性。例如，发明人的自相关矩阵将显示在一起工作的团队成员之间的高相关性。

请参见显示以下两位发明人之间高度相关性的自相关矩阵：Edmunds A 和 Schaetzer J。

发明人的自相关矩阵将显示由于在相同的记录中示为发明人而具有高度相关性的发明人。

注意：对于自相关矩阵，应当仅使用在大多数记录中具有多个值的字段（例如发明人、专利权人、作者和描述词都是很好的选择）。公开日期则不是一个好的选择，因为每个记录只有一个公开日期。

了解更多： [创建自相关矩阵](#)。

在自相关矩阵中，您可以执行以下操作：

- 缩放
- 排序
- [创建热力图](#)
- 填充颜色
- “填充”矩阵
- 选择多个单元
- 查找字符串
- 矩阵项列表

创建互相关矩阵

1. 单击功能区中的**分析**标签，然后单击**添加区域中的矩阵**
2. 单击“创建矩阵”界面中的**互相关**标签
3. 选择要用作矩阵中的**行和列**的列表（字段）；
注意：自相关矩阵始终是对称的（使用同一字段作为行和列）
4. 如果您选择含组的字段（此类字段名称旁边有一个加号），则可以选择是否要选择组；您可以通过下面两种不同的方式在矩阵中显示组：
 - **选择组/显示项：**对组内的各个列表项进行矩阵求值
 - **选择组/显示组：**将组作为一个整体进行矩阵求值
5. 如果您打算选择组，请选中要包含在矩阵中的组对应的复选框
6. 选择应基于哪些字段和/或组建立相关性；通常而言，共享许多相同“**互用**”值的行和列的项将具有较高的相关性值。

7. （可选）单击**更多选项**以选择您要以何种方式构建矩阵（**记录数量或频次**）

提示：我们建议在分析中主要运用“记录数量”；对于在一条记录中可能出现一次以上给定项的字段，**频次**则可能是适当的选择（例如，NLP 单词或词组）。

8. （可选）选择您需要的**相关函数**：**皮尔逊系数或余弦**

9. 单击**确定**

请注意，在您选择字段和组时，字段列表下面的表格将会更新以提供矩阵外观的预览。

互相关矩阵

互相关矩阵基于一个列表中的值，显示另一个列表中的项与项之间的关系。例如，使用描述词的作者互相关矩阵可以显示撰写同一文献的一组作者。再如，使用描述词的机构互相关矩阵可以显示撰写同一文献的组织机构。

创建互矩阵需要选择两个字段。第一个要选择的是将在矩阵中实际显示为行和列的项，通常是在列表视图中定义的一个字段或较小的一组列表项。要选择的第二个字段是对行和列项之间的关系进行分析的基础。

了解如何[创建互相关矩阵](#)。

在互相关矩阵中，您可以执行以下操作：

- 缩放
- 排序
- [创建热力图](#)
- 填充颜色
- “填充”矩阵
- 选择多个单元
- 查找字符串
- 矩阵项列表。

创建因子矩阵

1. 单击功能区中的**分析**标签，然后单击**添加区域中的因子矩阵**
2. 选择要分析的列表项

注意：从大量的列表项中创建因子矩阵需要花费大量计算时间。通常情况下，我们建议您先使用适量的项（例如，小于

100）以获得一种分析的“感觉”。您可创建一个列表项成组，从该组中创建因子矩阵。例如，创建名为“排名前 50”的一个组，该组含有 50

个最频繁出现的列表项。在**创建因子矩阵**对话框中，为您的因子矩阵选择“排名前 50”的组。

3. 选择要使用的因子的数量。打开**创建因子矩阵**对话框时所示的默认值是您的分析中列表主题词数目的平方根。因子的适当数量取决于分析中的数据和您的目的。该平方根是其中一种方法。另一种方法是取分析中所含列表项的数目的一半；该方法试图获得计算累积方差的某个阈值。
4. 选择是否旋转因子。旋转因子这一操作会设法提高这些因子与数据的贴合程度，使得它们更容易解释
5. 选择是否通过其特征值调整因子规模。默认是通过特征值进行调整
6. 单击**确定分析数据**

创建热力图

热力图会高亮显示矩阵中具有最高数值的单元，这样易于识别具有最强关系的主题词。

1. 在矩阵视图中单击右键，并选择**创建热力图**
2. 为最高值选择一种颜色并单击**确定**
3. *Derwent Data Analyzer*

会使用所选颜色的不同渐变梯度在矩阵中给单元上色；最小值最接近白色，最高值则完全饱和并高度可见

如要移除矩阵颜色，请右键单击矩阵并选择**移除颜色**。

要在默认情况下显示热力图，请参见热力图设置。

使用细节窗口发现趋势

细节窗口总览

细节窗口提供主视图中选定记录的详细信息。这些细节窗口会按视图中的选定项或节点，显示一个字段中各个项的共现信息。细节窗口有两种类型的视图：列表和图表。

列表型细节窗口有三列：第一列显示共现值，第二列显示“[差异标识箭头](#)”，而第三列则包含共现项的文本。而在图表型细节窗口中，除了共现值是以图表形式展现且不显示差异标识箭头以外，其余部分均与列表型细节窗口显示的数据相同。

您可以切换使用不同的视图类型、更改图表类型、在该列图表中缩放、复制和/或打印该数据，以及通过右键单击您要更改的细节窗口来打开菜单实施其他操作。每个菜单会在下文中进行介绍。

当您在细节窗口中单击某个项时，其相关记录会在标题视图中高亮显示。在以上示例中，用户在“Family Member Years”（同族专利年份）细节窗口中单击了年份“1998”，因此 1998 年的 43 个记录标题（来自共现矩阵中选定的 152 个标题）均在“标题”视图中高亮显示。

所有字段均可在细节窗口中查看 – 在每个细节窗口顶部的下拉菜单中选择字段即可。

添加新的细节窗口

与细节窗口相关的菜单指令位于功能区中的**查看**标签下。

1. 单击功能区中的**查看**标签
2. 单击“工作空间”区域中的**细节窗口**以显示或隐藏所有细节窗口
3. 单击**添加细节窗口**以添加新的细节窗口

细节窗口会与单个表单一起保存

细节窗口对于数据集中的每个表单都是唯一的。保存数据集时，还将保存每个表单的细节窗口。这包括字段和图表类型的选择内容。

有权访问已保存数据集的其他人员将看到您为每个表单选择的相同细节窗口，以帮助他们了解您的分析。这包括使用 Derwent Data Analyzer Reader 访问数据集的人员。

其他细节窗口主题

- [细节窗口 - 差异标识](#)
- [细节窗口 - 列表弹出式菜单](#)
- [细节窗口 - 图表弹出式菜单](#)
- 细节窗口中的记录/父项范围
- 细节窗口 - Meta 标签弹出式菜单
- [细节窗口 - 排序列表](#)
- 细节窗口 - 缩放图表
- [细节窗口 - 图表颜色](#)
- 细节窗口 - 分组的数据

细节窗口 - 列表弹出式菜单

通过右键单击细节窗口中的列表调出包含以下选项的菜单：

选项	描述
	由显示列表切换至显示下述图表之一：
	<ul style="list-style-type: none">• 垂直柱状图：切换细节窗口，以垂直柱状图显示数据。• 水平柱状图：切换细节窗口，以水平柱状图显示数据。• 饼图：切换细节窗口，以饼图显示数据。• 线图：切换细节窗口，以线图显示数据。
显示图表	
打印	打印列表。（ 警告 ：如果列表很长，可能会打印很多页。）
复制	复制列表选定（高亮）部分到剪贴板。
创建细节窗口	打开新的细节窗口。
允许行调整大小	选择此选项后，可以调整行高。

添加选定至组	将您选定的内容添加到现有组。
打开	如果项的数据类型是文件，则可以打开该项。

细节窗口 - 图表弹出式菜单

当您右键单击细节窗口中的图表时，会显示如下菜单选项：

选项	描述
显示列表	切换细节窗口以显示列表视图而非图表。 由细节窗口中显示的图表类型切换至下述图表之一： <ul style="list-style-type: none">垂直柱状图：切换细节窗口，以垂直柱状图显示数据。水平柱状图：切换细节窗口，以水平柱状图显示数据。
图表类型	<ul style="list-style-type: none">饼图：切换细节窗口，以饼图显示数据。线图：切换细节窗口，以线图显示数据。使用选定时间范围：将图表显示的时间范围限制为选定的范围。使用数据集时间范围：图表显示的时间范围为整个数据集覆盖的范围。

按照记录数量排序	按记录数量在图表中排序项目（降序排列）。
----------	----------------------

按照项标签排序	按标签字母顺序在图表中排序项目。
---------	------------------

缩小	如果您放大了一个图表，则此功能可以缩小该图表。如果您执行了多次放大操作，该视图一次会缩小一个级别。
----	---

缩小全部	此操作会一直缩小图表直到显示所有数据。
------	---------------------

页面设置	调出对话框以设置用于打印图表的多个选项。
------	----------------------

打印	打印图表。
----	-------

复制	复制图表到剪贴板上（用于粘贴到其他应用程序）。
----	-------------------------

保存为位图格式/

保存为 JPEG	调出对话框，将图表另存为位图 (*.bmp) 或 JPEG 格式的文件。
----------	--------------------------------------

图片格式

创建细节窗口	打开新的细节窗口。
--------	-----------

细节窗口 - 排序列表


在细节窗口中查看列表时，双击数据上面的条可以对列表进行排序（类似在主窗口中对列表排序）。排序时，光标会变成向下箭头。

细节窗口 - 差异标识

当细节窗口中的共现值大幅高于或大幅低于共现期望值时，即出现差异标识箭头。当箭头出现时，意味着细节视图中的共现值偏离了预期。箭头的数目（一个、两个或三个）表示共现值偏离预期的程度，三个箭头表示偏离程度最大。无箭头表示该值没有很大程度偏离预期或不能确定其预期值。

绿色上指箭头意味着共现值大幅高于预期值。红色下指箭头则意味着该值大幅低于预期值。需要注意的是，对于视图选定的具有零共现性的项，仅当其大幅低于预期时才会显示在细节窗口中。

细节窗口颜色

当选择矩阵视图中的整行或整列时，细节窗口会使用颜色来高亮显示与这个（这些）行和列相关联的项。通过单击  图标并选择**选项**来定义颜色。

然后，单击**栅格颜色**标签。使用“**列表**”（包含主视图中的矩阵）区域选择矩阵视图的细节窗口列表中显示的颜色。

矩阵与细节窗口颜色之间交互的示例


以黄色高亮显示的是仅与选定列共现的项，而以蓝色高亮显示的是仅与选定行共现的项。以绿色高亮显示的则是与行和列**两者**所选择的项均共现的项。注意，必须选择整个矩阵的行和/或列，才能在列表型细节窗口中以颜色进行高亮显示。

另请参见

- [细节窗口 - 图表颜色](#)

细节窗口 - 图表颜色

在细节窗口中查看图表时，可以选择用于该图表的颜色。

1. 单击  图标，并选择**选项**
2. 单击**栅格颜色**标签
3. 使用**图表区域**选择显示在矩阵视图的细节窗口列表中的颜色

注意：只有您重选图表类型（或显示于细节窗口的字段名），才会将更改应用于现有图表。

另请参见

- [细节窗口颜色](#)

使用图谱进行分析

使用图谱

当您在节点上移动光标时，会显示一个下拉列表。该列表显示该节点的有关信息或与该节点相关记录的信息。您可以通过双击节点使下拉列表“伸出”。如要“收回”下拉列表，再次双击该节点即可。

注意：如果下拉列表不“伸出”，请参见更改图谱显示偏好中有关节点高亮显示的特别说明。

当您单击节点时，标题窗口将显示该节点的标题。对于因子图谱，还会显示每一个记录的相对得分。得分（绝对值）越高，则该记录与因子越相关。您可以按照标题字母顺序或得分的数字大小来排列记录——只需单击栏位标题即可（“得分”或“项目数量、选择数量”）。

图例显示以下关于图谱的信息：

- 最上面的标题显示了图谱的类型：因子图谱、自相关图谱、互相关图谱或 PCD 图谱。
- 第一行显示所使用的第一个字段（和组）。
- 第二行显示所使用的第二个字段（和组）（用于互相关图谱）或因子的数量（用于因子和 PCD 图谱）。
- 第三行显示覆盖率。覆盖率表示图谱中的节点覆盖您的数据集所占的百分比。剩余的记录（即，100% 减去覆盖率）则不含任何聚集在您的分析中的主题词。
- 第四行显示用户选择用于所示“连线数量”的方法。

连线图例显示了每一范围中的相似度连线数量；圆括号中的数字表示处于相似度类别之中、但不显示在图谱上的连线数量。这两个数量总和即图谱相似度类别中所有连线的总数量。

当光标悬停在一条连线上时会显示其相似度（。

右键单击图谱时，将显示一个具有以下选项的弹出式菜单。

选项 描述

**选择
字段** 将字段添加到图谱后，该选项允许您选择哪些字段出现在节点的下拉列表中

**添加
字段
至图
谱** 当您第一次创建图谱时，只有对相似度进行定义的字段可供显示在图谱上。该选项则可让您选择其他字段添加至图谱。您选择的任何字段都将出现在“选择字段”菜单项中以供选择。最后选择的任意字段将自动设为在每个节点的下拉菜单中显示。

**显示
信息
框** 显示每个节点所对应的所有下拉框（列表）。

**隐藏
图例** 从图谱中隐藏图例。隐藏后该选项将变为“显示图例”。

缩放 放大或缩小图谱。还可以选择缩放百分比，选择“适合窗口”以在窗口中显示整个图谱，或通过输入您自己的缩放百分比来定制该设置。

复制到剪贴板 将图谱复制到剪贴板。之后则可粘贴到另一应用程序（例如，Microsoft Excel、Word 或 PowerPoint）。

导出至文件 将图谱保存为位图、JPEG、TGA 或 TIFF 文件。

重设图谱 使用 *Derwent Data Analyzer* 默认设置重新绘制图谱。

编辑偏好 通过选择字号、画布尺寸和其他属性来改变图谱的外观。这也提供了一种用于垂直和水平扩展节点的机制。

了解更多：更改图谱显示偏好

当您选择显示“年份”类型的字段时，会出现数据直方图。

当您选择显示“数字”类型的字段时，则会出现数据箱线图。

另请参见

- [更改图谱显示偏好](#)

因子图谱

Derwent Data Analyzer 可用于创建数据的可视图谱。因子图谱是“主成分分析 (PCA)”方法所得结果的图形表示。PCA

会找到频繁在数据集中一同出现的列表项。图谱中的每个节点代表主题词的一个聚类。节点之间的连线代表两个主题词聚类之间的相似程度。连线的粗细（或线型）表示相似的程度（如图例中所定义）- 即 0 和 1

之间的一个数字。为避免在视觉上造成混乱，只有相似度最强的部分才会被显示。

创建因子图谱

1. 在要分析的列表中创建一个组
 - 请不要将任何只出现几次的列表项包含在内。一般的经验建议是仅包含在十 (10) 个或更多记录中出现的列表项。包含不那么频繁出现的列表项可能导致分析失败。
 - 请不要将任何在大多数记录中都出现的列表项包含在内。
 - 根据数据集中记录的数量而定，包含的主题词不应少于 15 或 20 个，通常不应多于几百个。
2. 单击功能区中的**分析**标签，然后单击“添加”区域中的图谱

3. 单击**因子图谱**标签
4. 选择包含要在图谱中使用的组的**字段**
5. 选择要以何种方式在图谱中显示组：
 - **选择组/显示项**：对组内的各个列表项进行矩阵求值
 - **选择组/显示组**：将组作为一个整体进行矩阵求值
6. 选中要包含在图谱中的组的对应复选框
7. 指定要在分析中使用的**因子数量**。这会影响图谱中显示的节点数

提示：一般经验是建议在分析中使用主题词数目的平方根。当选择组时，这是默认值。

5. 如果您要为图谱中的每个节点创建组，请选择**为每个节点创建一个新组**

注意：如果您选择了“选择组”，则此选项不可用。

6. 单击**确定**创建图谱

请注意，在您选择字段和组时，字段列表下面的表格将会更新以提供矩阵外观的预览。

自相关图谱

自相关图谱显示一个列表中的项与项之间的关系。例如，作者的自相关图谱可以显示团队中的联合作者。描述词自相关图谱则可以显示由于在相同的记录中使用而具有高度相关性的描述词。

注意：对于自相关图谱，应当仅使用在大多数记录中具有多个值的字段（例如作者和描述词都是很好的选择）。而公开日期则不是一个好的选择，因为每个记录只有一个公开日期。

每个节点代表一个发明人。节点的大小反映与该发明人相关的记录数。这些节点的大小全部相同，因为这些发明人具有相似数量的记录数（按照与数据集中总记录数进行的对比）。正如因子图谱一样，这些连线反映节点之间的相似度。在以下示例中，连线的粗细与作者所发表文献的数量相关。为避免在视觉上造成混乱，只有相似度最强的部分才会被显示。

创建自相关图谱

1. 在您想要分析的列表中创建组

注意：在图谱中应包含足够多的主题词，但不宜过多。与因子图谱不同的是，您选择的全部项都将出现在自相关图谱中。通常，15 至 20 个主题词最适合填满图谱且仍然可读。

2. 单击功能区中的**分析**标签，然后单击“添加”区域中的**图谱**
3. 单击**自相关图谱**标签
4. 选择包含要在图谱中使用的组的**字段**
5. 选择要以何种方式在图谱中显示组：

选项	描述
选择组/显示项	对组内的各个列表项进行矩阵求值
选择组/显示组	将组作为一个整体进行矩阵求值

6. 单击**确定**创建图谱

请注意，在您选择字段和组时，字段列表下面的表格将会更新以提供矩阵外观的预览。

互相关图谱

互相关图谱基于一个列表中的值，显示另一个列表中的项与项之间的关系。例如，使用描述词的作者互相关图谱可以显示撰写同一文献的一组作者。再如，使用描述词的机构互相关图谱可以显示撰写同一文献的组织机构。

互相关图谱需要两个字段。第一个是将在图谱中显示为节点的项，通常是列表视图中的一组列表项。第二个字段将作为分析这些节点间关系的基础。

警告：互相关图谱中关系约束的严格程度略低于因子图谱和自相关图谱。这使得一些“一次性”关系能够在互相关图谱中描述。例如，如果作者 A 和作者 B 不是共同作者，但均与 C 为共同作者，则互相关图谱（字段 1 = 一个包括 A 和 B 的作者组，字段 2 = 所有作者）可以揭示出作者 A 和 B 之间存在关联，即使作者 C 并未显示在该图谱上。因此，在互相关图谱中，您应该注意进一步调查所示关系。应将“低相似度”关系视为“可能的”关系——在某些情况下，它们体现的是间接关系。

该示例中的关系表明这些机构正在研究类似的主题（由出版物中的描述词字段所定义）。为避免在视觉上造成混乱，只有相似度最强的部分才会被显示。

创建互相关图谱

1. 在您想要分析的列表中创建组

注意：在图谱中应包含足够多的主题词，但不宜过多。与因子图谱不同的是，您选择的全部项都将出现在互相关图谱中。通常，15 至 20 个主题词最适合填满一页图谱且仍然可读。

2. 单击功能区中的**分析**标签，然后单击“添加”区域中的**图谱**

3. 单击**互相关图谱**标签

4. 选择包含要在图谱中使用的组的**字段**

5. 选择要以何种方式在图谱中显示组：

- **选择组/显示项：**对组内的各个列表项进行矩阵求值
- **选择组/显示组：**将组作为一个整体进行矩阵求值

6. 选中要包含在图谱中的组的对应复选框


7. 请注意，图谱定义会随着您的选择而进行构造。这样可以预览正在创建的图谱的大小

8. 选择要用来关联各个图谱项（例如，摘要词组）的字段或组
9. 单击**确定生成图谱**

请注意，在您选择字段和组时，字段列表下面的表格将会更新以提供矩阵外观的预览。

记录分类

您可以使用记录分类为您的数据创建、管理并应用您自己的分类系统。添加新的分类类别时，将在您的数据集中创建一个新的字段。该字段是“记录标识”型字段，即该字段中的每个项都是一个唯一的记录标识。通常而言，每个记录都唯一对应一个项，除非两条记录彼此完全相同（这种情况比较少见）。每个分类成为该字段中的一个“组”。

您可以从“记录视图”界面中的  图标访问“管理记录分类”对话框。或者，也可单击功能区中的**分析**标签，然后单击**记录分类**。

构建和管理记录分类

要对记录应用分类，您必须先配置要应用的分类。您还可以进一步配置细分类别，以便对记录进行更精细的分类。

分类选项

您必须至少构建一个分类，然后才能使用自动分类。

选项	描述
创建	添加新的分类（此操作也会将一个新的字段添加到您的数据集）。
删除	删除选定的分类类别。
重命名	更改分类的名称（此操作也会重命名数据集中的字段）。
更改顺序	更改所选分类在列表中的位置。使用双箭头将分类移动到顶部或底部。
导入知识库	导入在不同的相似数据集中创建且要与该数据集一起使用的 知识库 。
导出知识库	导出您为此数据集创建的 知识库 ，以便与其他类似的数据集一起使用。
自动分类设置	复查此数据集的 自动分类设置 。
更改训练字段	选择要用于此数据集的训练字段。例如，您可能已经使用“进一步处理”或 NLP 来识别长文本字段（如权利要求）中的关键字或短语。

类别选项

您必须选择一个分类，然后才能查看为其分配的细分类别。

选项	描述
创建	添加新的分类（此操作也会将一个新的字段添加到您的数据集）。
删除	删除选定的分类类别。

重命名	更改分类的名称（此操作也会重命名数据集中的字段）。
创建子类别	在所选类别中创建一个子类别。
更改顺序	更改所选分类在列表中的位置。使用双箭头将分类移动到顶部或底部。
导入	从文本文件或现有字段导入分类列表。
导出	导出分类列表至文本文件。
添加	添加新的分类。
允许多选	允许将记录分配至一个以上的类别。否则，选择新的分类后将移除已分配的任何其他分类。这也可以应用于子类别。
创建“忽略”类别	自动为所选分类创建“忽略”类别。如果您不想在 Smart Trainer 中使用某些记录（例如，如果它们不相关），请将记录分配至此类别。这也可以应用于子类别。

从字段导入类别

您可以选择一个字段作为对特定分类进行类别细分的来源。例如，您可能对某个字段执行了自然语言处理，您可能创建了一系列组名，或者您可能从某个字段提取了主题词（如 DWPI 标题词）。

如果选择从字段导入类别，您必须：

1. 选择包含您要导入的类别的**数据集**
2. 选择包含要导入的类别的**字段**
3. 选择要将字段中的哪些**项用作类别**：使用所有项、使用选定的组中的项，或使用组名
4. 如果选择使用选定的组中的项或使用组名，请选择相应的**组**
5. 单击**导入**

注意：分配给子类别的记录将自动分配给与其关联的主类别。

分类在数据集中的显示方式

每个分类都会创建一个新字段。字段名是分类名称，前面有两个冒号（“::”），如以下示例所示。

在每个分类字段中，每个类别都有一个组名，如[此图像](#)中所示。

有两处地方可将记录分配至分类：[显示记录](#)与[标题视图](#)。

使用标题显示将记录分配到记录分类

注意：您必须配置至少一个记录分类才能启用此选项。

1. 在[标题视图](#)中选择（或多选）记录

提示：使用多选可同时对一批记录（例如同族专利中的一组专利）进行分类
2. 右键单击所选记录并选择**记录分类**，然后选择要应用的分类

3. 这样会显示“记录分类”界面；选择要应用于记录的分类和类别
4. 单击**关闭**

请查看[此处的记录分类菜单示例](#)。

[使用记录视图将记录分配到记录分类](#)

注意：您必须先配置至少一个记录分类，然后才会显示记录分类面板。

1. 打开[记录视图](#)
2. 选中要添加到记录的分类的复选框
3. 单击“**下一条记录**”箭头继续对记录进行分类

自动分类记录

在对记录进行分类时，机器学习算法会检查这些记录以识别相似的特性（文本中的段落、分类代码，等等）。该算法通过这些信息为数据集中未分类的记录预测其分类。

正如任何机器学习算法一样，提供的样本数据越多，自动分类的准确度也就越高。在对记录进行分类时，百分比指示符将显示该算法预测自动分类的准确程度。在预测的准确度达到可接受的级别时，即可开始进行自动分类。

1. 打开[记录视图](#)并按上文所述对记录进行分类
2. 在对记录进行分类时，百分比指示符会提高以指示自动分类的预测准确度
3. 在预测的准确度达到可接受的水平（例如 80%）时，双击百分比指示符即可开始进行自动记录分类；您也可以单击工具栏中的“记录分类”按钮
4. 在“自动分类记录”界面上，选择要自动应用的**分类**
5. 选择自动分类的**准确度标准**；可选择“高”或“低”的级别
 - **高准确度**可自动分类的记录较少，但预测的准确度较高
 - **低准确度**可对较多的记录进行分类，但预测的准确度较低
6. 单击**开始分类**，将分类自动应用于数据集中的未分类记录

自动分类设置

“自动分类设置”界面允许您微调自动分类算法的函数。通过此界面，您可以：

- 查看与每个分类关联的类别和记录数量，以及分类算法可以处理的置信度
- 导入或导出用于自动分类的知识库，以便将在某个数据集上完成的工作应用到另一个数据集
- 选择分类算法用来进行预测的训练字段

您也可以从此界面中开始进行自动分类或启动 Smart Trainer。

Smart Trainer

Smart Trainer 可帮助您减少为自动分类提供充足样本数据所需的时间。

Smart Trainer

会显示一系列记录，因此在理想情况下，您不会看到属于同一类别的一长列记录。Smart Trainer 将开始开发一个在对记录进行分类时，应将该记录归入哪个类别的模型。在该模型的开发过程中，Smart Trainer 会为您提供它认为应归入特定类别的记录。这些记录通常包含与训练字段中的主题词中位数距离在一个标准差范围内的一些主题词。

这可以帮助您在较少的时间内将更多的记录添加到每个类别，因为您不会看到属于同一类别的一长列记录。您分类的每条记录都会为自动分类算法提供更具有特色的数据，这可以减少在达到高预期分类准确度之前需要进行分类的记录数量。

1. 单击功能区中的**分析**标签
2. 单击“**添加**”区域中的**记录分类**
3. 添加适用于您的数据集的分类和类别
4. 单击**启动 Smart Trainer** 按钮
5. 将显示一条记录，其中包含您选为训练字段的那些字段；查看记录并选择适当的类别
6. 单击**查看下一条记录**按钮
7. 继续对记录进行分类，直到预期准确度达到可接受的水平；当您训练模型时，它会在达到足够高的置信度阈值后开始亮显建议的类别
8. 双击百分比指示符
9. 在“**自动分类记录**”界面上，选择要自动应用的**分类**
10. 选择自动分类的**准确度标准**；可选择“**高**”或“**低**”的级别
 - 高准确度可自动分类的记录较少，但预测的准确度较高
 - 低准确度可对较多的记录进行分类，但预测的准确度较低
11. 单击**开始分类**，将分类自动应用于数据集中的未分类记录

Smart Trainer 界面参考

Smart Trainer 提供了若干选项，可帮助您为自动分类提供充足的样本数据。

选项	描述
----	----

管理记录分类	打开 构建和管理记录分类 界面，您可以在其中随手快速添加更多分类或细分类别。
--------	--

自动分类设置	打开 自动分类设置 ，您可以在进行自动分类之前，先在此处更改自动分类的运行方式。
--------	--

记录分类	打开 自动分类记录 界面，您可以在其中设置准确度目标并启动自动分类。
------	--

查看先前记录	转到训练数据集中的上一条记录。
查看下一条记录	查看训练数据集中的下一条记录。
退出	退出 Smart Trainer。
查找相似记录	查找与当前记录相似的记录。有关详细信息，请参见查找相似记录。
打印	打印当前显示的字段。
复制	将选定文本复制到剪贴板。
全选	选择当前显示字段中的所有文本。
查找文本	查找当前显示字段中的文本。
显示原始记录	查看从数据源中导入的原始记录，即未选定任何字段或应用任何格式的记录。
显示字段	显示只含所选字段的表格视图。默认显示您选择的训练字段，但您还可以添加其他字段以提供上下文信息。
字段顺序	选择要显示的字段以及这些字段的显示顺序。您选择的训练字段将始终按选定顺序显示，但您还可以选择显示其他字段以提供更多上下文信息。 单击此界面中的“更改训练字段”按钮，即可选择更多或更换训练字段。
在“唯一ID”字段中添加到组	将当前记录集添加到用作唯一标识符的组。
文本换行	选择在原始记录视图中是要对文本进行换行以适应界面大小，还是要水平滚动浏览文本。
高亮关键词	打开或关闭对我的关键词的高亮显示。
设定关键词颜色	选择用于高亮显示关键词的颜色。
训练字段	显示选择用于训练自动分类工具的字段。查看这些字段有助于选择要对记录应用的分类。
其他字段	为提供与训练字段相关的上下文信息而添加的其他字段。Smart Trainer 不会使用这些字段来开展分类训练；这些字段仅供参考，旨在帮助您选择适当的分类。 您必须单击“字段顺序”按钮，才可添加其他字段。

知识库

在知识库中，自动分类可跟踪它从分类记录中学到的模式，因而它能够自动对剩余的记录进行分类。您可以导出知识库，以供将来在类似数据类型的数据集中使用。

当您导入知识库时，Derwent Data Analyzer

会进行检查，确保当前数据集具有的自动分类设置与您要使用的知识库中存储的自动分类设置相同：

- n-gram 模型：
- 应用词干
- 删除停用词
- 训练字段

如果这些设置中的任一设置与导入的知识库不匹配，就会显示一条警告消息，通知您将会发生变化的设置。单击“否”取消导入，或单击“是”接受更改并导入知识库。

如果出现以下情况，导入将失败：

- 知识库和数据集的源数据库不匹配，并且您在出现警告消息时单击“否”
- 数据集中不存在一个或多个知识库训练字段
- 您在出现不匹配设置警告时单击“否”
- 知识库或数据集已损坏

探索

桑基图

桑基图能直观呈现从一个值到另一个值的流动情况，最多涵盖 5

个字段。例如，您可能需要了解数据集中排名前 10

的专利权人如何流入到高层级分类，然后进一步了解这些分类中专利记录的当前状态（失效或有效）。

创建桑基图

1. 单击功能区中的**分析**标签
2. 单击“探索”区域中的**桑基图**按钮
3. 选择要用于桑基图的字段，最多可选择 **5 个字段**
 - 建议在桑基图中使用的字段包括：类别、组织、年份、授权机构以及失效/有效状态。
 - 使用组来限制各个字段中的项目数量
 - 每个字段中的项目数量不应超过 10 个
4. 单击**确定**以创建桑基图

浏览桑基图

- 选择任一条形，即可在[标题视图](#)中查看相关记录并在[细节窗口](#)中查看选定记录的汇总信息
- 选择两个项目之间的连线，即可查看两者共有的所有记录
- 与所有图表或可视化视图一样，单击桑基图左上角的**齿轮图标**（设置）可打开“可视化样式”菜单，在该菜单中，可以更改可视化效果的颜色、文本样式以及注释

了解更多：[更改可视化样式](#)

创建概念栅格

概念栅格基于对至少含 20 个关键主题词的字段进行的主成分分析

(PCA)，将记录分类为主概念和子概念。该算法重在寻找彼此不重叠、并且合起来后可以覆盖数据集当中尽可能多的记录的概念集合。在确定整个数据集所涉及的比较宽泛的一些高层级主概念之后，该算法会将每个主概念视为一个独立的数据集合，并在其中寻找该主概念范畴内所包含的子概念。

在概念栅格中，确定的每个主概念各显示为一列。在每个主概念下，子概念显示为行。主概念与相关子概念之间用线条连接。

用法提示

- 您必须选择一个至少含 20 个关键主题词的字段
- 该算法需要使用整理后的关键词列表，并假定字段中的每个词都有含义
 - 如果使用 NLP 提取“标题”或“摘要”中的词组，请先运行整理 NLP（自然语言处理）进程，然后再基于提取出来的主题词创建概念栅格
 - 彻底清理要使用的任何主题词列表，以删除所有不适用的词或停用词
- 不同的高层级主概念经常会包含相同的子概念；如果同一个子概念出现在不同的主概念下，则概念栅格会为此共同子概念分配同一种颜色，以便于进行跟踪

创建概念栅格

1. 单击功能区中的**分析**标签
2. 单击“计算”区域中的**创建概念栅格**按钮
3. 选择要用于概念栅格的字段；可整理 NLP 的文本字段最适合用于概念栅格
4. 单击**确定**以创建栅格

浏览概念栅格

概念栅格会在最上面的一行显示所有主概念。在每个主概念下，会列出与该概念相关的若干关键主题词。

主概念与其子概念之间用线条连接。将光标悬停在子概念上方时，会显示一个加号图标。单击这个加号图标，可查看与这个子概念相关的关键词。

由于每个高层级主概念相当于一个数据集合，子概念是其子数据集，因此同一个子概念可能会出现在多个列（即多个主概念）中。在这种情况下，概念栅格会为在多个主概念中出现的共同概念分配一种独特的颜色，以便于进行跟踪。

选择任一概念或子概念，即可在[标题视图](#)中查看相关记录并在[细节窗口](#)中查看选定记录的汇总信息。

与所有图表或可视化视图一样，单击概念栅格左上角的**齿轮图标**（设置）可打开“可视化样式”菜单，在该菜单中，可以更改可视化效果的颜色、文本样式以及注释。

了解更多：[更改可视化样式](#)

计算

新兴趋势指标

新兴趋势指标基于四个关键因子来跟踪数据集中主题词的演变过程：新颖性、持久性、社区化程度和增长。

根据美国政府所做的工作，新兴趋势指标可以帮助您识别技术领域中值得关注的新研究领域。

新兴趋势指标的工作原理

为计算新兴趋势指标，Derwent Data Analyzer 将检查从过去 10 年的文本字段（例如，DWPI 标题或摘要）中挖掘出的单词或词组。头三年是一个基准，算法会根据该基准来检测新出现的主题词。

对于每个主题词，算法都会根据这四个因子（新颖性、持久性、社区化程度和增长）为每个词或词组生成一个分数。得分越高，表明技术的出现次数越多。

这些得分适用于所有数据集，甚至适用于不同的技术。成熟的技术领域往往比新兴领域的得分要低，这是因为改进通常是渐进式的，而不是突破性的。

新兴趋势指标要求

要计算新兴趋势指标，您必须：

- 使用 DWPI 同族专利数据集
- 至少包含 10 年的数据
- 对文本字段（DWPI 标题或 DWPI 摘要）运行 NLP 词组
- 探索范围相对宽泛的技术领域，例如冰淇淋、太阳能板或石墨烯

计算新兴趋势指标

1. 单击功能区中的**分析**标签
2. 单击“计算”区域中的**计算新兴趋势指标**按钮

3. 单击选择“**主题词**”字段下拉列表，然后选择要从其中提取主题词的字段
 - 选择您已对其运行 NLP 单词或词组的文本字段
 - 为了获得最佳结果，请使用具有标准化数据的文本字段，例如“DWPI 标题”或“DWPI 摘要”
4. 单击选择“**年份**”字段下拉列表，然后选择要用于确定数据集年份的字段
 - 您的数据集必须至少包含 10 年的数据
 - 如果数据集中的最后一年尚未结束（例如，数据集包含当前年份），则请选中“忽略最后一年的数据”复选框
5. （可选）选择要使用的“组织”、“人员”、“位置”和/或“标题”字段；此处所做选择可为算法提供更多有用信息
6. （可选）选择为计算的新兴趋势指标生成[新兴趋势散点图](#)。
7. 单击**确定**，即可根据选定字段中的数据生成新兴趋势得分

计算日期差异

通过使用“计算日期差异”，可以自动计算两个日期之间的时间，并将该信息保存到新字段中。例如，您可能需要计算记录的优先权日和当前日期之间的差异，以识别可能已过期的专利，或计算优先权年份和未来某个日期之间的差异，以查看届时哪些专利可能会过期。结果将显示在一个以年、月、周或天为单位的新字段中。

计算不同日期之间的差异

1. 单击功能区中的**分析**标签
2. 单击“**数学统计**”区域中的**计算日期差异**按钮
3. 选择用于计算的**第一个日期**，可以是数据集中的日期字段，也可以是您选择的固定日期
4. 选择用于计算的**第二个日期**，可以是日期字段，也可以是固定日期
5. 选择计算时使用的**间隔时间单位**（年、月、周或天）
6. 选择一个**字段名称**以输出计算结果
7. 单击**确定**

结果将显示在一个新列表中。该列表中的每行将显示按照您选择的间隔时间单位计算出的数值。例如，如果您选择以年为单位，则行将显示第一个日期与第二个日期之间间隔的年数。如果您选择以月为单位，则行将显示计算出的月数。要将计算出的日期差异结果精确到日历日期，您可能需要进行进一步计算。

网络指标

网络衡量指标可帮助您评估发明人和文献作者的社会网络情况。如果需要评估组织机构内的专利发明人和跨组织机构分析科技文献作者，则此类指标特别有用。

计算网络指标

1. 单击功能区中的**分析**标签
2. 单击“计算”区域中的**网络指标**按钮
3. 单击**选择字段**下拉列表，然后选择要用于生成社会网络指标的字段
 - 选择包含人员的字段（“发明人”，“发明人 - DWPI”）
 - 为获得最佳结果，请使用已清理过的字段
4. 单击**确定**以创建包含社会网络指标的新列表

创建的网络指标

中心性指标会根据节点在网络图中的位置为节点分配数值或排名，这有助于识别在社会网络中具有影响力的个人。Derwent Data Analyzer 提供了三个通用的社会网络指标，和一个独有的 Keystone 指标——该指标基于三个通用指标计算得出。

指标	描述
中介中心性	根据最短路径衡量中心性。对于连通图中的每对顶点，顶点之间至少存在一条最短路径，使得该路径通过的边数最少（对于未加权图）或通过的边的权重之和最小（对于加权图）。每个顶点的中介中心性值即是通过该顶点的最短路径数。 了解更多： 中介中心性 (Wikipedia)
接近中心性	通过计算节点与图中所有其他节点之间最短路径长度之和的倒数，衡量节点在网络中的中心性。因此，节点越接近于中心，它与所有其他节点的距离就越近。 了解更多： 接近中心性 (Wikipedia)
度中心性	连接到顶点的边数。 了解更多： 度中心性 (Wikipedia)
Keystone	上述 3 个指标的调和平均数。Keystone 指标提供了一种简化的衡量方法，可用于评估某个人在您检索的研究领域中的社会网络情况。Keystone 指标值越高，则表明该人在研究团体中起到的与其他人建立联系和传播信息的作用越重要。

统计数据

根据数值字段计算有关字段的统计数据。您可计算最小值、最大值、平均值、中间值、众数、和，以及标准差。您还可以创建复合字段，例如，按引用情况统计的专利权人数据。

此工具类似于定制化报告表中的数字字段统计，但允许排序并且您可创建组（与其他列表方式一样）。

创建字段的统计数据

1. 单击功能区中的**分析**标签
2. 单击“**数学统计**”区域的**统计数据**按钮
3. 从**左侧**下拉菜单中选择要为其创建统计数据的**字段**
4. 从**右侧**下拉菜单中选择要用于创建统计数据的**数字字段**


Derwent Data Analyzer 会基于右侧数字字段中的数据，自动计算字段统计。

处理统计表

- 单击某一列按升序或降序对其进行排序
- 选择某一行以在“标题”和“细节”窗口中查看关联记录
 - 您可以使用 Ctrl+单击来选中多个行
- 显示的列可进行常用统计数据计算：和、平均值、最小值、中间值和最大值
- 每个表都将显示箱线图，以提供对计算的更多见解；将鼠标悬停在箱线图上即可查看其他详细信息，例如离群值和四分位数
- 您可以从统计数据表单中的选定行创建组；这将在源字段中创建组，并且会将所选项目添加到这些组中
 - 选择包含要添加到组中的项的行
 - 右键单击并选择“在分析字段中添加到组”
 - 选择要添加项的组，或单击“在分析字段中创建新组”以创建新组
 - 单击“确定”以将字段添加到选定组
- 右键单击统计数据字段并选择“在分析字段中创建四分位组”以自动为源字段中的每个四分位数创建组，并将该字段中的项分配到相应的组中
- 单击左上角的 Excel (XLS) 图标以 Excel 格式从统计数据表单中下载数据，以供进一步分析

自定义指标

自定义指标会根据您使用自定义指标公式生成器创建的公式，在统计数据表单中创建一个新列。

1. 单击统计数据表单中的**添加自定义指标列**按钮 ()
2. （可选）单击**保存的指标**下拉菜单以选择一个保存的指标作为着手点
3. 为将要创建的指标输入一个**新指标名**

- 4. 设置各种**条件**来计算您的指标 (有关详细信息，请参阅[设置条件选项](#))
- 5. 单击行右侧的**加号 (+)** 以添加更多成分；单击**减号 (-)** 以删除某个成分
- 6. 默认的运算顺序是从左到右，并遵循括号以及您在每个成分之间所选运算符的运算顺序；
您也可以更改运算顺序：
 - 1. 单击**编辑顺序**按钮
 - 2. 更改要放置括号的位置
 - 3. 单击**保存**
 - 4. 请注意，您无法从“编辑顺序”屏幕更改运算符或成分的顺序；您只能更改括号
- 7. 完成自定义指标后，单击**计算指标**将其保存到统计数据表单并在新列中生成分数；您还可以单击**保存模板**来保留指标以用于其他统计数据表单

在您处理指标时，右上角的**实时计算**会随着您所做更改而相应更新，以根据统计数据表单第 1 行中的信息显示您的自定义指标。如果您需要完全重新开始计算，请单击**重置指标**按钮删除所有编辑，并从空白的自定义指标重新开始。

“设置条件”选项

成分类型

选项	描述
记录	该成分对应表中的第一列，无需选择任何额外的字段，因为记录数是基于您在创建统计数据表单时选择的分析字段。
和、平均值、中间值、最小值、最大值、标准差	这些成分对应于表中的列。值是根据您在相应的“ 字段 ”下拉列表中选择的 字段 计算的，计算方式与在统计表中选择该字段时相同。
总记录数	该成分对应表中第一列的最后一行，无需选择任何额外的字段，因为记录数是静态的。
总和、总平均、总最小值、总最大值、总标准差	这些成分对应于表中各自列范围内的“ 总计 ”行。值是根据您在相应的“ 字段 ”下拉列表中选择的字段计算的，计算方式与在统计表中选择该字段时相同。
常数	此成分让您能够在单独的输入框中输入任意数字。

字段/常数

某些成分类型让您能够选择字段或用作值或者输入常数值。

标准化？

选择“标准化？”复选框，以将该成分所在行上的值转换为 0 到 1 的区间。

运算符

选择要在不同成分之间使用的数学运算：加 (+)、减 (-)、乘 (*)、除 (/)。请注意，最后一个运算符将被忽略，因为没有后续成分。

报告结果

图表和可视化

图表和可视化可为您数据集中的信息提供可视化图表的呈现形式。可以是用于显示列表中排名前 5 位项目的简单柱状图，也可以是用于显示分类代码的更复杂表现形式。

更改可视化样式

您可以更改文本和数据项（线、条形等）在图表中的显示方式。单击图表或可视化左上角的**齿轮图标**（设置）可打开“**可视化样式**”菜单，在该菜单中，可以更改可视化效果的颜色、文本样式以及注释。

您也可以选择可视化效果的默认颜色。

您为某个可视化效果选择的选项将仅适用于该特定的可视化效果。

特定于可视化效果的选项

第一个标签可设置特定于当前可视化效果的选项。例如，条形图将允许您设置条形的颜色并高亮显示颜色，词云将允许您选择用于各个词的调色板，等等。此外，大多数可视化效果还将允许您选择图表背景颜色。

要更改颜色，请单击要更改的颜色，然后从拾色器中选择新颜色。您可以从预定义的调色板中选择、根据渐变进行选择，或者输入 RBGA 值。

字体选项

为当前的可视化效果选择文本样式。不同的可视化效果具有不同类型的文本，但一般情况下，您可以更改字体、大小、样式和颜色。

注释

为可视化效果中包含的注释选择文本样式选项。

更改图表或可视化标题

在所有图表和可视化中，您都可以对显示的标题进行自定义。如果您要在报告中包含图表或图形可视化，此功能特别有用。

1. 在图表或图形中，单击默认标题
2. 编辑该文本，将其改为您的自定义标题
3. 单击图表上的其他任意位置以保存所做的编辑

只要您未删除（关闭）可视化的标签，自定义标题会一直保留下来。

每种图表和可视化的说明

1. [柱状图](#)
2. [条形图](#)
3. [饼图](#)
4. [线图](#)
5. [词云](#)
6. [圆环图](#)
7. [气泡图](#)
8. [透视工具](#)
9. [甘特图](#)
10. [蝶形图](#)
11. [矩阵查看器](#)
12. [散点图](#)
13. [聚类图谱](#)
14. [定制化报告表](#)
15. [世界地图](#)
16. [绘制分类图](#)
17. [绘制新兴趋势图](#)

柱状图

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“添加表单”区域中的**柱状图**按钮。

描述：根据列表中的所有项或选定项显示垂直柱状图。列表中的项为 X 轴，记录的数量为 Y 轴。对于大型列表，您可能需要选择其中的前 10 项或前 20

项，因为具有成百上千个项目的图表可能没有意义。或者，您可以为列表的不同分段创建组（如前 10 项、特定范围内的记录项等），然后为这些组创建图表。**了解更多：**创建组

选择一个条柱可在[标题视图](#)中查看关于这些记录的细节。您还可以从选定列中导出字段化的记录或原始记录。了解更多：[导出字段化的记录](#) | [将原始记录导出为文本文件](#)

使用要求：列表必须已打开。您可以选择要包含在视图中的项，也可以不选中任何项从而在列表中包含所有项。

条形图

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“添加表单”区域中的**条形图**按钮。

描述：根据列表中的所有项或选定项显示水平柱状图。列表中的项为 Y 轴，记录的数量为 X 轴。

对于大型列表，您可能需要选择其中的前 10 项或前 20 项，因为具有成百上千个项目的图表可能没有意义。或者，您可以为列表的不同分段创建组（如前 10 项、特定范围内的记录项等），然后为这些组创建图表。**了解更多：** [创建组](#)

选择一个条柱可在[标题视图](#)中查看关于这些记录的细节。

使用要求： 列表必须已打开。您可以选择要包含在视图中的项，也可以不选中任何项从而在列表中包含所有项。

饼图

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“添加表单”区域中的**饼图**按钮。

描述： 根据列表中的所有项或选定项显示环状图。图表中的每块区域表示列表或选定内容中某个特定项的记录的数量（相对于该列表或选定内容中所有项的总记录数量而言）。

您可以使用三种类型的环状图对列表进行可视化呈现。

环状图类型	描述
饼图	显示完全实心的环状图。
双层环状图	显示中间区域空白的环状图。
智能饼图	显示完全实心的环状图，其中各分段的重叠部分将在中间以百分比显示。

对于大型列表，您可能需要选择其中的前 10 项或前 20 项，因为具有成百上千个项目的图表可能没有意义。或者，您可以为列表的不同分段创建组（如前 10 项、特定范围内的记录项等），然后为这些组创建图表。**了解更多：** [创建组](#)

选择一个区域可在[标题视图](#)中查看关于这些记录的细节。

使用要求： 列表必须已打开。您可以选择要包含在视图中的项，也可以不选中任何项从而在列表中包含所有项。

线图

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“添加表单”区域中的**线图**按钮。

描述： 根据列表中的所有项或选定项显示线图。图表上的每个点表示列表中的一个项，图表上每个点的高度（Y 轴）表示该项的记录数。

对于大型列表，您可能需要选择其中的前 10 项或前 20 项，因为具有成百上千个项目的图表可能没有意义。**了解更多：** [创建组](#)

选择一个点可在[标题视图](#)中查看关于这些记录的细节。

使用要求： 列表必须已打开。您可以选择要包含在视图中的项，也可以不选中任何项从而在列表中包含所有项。

词云

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“添加表单”区域中的**词云**按钮。

描述：显示列表中的项在您数据集中的流程度视图。这对于单个词汇或短语（例如 IPC 词目或专利权人）非常有用。您可以为列表中的选定项创建视图，或者不选中任何项从而为整个列表创建视图。

视图中每个项的大小与其在您数据集中的出现频率成正比。较大的项出现频率更高；较小的项出现频率较低。

使用要求：列表必须已打开。您可以选择要包含在视图中的项，也可以不选中任何项从而在列表中包含所有项。

使用说明

- 将鼠标悬停在视图中的一个词上，可查看有多少个记录包含该项
- 视图中词的大小和排列会随着您调整 Derwent Data Analyzer 的大小（或调整 Derwent Data Analyzer 内面板的大小）而动态变化
- 如果其中包含超过 50 项，会显示一则警告消息，指出因为项过多而会降低视图的实用性

圆环图

描述：创建一组不同类别的环状图（圆环）。

要求：已打开数据集，其中至少包含一个字段，且该字段中包含的项不超过 7 个。

创建圆环图

1. 确保数据集至少包含一个字段，且该字段中包含的项不超过 7 个
 2. 单击报告功能区，然后单击圆环图
 3. 选择各圆环代表的**字段**；（可选）如果该字段具有多个唯一项，那么可以选择一个组（例如，选择一个前 20 排名的组）
 4. 选择要在每个圆环中用来表示类别的**字段**；类别用于对图表进行细分
- 注意：**所选字段包含的不重复项数不能超过 7 个
5. 单击**运行**

使用示例

您想要调查某些组织机构是否在某些特定市场中提交的申请更多。您可以为这些市场创建组，例如，北美、APAC、EMEA，然后针对数据集当中排名靠前的组织创建圆环图，以查看这些组织的申请重点。

1. 为包含专利授权机构的字段（例如，申请国家/地区）打开字段列表
2. 在列表中为上述市场区域创建组

了解更多：创建组

3. 将授权机构从列表添加到各组中

了解更多：在列表视图中添加项至组

4. 由组名创建新字段；这样将会提供含 3 个项的列表，以供您用于对各圆环进行细分

了解更多：由组名创建字段

5. 使用上述步骤创建圆环图，并使用以下字段：

- **对应每个图表的字段：**来自清理后的专利权人列表中排名靠前的专利权人
- **对应类别的字段：**您根据组名创建的字段

圆环图中包含的各个圆环，对应于排名靠前的专利权人组中的每个专利权人。每个圆环都将包含对应您创建的市场区域（例如，北美、APAC、EMEA）的段。

使用说明

- 单击图表标题可对其进行编辑
- 悬停于圆环或圆环段上可查看记录数量
- 单击圆环或圆环段可高亮并为这些记录更新“标题”和“细节”窗口

为圆环图创建和加载模板

（可选）您可以保存一组字段用作圆环图的模板。模板可帮助您快速创建常用圆环图。

1. 单击**报告功能区**，然后单击**圆环图**
2. 选择代表各圆环和类别的对应字段，如上所述
3. 输入模板的**名称**
4. 单击**保存模板和运行**以保存模板并查看圆环图，或者单击“**保存**”仅保存模板

保存模板后，您可以在创建圆环图时，单击“选择模板”下拉列表以选择保存的字段。

气泡图

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“**添加表单**”区域中的**气泡图**按钮。

描述：一个用气泡表示所选两个字段中数字的图表。较大的气泡表示数值较大。

（可选）您可以重新命名图表。单击图表标题，然后输入图表标题的文本。

单击气泡，可在“标题”窗口和“细节”窗口中查看这些专利的信息。

提示：您可以从选定的气泡中导出字段化的记录或原始记录。了解更多：导出字段化的记录 | 将原始记录导出为文本文件

使用要求：数据集必须已打开。

使用示例：根据最早优先权年份，以可视化图表形式显示排名前 10

位的企业每年申请专利的记录数，以发现企业进入市场的时间及其专利申请量的增长情况。

气泡图选项

创建气泡图时，必须选择以下选项：

选项	描述
----	----

加载模板	可以选择一个保存的模板以将这些字段选择应用到您的气泡图。
选择行字段	<p>选择要在图表中用作行（气泡）的字段。您可以选择任何字段，然后可以分析该字段中的所有项，该字段中的所有组，或该字段中的选定组。</p> <p>请注意，在处理包含大量项的字段或组时需要花费较长时间。</p>
选择列字段	<p>选择要在图表中用作列的字段。您可以将其限制为选择显示年份的字段，显示数字的字段，或者选择数据集中的任何字段。</p> <p>对于所选字段，您可以选择分析字段中的选定组、所有组或所有项。</p>
运行/保存模板	运行（创建）气泡图，且/或（可选）输入名称并将选定选项保存为模板。

气泡图创建界面上所选的默认选项是您在上一个气泡图中所用的选项。

气泡图模板

您可以将气泡图上字段的当前选择保存为模板，以便快速创建一致的报告。创建模板以后，您可以在任何数据集中使用该模板。

此外，您在打开气泡图创建界面时选择的选项与您为上一个气泡图选定的选项相同。

要创建气泡图模板，请单击功能区上的**气泡图按钮**，打开该图并按常规选择各个选项。然后，输入**模板名称**并单击**保存模板并运行**（同时执行保存模板和显示图表的操作），或者单击**保存**仅将所选内容保存为模板。

要在创建气泡图时使用模板，请单击**选择模板下拉菜单**，并选择您要使用的**模板**。所选选项将更改为模板中保存的选项。

按字段细分气泡（气泡饼图）

（可选）您可以按特定字段对气泡图中的气泡进行细分。例如，您可能想要将每个气泡进行细分，以了解其中的有效项、失效项和未确定项。

VizLink 图

描述：可为各种字段同时显示三种可视化的视图。可视化选项包括：表格、柱状图、条形图、饼图、折线图或世界地图。

要求：您必须打开一个数据集。

创建 VizLink 图

1. 单击**报告**标签，然后单击 **VizLink 图按钮**
2. 从每个下拉框中选择**字段**
3. 单击每个字段所需图表类型的图标；选项包括表格、柱状图、条形图、饼图、折线图或世界地图（用于包含国家/地区数据的字段）

复查并浏览 Vizlink 图

选择左侧图中的项目，以更新右侧图中显示的数据。将鼠标悬停在每个图中的数据点上，以查看与该数据点关联的记录数。

单击右侧图中的数据点，以查看[标题视图](#)中的相关记录。单击地图中的国家/地区可放大。双击地图可缩小。

透视工具

描述：可供您编辑的透视表，最多包含 10 个字段。

要求：数据集已打开。

创建透视表

1. 单击**报告**标签，并单击**透视工具**按钮
2. 最多选择 10 个要用于行、列和透视的字段（用于对行进行细分的字段）；您可以选择完整的字段或组
3. 单击“确定”以创建透视表；如果所选字段包含大量项，这可能需要一些时间

复查并浏览透视表

单击图表标题可对其进行编辑

您可在灰色框之间拖放自己的字段

- 透视表上方灰色框中的字段可用于确定列
- 透视表左侧相邻的灰色框中的字段可用于确定行
- 行框左侧灰色框中的字段在透视表中不可见，但可供将来使用；您可以利用此空间来减少分析的字段数量

如果透视工具字段选择对话框中选定的字段并未全部达到 100%

覆盖率，那么如果通过拖放操作从透视表中移除了字段，则表中显示的记录数与标题和细节视图中的记录数可能会不一致。原因在于这个记录数是根据 10 个字段的三维共现矩阵中的记录来进行计数的。

单击任意字段可选择该字段中哪些项会显示在图表中（应用过滤器）；如果未选中其中所有项，那么字段名称会以斜体显示。

单击表格中的某个单元格可在[标题视图](#)和[细节窗口](#)中查看相关记录。

行字段上方和列字段左侧的灰色框用于控制表格中数字的显示方式以及排序顺序。

- 单击下拉菜单可将数字显示为总记录数或百分比；
 - 百分比的计算是基于 Y 轴上的项（即图表左侧面板中的字段），作为 X 轴上的项（图表上方面板中的字段）的函数计算得出
- 单击上下箭头可对列进行排序；单击左右箭头可对行进行排序

左上角的下拉菜单用于选择您的透视数据视图。下表描述了可用的选项。

选项	描述
----	----

表	显示透视表的标准表格视图
---	--------------

热图	将同样的数据显示为表格，但会基于记录频次标记对单元格上色
----	------------------------------

基于透视表创建条形图	
------------	--

条形图	每个条形代表用于行的最低级别的字段。条形按用于表格中的列的字段进行分组。 选择任一条形，即可使用该条形中所含记录来更新“标题”视图和“细节”视图。单击图例中的某个项，即可在图中的所有组中隐藏或显示该条形。
-----	---

堆积条形图	显示的数据与条形图相同，但会将每个组中的多个条形依次堆积起来。
-------	---------------------------------

表格条形图	在透视表的每个单元格中显示的条形指示单元格中表示的信息所对应的记录的相对数量。
-------	---

线图	根据图表中的数据创建线图。当您将序列数据（例如优先权年份信息）用作透视表中的行时，最好使用线图。
----	--

散点图	创建散点图可帮助您可视化呈现行中数据的计数分布情况。
-----	----------------------------

多饼图	为您选择的数据显示多个饼图。列中的数据用于分别创建各个饼图。行中的数据用于创建饼图中的不同部分。
-----	--

Excel	
-------	--

导出格式	设定数据透视图格式，以便您可以轻松地将数据复制并粘贴到 Microsoft Excel 中。
------	--

甘特图

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“添加表单”区域中的**甘特图**按钮。

描述：显示机构甘特图，列示每个公司活跃的年份范围。允许使用两个年份字段 –

通常是优先权年份和同族专利年份。甘特图保存在 Derwent Data Analyzer

文件中。单击图表上的各个部分可在“标题”视图及任何“细节”窗口中显示其相关专利。使用向上和向下箭头可对甘特图行进行排序。

要创建图表，请执行以下步骤：

1. 选择一个组织类型字段（例如“专利权人/申请人”）
2. （可选）选择要分析的对象：“字段中的所有项”、“字段中的所有组”或“选择组”

提示：让您的图表重点聚焦在一小部分组织，例如，排名前 10 的企业

3. 选择年份字段；您可以使用一个或两个字段

提示：选择两个年份字段来比较优先权日期及公开日期或授权日期

（可选）您可以重新命名图表。单击图表标题，然后输入图表标题的文本。

使用要求：在访问工具之前，您必须创建了任意组。**了解更多：**创建组至少需要提供一个年份字段。两个选择中可使用相同的年份字段。

蝶形图

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“添加表单”区域中的**蝶形图**按钮。

描述：显示蝶形图，可帮助您根据所选字段中的信息深入而详细地比较列表中的两个项。您只能从列表中选择两个项进行此比较。

通过此图表中显示的条形，可根据选定字段中的数据比较选择的两个项。对于您在字段中选择的每个项，此图表中都包含一个相应行。每行中包含两个条形，这两个条形从中间轴水平延伸。查看任一行中两个条形的长度，即可了解所选的两个项之间的异同

要创建图表，请执行以下步骤：

1. 从列表中仅选择两个项（例如两个专利权人）
2. 单击功能区中的“报告”标签，然后单击“蝶形图”
3. 选择要比较的字段；对于包含大量项的字段，您可能想要使用较小的组
4. 单击“确定”即可生成蝶形图

示例

您想要比较数据集中前两名终属母公司所使用的标题词。打开“终属母公司”类别，然后从该列表中选择前两个项。单击**“蝶形图”**按钮并选择要使用的字段，例如“标题词 - DWPI”。单击**“确定”**，并等待图表生成。

对于您在该字段中选择的每个项，此图表中都会显示一个行。对于您选择的两个终属母公司，将显示两个列。每行中的水平条形将反映每个终属母公司所具有的记录的相对数量（相对于该行中相应项所具有的记录总数）。

矩阵查看器

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“添加表单”区域中的**矩阵查看器**按钮。

使用要求：您要以可视化效果呈现的矩阵必须已打开。（可选）您可以选择特定的单元、行和/或列。矩阵查看器将以可视化效果显示选定的信息（而不是矩阵中的所有信息）。

描述：在一个矩阵中显示各项之间关联强度的可视化呈现效果。例如，您可以探索不同企业之间的共同专利权人关系，或者按照不同的组织机构（按 IPC-4 字符码/DWPI 分类绘制的专利权人矩阵），以可视化形式呈现这些组织机构所研究的主要技术领域。

节点（圆圈）在矩阵中表示项（各个行和列）。节点大小表示包含该项的记录数。

橙色节点表示矩阵中的行。在具有两个不同字段的共现矩阵中，蓝色节点表示列。

提示：可视化图表的标题以相应颜色（橙色或蓝色）显示字段名称。

可以对节点执行交互操作：

- 将光标悬停在节点上可高亮显示到该节点的所有连接
- 单击一个节点可在标题视图中查看这些记录
- 单击并拖动节点可重新定位

各节点之间的连线表示这些节点之间的关联强度。连线的粗细表示关联的强弱。在您移动相关节点时，连线会自动随之移动。

矩阵查看器布局

矩阵查看器包含两个布局，可帮助您以可视化效果呈现矩阵。

布局	描述
力导向	显示一个互联网络，以可视化效果呈现矩阵中各项之间的关系。最初，节点会根据各项之间的关联强度自动排列，但是您可以根据需要重新排列节点顺序。 这是矩阵查看器中的默认布局。在分析单个字段时，最好是使用力导向布局（例如，同一专利中最常出现的分类代码）。
圈层式	在一个圈（或圈的一部分）中显示所有数据点，以可视化效果呈现矩阵中各项之间的关系。这种方式可以为每个节点提供一致的位置，而不会与其他节点重叠。 只有在分析行和列都使用相同字段的矩阵时，圈层式布局才可用（例如，合作研发专利的专利权人）。
多栏式	在不同栏中为矩阵中分析的每个字段显示各自的项。节点最初按记录数以降序显示。 在分析两个不同的字段时，最好是使用多栏式布局（例如，专利权人和优先权国家/地区）。

矩阵查看器过滤器

矩阵中可能包含仅在几个记录中出现的项，或关联非常弱的项。您可以通过过滤器设置最低阈值，仅允许达到该阈值的项或关联在视图中显示。这些阈值可帮助您仅在视图中显示最重要的数据。

要设置过滤器，请单击滑块并拖动，为该过滤器设置一个最低记录数

过滤器	描述
橙色圆圈	设置在矩阵查看器中，矩阵行中的某项要作为一个节点显示时所要求达到的最低记录数。
灰线	设置在矩阵查看器中，某项关联要作为一条连线显示时所要求达到的最低记录数。
蓝色圆圈	设置在矩阵查看器中，矩阵列中的某项要作为一个节点显示时所要求达到的最低记录数。

散点图

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“**添加表单**”区域中的**散点图**按钮。

描述：使用时间滑块创建散点图。散点图可显示一些点，来代表一个字段中的项数；通过将该字段与分别位于 x 轴和 y

轴上的最多另两个字段包含的项数进行对比绘制而成。点的大小表示项的相对数量。

各轴上的值可使用唯一项的总数，也可以使用每条记录唯一项的平均数目。时间滑块会计算 x 轴和 y 轴的值，直至选定的年份为止。

使用要求：数据集必须包含至少一个年份字段。

使用说明

- 将鼠标悬停在视图中的一个点上，可查看有多少个记录包含该项
- 选择一个点可在[标题视图](#)中查看关联记录
- 使用日期滑块可显示分析随时间的变化情况

聚类图谱

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“**添加表单**”区域中的**聚类图谱**按钮。

描述：显示动态图谱，以展示列表中各个项之间的连线和重叠情况。例如，您可能想要分析某个组织的主要发明人，从而确定其研究团队。

您可以根据组或整个列表创建聚类图谱。

此聚类图显示所选字段或组中的每一项（例如，发明人），且以不同颜色标识以方便查看。与每一项连线的“节点”内的黄色球表示与该项关联的记录数（例如，专利数量）。节点类型分为两种：

- **单节点**包含与单个项目连线的记录。单节点总是与单个项目连线。
- **组节点**包含由多个项目组合创建的记录。组节点可能与两个或以上的发明人连线。

项目与节点之间通过连线相互关联。一个项目可能与一个或多个节点连线；同样地，一个节点也可能与一个或多个项目连线。

有时，两个（或更多）项目会连线到同样几个节点中的若干节点（但是这些节点彼此并不相同）。几个项目以不同的组合方式与不同的记录相关联，这样就会产生互连节点的一种“**聚类**”。

使用要求：该小程序可能需要较长一段时间来加载大型列表。

参见聚类图谱示例：业务需求：挖掘创新团队

查看和操作聚类图谱

将鼠标悬停在某个项的标签上方，可高亮显示与该项关联的节点。单击某个项的标签，可在聚类图谱中突出显示与该项相关的节点；其他节点和项标签在图谱中则会淡化，以突出您要重点关注的內容。

当您重点关注某个特定项的标签时，还可以将鼠标悬停在另一个标签上方，以突出（而不是淡化）显示另一个项的相关节点。这样可帮助您以可视化效果呈现这两项之间的关联情况（例如两个发明人之间的合作关系）。

将鼠标悬停在某个节点上可高亮显示与该节点关联的所有项标签。单击某个节点，可在聚类图谱中突出显示该节点；其他节点和项标签在图谱中则会淡化，以突出您要重点关注的內容。

您可以选择多个节点。按住键盘上的 Ctrl 键，并单击要选择的节点。

单击并拖动节点或项标签，可重新调整它们在聚类图谱中的位置。这样，您可以对聚类图谱进行调整，从而更加轻松地评估关联情况。

提示：您可以从选定的节点中导出字段化的记录或原始记录。了解更多：导出字段化的记录 | 将原始记录导出为文本文件

聚类图谱过滤器

您可以选择按不同的方式对聚类图谱进行过滤，以聚焦在特定的方面。

过滤器 描述

设置与要在图谱中显示的项标签相关联的最小和最大节点数量。

标签节点范围 例如，您可能想要重点研究节点数超出特定阈值的项目。或者，您可能想要将节点数过多的项目排除在分析范围之外（例如，在一个发明人聚类图谱中，将每个专利上都出现的一位研发主管人排除在外）。

这些数值与项目列表旁显示的数值相同。调整此过滤器将重置其他数字型过滤器。

记录节点范围 设置必须要包含在图谱中的节点所对应的最小和最大记录数。

例如，您可能不想评估只有一个记录的节点。

设置要在图谱中显示的节点所对应的最小和最大连线数。

网络链接范围 例如，您可以将连线的最小值设为

2，以突出显示在两个项目之间存在重叠的记录（例如，移除只列出了一个发明人的专利）。

显示/隐藏单个项 选中单个项名称旁的复选框，可在图谱中显示/隐藏这些项。

例如，您可能想要隐藏这样的小型发明人团队：该团队与您的其他过滤器相匹配，但仅在一个项内部开展相互合作（而不与公司其他部门合作）。

定制化报告表

此报告可通过单击功能区上的**报告**标签访问；单击“添加表单”区域中的**定制化报告表**按钮。

描述：创建一个交互式图形视图报告表，作为 Derwent Data Analyzer

文件中的一个表单或作为一个 Excel

文件。该表中的行可以是任何字段，但通常会选择组织机构（专利权人）或个人（发明人）。还可创建模板用于经常运行的配置文件。

了解更多信息：定制化报告表

使用要求：已打开数据集。

世界地图

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“添加表单”区域中的**世界地图**按钮。

描述：显示最多两个世界地图（一个地图在另一个地图之上），以说明数据集中的专利是在哪些国家/地区公开的。地图中使用了多种颜色进行区分，以指明在每个国家/地区所公开专利的相对数量。单击某个国家/地区可在“标题”视图中显示该国家/地区授予的专利。

一份地图会始终显示。“优先权国家/地区”为默认字段，但是您可以选择任何可用于此地图的字段：“申请国家/地区”、“基本专利国家/地区”、“同族专利国家/地区”、“优先权国家/地区”、“最早优先权国家/地区”、“公开国家/地区代码”。

您也可以选择第二个字段，以使用不同的国家/地区字段选项显示另一份地图。通过这种方式可以比较这两个字段之间的覆盖率（例如优先权国家/地区 and 同族专利国家/地区）。

非政府机构（例如，WIPO，EPO）在地图上由如下符号表示：

组织机构

符号

ARIPO（非洲地区工业产权组织）



CDVA



EPO（欧洲专利局）



欧亚专利组织



国际专利协会



OAPI（非洲地区知识产权组织）



国际技术公开



WIPO（世界知识产权组织）



使用要求：数据集必须已打开。

绘制分类图

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“**添加表单**”区域中的**绘制分类图**按钮。

描述：显示可帮助浏览指定给选定文献的分类代码的视图。您可以选择视图的布局和要在该布局中呈现的数据。

使用要求：数据集必须已打开。

布局

有三种布局可帮助浏览分类代码。这些布局的功能相似，但显示信息的方式不同。所有这些布局包括：

- 以层级结构的方式显示选定文献中的代码或符号
- 使用不同的颜色呈现分类的不同部分
- 在每一部分中，颜色越浅，呈现的分类越细致、详尽
- 使用大小表示某种分类的文献数量
- 显示特定分类的记录数量以便查看包含该分类及其定义的文献数量
- 允许单击某个区域以便在标题视图和细节窗口查看包含该分类的文献

每个视图的描述如下：

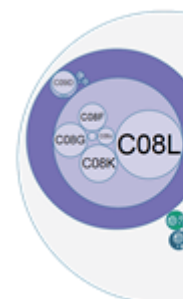
视图

描述

示例

聚类
圆
群

在同心圆中显示分类。较大的圆表示级别较高的分类，而这些圆中的小圆则表示级别较低的相关分类。
每个圆的大小表示选定内容中包含该分类的文献数量。



树形图 在表格中显示分类。每种颜色的表格标题表示较高级别的分类，各个单元表示分析中最低级别的分类。
单元（或标题和相关单元）的总面积表示包含该分类的文献数量。



光谱图 将分类符号显示为环状图的圆环和切片。内部的圆环表示较高级别的分类，外部的圆环表示较低级别的分类。由深（内部）到浅（外部）的颜色变化可帮助显示分类的层级结构。
同一圆环中每个切片相对于其他切片的面积表示包含该分类的文献数量。



放大和缩小

要放大聚类图谱的某一部分，请双击相应区域。若要缩小，请双击上一级区域。例如，单击"聚类圆群"中较小的圆圈可放大，而单击较大的圆圈可缩小。
您可以随时双击标题栏以重置为默认图谱。

数据

选择要在视图中呈现的字段和字段表示的分类类型。“字段”下拉列表中的数据选项取决于数据集中包括的字段。目前，Derwent Data Analyzer 可以处理：

- CPC 码（当前、原始和/或主要）
- DWPI 分类代码
- DWPI 手工代码
- 完整 IPC 码
- 小类（四个字符）级别的 IPC 码
- IPC 词目

您可以在分析中选择包括定义。包括定义可以更深入地分析选定分类。

警告：字段和分类类型必须选择相同的分类类型（CPC 或 IPC）。

绘制新兴趋势图

要访问此可视化效果，请单击功能区上的**报告**标签，然后单击“添加表单”区域中的**绘制新兴趋势图**按钮。此外，在计算新兴趋势指标时，可以自动创建新兴趋势图。

您必须曾使用[计算新兴趋势指标](#)生成过新兴趋势统计数据，然后才能使用这些统计数据绘制图表。

描述：显示四个散点图，分别是主题词、组织、人员、位置散点图。这四个散点图可根据这四个类别中每个项的[新兴趋势得分](#)和记录数，展示这些类别中排名最前的项之间的关联。每个项（主题词、组织、人员、位置）在散点图中显示为点。

每个散点图允许您选择为相应类别计算的新兴趋势指标。例如，如果您根据清理后的专利权人列表和优化的专利权人计算了新兴趋势指标，则可以选择要在散点图中为组织显示哪个指标。

选择散点图中的任意项后，可在标题窗口中查看记录或在细节窗口中查看其他信息。

导出图像

您可以将您在 *Derwent Data Analyzer*

中创建的图表和图形快速导出为图像，以用于向团队或高层管理人员提交的报告。如果要将这些图像用在打印版的材料中，您还可以选择导出它们的高分辨率版本。

1. 创建可视化的分析图（例如柱状图、饼图、世界地图等）
2. 单击功能区中的**数据整理**标签，然后单击“导出”区域中的**导出图像**
3. 选择导出的格式
4. 选择导出文件的保存位置

导出图像格式

选项	描述
截图 (jpg)	将可视内容的当前可见部分保存为 JPG 文件。
导出完整图像 (png)	将可视内容的完整图像保存为 PNG 文件。
导出高分辨率图像	创建适合在打印版材料中使用的高分辨率 PDF。
导出 HTML	创建可视内容的 HTML 文件，其中包含任何可供使用的交互元素。例如，您可以在世界地图上放大，在饼图的不同环状图类型之间切换，等等。

公司比较报告

要求使用 Microsoft Excel 2003 或以上版本

Derwent Data Analyzer 具有三个强大的报告模块，可将数据变换成一系列 Microsoft Excel 工作表，随时准备用于分析人员的评述。“公司比较报告”将基于 2-5 家公司的记录组合生成详细的比较报告。

步骤 1：数据采集

公司比较报告要求您的数据集中包含要比较的公司的所有必要记录。

步骤 2：准备字段

公司比较报告要求您选择 5

个不同类别的字段。在运行报告之前，请确保数据集中的数据干净清晰。

选项	描述
公司（例如专利权人、组织机构）	报告会以最频繁出现的公司默认为目标公司。请确保其出现在所有记录中。这个字段中的所有其他名称都将被视为合作公司；如果将个人的姓名移除（例如，同样也列为专利权人的发明人），以及如果考虑到拼写以及重组和并购活动而将公司名称进行过清理，将得到最有效的结果。
人员（例如发明人、作者）	报告将识别最多产的人以及他们与谁共事。数据清理的重要性在于确保搜集到每个人的全部活动范围。
年份（例如优先权年份、公开年份）	当有不同的年份字段可用时（例如专利），字段的选择将对结果产生重大影响。例如，最早优先权年份显示了各个创新曝光的时间趋势，但也会反映在公开年份上会有 18 个月的延迟；同族专利公开年份虽然更近，但它体现的是在一个专利家族架构下，每个专利在整个专利系统而不是个别技术中的活跃度。选用不同字段多次生成不同报告，或许会有所帮助。
国家/地区（例如优先权国家/地区、公开年份）	需要注意所选字段的影响。例如，“优先权国家/地区”字段将更清楚地显示出发明所在地，而“同族专利公开年份”或“指定国家/地区”则会显示专利所有人在哪个地区寻求或没有寻求专利权保护。选用不同年份或国家/地区的排列组合多次生成不同报告，或许会有所帮助。
技术（例如 NLP 词组、分类）	根据您希望从报告中获取的细节程度而定，可以选择精细技术字段（如手工代码、单词、词组）、适中技术字段（如德温特分类，IPC 4 位代码）或粗略技术字段（如聚类单词、词组）。对于分类方案，建议先运行一个叙词表，将代码转换成有意义的文本。

步骤 3：运行宏指令

单击功能区中的**报告**标签，然后单击“Derwent Data Analyzer

报告”区域中的**公司比较**按钮。根据记录的数量和所选字段中的主题词数量而定，通常需要 5 到 10 分钟来完成报告（例如德温特分类就比手工代码运行得更快）。

步骤 4：解释结果

所得输出结果将是一系列 Microsoft Excel

工作表，其中包含各种图表和视图报告表。在这些图表背后的数值数据全部被包含在工作簿中，

但处于隐藏状态，可让您对最终结果进行定制。要取消隐藏数据，请在 Excel 的菜单中，依次选择**格式、工作表、取消隐藏**。

公司分析报告

要求使用 Microsoft Excel 2003 或以上版本

Derwent Data Analyzer 具有三个强大的报告模块，可将数据变换成一系列 Microsoft Excel 工作表，随时准备用于分析人员的评述。提供的“公司报告”将基于公司的记录组合生成详细的挖掘信息，以便查看合作者、技术趋势以及核心员工。

步骤 1：数据采集

公司报告要求：

- 包含目标公司所有数据的数据集（如果记录不全，则输出的报告内容可能有误导性）。
- 公司列表中排在第一位的即为目标公司，它出现在所有记录中（但仍然会有很多公司在列表中，因为它们与目标公司有合作关系）。

如果不是这样，请通过列表清理功能合并公司名称，或选择列表中排名最高的公司，单击功能区“主页”标签中的**创建子数据集**，然后单击对话框中的**确定**。这会产生一个新的只包含选定记录的 .vpt 文件。

步骤 2：准备字段

公司报告要求您选择 5

个不同类别中的字段。在运行报告之前，请确保数据集中的数据干净清晰。

选项	描述
公司（例如专利权人、组织机构）	报告会最频繁出现的公司默认为目标公司。请确保其出现在所有记录中。这个字段中的所有其他名称都将被视为合作公司；如果将个人的姓名移除（例如，同样也列为专利权人的发明人），以及如果考虑到拼写以及重组和并购活动而将公司名称进行过清理，将得到最有效的结果。
人员（例如发明人、作者）	报告将识别最多产的人以及他们与谁共事。数据清理的重要性在于确保搜集到每个人的全部活动范围。
年份（例如优先权年份、公开年份）	当有不同的年份字段可用时（例如专利），字段的选择将对结果产生重大影响。例如，最早优先权年份显示了各个创新曝光的时间趋势，但也会反映在公开年份上会有 18 个月的延迟；同族专利公开年份虽然更近，但它体现的是在一个专利家族架构下，每个专利在整个专利系统而不是个别技术中的活跃度。选用不同字段多次生成不同报告，或许会有所帮助。

国家/地区（ 需要注意所选字段的影响。例如，“优先权国家/地区”字段将更清楚地显示出发明所

例如专利权国 在地，而“同族专利公开年份”或“指定国家/地区”则会显示专利所有人在哪个地区寻家/地区、公 求或没有寻求专利权保护。选用不同年份或国家/地区的排列组合多次生成不同报开年份) 告，或许会有所帮助。

技术 (例如 NLP 词组、分类) 根据您希望从报告中获取的细节程度而定，可以选择精细技术字段 (如手工代码、单词、词组)、适中技术字段 (如德温特分类，IPC 4 位代码) 或粗略技术字段 (如聚类单词、词组)。对于分类方案，建议先运行一个叙词表，将代码转换成有意义的文本。

报告向导将基于主数据类型，为每个类别选择一个标准字段集。

步骤 3：运行宏指令

单击功能区中的**报告**标签，然后单击“Derwent Data Analyzer 报告”区域中的**公司报告**按钮。根据记录的数量和所选字段中的主题词数量而定，通常需要 5 到 15 分钟来完成报告 (例如德温特分类就比手工代码运行得更快)。

步骤 4：解释结果

所得输出结果将是一系列 Microsoft Excel

工作表，其中包含各种图表和视图报告表。在此背后的数值数据全部被包含在工作簿中，但处于隐藏状态，可让您对最终结果进行定制。要取消隐藏，请在 Excel 的菜单中，依次单击**格式、工作表、取消隐藏**。

技术分析报告

要求使用 Microsoft Excel 2003 或以上版本。

Derwent Data Analyzer 具有三个强大的报告模块，可将数据变换成一系列 Microsoft Excel 工作表，随时准备用于分析人员的评述。提供的“公司报告”将基于公司的记录组合生成详细的挖掘信息，以便查看合作者、技术趋势以及核心员工。

步骤 1：数据采集

技术报告假定数据集包含了所选技术的全部必要记录。

步骤 2：准备字段

公司报告要求您选择 5

个不同类别中的字段。在运行报告之前，请确保数据集中的数据干净清晰。

选项	描述
----	----

公司 (例如专 利权人、组织 机构)	报告会最频繁出现的公司默认为目标公司。请确保其出现在所有记录中。这个字段中的所有其他名称都将被视为合作公司；如果将个人的姓名移除 (例如，同样也列为专利权人的发明人)，以及如果考虑到拼写以及重组和并购活动而将公司名称
--------------------	--

进行过清理，将得到最有效的结果。

人员（例如发明人、作者） 报告将识别最多产的人以及他们与谁共事。数据清理的重要性在于确保搜集到每个人的全部活动范围。

年份（例如优先权年份、公开年份） 当有不同的年份字段可用时（例如专利），字段的选择将对结果产生重大影响。例如，最早优先权年份显示了各个创新曝光的时间趋势，但也会反映在公开年份上会有 18 个月的延迟；同族专利公开年份虽然更近，但它体现的是在一个专利家族架构下，每个专利在整个专利系统而不是个别技术中的活跃度。选用不同字段多次生成不同报告，或许会有所帮助。

国家/地区（例如优先权国家/地区、公开年份） 需要注意所选字段的影响。例如，“优先权国家/地区”字段将更清楚地显示出发明所在地，而“同族专利公开年份”或“指定国家/地区”则会显示专利所有人在哪个地区寻求或没有寻求专利权保护。选用不同年份或国家/地区的排列组合多次生成不同报告，或许会有所帮助。

技术（例如 NLP 词组、分类） 根据您希望从报告中获取的细节程度而定，可以选择精细技术字段（如手工代码、单词、词组）、适中技术字段（如德温特分类，IPC 4 位代码）或粗略技术字段（如聚类单词、词组）。对于分类方案，建议先运行一个叙词表，将代码转换成有意义的文本。

报告向导将基于主数据类型，为每个类别选择一个标准字段集。

步骤 3：运行宏指令

单击功能区上的**报告**标签，然后单击“Derwent Data Analyzer

报告”区域中的**技术报告**按钮。根据记录的数量和所选字段中的主题词数量而定，通常需要 10 到 15 分钟来完成报告（例如德温特分类就比手工代码运行得更快）。

步骤 4：解释结果

所得输出结果将是一系列 Microsoft Excel

工作表，其中包含各种图表和视图报告表。在此背后的数值数据全部包含在工作簿中，但处于隐藏状态，可让您对最终结果进行定制。要取消隐藏，请在 Excel 的菜单中，依次选择**格式、工作表、取消隐藏**。

记录视图

通过双击标题视图中的标题，访问**记录视图窗口**。该**记录视图窗口**会显示一个或多个记录。

以下按钮按以下顺序从左到右显示在该视图的顶部：

选项	描述
----	----

管理记录分类	打开 管理记录分类 界面，您可以在其中添加、编辑或删除分类和类别。您还可以访问自动分类设置，并从该界面启动 Smart Trainer。
--------	--

自动分类设置	微调自动分类算法的函数。了解更多： 自动分类设置
--------	--

自动分类记录	开始进行自动记录分类。了解更多： 自动分类记录
--------	---

查找相似记录	在项目文件中查找与当前记录类似的记录，然后按相似度最高到最低的顺序将其显示在新窗口中。将根据选定字段的内容向每个记录分配相似度百分比。 了解更多： 查找相似记录
--------	--

打印	以当前视图格式（“原始”或“字段”）打印记录。
----	-------------------------

复制	复制记录中高亮（选中的）部分到剪贴板（用于粘贴到另一个应用程序）。
----	-----------------------------------

全选	选择整个记录（用于复制到剪贴板）。
----	-------------------

查找	打开 查找 对话框，查找文本。
----	------------------------

显示原始记录	从 字段 视图切换到 原始记录 视图。
--------	-----------------------------------

显示字段	从 原始记录 视图切换到 字段 视图，显示分析的字段。
------	---

字段顺序	打开“ 字段顺序 ”对话框。使用此对话框可更改“ 字段视图 ”中的 字段 显示顺序。 了解更多： 选择记录视图和导出字段并排序
------	--

在“唯一 ID”字段中添加组	在选定的“ 唯一 ID ”字段中将当前记录添加到组。
----------------	-----------------------------------

文本换行	对文字进行换行。仅适用于原始记录视图。
------	---------------------

固定字体	以固定宽度的字体显示记录。这会改进某些记录格式的文字清晰度。
------	--------------------------------

高亮关键词	如果用户创建了“ 我的关键词 ”列表，则此按钮将打开或关闭“ 记录视图 ”中对那些主题词的高亮显示。
-------	--

设定关键词颜色	允许您将所选颜色分配到“ 关键词列表 ”中的主题词。
---------	-----------------------------------

查看先前 在标题视图中显示上一个记录（如果在标题视图中选中了多个记录，此按钮不可用）记录。

查看下一 在标题视图中显示下一个记录（如果在标题视图中选中了多个记录，此按钮不可用）条记录。

退出 关闭显示记录。

在显示记录窗口的底部，会出现如下内容：

- **关于此记录的备注：**（也称为“可编辑备注”，用于某些导入/导出功能。）为查看的记录添加注释。添加注释的同时还会在标题视图中的该记录旁边添加一个图标。
- **从新数据集中移除：**将查看的记录标记为在创建新数据集时移除。标记要移除的记录时，还会在标题视图中的该记录旁边添加一个图标。

注意：“移除”功能仅在您随后执行创建新数据集的操作时起作用（例如：创建子数据集、导出原始记录、导出字段化的记录、移除冗余记录，以及数据集合并）。对于创建子数据集、导出原始记录、导出字段化的记录而言，会出现一个复选框选择是否移除标记的记录。对于移除冗余记录和数据集合并功能而言，如果在操作中有任意记录被标记为“从新数据集中移除”，您将看到如下确认请求：“本操作包含已被标记为丢弃的文本。您是否想要丢弃这些记录？”如果回答是，那么被标记的记录将会从新数据集中移除。如果回答否，那么“移除”标记将被忽略。

您可以通过双击标题视图中的另一个标题查看另一个记录，而不需要关闭记录视图。您也可以通过使用**上一个/下一个**箭头来浏览标题视图中的每个记录。

另请参见

- 如何显示记录
- [标题视图](#)

DDA 使用案例

业务需求：顶级企业分析

您想了解的问题

- 基于专利组合的规模，哪些企业是您所关注领域内的顶级企业？
- 这些顶级企业是否代表了您所关注领域内的专利分布全景？
- 是否有隐藏的企业已做好引进专利授权或收购的准备？
- 您关注领域中的专利持有人集中在哪些地方？

您在哪里能得到答案

在垂直柱状图中，展示了排名前 10 的企业（基于专利组合的规模）及他们各自拥有专利的数量：

而在专利权人分组的环状图中，展示了每个组所持有专利数量占总数的百分比：

如何创建视图

假设：您所使用的数据集与您希望分析的所有公司和所有技术紧密相关。

创建垂直柱状图

1. 从总览界面打开一个清理过的专利权人列表。

了解有关清理列表的更多信息：数据清理简介

2. 针对要分析的每个类型的专利权人分别创建组。在本例中，分别为企业、学术机构、政府机构和个人专利权人创建组。

了解有关创建组的更多信息：组

提示：您可以使用基于叙词表的组工具（位于功能区“数据整理”标签的组区域中）快速创建企业、政府机构、学术机构或代表个人的组。运行叙词表后，请验证这些组以确保所有专利权人均正确分组。

3. 双击代表**企业**组的列。这样即可让企业组中的所有专利权人显示于列表顶部，并按照专利数量降序排列。
4. 为不同级别的企业专利权人创建新的组。您可以使用各种标准进行分级（排名前 10、每个企业持有的专利数量等）。本例中为排名前 10

的公司创建一个组。其余的企业专利权人则根据专利数量分组，包括一个组专门容纳专利组合中仅有一个专利的企业。

5. 单击代表**排名前 10 公司**组的列以选定该列，然后在工具栏上单击**垂直柱状图**按钮。

此时将显示一个垂直柱状图，根据您定义的参数显示排名前 10

的企业。将光标悬停在一个柱上即可看到各企业持有专利的确切数目。单击图表上的一个柱即可在标题视图中查看该企业的专利。

您可以使用屏幕截图程序，将该图表应用于演示文稿和报告中。

创建环状图

1. 返回到清理过的专利权人列表。
2. 单击功能区中的“数据整理”标签，然后依次单击**字段、由组名创建字段**，以便使用您之前所创建组的名称来创建新的列表。将显示一个对话框。
3. 选择要包含在图表中的组（每个层级的公司组、个人专利权人、政府机构以及学术机构），如果需要可将字段重命名，然后单击**确定**

警告：如果您将“公司”组划分为不同层级，那么请在新列表中仅包含这些不同层级的组，而非整个“公司”组。如果包括整个“公司”组，新列表将双倍计算这些公司所持有的记录数。

4. 此时将显示新列表。选择列表中的所有行并单击功能区中的“报告”标签，然后单击**环状图**按钮。

该环状图将显示您定义的每个组。每个组在环上都占有一个区域，表示其占数据集总记录数的百分比。将光标悬停在图表中的某个部分可显示该组所持有记录的确切数目。单击图表上的某个部分则可在标题视图中浏览该组的专利。

您可以保存该图像，或者将该图表应用于演示文稿和报告中。

如何获得答案

通过查阅上述柱状图以及对比环状图上各个组的区域，获得您需要答案。

上述柱状图按照专利组合的规模显示了排名前 10 的公司。

通过对比每个柱可以看出，在排名前 10

的公司中哪些公司拥有专利最多，这可以表明谁是该领域的顶级企业。第 1 名与第 10 名的公司之间如果相差悬殊，则可能表明该领域主要由少数几家大公司把持。

环状图则以占整个数据集总专利数的百分比方式，显示您所定义各个组分别持有多少专利。一个组在图表中占有的区域越大，则该组的专利数就越多。

通过对比各个组占有的区域，来获得您所关注的领域中专利分布全景的线索。

- 如果排名前 10 的公司控制着绝大多数的专利，那么这些公司可能会积极地保护自己的知识产权。您在申请专利的时候就要小心谨慎以防侵权诉讼。
- 如果专利主要由个体发明人和专利组合规模较小的公司所控制，那么这些专利持有人可能会寻找投资资金。因此这个组可能是引入专利授权或收购的“黄金目标”。

需要进一步分析的问题

分析是一个反复比较的过程。反复审视这些图表会提出额外的问题，许多此类问题可由 Derwent Data Analyzer 进行深度分析而得到解答。以下是此类问题中的几个示例：

- 您所关注领域的发展速度多快？是缓慢增长还是飞速增长？
- 在顶级企业当中，哪些企业更加积极地寻求新技术的研发？
- 在您所关注的领域中，顶级企业持有多少已经授权的专利而非专利申请？这说明他们现在与过去的工作重心是什么？
- 顶级企业在保护自己的知识产权方面是否成功？

业务需求：评估竞争者在寻求新技术方面的积极程度

您想了解的问题

- 在您关注的领域中，竞争对手们是如何积极地寻求获取新技术？

您在哪里能得到答案

通过堆积柱状图获得您需要答案。

堆积柱状图展示了一个企业的专利申请量与专利授权量的比例。在您关注的领域中，拥有大量专利申请的企业比拥有少量专利申请的企业在推动新技术研发方面更加积极。

如何进行操作

1. 打开清理过的专利权人列表
了解更多：数据清理简介
2. 针对您想了解的竞争对手创建一个组
了解更多：创建组
3. 打开公开号列表，并针对专利申请和授权专利分别创建组
提示：您可以通过“使用叙词表创建组”功能创建这些组。所述叙词表位于宏指令 (Macros) 文件夹中的资源 (Resources) 文件夹
4. 在“主页”标签中，单击**矩阵按钮**
5. 在“行”的部分，打开公开号列表并单击**选择组/显示组**。在“选择组”列表中，选择先前创建的专利申请组和授权专利组

6. 在“列”的部分，打开您使用的清理过的专利权人列表，并单击**选择组/显示项**。在“选择组”列表中，选择先前创建的竞争对手组

注意：“选择组/显示项”允许您在列表中选择您之前创建的组，但会在矩阵中列出该组内的每一个条目

7. 确保矩阵设置为基于记录数的共现矩阵，然后单击**确定**
8. 将显示一个以竞争企业为行、以专利申请组和授权专利组为列的矩阵
9. 选择矩阵中的所有行并单击功能区中的**分析**标签，然后依次单击**脚本、报告 - 矩阵输出到 Excel 并可视化**

此时将打开一个新的 Excel

工作表，其中包含来自该矩阵的数据。默认显示三维柱状图。将图表布局改为堆积列并调整布局，使其显示您想要的数据。

如何获得答案

堆积柱状图中的每个柱代表一个企业。在该柱中，以不同颜色的区域代表授权专利数量和专利申请量。每个区域的面积与该企业持有的授权专利数量或专利申请量成正比。

该图表提供了一种可视化呈现效果，表明每个企业目前拥有多少专利申请和授权专利。在您关注的领域中，专利申请数量越多则表示在追求技术创新方面越积极。

由于不同企业的专利总量都不尽相同，因此使用堆积柱状图（而不是使用 100%

堆积图）可以显示哪些企业持有的专利更多。每个区域上的数字可以显示该企业专利申请和授权专利的确切数量。

您可以复制并粘贴此 Excel 图表到报告和演示文稿中。

业务需求：找到学术界合作伙伴

您想了解的问题

- 哪些公司与学术机构建立了合作研究关系？
- 通过这些合作研发了哪些专利？

您在哪里能得到答案

共现矩阵可以列出专利权人为学术机构和企业的专利。

如何创建视图

假设：您所使用的数据集与您希望分析的所有公司和所有技术紧密相关。

1. 从总览界面打开一个清理过的专利权人列表。

了解有关数据清理的更多信息：数据清理简介

2. 针对企业和学术机构专利权人分别创建组。

提示：您可以使用“基于叙词表的组”工具（可在功能区“数据整理”标签的“组”区域中找到），快速创建企业、政府机构、学术机构和代表个人的组。确认这些组的准确性之后，您可以通过新的“公司”组对列表进行排序，以便基于您的标准（排名前10、专利数量等）快速为公司创建层级。

3. 单击“主页”标签，然后单击**矩阵按钮**。
4. 在行的部分，打开您之前使用的清理过的专利权人列表，并单击**选择组/显示项**。在“选择组”列表中，选择先前创建的学术机构组。

注意：“选择组/显示项”允许您在列表中选择您之前创建的组，但会在矩阵中列出该组内的每一个条目。

5. 在“列”的部分，打开您使用的清理过的专利权人列表，并单击**选择组/显示项**。在“选择组”列表中，选择先前创建的公司组。
6. 确保矩阵设置为基于记录数的共现矩阵，然后单击**确定**。
7. 此时将显示一个共现矩阵，其中包含所有曾经合作过的大学和企业。

共现矩阵中有数据的单元表示两个组或列表相交。在这种情况下，交叉点可以显示已经发生的合作以及该合作创造了多少专利。

如何获得答案

通过查看矩阵中的列并选择列中的数字可获得需要的答案。

通过上述步骤，公司被放置在矩阵的列中。由于矩阵会自动过滤掉没有合作关系的企业和大学，因此矩阵中列出的企业均和大学有过合作关系。

列中的数字代表与相应行中的大学合作共同研发的专利数目。在列中选择一个数字，即可在标题视图中显示该次合作所产生的专利。

您可以在标题视图中双击一个专利以显示其细节。需要注意的是，公司和学术机构两者都列在专利记录的专利权人字段中。通过记录视图，您可以看到 Derwent Innovation 中该记录的最新信息以及 PDF 格式的原始专利文献。

需要进一步分析的问题

分析是一个反复比较的过程。反复审视此矩阵会提出额外的问题，许多此类问题可由 Derwent Data Analyzer 进行深度分析而得到解答。以下是此类问题中的几个示例：

- 这些公司和学术机构多久合作一次？
- 由这些合作产生的创新有多少已经获得专利授权（与仍在申请阶段相对比）？
- 用于合作的资金是来源于基金资助还是私人资金？

- 有没有任何论文作为合作研发的结果出版？
- 这些创新是否在市场上取得了成功？

业务需求：分析成功的知识产权保护

您想要得到解答的问题

- 竞争对手们在保护其知识产权方面做得有多成功？

您在哪里能得到答案

通过授权专利与专利申请的比率获得您需要的答案。

请参见整个数据集在环状图中体现的这一比率，该图显示了所有授权专利和所有待定专利申请：

请参见不同层级的专利权人分别在共现矩阵以及从该矩阵绘制的 100%

堆积柱状图中体现的这一比率：

请参见排名前 10 的企业在 100%

堆积柱状图中体现的这一比率，该图显示了每家企业各有多少授权专利和专利申请：

如何创建视图

假设

- 您所使用的数据集与您希望分析的所有发明人和所有技术紧密相关。
- 您的数据仅包含在专利授予机构尚未过期的有效专利。（这根据是否存在非活动 INPADOC 法律状态代码来确定。）
- 您已经针对不同层级的专利权人（例如，排名前 10 的企业、学术机构、个人专利持有者等）分别创建了组。

将“公开号”列表分为“非活动”、“专利申请”、“授权专利”和“待定”组。

1. 从总览界面中，打开公开号列表
2. 针对专利申请和授权专利分别创建组。您可以通过**使用叙词表创建组**功能执行此操作，也可以基于专利文献类型识别代码手动创建这些组。

了解更多有关组的信息：创建组

3. 依次单击**字段**、**由组名创建字段**。
4. 选择您刚才创建的授权专利和待定专利申请组。
5. 为新列表输入**新字段名称**。

6. 单击**确定**。

此时会显示新列表，其中包含授权专利和专利申请项。

创建环状图

1. 在授权专利和待定专利申请列表中，同时选择两项。
2. 单击工具栏上的**环状图**按钮。

此时会显示环状图，其中显示了专利申请和授权专利分别在数据集中所占的百分比。

创建用于比较专利申请和授权专利的矩阵和图表

1. 单击工具栏上的**矩阵**按钮。
2. 在行的部分，选择您为专利申请和授权专利创建的列表。
3. 在列的部分，展开您之前使用的清理过的专利权人列表，并单击**选择组/显示组**。为您要检查的企业层级选择对应的组。

注意：“选择组/显示组”允许您在列表中选择您之前创建的组并在矩阵中列出每个组。

要创建使用组中各项（例如，排名前 10

的企业中的各个企业）的矩阵，请使用**选择组/显示项**并选择您要检查其项的组。

4. 确保矩阵设置为基于记录数的共现矩阵，然后单击**确定**。
5. 将显示一个以竞争企业为行、以专利申请组和授权专利组为列的矩阵。
6. 选择矩阵中的所有行，依次单击**脚本、矩阵输出到 Excel 并可视化**。

此时将打开一个新的 Excel

工作表，其中包含来自该矩阵的数据。默认显示三维柱状图。将图表布局更改为 100%

堆积列并调整布局，使每个柱部分的数字都显示出来。

100%

堆积柱状图显示每家企业在您所关注的领域当前持有的授权专利和专利申请数量。授权专利和专利申请分别以不同颜色的柱区域表示，由此很轻松就能看出授权专利与专利申请的比率。

由于各层级（企业）的专利（授权专利和专利申请）总数不尽相同，使用 100%

堆积柱状图可标准化这些差异，使您重点关注比率。每个区域上的数字可以确切显示该企业在您所关注的领域持有的授权专利或专利申请数量。

您可以**复制并粘贴**此图表到报告和演示文稿中。

如何获得答案

通过比较整个数据集、每个专利权人层级和排名前 10

企业的授权专利与专利申请比率，可获得您需要答案。

初始环状图显示了您的数据集中授权专利与专利申请的总比率。这可以用作与其他比率相比较的“平均值”。

专利权人层级的矩阵和图表显示您所关注领域的每个区域内的趋势。将每个专利权人层级与总体市场比率进行比较，可显示该层级在保护其知识产权方面做得有多成功，因为授权专利越多，表示保护得越成功。

在此例中，学术机构和发明人分配的专利具有较高的专利申请与授权专利比率（约为 3:1）。这可能表明：

- 这些专利权人正在开展更多研究并提交更多专利申请。
- 这些专利权人正在研发技术，以期吸引投资者。

相反，拥有中型专利组合规模的政府机构和企业则具有较低的专利申请与授权专利比率（约为 1:1）。这可能表明：

- 这些专利权人在研发方面的花费较少。
- 这些专利权人持有许多经常被引用的高价值专利，他们主要授予技术许可。

对于排名前 10

企业的矩阵/图表，可以应用相同的分析。您可以得出许多相同的结论，但是关注各个企业可使您获得额外的发现。

在此例中，Nippon Zettoc KK 拥有的授权专利显著多于专利申请。这可能表明：

- 他们正在退出市场。
- 他们正陷入困境，没有可用于研发的资金。
- 他们倾向于广泛提交专利以覆盖整个发明领域。

但是，Lion Corp 和 Henkel KGAA 则具有极高的专利申请与授权专利比率。这可能表明：

- 他们有大量资金可用于研发。
- 他们针对技术的各个细小环节都提交专利申请。
- 他们正在研发专利行政保护，但尚未看到工作成果。如果他们的专利申请成功获得专利授权，他们将有潜力成为市场主导者。

需要进一步分析的问题

分析是一个反复比较的过程。反复审视这些图表和矩阵会提出额外的问题，许多此类问题可由 Derwent Data Analyzer 进行深度分析而得到解答。以下是此类问题中的几个示例：

- 对于拥有许多授权专利或许多专利申请的企业，这些专利涉及的范围是非常窄还是非常广？这可以通过检查分配给这些专利的 DWPI 手工代码进行确定。
- 这些专利是否仍然有效？是否有部分授权专利已失效？是否有部分专利申请已遭到拒绝？深度分析可以细化图表并提供企业活动专利的概况。

- 哪些专利被高度引用？哪些企业拥有这些专利，这同专利申请与授权专利的比率之间有何关系？他们的专利是否经常对外授权，从而主导该领域的技术？
- 对于专利组合规模较小的企业以及由发明人拥有的专利，这些专利是否涉及新的领域？它们是否与顶级企业的研发重叠？

业务需求：分析专利申请趋势

您想了解的问题

- 过去 10 年间，排名前 10 的企业在您所关注的领域内，提交了多少专利申请？
- 如何对这些企业在专利申请数量方面进行相互比较？

您在哪里能得到答案

通过堆积柱状图获得您需要答案。

堆积柱状图提供了一种可视化呈现效果，显示了源数据中每个企业在给定年份提交的专利申请量。每个企业都以不同颜色的区域表示，由此很轻松就能看出申请趋势并比较各个企业的申请状况。

如何进行操作

1. 打开清理过的专利权人列表。

了解更多信息：数据清理简介

2. 为排名前 10 的专利权人创建一个组。

了解更多信息：创建组

3. 打开优先权年份（最早）列表，并根据所需时间跨度创建一个组。在本例中，我们采用的时间跨度为十年。

提示：“优先权年份（最早）”是指，专利申请最早提交的那一年。此字段用于显示专利活动和相关研究在何时发生，对确定先有技术起决定性作用。

4. 在“主页”标签中，单击**矩阵按钮**。
5. 在“行”的部分，打开您之前使用的清理过的专利权人列表，并单击**选择组/显示项**。在“选择组”列表中，选择先前创建的排名前 10 的专利权人列表。

注意：“选择组/显示项”允许您在列表中选择您之前创建的组，但会在矩阵中列出该组内的每一个条目。

6. 在“列”的部分，打开优先权年份（最早）列表，并单击**选择组/显示项**。在“选择组”列表中，选择先前创建的组。
7. 确保矩阵设置为基于记录数的共现矩阵，然后单击**确定**。
8. 此时将显示一个以排名前 10 的专利权人为行、以所选年份为列的矩阵。按年份对该矩阵排序。
9. 选择矩阵中的所有行并单击功能区中的**分析**标签，然后依次单击**脚本、报告 - 矩阵输出到 Excel 并可视化**

此时将打开一个新的 Excel

工作表，其中包含来自该矩阵的数据。默认显示三维柱状图。将图表布局更改为堆积列并调整布局，使每个部分的数字都显示出来。

如何获得答案

在上述堆积柱状图中，每年显示一个柱。在柱中，不同的颜色代表不同的公司。柱中每个公司所占区域的大小与该公司所公开专利的数量成正比；区域越大意味着公开的专利越多。每个区域上面的数字表明该公司申请专利的实际数量。柱的高度表明在给定年份申请专利的总量。

由于专利申请的总数每年都会有变化，因此使用堆积柱状图（而不是使用 100% 堆积图）可以更好地体现每年专利申请的活动情况。每个区域上的数字可以显示公司在该年申请专利的确切数量。

您可以复制并粘贴此图表到报告和演示文稿中。

业务需求：了解市场上专利组合的增长率

您想了解的问题

- 专利分布全景中可观察到的趋势是什么？
- 市场是否从小企业在专利领域占据主导而转型为由大型企业持有多数专利，或反之？亦或双方分据持平？
- 顶级企业中，谁曾为其专利组合的增长进行过投资？
- 是否有个别的企业加快或减缓了其新技术开发？

您在哪里能得到答案

通过矩阵显示每年按各层级分组的专利权人（学术机构、个人、排名前 10 的企业等）提出的专利申请量，并以堆积柱状图的形式显示结果：

在矩阵和堆积柱状图中显示排名前 5 的企业每年的专利申请量。

如何创建视图

假设

- 您所使用的数据集与您希望分析的所有发明人、所有技术和所有年份紧密相关。
- 您已经针对不同层级的专利权人（例如，排名前 5 的企业、学术机构、个人专利持有者等）分别创建了组。

生成矩阵

1. 在“主页”标签中，单击**矩阵按钮**
2. 在行的部分，打开您之前使用的清理过的专利权人列表，并单击**选择组/显示组**。从选择组列表中选择您要分析的组。

注意：“选择组/显示组”允许您在列表中选择您之前创建的组并在矩阵中显示每个组。您可以转而使用“选择组/显示项”，选择一个特定的组（例如，排名前 5 的企业）并显示其在矩阵中的项目。

3. 在列的部分，选择**公开年份列表**。
4. 确保矩阵设置为基于记录数的共现矩阵，然后单击**确定**。

此时将显示一个以组（或某个组中的项）为行、以年份为列的矩阵。按年份对该矩阵排序。

创建图表

选择矩阵中的所有行。单击功能区中的**分析**标签，然后依次单击**脚本、报告 - 矩阵输出到 Excel 并可视化**。

此时将打开一个新的 Excel

工作表，其中包含来自该矩阵的数据。默认显示三维柱状图。将图表布局更改为堆积列并调整布局，使每个部分的数字都显示出来。

在上述堆积柱状图中，每个年份显示一个柱，以此表示专利申请的总数，且不同的颜色代表不同的企业。每个企业区域面积的大小代表与该企业所公开专利的数量成正比。每个区域上的数字为企业申请专利的实际数量。

由于专利申请的总数每年都会有变化，因此使用堆积柱状图（而不是使用 100% 堆积图）可以更好地体现每年专利申请的活动的情况。每个区域上的数字可以显示公司在该年申请专利的确切数量。

您可以复制并粘贴此图表到报告和演示文稿中。

如何获得答案

通过在您的矩阵和图表中按照专利权人层级或企业来对比其每年的专利公开量，可以获得需要的答案。

当您查看专利权人层级时，可以看到在给定的年份中哪一层级申请了大多数的专利。通过浏览一个特定的层级（例如，排名前 5

的企业或发明人专利权人），可以看到占总申请数量比重较大的组。这有助于发现是否有新的参与方（小型企业或个体发明人）正加入该领域，或者大型企业是否正在巩固其在专利全景中的主导地位。

在此例中，公开比例与所处时间周期保持相当一致。

查看个体专利权人（例如，排名前 5

的企业）在一段时间内的表现，可提供线索了解到每个企业的战略组成，这可能有助于预测其未来动向。在本例中，Lion Corp 和 Colgate

似乎都在大力推行专利公开活动，这可能意味着这两家企业都在用自己的创新来响应彼此的创新。

需要进一步分析的问题

分析是一个反复比较的过程。反复审视这些矩阵和图表会提出额外的问题，许多此类问题可由 Derwent Data Analyzer 进行深度分析而得到解答。以下是此类问题中的几个示例：

- 每年有多少公开的专利仍旧活跃？又有多少被驳回和/或没有继续维持？
- 查看 DWPI 手工代码，公开的专利涉及范围是宽还是窄？
- 关于专利申请逐步成为授权专利的过程，哪些公司具有更好的追踪记录？

业务需求：确定取得相似技术专利的公司

您想了解的问题

- 顶级企业是否正在研发相似的技术？
- 某一特定研究领域是否已有过多参与者？
- 顶级企业在技术研发方面有多少重叠？

您在哪里能得到答案

互相关图谱基于分配给专利的 DWPI 手工代码，显示针对相似技术公开了专利的企业。（DWPI 手工代码根据涵盖的技术对专利进行分类。）

如何创建视图

假设：您所使用的数据集与您希望分析的所有公司紧密相关。并且您的数据集包含手工代码。

1. 从总览界面打开一个清理过的专利权人列表。

了解有关清理数据的更多信息：数据清理简介

2. 为排名前 30 的专利权人创建一个组。

了解有关创建组的更多信息：创建组

3. 单击功能区中的“分析”标签，然后单击“添加”区域中的图谱按钮。
4. 选择**互相关图谱**，然后单击**下一步**。
5. 展开您在步骤 1 中打开的清理过的专利权人列表，并选择您先前创建的排名前 30 专利权人的组。单击**下一步**。
6. 选择**手工代码字段**。这是将被用于执行互相关的字段。

提示：您可以创建一个清理过的手工代码字段版本以使数据规范化，并将代码定义包含在图谱中。

7. 单击**完成**。此时会显示图谱。

每个企业都由图谱上的一个节点表示。为了使特定的节点更为突出，您可以改变它们在图谱上的颜色或者位置。要改变节点颜色，请右键单击节点，选择**节点颜色**，然后挑选一个新的颜色。要改变节点位置，请右键单击节点并选择**移动节点**。

您可以导出图谱以便在报告或演示文稿中使用：右键单击图谱，选择**导出至文件**，然后选择保存图谱的文件格式。

如何获得答案

通过浏览图表中的节点以及节点之间的连线，获得您需要的答案。

上述图谱将一家企业所持有每个专利记录的 DWPI

手工代码与列表上另一家企业所持有的手工代码进行比较。如果两家企业持有具有相同手工代码的专利，则用一条直线将他们相连。企业之间的连线越粗，表示其手工代码越匹配，即这些企业的专利涵盖的相似技术越多。

该图谱可能会显示出一些互联企业的聚类，这表明这些企业都致力于研发相似的技术。相反地，该图表也可能会显示企业之间没有任何联系：这些企业与其他企业之间共享的手工代码极少，也就意味着他们的技术研发重心是不同的。

需要进一步分析的问题

分析是一个反复比较的过程。反复审视此图谱会提出额外的问题，许多此类问题可由 Derwent Data Analyzer 进行深度分析而得到解答。以下是此类问题中的几个示例：

- 对于那些同其他任何企业都没有联系的企业，他们是持有能体现出新的关联关系的技术还是应该将其从分析中排除？
- 对于那些具有诸多相似性的企业，手工代码是否依附于授权专利或专利申请？
- 专利技术重叠的出现应归因于相似但不同的发明还是潜在的侵权？
- 哪些专利中的先有技术可以让其中的部分（或全部）专利失效？

常见问题解答

常规

- [怎样使分析运行更快？](#)
- [为什么我的文件不包含某个特定字段？](#)
- [可以分析全然不同的信息（例如专利数据集与科技文献数据集）吗？](#)
- [如何创建不同数据集共有人员的列表（发明人、作者）？](#)
- [可以导入报刊文章进行分析吗？](#)

数据清理

- [如何创建“停词表”并从列表中移除某些单词？](#)
- [如何提取多个字段？](#)
- [如何合并两个数据集？](#)
- [如何处理重复记录？](#)
- [如何从专利权人列表中移除个人姓名？](#)
- [如何在“列表清理”功能中利用叙词表？](#)
- [可以分配专利字段（如 Aureka）吗？](#)
- [可以添加我自己的分类吗？](#)

分析

- [我应该使用哪种图谱类型？](#)
- [怎样解读图谱？](#)
- [为什么图谱运行失败？](#)
- [互相关图谱使用两个字段确定绘图 - 但是，既然图谱上只对一个字段进行描述，那么使用两个字段的真正用意是什么？](#)
- [如何区分互相关分析（基于 2 个字段）与自相关分析（基于 1 个字段）？](#)

常见问题解答 – 常规

怎样使分析运行更快？

运行速度取决于电脑的运算能力、数据集的记录数和数据集的字段数。对其中一个或多个因素进行如下变化都将提高速度：

- 确保数据集只含有您需要的记录（例如，要分析过去 3 年的活动，就不需要对象区域的整个历史数据）。使用“创建子数据集”指令可降低大小
- 确保数据集只包含有值的字段 - 使用“删除字段”指令可以删除不需要的字段（您随时都可以在以后使用“导入更多字段”指令，但是当然，不包含您清理过的那些字段）

- **增加内存和提高处理器速度**会有所帮助，使用第二块硬盘也是如此。若使用笔记本电脑，请确保其使用电源工作，而且不是在节能配置下运行（通常，处于节能配置时，数学协处理器会被禁用）

为什么我的文件不包含某个特定字段？

该字段在您导出时所用的原始数据中是否存在？

- **如果是**，则使用“导入更多字段”功能重新获取该字段（它也许是一个二级字段，因此请在出现提示时勾选此框）
- **如果不是**，那么您需要执行另一次导出，确保这次导出时该字段出现在原始数据中
- **如果该字段在原始数据中**，但没有列在导入过滤器中，请联系 Clarivate Analytics 的支持平台

可以分析全然不同的信息（例如专利数据集与科技文献数据集）吗？

可以。分析工具和数据清理工具与数据类无关，因此可以同等应用。下面是特别值得提及的两个功能：

- **列表比较**：允许在来自不同文件的两个列表之间进行比较，甚至是比较来源于不同数据库的数据。例如，将来自专利和科技文献的主题信息进行对比，从而识别具有共同性或唯一性的区域。
- **数据集合并**：允许将两个不同的数据集整合成一个，而不用考虑数据类型，使用户可以“贴合”相似字段（例如，文献的作者与专利的发明人）。

如何创建不同数据集共有人员的列表（发明人、作者）？

使用列表比较功能 -

此功能将对来自相同或不同数据集的两个列表。请注意，两个数据集都需打开，并且请仔细命名文件，以便您能够确切区分这两个文件。

1. 在文件 1 中，创建作者列表
2. 使用列表比较，选择将“全列表”与文件 2 中的“发明人”字段进行对比。选择比较类型为“共有项”，然后将要创建的组标注为“共有人员”。如有必要，请使用模糊匹配选项（即，如果预期会存在一些微小差异，那么这些差异是可以接受的）
3. 一经运行，将会添加一个组到文件 1 的作者列表中，其中两个文件共有的姓名都将被选中

可以导入报刊文章进行分析吗？

可以。实际上，任何结构化数据都可以分析，不过每种不同类型的数据将需要新的导入过滤器。导入过滤器包含一些规则，可用来识别数据中的个别记录、字段和项目；当然，这些规则会依数

据类型而异。您也可以使用导入引擎编辑器自行构建（这是一个高级主题），或者告知您的 Clarivate Analytics 客户经理，为您构建一个。

常见问题解答 - 数据清理

如何创建“停词表”并从列表中移除某些单词？

1. 创建一个叙词表文件，该文件包含您希望移除的所有主题词，且这些主题词都置于一个标题下（如 *****停词表*****）
2. 打开要从中移除主题词的列表
3. 使用“基于叙词表的组”功能，选择停词表-叙词表文件，并将结果添加到一个具有别名的新组里
4. 所有的停词现在已经出现在一个名为“*****停词表*****”的组里
5. 现在使用“由组中项创建字段”，选择“*****停词表*****”组，使用“未确认项”并取消选择该选项以保存组
6. 现在，即已创建了一个新的没有停词项的字段

随着将来的不断使用，您可能希望添加新内容到停词表 -

这可以通过在叙词表编辑器中添加额外的主题词、或者通过将其添加到组并使用“基于组创建叙词表”功能来完成。上述步骤 1-4 只需手动将列表中的主题词添加到组即可完成。

如何提取多个字段？

使用“导入更多字段”功能，以及选择“显示二级字段”选项。

如何合并两个数据集？

使用“数据集合并”功能 -

您需要打开两个数据集。如果它们具有相同的数据类型，请单击“添加全部字段”以便两个数据集自动合并字段。如果二者是不同类型，则通过拖放操作进行手动合并（例如，将发明人和作者进行合并，或者将机构与专利权人进行合并）。合并之后，您可能需要使用“移除冗余记录”功能。

如何处理重复记录？

您可以移除重复记录（只保存重复记录的一个实例，保留该实例最高字符数）或合并重复记录（将所有重复信息合并到一个记录 - 适用于 Aureka 报告输出）。您可以选择一个或多个要确定其重复状况的字段。

如何从专利权人列表中移除个人姓名？

使用 AcadGovCorpIndiv

叙词表来创建组，然后使用“由组中项创建字段”以创建一个新的字段，在该字段中包含该组不含的项（选择未标注）。

如何在“列表清理”功能中利用叙词表？

在完成列表清理时选择“保存为叙词表”，将您的处理结果转换成叙词表。此操作会将您创建的分级结构保存为一组规则，以便在将来使用叙词表功能时应用这些规则。您可以将其添加为新的叙词表，也可以将其合并到一个现有叙词表（例如，对使用中的公司叙词表进行更新）。

可以分配专利字段（如 Aureka）吗？

可以。新的字段可以从组、组名或分类中创建。

可以添加我自己的分类吗？

可以。有关添加内部分类方案到记录的详细信息，请参见帮助 (F1)。

常见问题解答 - 分析

我应该使用哪种图谱类型？

- 自相关图谱 - 查看共同发布者（例如，发明人网络图谱）
- 互相关图谱 -
查看相似领域中的工作者，基于二级字段的主题词（例如，使用基于手工代码的互相关来查看在相似领域投入工作的专利权人）
- 因子图谱 - 技术聚类，特别适用于 NLP 文本字段

怎样解读图谱？

- 图谱上用“节点”（蓝色小球）表示每个项
- 节点大小表示数据中的记录比例，即，节点越大意味着记录越多
- 节点之间的连线表示亲缘关系 - 连线越粗，则亲缘关系越强
- 通常，那些相互较为接近的节点，比其他较为分开的相关程度更高

为什么图谱运行失败？

最常见的绘图问题是在试图用出现次数少于 3

次的主题词创建因子图谱时出现的。进行此计算所需的算法（“主成分分析”法）需要出现 3 次或更多次的项。

互相关图谱使用两个字段确定绘图 - 但是，既然图谱上只对一个字段进行描述，那么使用两个字段的真正用意是什么？

这些图谱表明了基于各项在二级字段中的出现次数，它们在一主字段中的相互关系，如专利权人基于他们所使用的 IPC

的相似度：这样每个节点表示一位专利权人，但是他们之间的连线和间距则是基于 IPC 进行计算。

如何区分互相关分析（基于 2 个字段）与自相关分析（基于 1 个字段）？

互相关分析基于二级字段（例如，一个技术字段）显示相似领域中的工作者。自相关分析则显示共同发布者，即共同出现在所分析的字段中。