

专利课题前沿论文研究进展

2018.11.05 方建勇

提示：采用手机 safari 微软翻译技术

1. 向量空间模型中的文本相似性：一个比较研究

作者:omid shahmirzadi, adam lugowski, kenneth younge

摘要: 语义文本相似性的自动测量是自然语言处理中的一项重要任务。在本文中，我们评估了不同的向量空间模型的性能来执行此任务。我们解决了构建专利与专利相似性的实际问题，并比较了 tfidf (和相关扩展)、主题模型（例如潜在语义索引）和神经模型（例如，段落向量）。与预期相反，只有在：1) 对目标文本进行压缩时，文本嵌入方法的附加计算成本才是合理的；和 2) 相似度的比较是微不足道的。否则，tfidf 在其他情况下的表现令人惊讶：特别是在更长和更技术性的文本中，或在最近的邻居之间进行更细粒度的区分。出乎意料的是，对 tfidf 方法的扩展，如增加名词短语或递增计算术语权重，在我们的上下文中没有帮助。少

2018 年 9 月 24 日提交;最初宣布 2018 年 10 月。

2. 基于 ssv 出的偏头痛患者固有模糊熵的提取

作者:曹泽宏,林金腾,赖冠林, 高丽伟,钟泰王, 傅钟玲, 王树军

摘要: 固有的模糊熵是对脑电图 (eeg) 复杂性的客观测量, 反映了大脑系统的鲁棒性。在本研究中, 我们提出了一个新的应用多尺度相对固有模糊熵使用重复稳态视觉诱发电位 (ssvep) 来研究脑电图复杂性的变化之间的两个偏头痛阶段, 即国际间 (基线) 和前 (偏头痛发作前) 阶段。在以下两个条件下, 我们使用了带有 o1、oz、o2 和 fpz 电极的可穿戴头带脑电图装置, 收集 80 名参与者 (40 名偏头痛患者和 40 名健康对照组 [hc]) 的脑电图信号: 在休眠状态和具有 5 个 15 赫兹光的 ssp 中刺激。我们发现, 在 hc 和患者的术前阶段, 枕部脑电图熵显著增加, hc 和患者的刺激时间都在增加, 但在检查前阶段的患者则出现了相反的趋势。在第一 ssvep 中, hc 的枕部脑电图熵明显高于术前阶段的专利 (fdr 调整 $p < 0.05$)。对于第 1 和第 5 次 ssvep 之间脑电图熵的过渡方差, 在检查前阶段的患者在检查间期的值明显低于患者 (fdr 调整 $p < 0.05$)。此外, 在分类模型中, adaboost 集成学习显示, 在间和前置相的分类中, aucc 的准确度为 81%, 而 auc 的准确度为 0.87。相反, 使用其他竞争熵模型, 包括同一数据集上的近似熵、样本熵和模糊熵, 各组或会话的脑电图熵没有差异。总之, 固有的模糊熵在视觉刺激环境中提供了新的应用, 并有可能为偏头痛患者提供预检查警报。少

2018 年 9 月 20 日提交;v1 于 2018 年 9 月 18 日提交;最初宣布 2018 年 9 月。

3. 级联引文扩展

作者:[陈朝梅](#)

摘要: 数字科学的维度被设想为下一代研究和发现平台, 以便更好、更高效地获取交叉引用的学术出版物、赠款、**专利**和临床试验。作为日益开放的引文资源的新补充, 它提供了机会, 可能使研究人员、决策者和公众的各种科学出版物的利益攸关方受益。本文从级联引文扩展的角度探讨和论证了一些实用的潜力。给定一组出版物, 级联引文扩展过程可以连续应用于一组文章, 以便通过引文链接将覆盖面扩展到越来越相关的文章。虽然概念的起源可以追溯到加菲猫的引文索引, 但直到最近, 它在很大程度上都被限制在少数能够不受限制地访问引文数据库的人身上, 该数据库足够大, 可以维持这种迭代扩展。在开放的维度 api 的基础上, 我们在 citespace 中集成了级联引文扩展功能, 并演示了如何从这些新功能中受益。总之, 层叠引文扩展有可能提高我们对科学知识结构和动态的认识。少

2018 年 5 月 31 日提交;最初宣布 2018 年 6 月。

4. 针对 gpu 优化的并行卢万社区检测

作者:[理查德·福斯特](#)

摘要: 社区检测现在是许多基于图形的应用程序中的一项重要操作。它被用来揭示现实世界网络中存在的群体, 而不对一组社区施加先前的规模或基数限制。尽管有潜力, 但对并行计算机的支持相当有限。原因主要是算法的不规则性, 而潜在的启发式则意味着顺序性质。本文提出了一种基于 gpu 的卢万方法并行化版本。卢万方法是一种多相迭代启发式模块化优化方法。它最初是由 blondel 等人 (2008 年) 开发的, 由于它能够以快速和内存高效的方式检测高模块化社区分区, 因此变得越来越流行。使用的平行启发式, 首先由郝 lu 等人 (2015 年) 介绍。由于 louvain 方法本质上是连续的, 因此它限制了可扩展使用的可能性。由于提出了并行启发式方法, 我观察了此方法在 gpu 上的行为方式。对于评估, 我在 geforce gtx 980 gpu 上使用 cuda 实施了启发式方法, 并对 cern 开发的合作发现项目中的我使用过的组织环境进行了测试, 该项目涉及**专利**和出版物, 以可视化在技术上的联系, 在它的合作者之间。与在具有使用 gpu 的同一台计算机上运行的并行 louvain 实现相比, cuda 实现能够产生与 cpu 生成的结果相当的社区输出, 同时提供多达 30 个使用的绝对速度地力 gtx 980 消费级 gpu。少

2018 年 5 月 28 日提交;最初宣布 2018 年 5 月。

5. 支持智能系统研究设计的信息技术和综合工具

作者:[oleksandr palagin](#), [mykola petrenko](#), [Kyylo malakhov](#)

摘要: 在使用跨学科和本体管理概念的基础上, 建设信息技术的方法基础、实施智能系统研究相关设计的正式模型和工具被考虑。旨在建立具有知识、科学和技术创造力的工作系统, 包括开发新技术的对象。旨在建立具有知识、科学和技术创造力的工作系统, 包括开发新技术的对象。结果。开发了智能系统研究设计的信息技术。考虑了一个语言处理器的创建和**专利**, 以处理大量文本数据, 以便随后提取知识的一个例子。结论。开发科学研究设计系统的方法基础包括信息处理认知过程的模型、技术和工具, 包括新工程对象的设计、自动化开发知识领域的本体生物学知识库。最后是跨学科和跨学科性质综合科学研究的智能技术和系统的基本组成部分。本体系统集成过程使得有可能研究对象知识的相互作用, 作为学科交汇处的对象, 以及融合的集群, 为新知识和新知识的交叉提供了机会。科学理论的发展。关键词: 研究设计, 智能系统, 本体管理, 本体。少

2018 年 4 月 13 日提交;最初宣布 2018 年 5 月。

6. 当信号处于噪声中时: **diffix** 粘性噪声的极限

作者: [andrea gadotti](#), [Florimond houssiau](#), [lucrocher](#) , [yves-山铁德 montjoye](#)

摘要: 在隐私和效用之间找到平衡, 让研究人员和企业人们在保护人们隐私的同时, 让数据得到充分利用, 是我们今天面临的最大挑

战之一。大量研究表明, 传统的匿名(或去识别)模式的局限性促使使用问答或基于查询的系统。diffix 是由 aircloak 开发的基于查询的系统, 它使用 "粘性噪声" 的概念来保护人们的隐私。我们在这里对 diffix 进行攻击, 利用其粘性噪声的结构推断数据集中人员的私人属性。我们认为, 这种漏洞是严重的, 使我们能够准确地推断几乎没有背景知识的用户的私人信息。虽然我们同意 diffix 的创作者的观点, 即我们需要重新审视构建实用的隐私保护系统, 但我们认为, 这需要分层的安全方法和完全开放的工具和讨论。**专利**和专有代码不可能足以真正帮助我们在数据的巨大潜力和基本的隐私权之间找到平衡。少

2018 年 7 月 13 日提交;v1 于 2018 年 4 月 18 日提交;**最初宣布** 2018 年 4 月。

7. 生物医学研究的科技影响比较

作者:[青克](#)

摘要: 传统上, 学术论文从其他论文收到的引文数量被用作其科学影响的代名词。然而, 引文可以来自科学界以外的领域, 这样的例子之一是通过**专利技术**——纸可以被**专利**引用, 从而产生技术影响。虽然对论文的科学影响进行了广泛的研究, 但技术方面在文献中的了解程度仍然较低。在这里, 我们的目标是通过提交一份关于美国**专利**和其他论文如何随着时间的推移引用 919, 000 篇

生物医学论文的比较研究来填补这一空白。我们观察到**专利**引文和论文引用之间存在正相关关系,但无论是引用次数最多的论文,还是延迟识别的论文中,这两个领域之间几乎没有重叠。我们还发现,这两种类型的引文表现出明显的时间变化,**专利**引文落后于论文引文的中位 6 年的大多数论文。我们的工作有助于了解论文的技术影响。少

2018 年 7 月 3 日提交;v1 于 2018 年 4 月 11 日提交;最初宣布 2018 年 4 月。

8. 研究与开发工作站环境: 当前研究信息系统的新类

作者:oleksandr palagin, vitalii velychko , Kyrlyo malakhov, oleksandr shchurov

摘要: 在科学研究领域现代技术发展的背景下,出现了新的研究信息系统类和相关的智能信息技术。它被称为—研究和开发工作站环境 (rdwe)—科学研究和开发生命周期支持的面向问题的综合信息系统。本文介绍了 rdwe 类系统的设计和开发基础知识。本文将 rdwe 类系统的广义信息模型表示为一个三元组复合 web 服务,其中包括: 一组原子 web 服务,每个服务都可以设计和开发为微服务或桌面应用程序,从而允许单独用作独立软件;一套功能,功能填充的研究和开发工作站环境;实现复合 web 服务功能所需的原子 web 服务子集。根据 rdwe 类的基本信息模型,支持本体工程领域研究的系统——任意领域应用本体的自动化构

建、科学和技术创造力——开发了乌克兰发明专利申请文件的自动编制。它被称为——个人研究信息系统。这类系统的一个显著特点是，通过结合各种功能服务并在云集成环境中增加新的功能服务，有可能将其问题定向到各种科学活动中。我们的主要工作成果侧重于提高科学家在任意领域的研究和开发生命周期的有效性。

少

2018 年 7 月 25 日提交;v1 于 2018 年 3 月 15 日提交;最初宣布 2018 年 3 月。

9. 19 年后 cayley-puser 公钥密码系统的回顾与展望

作者:[道格拉斯·斯丁森](#)

摘要: . flannery 于 1999 年提出。我将介绍对它的两次攻击，其中一次显然是新的。我还将研究 2008 年斯拉夫获得专利的 cayley-puser 算法的一个变种，并表明它也是不安全的。

2018 年 3 月 13 日提交;最初宣布 2018 年 3 月。

10. 个人研究信息系统。关于探索发明专利类似物的研究方法

作者:[o. v. palagin](#), [k. s. malakhov](#), [v. yu. velychko](#), [o. s. shchurov](#)

文摘:本文介绍了个人研究信息系统专利模拟的信息模型和检索方法。

2018 年 3 月 5 日提交;v1 于 2018 年 2 月 21 日提交;最初宣布 2018 年 2 月。

11. 利用深度学习从文献中提取分子结构

作者 :joshua staker, kyle marshall, robert abel, carolyn mcquaw

摘要: 从文档中提取化学结构仍然是一个难题, 因为在分割过程中对结构的误报和预测结构的错误。目前的方法依赖于手工制定的规则和子程序, 这些规则和子程序的总体表现相当好, 但仍然经常会遇到识别率尚不令人满意和系统性改进具有挑战性的情况。影响当前方法性能的复杂性包括各种软件用于渲染结构的视觉样式的多样性、临时注释的频繁使用以及与图像质量相关的其他挑战, 包括分辨率和噪声.我们在这里提出了端到端的深度学习解决方案, 既从文档中分割分子结构, 也预测化学结构从这些分割图像。这种基于深度学习的方法不需要任何手工制作的功能, 直接从数据中学习, 并且对图像质量和样式的变化具有鲁棒性。使用本文描述的深度学习方法, 我们表明, 在对含有中等大小分子的低分辨率图像进行分割和预测方面, 有可能在期刊文章和专利中得到很好的分割和预测。少

2018 年 2 月 13 日提交;最初宣布 2018 年 2 月。

12. 动态语义空间中的无监督语感消歧

作者: [jean-françois delpech](#)

摘要: 本文主要关注的是在动态语义空间中快速、自动分辨语感的能力, 在动态语义空间中, 新术语和新感官频繁出现。这些空间是由不断发展的数据集 (如维基百科、**专利**授予和应用程序存储库, 或用于技术辅助审查和电子发现的大量法律文件) "动态" 构建的。这种即时性排除了监督以及使用先验训练集的可能性。我们证明了一个术语的各种意义可以用一个简单的聚类算法自动显现出来, 每个意义都是语义空间中的一个向量。虽然我们只考虑使用随机向量构建的语义空间, 但这种算法应该适用于任何类型的嵌入, 前提是术语之间有意义的相似之处可以计算, 并至少满足关闭的术语的两个基本条件意义有很高的相似性, 具有无关意义的术语几乎为零相似性。少

2018 年 2 月 16 日提交;v1 于 2018 年 2 月 7 日提交;**最初宣布** 2018 年 2 月。

13. 光伏材料的基础专利数据挖掘: 专利引文光谱的应用

作者: [jordan comins](#), [loet leydesdorff](#)

文摘: 我们将**专利**引文光谱 (pcs)——最初开发为参考出版年光谱, 用于研究科学文献中的地标和里程碑——将专利文献分为九种合作**专利分类**(cpc) 的 y 子类, 描述材料光伏技术。对于这项研究, 我们扩展了例程, 并选择在 patentsview 使用高级搜索查询。基

于检索到的**专利**所引用的**专利**的纵向分布的两个规范化, 例程 (<http://www.leydesdorff.net/comins/pcs/index.html>) 提供了对字符串中指定的主题的**基础专利**。在这 9 起案件中, 有 5 起案件, 我们发现了常规所表明的**专利**基本性质的佐证。少

2018 年 4 月 10 日提交;v1 于 2018 年 1 月 29 日提交;**最初宣布** 2018 年 1 月。

14. 神经机器翻译对人的后编辑在线适应的用户研究

作者:[sariya karimova](#), [patrick Simianer](#), [stefan Riezler](#)

摘要: 神经机器翻译 (nmt) 的优点已被广泛验证, 适用于不同语言和书写语言领域的几种语言对的离线翻译。然而, 迄今为止, 通过对人类后期编辑的适应对 nmt 的交互式学习的研究还仅限于仿真实验。我们提出了第一个用户研究的在线改编 nmt 用户后期编辑在**专利**翻译领域。我们的研究涉及 29 名人类主体 (翻译学生), 他们的后期编辑工作和翻译质量是在一个人工后编辑和一个集成在线自适应学习算法的机器翻译系统的大约 4, 500 次互动中进行的。我们的实验结果显示, 根据 hter、hbleu 和 ksmr 等多个评估指标, nmt 中的在线适应导致人工后期编辑工作显著减少。此外, 我们发现 nmt 输出和授予**专利**的专业翻译之间的 bleuter 有了显著改进, 为交互式设置中在线自适应 nmt 的优势提供了进一步的证据。少

2018 年 9 月 18 日提交;v1 于 2017 年 12 月 13 日提交;最初宣布 2017 年 12 月。

15. 文本处理中机器学习中的质量效率权衡

作者:[ricardo Baeza-Yates](#), [zeinab liaghat](#)

摘要: 数据挖掘、机器学习和自然语言处理是强大的技术, 可用于从大型文本中提取信息。根据手头的任务或问题, 可以使用许多不同的方法。现有的方法正在不断地优化, 但并非所有这些方法都经过了测试, 并在一组可以使用监督机器学习算法解决的问题中进行了比较。问题是, 如果我们将训练数据大小从 100 mb 增加到 1 gb 以上, 方法的质量会发生什么变化? 此外, 当数据处理速度下降时, 质量收益是否值得? 我们能不能通过能够处理更多的数据来换取时间效率和质量损失? 考虑到培训数据大小、学习时间和获得的质量等方面的权衡, 我们试图以一般的方式回答这些问题。我们提出了一个性能权衡框架, 并将其应用于三个重要的文本处理问题: 命名实体识别、情感分析和文档分类。之所以选择这些问题, 也是因为它们具有不同级别的对象粒度: 单词、段落和文档。对于每个问题, 我们选择了几种受监督的机器学习算法, 并评估了它们在大型公开数据集 (新闻、评论、**专利**) 上的权衡。为了探索这些权衡, 我们使用了大小不断增加的不同数据子集, 从 50 mb 到几个 gb 不等。我们还考虑了数据集和评估技术的影响。我们发现, 结果没有明显的变化, 大多数时候最好的

算法是最快的。但是，我们也表明，小数据（比如小于 100 mb）的结果与大数据的结果不同，在这种情况下，最好的算法很难确定。少

2017 年 11 月 7 日提交;最初宣布 2017 年 11 月。

16. 通过网络中心性早期识别重要专利

作者:[manuel sebastian mariani](#), [matus medo](#), [françois Lafond](#)

摘要: 技术预测中最具挑战性的问题之一是尽早查明那些有可能导致我们社会发生根本变化的技术。在本文中，我们使用美国**专利引文网络**（1926–2010）来测试我们通过引文网络分析及早期确定具有历史意义的**专利**列表的能力。我们表明，为了在这些**专利**发布后不久有效地发现它们，我们需要超越原始引文计数，同时考虑引文网络拓扑和时间信息。特别是，一种年龄归一化的**专利**中心度测量，称为重新划分的网页排名，允许我们比引文计数和年龄排名分数更早地确定重要的**专利**。此外，我们发现，虽然高影响力**专利**往往依赖于其他高影响力**专利**，其方式与科学论文相似，但**专利**的引用动力学明显低于论文，这使得早期确定重要**专利**比重要论文更具挑战性。少

2017 年 10 月 25 日提交;最初宣布 2017 年 10 月。

17. 专利引文光谱 (pcs): 地标性专利的算法检索

作者:jordan a comins, stephanie a carmack, loet leydesdorff

文摘: 生物医学研究与开发 (r & d) 中专利景观建设的一个重要组成部分是确定最重要的**专利**。迄今为止, 鉴定种子**专利**需要生物医学领域的专题专家。在这份简短的通信中, 我们报告了一种分析方法和工具, 即**专利**引文光谱 (pcs), 用于快速识别用户指定的生物医学创新领域的具有里程碑意义的**专利**。pcs 对大量**专利**中引用的参考资料进行了挖掘, 并对历史上影响最大的先前工作进行了估计。pcs 的有效性体现在与临床相关的生物医学创新的两个案例研究中: (1) ma 干扰和 (2) 胆固醇。挖掘和分析了 4, 065 引用了与 ma 干扰**专利**有关的参考资料, 并正确地确定了 mai 知识产权专题专家独立报告的这一技术的基础**专利**。其次, pcs 被应用于处理胆固醇的一组广泛的**专利**——这一案例研究选择是为了反映更一般的**专利**检索查询, 而不是专家。通过 11, 326 项引用了 pcs, 并确定了利普托的**开创性专利**, 利普托是治疗高胆固醇的突破性药物, 也是雷帕塔的**专利**。这些案例表明, pcs 为确定生物医学创新和治疗领域的**精益专利**提供了一种有用的方法。交互式工具可在: www.leydesdorff.net/pcs/ 免费使用。

少

2017 年 10 月 14 日提交;v1 于 2017 年 10 月 9 日提交;**最初宣布** 2017 年 10 月。

18. 学术知识发现中的有限理性

作者:k 丽娜 lerman, nathan h 高达 wu

摘要: 在一个信息丰富的世界里, 人们的时间和注意力必须在迅速变化的信息来源和对他们提出的各种任务之间进行划分。人们如何决定在自己的工作中阅读和使用的众多来源中的哪一个, 如科学文章或**专利**, 会影响学术知识的传播和创新的采用。我们分析了人们对物理评论期刊的引文网络、美国**专利**和法律意见传播哪些信息的选择。我们观察行为中符合人类有限理性的规律性: 人们不应该评估所有可用的选择, 而是依靠简单的认知启发式来决定要关注哪些信息。我们证明了这些启发式偏见的选择, 使人们优先传播更容易发现的信息, 通常是因为它更新或更受欢迎。然而, 我们没有发现任何证据表明, 流行的来源有助于扩大信息的传播, 使其更加突出。本文为有限理性在社会传播中的注意力分配决策中发挥着至关重要的作用提供了新的证据。少

2017 年 9 月 30 日提交;最初宣布 2017 年 10 月。

19. 苹果在 macos 10.12 上实施差异隐私中的隐私损失

作者:唐军, Aleksandra korolova, 白晓龙, 王学强, 王晓峰

摘要: 2016 年 6 月, apple 宣布将为部分用户数据收集部署差异隐私, 以确保用户数据的隐私, 甚至是来自 apple 的数据的隐私。苹果做法的细节仍然很少。尽管此后出现了一些**专利**暗示了可用于实现差分隐私的算法, 但它们没有准确解释对隐私参数选择所

采取的方法。这种选择以及隐私预算使用和管理的总体方法是理解任何不同隐私部署所提供的隐私保护的关键问题。在这项工作中，通过实验、静态和动态代码分析的结合，对 macos sierra (10.12 版) 的实现进行了分析，阐明了 apple 在隐私预算管理方面做出的选择。我们发现并描述了 apple 用于不同私有数据处理的设置，包括整体数据管道、用于每个数据的不同私有扰动的参数以及将此类数据发送到 apple 的服务器。我们发现，尽管 apple 的部署确保了提交到其服务器的每个数据的（差异）隐私损失是 1 或 2，系统允许的整体隐私损失要高得多，高达 16 每天四个最初宣布的应用的表情符号，新词，迪普林和查找提示。此外，apple 每天更新可用的隐私预算，这可能会导致用户选择加入这四个应用程序的不同私人数据收集后的 16 天可能会损失 16 倍的隐私。我们主张，为了要求不同的私人数据收集的全部好处，苹果必须给予其实施的充分透明度，使用户能够在与隐私损失相关的领域作出选择，并设置有意义的默认值允许的隐私损失。少

2017 年 9 月 11 日提交;vi 于 2017 年 9 月 8 日提交;**最初宣布** 2017 年 9 月。

20. 解开增长、老龄化和通货膨胀的动态，引用特定研究领域的科学文章

作者:k. w. higham, m. 府式 , a. b.jaffe, u. zülicke

文摘: 我们分析了从美国物理学会 (pra、prb、prd、pre 和 prl) 期刊上获得的文章所获得的引文的时间演变。随着时间的推移, 在每份期刊上发表的论文数量的观察变化被认为是由规范化造成的可遗传性的外来原因变化。适当的通货膨胀调整引文率被发现是可分离的, 形成一个优先处理型的生长核和一个纯粹的过时相关的 (即, 单调地减少作为时间的函数, 因为出版) 老化函数。与不同期刊相关的生长核和老化函数的经验提取参数的变化指向引文强度和知识流的研究场特异性特征。与类似结果的类似结果的引用动态的技术分类的专利队列提供了更深入的见解, 以引用行为表明的信息传播的基本原则。少

2017 年 10 月 16 日提交;v1 于 2017 年 8 月 18 日提交;**最初宣布** 2017 年 8 月。

21. 三面齿轮动力分体式混合动力总成的优化设计

作者:庄伟超,张晓武, 赵丁,彭辉,王连谋

摘要: 今天的许多动力分裂混合动力电动汽车 (hev) 使用行星齿轮 (pg) 连接动力总成元件在一起。最近的功率分裂 hev 倾向于使用两个 pg, 其中一些有多种模式, 以实现更好的燃油经济性和驾驶性能。展望未来, 混合动力总成技术必须得到加强, 以设计混合轻型卡车。对于轻型卡车, 对多模式和更多 pg 的需求更强, 以实现所需的性能。为了系统地探索三种 pg 多模 hev 的所有

可能设计，提出了一种有效的搜索和优化方法。对使用三个 pg 的现有配置的所有可能的离合器拓扑和模式进行了详尽搜索。启动性能首先用于筛选出无法满足所需发射性能的设计。然后采用近乎最优和计算效率高的能源管理策略来确定实现良好燃油经济性的设计。拟议的设计过程成功地确定了 8 种设计，实现了更好的发射性能和更好的燃油经济性，同时使用的离合器数量少于基准和专利设计。少

2017 年 7 月 31 日提交;最初宣布 2017 年 8 月。

22. 跨技术领域的知识流动动态模式：经验结果和联系预测

作者: [jieun kim](#), [christopher l. magee](#)

摘要: 本研究的目的是研究知识溢出的结构和演变跨越技术领域。具体而言，确定了 29 个技术领域之间的知识流动动态模式，以 8 个不同时期的**专利**引文来衡量，并测试了链接预测的能力，以预测这些领域的演变跨领域**专利**网络。使用 katz 指标的预测取得的总体成功意味着，有一种趋势是主要在以前联系的技术领域内产生更多的知识流动。本研究通过描述动态技术网络中的结构变化和进化行为，并为预测未来技术知识流动的出现提供依据，为创新研究做出了贡献。少

2017 年 6 月 21 日提交;最初宣布 2017 年 6 月。

23. 通过类比挖掘加速创新

作者: [tom hope](#), [joel chan](#) , [aniket kittur](#), [dafna shahaf](#)

摘要: 大型创意存储库（例如美国**专利**数据库）的可用性可以通过为人们提供类似问题的解决方案的灵感，显著加快创新和发现。然而，在这些大型、混乱的真实世界存储库中找到有用的类比，对于人工方法或自动化方法来说仍然是一个持续的挑战。以前的方法包括昂贵的手工创建的数据库，这些数据库具有较高的关系结构（例如谓词演算表示），但非常稀疏。更简单的机器学习相似度量可以扩展到大型的自然语言数据集，但难以解释结构相似性，这是类比的核心。在本文中，我们探讨了学习简单的结构表示的可行性和价值，特别是 "问题模式"，它指定了产品的用途和实现该目的的机制。我们的方法结合众包和递归神经网络，从产品描述中提取目的和机制矢量表示。我们证明了这些学习的向量使我们能够找到比传统的信息检索方法更高的精度和召回的类比。在一个构思实验中，与传统方法检索到的类比相比，我们模型检索到的类比显著增加了人们产生创造性思想的可能性。我们的研究结果表明，在规模上实现计算类比的一个很有希望的方法是学习和利用较弱的结构表示。少

2017 年 6 月 17 日提交;最初宣布 2017 年 6 月。

24. 基于神经序列模型的反合成反应预测

作者:[bowen liu](#), [bharath ramsundar](#), [prasad kawthekar](#), [jade shi](#), [joseph gomes](#), [quang luu nguyen](#), [stephenho](#), [jack sloane](#), [paul wender](#), [vijay pande](#)

摘要: 我们描述了一个完全数据驱动模型，该模型学习执行反合成反应预测任务，该模型被视为序列到序列映射问题。端到端训练模型具有由两个递归神经网络组成的编码器-解码器体系结构，这在解决其他序列到序列的预测任务（如机器翻译）方面已经显示出巨大的成功。该模型是根据美国专利文献中的 50,000 个实验反应例子进行训练的，这些实验反应例子跨越了医药化学家常用的 10 种广泛的反应类型。我们发现，我们的模型可与基于规则的专家系统基线模型进行比较，并克服了与基于规则的专家系统和任何包含基于规则的专家系统组件的机器学习方法相关的某些限制。我们的模型为解决计算反合成分析这一具有挑战性的问题提供了重要的第一步。少

2017 年 6 月 6 日提交;最初宣布 2017 年 6 月。

25. 应用评估信息度量: 第 1 部分

作者:[henk f. moed](#)

摘要: 本手稿是将于 2017 年夏天由 springer 出版的《应用评估指标》一书第 1 部分（一般介绍和概要）的预印版本。这本书介绍了应用评价信息度量领域，是为来自科学和学术各领域的感兴趣的学者和学生撰写的。它勾画了这个领域的历史、最近的成就，以

及它的潜力和局限性。阐述了多维研究绩效的概念,并讨论了 28 个基于城市、**专利**、声誉和高度指标的利弊。此外,它还将定量研究评估作为一项评价科学,重点介绍了超度量因素在制定指标方面的作用及其应用的政策背景。它还讨论了前进的道路,无论是为用户还是为信息度量工具的开发人员。少

2017 年 5 月 17 日提交;最初宣布 2017 年 5 月。

26. 用神经机器翻译翻译技术术语的大词汇翻译专利句子

作者:[紫龙](#),[武仁乌图罗](#),[三桥友春](#), 山本美雄

摘要: 神经机器翻译 (nmt) 是一种新的机器翻译方法,其效果与统计机器翻译 (smt) 等传统方法相当。尽管 nmt 最近取得了成功,但它无法处理更多的词汇,因为训练复杂性和解码复杂性随着目标单词数量的增加而相应增加。在翻译专利文件时,这个问题变得更加严重,因为专利文件包含了许多很少观察到的技术术语。在 nmt 中,词汇中的单词由一个未知的令牌表示。在本文中,我们提出了一种方法,使 nmt 能够翻译专利句子,包括大量的技术术语词汇。我们对一个关于双语数据的 nmt 系统进行培训,其中技术术语被技术术语令牌取代;这使得它可以翻译除技术术语以外的大部分源句子。此外,我们将其用作解码器,用技术术语令牌翻译源句,用使用 smt 的技术术语翻译替换令牌。我们还使用它根据 smt 分数和 smt 分数的平均值来重新翻译 1,000 个

最佳 smt 翻译。用技术术语令牌对翻译后的句子进行重新标记。我们对日中专利句的实验表明, 与传统的 smt 系统相比, 拟议的 nmt 系统实现了高达 3.1 BLEU 点和 2.3 ribes 点的实质性改进, 并提高了约 0.6 BLEU 点。0.8 ribes 点超过一个等效的 nmt 系统, 而不使用我们提出的技术。少

2017 年 4 月 14 日提交;最初宣布 2017 年 4 月。

27. 分枝熵选大词汇的神经机器翻译模型

作者:紫龙,木村龙一郎,武仁宇郎,三桥友春, 山本美雄

摘要: 神经机器翻译 (nmt) 是一种新的机器翻译方法, 其效果与统计机器翻译 (smt) 等传统方法相当。尽管 nmt 最近取得了成功, 但它无法处理更多的词汇, 因为训练的复杂性和解码的复杂性随着目标词的数量增加而相应增加。在翻译专利文件时, 这个问题变得更加严重, 因为专利文件包含了许多很少观察到的技术术语。在本文中, 我们建议使用分支熵的统计方法来选择包含词汇外单词的短语。这使得拟议的 nmt 系统可以应用于任何语言对的翻译任务, 而不需要任何特定语言的技术术语识别知识。然后, 在培训期间用代币替换选定的短语, 并在翻译后被 smt 的短语翻译表取代。对日式到中文、中、日、英文和英语到日语专利的评估句子翻译证明了用分支熵选择短语的有效性, 在没有我们提出的技术的情况下, 所提出的 nmt 模型比基线 nmt 模型

有了实质性的改进。此外, 在没有我们提出的技术的情况下, 基线 nmt 模型的翻译不足的翻译误差数通过所提出的 nmt 模型减少到一半左右。少

2017 年 9 月 6 日提交;v1 于 2017 年 4 月 14 日提交;**最初宣布** 2017 年 4 月。

28. 测绘专利分类: 投资组合与统计分析, 以及优势与劣势的比较

作者:[loet leydesdorff](#), [dieter franz kogler](#), [bowen yan](#)

文摘: 由欧洲和美国**专利局**共同开发的合作**专利分类** (cpc) 为制图和投资组合分析提供了新的基础。此更新为重新思考参数选择提供了机会。新地图与以前的地图有很大不同, 尽管这在目视检查中可能并不总是显而易见的。由于这些地图是基于索引术语的统计结构, 它们的质量--不同于效用--只能被无序地控制。我们在网上提供嵌套地图和投资组合覆盖的例程, 并进一步进行统计分析。我们为 "差异地图" 添加了一个新工具, 通过比较 2016 年授予诺华和 msd 的**专利组合**来说明这一工具。少

2017 年 10 月 14 日提交;v1 于 2017 年 2 月 24 日提交;**最初宣布** 2017 年 2 月。

29. 在基因组序列中寻找 rna 第三系运动的一种计算方法

作者:[kevin byron](#), [jason t. l. wang](#)

文摘: 在 dna、ma 和蛋白质中的动机发现在生命科学研究中起着重要的作用。有关生物分子数据中的主题发现的**最新专利**记录在 dna **专利**数据库中, 该数据库为决策者和对基因组学、遗传学和遗传学等领域感兴趣的公众提供了资源。生物。本文提出了一种在基因组序列中挖掘 ma 三级图案的计算方法。具体而言, 我们描述了一种方法, 称为 cs 姆矿工, 用于发现 ma 同轴螺旋堆叠在基因组中。同轴螺旋堆叠发生在 ma 三级结构中, 两个独立的螺旋元素形成一个伪连续的螺旋, 并为整个分子提供热力学稳定性。实验结果证明了我们方法的有效性。少

2017 年 1 月 2 日提交;最初宣布 2017 年 1 月。

30. 专利检索: 文献综述

作者:[walid shalaby](#), [wlodek zadrozny](#)

摘要: 随着每年提交的**专利**申请数量不断增加, 对管理如此大量数据的有效和高效系统的需求不可避免地变得很重要。**专利技术**检索 (pr) 被认为是几乎所有**专利**分析任务的支柱。pr 是信息检索 (ir) 的一个子领域, 它涉及开发技术和方法, 以便根据给定的检索请求有效和高效地检索相关的**专利**文档。本文对公关的方法和途径进行了全面的综述。显然, 如果不经深思熟虑的领域适应和定制, 在网络搜索等 ir 应用程序中最近的成功和成熟是无法直接转移到 pr 的。此外, 在自动公关方面最先进的性能仍然

在平均水平左右。这些观察促使人们需要交互式搜索工具，这些工具以最小的努力为**专利**专业人员提供认知帮助。这些工具也必须与**专利**专业人员同时开发，同时考虑他们的做法和期望。我们还涉及公关的相关任务，如**专利**评估、诉讼、许可，并强调这些领域的计算科学家的潜在机会和开放方向。少

2017 年 1 月 2 日提交;最初宣布 2017 年 1 月。

31. 神经机器翻译与有机化学反应预测的联系

作者: [juno nam](#), [jurae kim](#)

摘要: 寻找化学反应的主要产物是有机化学的重要问题之一。本文介绍了一种将神经机器翻译模型应用于有机化学预测的方法。为了将 "反应物和试剂" 转化为 "产品", 构建了一个基于门控递归单元序列模型和一个解析器, 用于从反应 smiles 字符串生成模型的输入令牌。训练集由**专利**数据库中的反应和手动生成的反应组成, 这些反应应用于韦德有机化学教科书中的基本反应。通过实例和问题对培训模型进行了测试。预测过程不需要手动编码规则 (例如, sarts 转换) 来预测产品, 因此它只需要足够的训练反应集就可以学习新类型的反应。少

2016 年 12 月 29 日提交;最初宣布 2016 年 12 月。

32. 开始使用 patstat 注册

作者: [gaetan de rassenfosse](#), [martin kracker](#) , [gianluca tarasconi](#)

文摘: 本文对 path 注册数据库进行了技术介绍, 其中包含了欧洲专利局处理的专利申请的书目、程序和法律地位数据。它提供了八个 mysql 查询, 涵盖了数据库中一些最相关的方面, 用于研究目的。它针对熟悉 path 数据库和 mysql 语言的学术研究人员和从业人员。少

2016 年 12 月 22 日提交;最初宣布 2016 年 12 月。

33. 城市专利组合分析: 统计与技术创新地图

作者: [dieter franz kogler](#), [gaston heimeriks](#), [loet leydesdorff](#)

摘要: 城市是知识经济的引擎, 因为它们是知识生产活动的主要场所, 随后决定了技术变革和经济增长的速度和方向。专利提供了丰富的信息, 用于分析特定地方的知识专业化, 如技术细节和有关发明人和实体的信息, 包括地址信息。每个专利文件上的技术代码都表明了特定发明的基础技术知识的专业化和范围。在本文中, 我们介绍了专利方面的组合分析工具, 为城市的技术专业化提供了见解。利用美国专利商标局 (uspto) 网站 (<http://www.uspto.gov>) 和 专 用 工 具 (<http://www.leydesdorff.net/portfolio>) 的数据绘制和分析城市专利组合, 可用于分析城市间创造性活动的专业化模式。研究结

果使决策者和其他利益攸关方能够确定进一步知识发展和 "智能专业化" 战略的有希望的领域。少

2016 年 12 月 17 日提交;最初宣布 2016 年 12 月。

34. 实现中国专利的准确词汇分割

作者:思力,薛念文

摘要: 专利是政府授予发明人的发明的财产权。发明是解决特定技术问题的办法。因此,专利往往集中了日常语言中罕见的科学和技术术语。在现有日常语言数据集上训练的中文分词模型由于无法有效识别这些科学和技术术语而表现不佳。本文描述了一种实用的专利汉语词分割方法, 通过从 142 项专利的手动分割语料库中提取特征, 训练基于字符的半监督序列标记模型,从中国树岸提取的信息得到了增强。实验表明, 在一个有效的测试集中, 我们的模型的精度达到了 95.08% (f1 分数), 在开发集中达到了 99.59%, 而在开发集中, 如果在中国树岸训练, 则 f1 分数为 95.08。我们还对现有的一些领域适应技术进行了实验, 结果表明, 目标域数据的数量和所选择的特征对域适应技术的性能有影响。少

2016 年 11 月 30 日提交;最初宣布 2016 年 11 月。

35. 使用 daala 视频编解码器的感知驱动视频编码

作者: [yushincho](#), [thomas j. daede](#), [nathan e.egge](#) , [guillaume martres](#), [tristanmatthews](#), [christopher montgomery](#), [timothy b. 特里伯里](#), [jean-marc valin](#)

摘要: daala 项目是一个免版税的视频编解码器, 试图与最好的专利支持编解码器竞争。我们的部分策略是用替代方法替换传统视频编解码器的核心工具, 其中许多方法旨在考虑感知方面, 而不是针对 psnr 等简单指标进行优化。本文记录了我们在这些工具方面的一些经验, 哪些工具有效, 哪些工具无效。我们评估哪些工具易于集成到更传统的编解码器设计中, 并在开放媒体联盟正在开发的编解码器的上下文中显示结果。少

2016 年 10 月 8 日提交;最初宣布 2016 年 10 月。

36. 基于仿真模型的技术不连续性定量识别

作者: [hyunseok park](#), [christopher l. magee](#)

摘要: 本文的目的是开发和测试度量, 以定量识别知识网络中的技术不连续性。我们在创新理论的基础上开发了五个度量, 并通过基于仿真模型的知识网络和假设设计的不连续性对指标进行了测试。设计的不连续性被建模为一个节点, 它结合了两个不同的知识流, 其知识在知识网络中占主导地位。通过如何将设计的不连续性与知识网络上的其他节点区分开来, 对所提出的指标的性能进行了评价。仿真结果表明, 收敛主路径的持久性时间 # 为识

别设计的不连续性提供了最佳性能: 设计的不连续性被确定为前 3 名专利之一, 概率为 96 ~ 99%。公制 5, 它是, 根据一个域的大小, 12 ~ 34% 优于性能的第二最好的指标。除了模拟分析之外, 我们还使用磁性信息存储域的**专利**集测试了度量值。在定性研究的基础上, 选择了与该领域著名突破技术——超磁阻 (gmr) 自旋阀传感器有关的三项代表性**专利**, 并通过对各项指标进行了测试。指标将选定的**专利**标识为排名靠前的**专利**。实证结果完全支持仿真结果, 因此建议收敛主干道的持久性时间 #, 以识别任何技术的技术不连续性。少

2016 年 9 月 13 日提交;最初宣布 2016 年 9 月。

37. 追踪技术发展轨迹: 一种遗传知识持久性主干道方法

作者: [hyunseok park](#), [christopher l. magee](#)

文摘: 本文的目的是提出一种利用**专利引用**来识别技术领域主要路径的新方法。以前使用主路径分析的方法大大提高了我们对实际技术轨迹的理解, 但仍有一些局限性。它们很有可能从已确定的主要途径中错过一些主要**专利**;然而, 其主要路径的高网络复杂性使得对轨迹的定性跟踪存在问题。该方法从基于标准遗传知识持久性算法的高持久性**专利**中搜索向后和向前路径。我们通过将新方法应用于海水淡化和太阳能光伏域进行了测试, 并使用先验方法将结果与相同域的输出进行了比较。实证结果表明, 该方法

克服了上述缺点, 定义了几乎 10 倍不复杂的主要路径, 同时包含了比现有方法定义的主路径网络更多的相关重要知识。少

2016 年 8 月 26 日提交;最初宣布 2016 年 8 月。

38. daala: 利用非常规技术构建新一代视频编解码器

作者: [jean-marc valin](#), [timothy b. 参述](#), [nathan e. egge](#), [thomas daede](#), [yushin cho](#), [christopher montgomery](#), [michael bebenita](#)

摘要: daala 是一款全新的免版税视频编解码器, 试图与最先进的带版税的编解码器竞争。要做到这一点, 它必须实现良好的压缩, 同时避免他们所有专利技术。我们使用尽可能不同于传统方法的技术来实现这一目标。本文介绍了 daala 背后的技术, 并讨论了它在开放媒体联盟新创建的 av1 编解码器中的位置。我们表明, daala 正在接近更成熟、最先进的视频编解码器的性能水平, 并可为改进 av1 做出贡献。少

2016 年 8 月 5 日提交;最初宣布 2016 年 8 月。

39. 埃博拉研究基础、临床和专利知识的霸权结构: 美军简化主义倡议

作者: [david fajardo-ortiz](#), [josw ortega-sanchez-de-tagle](#), [vorch-m castano](#)

摘要 背景: 在本文中, 我们提出了一种方法, 以了解埃博拉的基础知识、临床知识和**专利**知识是如何组织和相互交流的, 以及是什么主要因素可能决定了埃博拉知识翻译过程的演变。这种疾病。方法: 引文网络分析的结合;对医学标题主题 (mesh) 和基因本体论 (go) 术语进行了分析, 并对**专利**和科学文献进行了定量含量分析, 旨在绘制埃博拉研究的组织结构图。结果: 我们发现了六个假定的研究前沿 (即高相互关联的论文集群)。三个方面是埃博拉病毒结构蛋白的基础研究: 糖蛋白、vp40 和 vp35。关于发病机制的基础研究论文有第四个研究前沿, 是埃博拉研究的组织中心。第五个研究领域是以疫苗和糖蛋白为重点的临床前研究。最后, 确定了与疾病爆发有关的临床流行病学研究前沿。**专利**家族的网络结构表明, 主要设计是将埃博拉病毒蛋白作为疫苗和其他免疫治疗的目标。因此,**专利**网络组织类似于科学文献的组织。具体而言, 关于埃博拉的知识将从更高的 (临床流行病学) 流向中介 (蜂窝组织), 流向组织水平较低 (分子相互作用)。结论: 我们的研究表明, 埃博拉研究的一种强有力的简化方法可能受到这种疾病杀伤力的影响。另一方面,**专利**家族网络的所有权状况以及与美国陆军的主要研究关系表明, 该军事机构大力参与埃博拉研究。少

2016 年 4 月 12 日提交;最初宣布 2016 年 7 月。

40. 生物 dna 中数据存储的重复校正码

作者 :Siddharth jain, farzad farnoud, moshe schwartz, jehoshua bruck

摘要: 将数据存储在生物体 dna 中的能力在许多领域都有应用, 包括合成生物学和专利转基因生物的水印。存储在此介质中的数据可能会受到各种突变(如点突变、indels 和串联重复)引起的错误的影响, 这些突变需要进行更正以保持数据完整性。在本文中, 我们为串联复制引起的错误提供了纠错代码, 这些重复创建了序列块的副本, 并以串联方式(即在原始副本旁边)插入。特别是, 我们提出了两个代码系列, 用于纠正由于固定长度的 tandem-重复而导致的错误, 第一个系列可以纠正任意数量的错误, 而第二个系列可以纠正有限数量的错误。我们还研究了将长度的串联重复校正到给定常数的代码 K , 我们主要关注的是 $k=2,3$ 个。最后, 我们提供了串联重复中允许的长度集的完整分类, 这些长度集合为所有序列提供了一个唯一的根。少

2016 年 6 月 1 日提交;最初宣布 2016 年 6 月。

41. 网络中的多核-外围结构

作者:严宝文,罗建熙

文摘: 许多真实世界的网络表现出多面体-外围结构, 在多个核心中密集连接的顶点, 周围环绕着稀疏连接顶点的一般外围。多面体-外围结构的识别可以了解各种真实世界网络的结构和功能

提供新的镜头。本文定义了多核-外围结构, 并介绍了一种识别一般网络中多个核心和外围的最优划分的算法。我们通过将其应用于一个著名的社交网络和**专利技术**网络来演示我们的算法的性能, 这些网络的最佳特点是多核-外围结构。分析还揭示了我们的多核-外围检测算法和两种最先进的检测单核-外围结构和社区结构的算法之间的区别。少

2018 年 8 月 26 日提交;v1 于 2016 年 5 月 11 日提交;**最初宣布** 2016 年 5 月。

42. 改进自动专利申请分析: 数据集、系统和实验

作者:mgkehu, david cinciruk, john maclaren walsh

摘要: 现成的自然语言处理软件在分析**专利申请**时表现不佳, 因为它们使用的是相对于从新闻文章和通常用于培训此软件的网站上构建的语料库的不规则语言。阻止了积累足够大的数据集以完全重新训练用于**专利申请**的解析器的广泛而昂贵的过程, 这是一种将现有的自然语言处理软件与专利索赔相适应的方法。提出了语音标记校正的受力部分。讨论了亚马逊机械土耳其人收集活动, 以生成公共语料库, 培训这种改进的索赔分析系统, 确定在该活动中吸取的经验教训, 这些经验教训可用于与 amt 的未来 nlp 数据集收集活动。利用该语料库和其他**专利权利集**的实验测量通过声明解析系统获得的解析性能改进。最后, 通过实验证明了在

新的权利分析系统为新的**权利**分析系统时改进的自动专利主体分类, 证明了改进**后**的权利要求分析系统在其他专利处理申请中的效用。用于生成要素。少

2016 年 5 月 5 日提交;最初宣布 2016 年 5 月。

43. 技术领域的分解与分析, 以更好地理解技术结构

作者:[郭欣](#) , [玄石公园](#),[克里斯托弗 l. magee](#)

摘要: 专利是与技术变革有关的最完整的信息来源之一。本文介绍了美国**专利**分析领域三个月的研究. 该方法包括使用搜索词来定位最具代表性的国际和美国**专利**类别, 并确定这些类别的重叠, 以得出最后一组**专利**, 并使用开发的预测模型由本森和马吉计算技术领域的技术改进率。我的研究重点是生物化学药理学技术领域, 并为该领域的技术领域和子领域选择相关**专利**。目标是更好地了解技术领域的结构, 了解域及其子域的进展速度有多快。我使用的方法是由本森和 magee 开发的, 它被称为分类重叠方法 1, 它提供了一种可靠的、基本自动化的方法, 将**专利**数据库分解成可以理解的技术领域, 在这些领域可以取得进展测量。少

2016 年 4 月 19 日提交;最初宣布 2016 年 4 月。

44. 在专利中引用的睡眠美女: 是否也有发明宿舍?

作者:[anthony f. j. van raan](#)

摘要:《科学中的睡美人》是一本长时间不被人注意的出版物, 后来, 几乎突然间, 吸引了很多人的注意 (被王子唤醒)。在我们前面的研究中, 我们发现大约一半的睡眠美女是以应用为导向的, 因此是潜在的睡眠创新。在本文中, 我们研究了一个新的主题: 在**专利**中引用的睡眠之美。通过这种方式, 我们探索发明宿舍的存在。我们发现, **专利**引文可能发生在觉醒之前或之后, 睡眠的深度, 即睡眠期间的引文率, 并不能预测睡美以后的科学或技术影响。发明者作者的自我引用只发生在**专利**中引用的少数睡眠美女中, 但其他类型的发明者-作者链接发生的频率更高。我们分析他们是否处理新的主题, 通过测量整个科学文献中的时间依赖演变的数量有关的两个明确定义的主题, 以及更广泛的研究主题的睡美期间和在睡觉后。我们通过分析第一组引用睡美人的论文, 专注于觉醒。接下来, 我们创建与主题有关的概念图和引用论文的概念图, 时间在觉醒之后的一段时间内, 以及最近的一段时间内。最后, 我们对睡美人的引用和引用关系进行了广泛的评价。我们发现, 可调谐的共同引文分析是一个强大的工具, 以发现王子和其他重要的应用工作直接相关的睡美人, 例如作者写的论文引用睡美人在这两项**专利**其中, 他们是发明者, 以及在他们的科学论文。少

2016 年 4 月 19 日提交;最初宣布 2016 年 4 月。

45. 利用波束形成和粒子过滤对声源进行可靠的三维定位和跟踪

作者:jean-marc valin, françois michaud, jean rouat

文摘: 本文提出了一种新的使用 8 个麦克风数组 (美国正在申请专利) 的鲁棒声源定位和跟踪方法。该方法采用了基于可靠性加权相位变换 (rwphat) 的定向波束发生器和基于粒子滤波器的跟踪算法。拟议的系统能够估计来源的方向和距离。在视频会议的背景下, 对方向的估计精度高于 1 度, 而距离在 10% 的 rms 范围内准确, 在嘈杂的环境中显示对最多三个同时移动的扬声器的跟踪。少

2016 年 2 月 27 日提交;最初宣布 2016 年 4 月。

46. daala: 一个感知驱动的下一代视频编解码器

作者 : [thomas j. daede](#), [nathan e. egge](#), [jean-marc valin](#), [guillaumemartres](#), [timothy b. 特里伯里](#)

摘要: daala 项目是一个免版税的视频编解码器, 试图与最好的专利支持编解码器竞争。我们的部分策略是用替代方法替换传统视频编解码器的核心工具, 其中许多方法旨在考虑感知方面, 而不是针对 psnr 等简单指标进行优化。本文记录了我们在这些工具方面的一些经验, 哪些工具有效, 哪些工具无效, 以及我们从这些工具中学到了什么。其结果是一个编解码器, 它与静止图像上的 hevc 比较有利, 并且也在视频的路径上这样做。少

2016 年 3 月 9 日提交;最初宣布 2016 年 3 月。

47.一种不规则的多核和 gpu 处理器体系结构图算法的研究与评价

作者:瓦伦·纳格帕尔

摘要: 计算应用程序的一个领域是不规则的应用程序, 它对芯片多处理器 (cmp) 的性能可扩展性提出了重大挑战。与计算绑定应用程序相比, 此类应用程序的计算很少, 内存访问模式不可预知。由于处理器和内存性能之间的差距继续存在, 隐藏和缩小这一差距的困难是导致 cmp 上这些应用性能不佳的重要因素之一。本文的目的是克服一种称为三合会普查的不规则图形算法在性能加速过程中带来的许多此类挑战。我们在两个截然不同的芯片多处理器上加速了 triad 普查算法: 双插槽英特尔至强多核 (每个插槽 8 个硬件线程) 和 240 处理器核心 nvidia tesla c1060 GPGPU(128 每个内核的硬件线程)。在英特尔多核至强系统上获得的实验结果表明, 对于真实世界图形数据集, 性能的速度最高为 56x、平均 33x 和最小 8.3x。在 nvidia tesla c1060 gpgpu 上, 我们几乎能够匹配多个核的结果—最大 588.4 倍, 平均 32.倍和 4.2 倍最小速度。在原始性能方面, 对于称为"专利网络" 的图形数据集, 我们在英特尔至强 Multicore(16 hw 线程) 上的结果比以前 cray XMT(16 w 线程) 的结果快 1.27 倍, 而在 gpgpu 上实现的结果则相对较快 slower(0.72x)。据我们所知, 该算法仅在名为 cray xmt 的超计算机类计算机上加速, 没有任何工作可以

证明该算法在成本相对较低的多核和 gpgpu 上的性能评估和比较平台。少

2016 年 3 月 8 日提交;最初宣布 2016 年 3 月。

48. 对新兴技术趋势的认可。美国专利引文网络引文的经典研究

作者 :[péter bruck](#), [istván réthy](#), [judit szente](#), [jan tobochnik](#), [péter erdi](#)

摘要: 通过采用基于城市的专利递归排名方法, 可以追溯到新的技术领域的演变。具体而言, 它表明, 激光/喷墨打印机技术是由两种现有技术的重组产生的: 顺序打印和静态图像制作。来自不同 "前兆" 类别的引用动态揭示了新领域出现的机制, 并为预测未来的技术发展提供了可能性。对于**专利网络**, pagerank 阻尼因子的最佳值接近 0.5;应用 d 顺 0.85 会导致不可接受的排名结果。少

2016 年 2 月 25 日提交;最初宣布 2016 年 2 月。

49. 用于 uspto 发明者名称消歧的随机森林 dbscan

作者:[kunho kim](#), [madian khabsa](#), [c. lee giles](#)

摘要: 名称消除歧义和随后的名称混合对于正确处理数字库或其他数据库中的人员姓名查询至关重要。它将每个独特的人与数据库中的所有其他记录区分开来。利用早期作者姓名消歧工作的方

法和特点,研究了**专利**数据库中的发明人名称消歧问题,并提出了适合**专利**数据库的特征集。选择了一个随机林作为配对链接分类器,因为它们的性能优于朴素贝叶斯、逻辑回归、支持向量机(svm)、条件推理树和决策树。阻塞大小,非常重要的缩放,是根据实验确定的特点的重要性的选择。利用随机林分类器导出的距离函数,采用 dbscan 算法对记录进行聚类分析。对于额外的可伸缩性,群集是并行的。对 uspto 专利数据库的测试表明,我们的方法成功地消除了 1200 万发明者提到在 6.5 小时内的歧义。对 uspto patentsview 发明者名称消除歧义竞争的数据集的评估表明,我们的算法在竞争中优于所有算法。少

2017 年 9 月 14 日提交;v1 于 2016 年 2 月 4 日提交;**最初宣布** 2016 年 2 月。

50. 引用引用本福特定律的文章: 本福特分析

作者:[tariq ahmad mir](#)

摘要: 大数据中数字的第一个有效数字的出现往往受称为本福特定律 (bl) 的对数减少分布的制约,这首先由 s. newcombe (sn) 报告,几十年后由 f. benford (fb) 独立报告。由于法律的反直觉,几十年来,法律被忽略,仅仅是一种奇怪的观察。然而,其显著复苏的一个迹象是 sn/fb 的论文收到的引文数量大幅增加。法律已经走了很长的路,从默默无闻到现在成为书籍、同行评审论文、

专利、博客和新闻的固定主题。在这里，我们使用 google 学者 (gs) 收集引用 sn/fb 原文的文章所收到的引用次数的数据，然后调查这些引用数据的主要数字是否按照他们发现的法律分布。我们发现，关于 bl 的文献引用数据与法律的预测是显著一致的。少

2016 年 3 月 19 日提交;v1 于 2016 年 2 月 3 日提交;最初宣布 2016 年 2 月。

51. 文档图像分类, 对专利图像的应用有具体的看法

作者:[gabriela csurka](#)

摘要: 本文的主要研究重点是文档图像的分类和检索, 分析和比较基于 run 价直方图 (rl) 和基于 fisher vector (fv) 的图像表示的不同参数。我们使用不同的文档图像数据集进行详尽的实验研究, 包括 marg 基准、基于客户数据构建的两个数据集以及 clef-ip 2011 专利图像分类任务中的图像。这项研究的目的是提供如何最好地选择参数的指导, 以便相同的功能在不同的任务中表现良好。作为此类需求的一个示例, 我们描述了 clef-ip 2011 基于图像的专利检索任务, 在该任务中, 我们使用相同的图像表示来预测图像类型并检索相关专利。少

2016 年 1 月 13 日提交;最初宣布 2016 年 1 月。

52. 利用高分辨率地理位置数据消除专利发明人和受让人的歧义

作者: [greg morrison](#), [masimo riccaboni](#), [fabio pammolli](#)

摘要: 专利数据是通过引用、共同发明和共同**转让新专利**网络获得创新和技术演变的重要信息来源。从这些数据中提取有用信息的一个主要障碍是名称消除歧义问题: 将个人或机构的替代拼写与单个标识符联系起来, 以唯一地确定参与创建技术的各方。在本文中, 我们描述了一种新的算法, 该算法利用高分辨率的地理位置来消除发明人和受让人在**欧洲专利局(epo)**下发现的 360 多万项**专利**上的歧义。合作条约 (pct), 以及美国**专利商标局(uspto)**。我们证明, 与波士顿和巴黎的 epo 受让人名称的手动消除歧义相比, 我们的算法具有较高的精度和召回性, 并表明它在可链接到高分辨率地址 (但较差) 的美国专利商标局发明者姓名基准方面表现良好 (但较差) 为从未提供高质量地址的发明者提供)。这项工作的最重要的好处是高质量的受让人消除歧义与全球覆盖, 再加上发明者消除歧义, 是与其他国家的最先进的方法竞争。据我们所知, 这是在这三个主要**专利**收藏中, 发明人和受让人名称的最广泛和最准确的同时消除和交叉链接, 占专利的很大一部分。少

2015 年 12 月 13 日提交;最初宣布 2016 年 1 月。

53. 学术创业的全球化? 大学专利分解的近期增长 (2009–2014)

作者:loet leydesdorff, henry etzkowitz, duncan kushnir

摘要: 学术界对美国**专利**的贡献日益全球化。在学术创业国际化和 2008 年学术创业国际化的推动下, 1998 年至 2008 年, 大学**专利**在所有**专利**中所占份额上升的长期趋势又恢复了。美国大学技术转让的持续存在。我们从国家和**专利**类别的角度分列了美国**专利**和商标组织 (uspto) 大学**专利**最近的增长。2009–2014 期间, 美国的外国**专利**几乎翻了一番, 主要原因是台湾、韩国、中国和日本的大**学**申请**专利**。这些国家在**专利**组合方面与美国竞争, 而大多数欧洲国家——除英国外——都有更具体的投资组合, 主要是在生物医学领域。就中国而言, 清华大学拥有美国专利商标局 63% 的大学**专利**, 其次是法赫德国王大学, 占全国投资组合的 55.2%。少

2015 年 12 月 14 日提交;最初宣布 2015 年 12 月。

54. 基于认证的制成品假冒解决方案

作者:约瑟夫·基尔库伦

摘要: 假冒制成品的表现为对知识产权、**专利**、版权等盗窃, 并伴有身份盗窃。身份盗窃的目的是为知识产权盗窃提供便利。没有它, 知识产权盗窃将是显而易见的, 产品将被没收和销毁。然后为这两类制成品制定了身份查验解决方案, 以防止身份盗窃, 即可以接受破坏性筛选战略的货物和不能受到破坏性筛选战略的货

物,例如药品和货币,分别。所开发的解决方案被发现类似于数字签名。药品上的防篡改包装类似于加密,因为它可以防止 mallory 干扰产品。打破防篡改包装是单向功能。在包装内隐藏一个一次性密码,可用于通过互联网对产品进行身份验证。身份验证网站的名称必须是常识,就像用于验证数字签名的公钥一样。否则,造假者将指定自己的认证网站。这种解决方案可以改变货币,即单向功能,相当于打开防篡改包装,成为货币的制造方法。少

2016年10月3日提交;v1于2015年11月27日提交;**最初宣布** 2015年12月。

55. 上下文敏感文章排序与引文语境分析

作者:metin doslu , haluk o. b 冰 ol

摘要: 在特定的上下文中很难检测到重要的文章。基于全文搜索的信息检索技术可能不准确,无法识别主要主题,无法说明文章的重要性。生成引文网络是查找最受欢迎文章的好方法,但这种方法不了解上下文。引文标记周围的文字一般是被引用文章的一个很好的总结。因此,引文语境分析提供了一个机会,可以利用群众的智慧,以上下文敏感的方式检测重要文章。在这项工作中,我们分析引文上下文,以便为给定的主题正确地对文章进行排名。该模型利用引文上下文,建立了基于目标主题的定向加权引文网络。如果引文上下文包含与目标主题相关的术语,我们将在两篇文章

之间创建一个定向和加权边。然后我们应用常见的排名算法,以便在这个新创建的网络中找到重要的文章。我们表明,该方法成功地检测到了给定主题中最突出的文章的一个很好的子集。这种方法的最大贡献是,我们能够为给定的搜索词确定重要的文章,即使这些文章不包含这个搜索词。此技术可用于其他链接文档,包括网页、法律文档和**专利**。少

2016年4月26日提交;v1于2015年11月16日提交;**最初宣布** 2015年11月。

56. 知识翻译的层次结构: 从健康问题到特殊药物设计

作者:david fajardo, victor castano

摘要: 讨论了一种将新分子实体和纳米材料转化为药物替代品(即知识翻译、kt)的复杂性的创新方法。首先,回顾了生物医学知识的组织和翻译的一些关键概念(范式、同音、功率定律分布、层次结构、模块化和研究方面)。然后,我们提出了一个研究跨学科交际复杂性的药物发现知识翻译模型。具体而言,我们涉及两个高度相关的方面:(1)成功的 kt 需要出现有组织的跨学科研究机构;(2)这些知识机构的等级和模块化拓扑组织。我们重点研究了以前发表的一套关于 kt 的研究,这些研究依靠的是网络分析和计算机辅助分析对科学文献和**专利**内容的结合。选定的研究提供了两个互补的观点:知识需求(宫颈癌和埃博拉出血热)和知

识供应（脂质体和纳米颗粒治疗癌症和范式多西尔，第一个纳米块。批准）。少

2015年10月26日提交;v1于2015年10月20日提交;**最初宣布** 2015年10月。

57. 神经模糊算法 (nfa) 模型和评估工具

作者:[何志平](#), [luiz fernando capretz](#), [xishi huang](#), [jing ren](#)

摘要: 成本估算、质量估算、风险分析等准确估算是管理中的一个主要问题。我们提出了一个**正在申请专利**的软计算框架来解决这个具有挑战性的问题。我们的通用框架独立于估计的性质和类型。它由神经网络、模糊逻辑和算法估计模型组成。我们以构造成本模型 (cocomo)、方差分析 (anova) 和功能点分析为算法模型, 验证了神经模糊算法模型 (nfa) 在工业项目软件成本估算中的准确性数据。与单独使用算法模型相比, 我们的模型产生更准确的估计。我们还讨论了实现 nfa 模型的工具的原型。最后, 我们的路线图和方向, 以丰富模型, 以应对不同的估计挑战。少

2015年7月31日提交;**最初宣布** 2015年8月。

58. 社会网络分析在工业自由开放源码软件生态系统中的应用经验

作者:[jose Teixeira](#), [gregorio robles](#), [jesús gonzález-barahona](#)

摘要: 许多软件项目不再由一个组织在内部完成。相反, 我们正处于一个新的时代, 软件是由个人和组织组成的网络社区开发的, 他们之间的关系建立在共同利益的基础上。矛盾的是, 最近的研究表明, 软件开发实际上可以由竞争对手共同开发。例如, 据悉, 移动设备制造商苹果和三星在法庭上进行昂贵的专利战争的同时, 一直在开源项目上进行合作。采用案例研究的方法, 我们探讨竞争对手如何在开源领域进行协作, 采用多种方法, 将存档数据 (qa) 的定性分析与挖掘软件存储库 (msr) 和社会网络分析 (sna) 相结合。在探索 openstack 生态系统内的协作过程的同时, 我们的研究通过探索高网络开源生态系统中的群体、子社区和业务模型的作用, 为软件工程研究做出了贡献。令人惊讶的结果表明, 对同一收入模式的竞争 (即经营相互冲突的商业模式) 并不必要影响生态系统内的协作。此外, 在检测 openstack 社区的不同子社区时, 我们发现开发人员与同一公司 (即同音) 的开发人员合作的预期社会趋势并不存在 openstack 生态系统中。此外, 在解决一个新颖的、复杂的、未探索的开源案例的同时, 这项研究也为合作战略和高科技创业方面的管理文献做出了贡献, 并对企业内部的异质行为者如何进行了丰富的描述。高网络生态系统 (涉及个人、初创企业、老牌公司和公共组织) 共同开发了一个复杂的基础设施, 用于开源领域的大数据。少

2016 年 11 月 22 日提交;v1 于 2015 年 7 月 4 日提交;**最初宣布** 2015 年 7 月。

59. 语用学在法律规范代表中的作用

作者:shashishekar ramakrishna, lukasz gorski, adrian paschke

摘要: 尽管某一法律规定 "明显明确", 但其适用可能导致结果不完全符合法规的语义层面。法律案文中的模糊性是故意引起的, 以适应应适用这种规范的所有可能情况, 从而使语用学的作用也是代表法律规范和推理的一个重要方面。本文所考虑的语用学概念并不集中在与司法决策相关的方面。本文旨在阐明法律语言学中语用学的几个方面, 主要从知识表示的角度来阐述**专利法**的领域。本文所提出的哲学讨论是以格里斯和马莫的法律理论为基础的。少

2015 年 7 月 8 日提交;最初宣布 2015 年 7 月。

60. 时间和引文网络

作者:james r. clough, tim s. evans

摘要: 引文网络产生于一些不同的社会制度, 如学术界 (发表论文)、企业 (通过**专利**) 和法律 (通过法律判决)。引文代表着信息的传递, 因此研究引文网络的结构将有助于我们了解知识是如何传递的。引文网络与其他网络的区别在于时间;文档只能引用较旧的文档。我们建议现行的网络措施不应考虑时间所施加的强大限制。我们将用两种类型的因果感知分析来说明我们的方法。我们

将我们的方法应用于由关于 arxiv 的学术论文形成的引文网络、美国**专利**和美国最高法院的判决。我们的工具可以揭示, 通过标准网络度量似乎具有非常相似结构的引文网络, 结果具有明显不同的属性。我们的结论表明, 参考书目中的许多论文与工作没有直接关系, 我们可以提供一个重要引文的简单指标。我们建议我们的方法可能会突出对跨学科研究更感兴趣的论文。我们还量化了不同领域研究方向多样性的差异。少

2015 年 7 月 6 日提交;最初宣布 2015 年 7 月。

61. 对发明人和组织多样化路径可视化的专利地图进行筛选

作者:[严宝文](#),[罗建熙](#)

文摘: 在信息科学文献中, 最近的研究利用**专利**数据库和**专利**分类信息构建专利技术类的网络图。在这样的**专利技术**地图中, 几乎所有对技术类都是相连的, 而它们之间的大部分连接都是极其薄弱的。这一观察表明, 通过消除薄弱环节来过滤**专利**网络图的可能性。但是, 删除链接可能会降低网络对发明者或组织多样化的解释力。网络链接可以解释发明人和发明组织的**专利**组合多样化路径。我们测量了**专利**网络图的多样化解释力, 提出了一种客观地选择解释力和消除薄弱环节之间的最优权衡的方法。与以往基于任意阈值的滤波方法相比, 该方法可以消除一定程度的任意性, 并确定了以前在最佳权衡之外创建筛选器的过滤方法。该地

图旨在帮助对发明者、组织和其他创新推动者的技术多样化进行网络可视化分析, 并进行潜在的前瞻性分析。向多产的发明家(伦纳德·福布斯)和公司(谷歌)提出这样的应用。少

2016 年 6 月 25 日提交;v1 于 2015 年 7 月 6 日提交;**最初宣布** 2015 年 7 月。

62. 孩子是人类的父亲: 预见早期的成功

作者:李良岳,唐汉航

摘要: 了解推动高影响力科学工作的动态机制(如研究论文、专利)是一个长期争论的课题, 具有许多重要的影响, 从个人职业发展和招聘搜索, 到研究资源的管辖范围。最近在科学成功的特征描述和建模方面取得的进展使得有可能预测科学工作的长期影响, 其中数据挖掘技术, 特别是监督学习, 发挥着至关重要的作用。尽管取得了很大进展, 但在预测长期科学影响方面的若干关键算法挑战在很大程度上仍然是开放的。在本文中, 我们提出了一个联合预测模型来预测早期阶段的长期科学影响, 同时解决了这些开放的挑战, 包括学术特征设计、非线性、领域异质性和动力学。特别是, 我们将其表述为一个正则化的优化问题, 并提出了有效的、可扩展的算法来解决它。我们对大型、真实的学术数据集进行了广泛的实证评估, 以验证我们的方法的有效性和效率。少

2015 年 7 月 31 日提交;v1 于 2015 年 4 月 3 日提交;最初宣布 2015 年 4 月。

63. 测量专利测绘的技术距离

作者:严宝文,罗建熙

文摘: 最近在信息科学文献中的著作介绍了利用**专利**数据库和**专利**分类信息构建技术领域网络图的案例,旨在帮助竞争情报分析和创新决策。构建这样的**专利**网络需要适当衡量专利分类系统中不同类别**专利**之间的距离。尽管文献中存在着各种距离措施,但如何持续评估和比较,以及选择哪些方法构建**专利**技术网络图,尚不清楚。这种模糊性限制了这类技术地图的开发和应用。在此,我们建议通过分析所产生的**专利**网络图的结构特性的差异和相似性,对替代距离度量进行比较,并确定优越距离测量。利用 1976 年至 2006 年的美国**专利**数据和国际**专利**分类系统,我们比较了 12 项具有代表性的距离措施,这些措施量化了场间知识库的接近度、跨场多样化创新代理人的可能性或频率,以及专利类别在同一**专利**中的共同出现。我们的比较分析表明,基于规范化共同参考和发明者多样化可能性度量的**专利**技术网络图是最佳的代表。

少

2015 年 9 月 16 日提交;v1 于 2015 年 3 月 9 日提交;最初宣布 2015 年 3 月。

64. 创新系统中的知识生成效率: 结构效应与时间效应的关系

作者: [inga ivanova](#), [loet leydesdorff](#)

摘要: 利用美国每百万居民专利的时间序列, 可以区分知识生成周期。这些周期部分地与孔德拉蒂夫长波重合。它们之间坡度的变化表明了知识生成范式中的不连续性。知识生成范式可以用相互作用的维度来建模 (例如, 在大学-行业-政府关系中), 这些维度限制了创新系统的最大效率。模型中参数的最大值与经验波的回归系数的顺序相同。维度增加的机制被指定为自我组织, 导致现有关系被打破为一个类似分形的网络的更加多样化的结构。这种打破可以模拟类似于 2d 和 3d (科赫) 雪花。知识生成的推动导致了新兴技术, 预计这些技术将比以前更加多样化, 生命周期也会缩短。知识生成周期的时间跨度也可以用斐波那契数字来分析。这种观点允许预测未来可能的范式变化的预期日期。在政策影响方面, 这表明重点已从制造技术转向开发新的组织技术和人类互动形式少

2015 年 2 月 18 日提交;最初宣布 2015 年 2 月。

65. fstp 事实法律规范验证的规则推理

作者: [naouel karam](#), [shashishekar ramakrishna](#), [adrian paschke](#)

摘要:非显性或创造性步骤是大多数**专利法**体系中对**专利性**的一般要求。一项发明应在其现有技术之外的足够距离上, 以便获得**专利**。本文概述了法律规范的提出方法。

2014 年 12 月 5 日提交;最初宣布 2014 年 12 月。

66. 在 web2.0 中, npe 作为开放信息系统使用的专利—两个小型案例研究

作者:[abdelkader baaziz](#), [luc quoniam](#)

摘要: 围绕**专利**的信息系统是复杂的, 他们的研究加上 "开箱即用" 的创造性愿景, 克服了**专利**的严格基本功能。自新**专利**的编写以来, 我们多次指导围绕基于**信息**的**专利**范围的研究;现有**专利**的失效、增值信息的创建及其与其他信息系统的联系。传统的基于大量投资的研发是一种技术转让。但是,**专利**信息也是技术转让、创新和创造力的另一个有力工具。事实上, 从学术角度对**专利**进行研究, 虽然并不总是只关注财政收入, 但可以被视为一种 "非实践实体" (npe) 活动的一种形式, 被正确或错误地称为 "**专利巨魔**".我们会明白为什么这种活动的 "**专利巨魔**" 一词是有争议和不合适的。我们将在本文中描述的研究属于这一背景。我们展示了两个关于在新兴国家**有效利用专利**信息的案例研究, 第一个涉及巴西的制药业, 第二个涉及阿尔及利亚的石油工业。少

2014 年 11 月 26 日提交;最初宣布 2014 年 11 月。

67. 专利分类的多样性能证明技术生命周期吗？多样性的关键作用

作者:[loet leydesdorff](#)

摘要: 在先前对九种光伏电池材料技术的**专利分类**进行的研究中, leydesdorff 等人 (2015 年) 报告了 rao–starrling 多样性纵向发展的周期性模式。我们建议, 这些周期性模式可以用来表明技术生命周期。然而, 分解后, 周期完全是由于分类种类的增加和减少, 而不是由于测量为 (1–余弦) 的差距或技术距离。因此, 一个单频分量可以显示在周期图中。此外, 周期性模式与各自技术的发明者数量有关。有时品种的增加会导致发明者数量的增加, 但在早期阶段——当技术还在建设中的时候——也可能是相反的情况。由于周期的发展因此似乎独立于**专利**之间的技术距离,**专利地图**方面的可视化可以被认为是解决了一组分析不同的研究问题。少

2015 年 7 月 5 日提交;**v1** 于 2014 年 11 月 11 日提交;**最初宣布** 2014 年 11 月。

68. 浅谈导致超类型进化曲线的科技研究影响的增长

作者:[marco campani](#), [ruggero vaglio](#)

摘要: gartner 炒作曲线的经验和理论论证是技术生命周期分析领域中一个非常相关的开放问题。本文的研究范围是介绍一个简单的模型, 描述科技研究的增长影响, 在具体情况下, 科学是推

动技术发展的新理念的主要来源, 导致超类型的进化曲线。该模型的主要思想是, 在第一阶段, 一个新的特定领域(可以用出版物数量来衡量)的科学兴趣的增长基本上遵循经典的物流增长曲线。在第二阶段, 从更晚的触发时间开始, 基于科学思想(可以用**专利矿床**来衡量)的技术发展可以被描述为第一条曲线的积分(在数学意义上), 因为技术是基于科学知识的整体积累。该模型通过对有机发光二极管(oled)科学研究和**技术**以及其他新兴技术的出版和专利存款率的文献计量分析进行了测试。少

2014 年 10 月 31 日提交;最初宣布 2014 年 10 月。

69. 检测专利中的数字和零件标签: 基于竞争的图像处理算法的发展

作者 :christoph riedl, richard zanibbi, marti a.hearst , siyuzhu, michael menietti, jason crusan, ivan metelsky, karim r. lakhani

摘要: 我们报告了为期一个月的在线竞赛的结果, 在这场竞赛中, 参与者开发了算法来增强美国**专利**商标局(uspto)公布的**专利**文件的数字版本。目标是检测美国**专利**图纸页中的数字和零件标签。这次挑战吸引了 232 个小组, 共两个, 其中 70 个小组(30%)提交了解决方案。团队总共提交了 1, 797 个在竞争服务器上编译的解决方案。与会者报告说, 他们平均花费了 63 个小时来开发他们的解决方案, 结果总共有 5 591 个小时的开发时间。手动标记

的 306 项专利数据集用于培训、在线系统测试和评估。介绍了前 5 名系统的设计和性能, 以及赛后开发的一个系统, 该系统说明获胜团队在严格的时间和计算限制下产生了近乎最先进的结果。在第一名系统中, 数字区域检测的谐振和精度 (f-measure) 的谐波均值为 88.57, 正确识别数字标题的数字区域为只不过 8.81%, 零件标签检测和字符识别的谐波均值为 78.81。通过在线 uci 机器学习存储库提供竞赛数据和软件, 以激励图像处理社区的后续工作。少

2014 年 11 月 11 日提交;v1 于 2014 年 10 月 24 日提交;**最初宣布** 2014 年 10 月。

70. 反复合作所带来的创造性困境的证据

作者:hiroyasu inoue

摘要: 我们关注的是发明项目中的重复合作如何影响绩效。重复合作有两个相互矛盾的方面。一个积极的方面是团队发展或经验, 消极的方面是团队退化或衰落。由于这两种矛盾现象都被观察到, 发明者对于是否应该继续在一个团队中合作, 他们有一个两难的问题。以前没有对这一困境进行定量分析。我们利用来自日本和美国的**专利**数据, 对创意项目中的困境进行定量和广泛的分析。我们确认了三个预测, 以定量地验证这一困境的存在。第一个预测是, 团队获得的**专利**越大, 团队合作的时间就越长。第二个预测

是, 在一个团队做出一项了不起的发明后, 连续**专利**的影响会减少, 而这一发明是通过**专利**的影响来衡量的。第三个预测是, 对新球队影响的期望大于对过去相同球队成功的期望。我们发现这些预测在日本和美国公布的**专利**中得到了验证。在这三个预测的基础上, 我们可以定量地验证创意项目中的困境。我们还提出了退化的预防战略。一个是发展技术多样性, 另一个是在团队中发展发明者多样性。我们发现, 通过使用数据进行验证, 这两种策略都是有效的。少

2015年8月18日提交;v1 于 2014年9月22日提交;**最初宣布** 2014年9月。

71. 基于图的主题模型的交互式视觉探索

作者: [samuel rönqvist](#), [xioluwang](#), [peter sarlin](#)

摘要: 概率主题建模是一个流行且功能强大的工具系列, 用于在大型非结构化文本文档集中发现主题结构。虽然人们对建模算法及其各种扩展的关注是非常多的, 但很少有研究涉及如何以有意义的方式呈现或可视化主题模型。本文提出了一种利用图形直观地传达主题结构和意义的新型设计。通过描述性关键句连接主题节点, 图形表示揭示了主题的相似性、主题意义和共享的、模糊的关键字。同时, 该图可用于信息检索目的, 按主题或主题子集查

找文档。为了说明该设计的效用, 我们说明了它在组织和探索金融**专利**语料库中的用途。少

2014 年 11 月 27 日提交;v1 于 2014 年 9 月 19 日提交;**最初宣布** 2014 年 9 月。

72. 概念搜索的众源模糊和面向分类

作者 :richard absalom, marcus luzak-roschch, dap hartmann, aske plat

摘要: 寻找科学和技术的概念往往是一项艰巨的任务。为便利概念搜索, 创建了不同类型的人为元数据, 以界定科学和技术披露的内容。国际**专利**分类 (ipc) 和 medline 的 mesh 等分类方案是经过结构和控制的, 但需要训练有素的专家和中央管理层来限制歧义 (mork, 2013 年)。虽然可以处理民间组织的非结构化标签, 以产生一定程度的结构 (kalendar, 2010 年; 卡兰皮纳斯, 2012 年; sarasua, 2012 年; bragg, 2013 年) 人群享有的自由通常会导致精度降低 (stock 2007)。现有的分类方案缺乏灵活性和模糊性。由于人类比计算机更了解语言、推理、暗示、抽象以及概念, 因此我们建议利用人群的集体智慧。为此, 我们提出了一种新的分类方案, 该方案具有足够直观的方法, 便于人群使用, 但功能强大, 可以方便类推搜索, 并且具有足够的灵活性来处理歧义。该系统将加强现有的分类信息。公民科学工作 (良好, 2013 年) 将与语义网络和计算机智能联系起来, 通过提高授予**专利**的质量、减少两

面派研究和刺激面向问题的解决方案来支持创新设计。我们的设计原型正在准备中。众包模糊和刻面分类方案将允许更好地搜索概念, 并更好地获得现有的科学技术。少

2014 年 6 月 30 日提交;最初宣布 2014 年 6 月。

73.对新安全威胁类的反应

作者:[yuval elovici](#), [lior rokach](#)

摘要: 每一个新确定的安全威胁类别都会引发科学界和专业界的新的研究和开发工作。在这项研究中, 我们调查科学和专业团体对新确定的威胁类别的反应速度, 因为它反映在长期的**专利**、科学文章和专业出版物的数量上。研究了以下威胁类别: 网络钓鱼;sql 注入;僵尸网络;分布式拒绝服务;和高级持久威胁。我们的研究表明, 在大多数情况下, 科学界需要一年的时间, 行业需要两年多的时间才能对拥有**专利**的新威胁类别做出反应。由于新产品遵循**专利**, 可以合理地预期, 将有一个大约两到三年的窗口, 在这个窗口中, 没有有效的产品来应对新的威胁类别。少

2014 年 6 月 11 日提交;最初宣布 2014 年 6 月。

74.利用半正式的 kr 弥合法律执业者和知识工程师之间的差距

作者:[shashishekar ramakrishna](#), [adrian paschke](#)

摘要: 在 omg 语义和商务词汇表示 (sbvr) 中, 提出了在业务规则表示中使用结构化英语作为非技术用户的计算独立知识表示格式的方法。在法律领域, 我们也面临着类似的问题。正式表示语言, 如 oasis 法律规则和法律本体 (lkif、合法的 owl2 本体等), 支持技术知识工程师和自动化推理。但是, 它们几乎不能被没有计算机科学背景的法律领域专家直接使用。在本文中, 我们针对法律领域调整了 sbvr 结构化英语方法, 并实现了一个概念证明, 称为 kr4iplaw, 它使法律领域专家能够在计算独立的情况下代表他们在结构化英语中的知识, 从而, 他们, 更有用的方式。这种方法的好处是, 结构化英语方法的基本预定义语义使得转换为正式语言 (如 oasis legal 生殖 ml 和 owl2 本体) 成为可能。我们在**专利法**领域的方法是一个典范。少

2014 年 5 月 31 日提交;最初宣布 2014 年 6 月。

75. 引文网络中模式查找方法的案例研究

作者: [ben pringle](#), [mukai krishnamoorthy](#), [kenneth simons](#)

文摘: 为了回答有关该领域的问题, 使用多种方法对包括 led **专利**网络及其元数据在内的数据集进行分析。我们有兴趣找到元数据和网络结构之间的关系;例如, 网络中的核心**专利**是由较大的公司还是由较小的公司产生的? 我们首先探索没有任何元数据的网

络结构, 在引文分析中应用已知技术和简单的聚类方案。然后将这些技术与元数据分析相结合, 得出有关数据集的初步结论。少

2014 年 5 月 22 日提交;最初宣布 2014 年 5 月。

76. 带有示例查询的 patstat 数据库简介

作者:[gaetan de rassenfosse](#), [helene dennis](#) , [geert boedt](#)

文摘:本文介绍了 patstat 专利数据库。它提供了与研究目的相关并涵盖最重要数据表的十个热门查询的指导示例。它的对象是愿意学习数据库基础知识的学术研究人员和从业人员。

2014 年 4 月 2 日提交;最初宣布 2014 年 4 月。

77. 关于显示和调整窗口大小的发明

作者:[umakant mishra](#)

摘要: 在 gui 环境中, windows 的使用频率相当高。使用窗口的最大优点是每个窗口都创建一个虚拟屏幕空间。因此, 尽管物理屏幕空间限制为几英寸, 但使用窗口可以创建无限的屏幕空间来显示无数项目。窗口的使用便于用户在不同的窗口中同时打开多个程序或文档并与之交互。有时, 单个程序也可能打开多个窗口来显示各种项目。用户可以调整窗口的大小, 并根据需要将其位置时间移动到时间。但是, 窗口的大小、外观、位置、颜色、可见

性、可大小等方面存在一些问题。例如，窗口应该有一个最小和最大的大小，拖动和调整窗口的大小应该很容易，一个窗口不应该被另一个窗口遮挡，窗口应该调整其大小和位置，以配合分辨率的变化和显示环境等。本文阐述了美国**专利**数据库的**六项专利**，解决了与窗口显示和大小调整有关的问题。这些发明包括自动调整窗口大小和重新定位窗口，在有限的显示区域内显示窗口以适应的替代模式，结合空间和时间方法来调整窗口的大小，锁定和最小化窗口，以及使窗口在特定情况下不可见。少

2014 年 4 月 27 日提交;最初宣布 2014 年 4 月。

78. 选择 gui 元素的说明

作者:[umakant mishra](#)

摘要: 选择对象或元素是任何图形用户界面中的基本操作。在对对象执行任何操作（如拖动、复制、打开、删除等）之前，必须选择该对象。gui 可能提供选择任何单个对象甚至多个对象的功能。选择多个对象的功能可以为 gui 提供巨大的功能，因为用户可以一次对多个对象执行复杂的操作。但是，选择的过程并不像 gui 的用户所显示的那样简单。在某些情况下，选择机制的内部逻辑可能非常复杂。本文介绍了与选择机制相关的一些基本困难，并阐述了从美国**专利**数据库中选择的不同发明所提供的解决方案。少

2014 年 4 月 27 日提交;最初宣布 2014 年 4 月。

79. 在图形用户界面中使用颜色的发明

作者:umakant mishra

摘要: 颜色是任何图形用户界面 (gui) 的一个重要方面。颜色是用来使 gui 有吸引力和有意义的。但使用颜色也有困难。颜色使用不当会导致不利影响。错误的颜色在错误的位置可能会使 gui 看起来笨拙和混乱。除了美学问题,还有许多其他问题涉及的颜色太多。与颜色的使用有关的矛盾之一是 "gui 的颜色应该是可定制的, 以适应用户的喜好。但同时它也不应该是可定制的, 因为这会给其他用户带来烦恼和混乱。与使用颜色相关的另一个收缩是 "用户应显示所有 1600 万种颜色, 以精确选择所需的颜色, 但从另一个角度来看, 用户不应显示所有 1600 万种颜色, 因为这会造成混乱和困难选择。本文分析了从美国**专利**数据库中挑选出来的一些发明, 并说明了发明者是如何解决与图形用户界面中使用颜色有关的各种矛盾的。少

2014 年 4 月 27 日提交;最初宣布 2014 年 4 月。

80. 图形用户界面中表达情绪的发明

作者:umakant mishra

摘要: 传统的 gui 更机械, 不识别或沟通情绪。现代 gui 试图推断用户可能的情绪状态和个性, 并通过相应的情绪状态进行沟通。情绪以图形图标、声音、图片和其他方式表达。研究发现, 这些情绪在通信软件、交互式学习系统、机器人和其他自适应环境中尤其有用。已经开发了各种机制来通过图形用户界面表达情感。本文阐述了从美国**专利**数据库中挑选出来的一些有趣的发明。少

2014 年 4 月 27 日提交;最初宣布 2014 年 4 月。

81. 图形用户界面中使用的树导航器的发明

作者:[umakant mishra](#)

摘要: 树视图或树导航器用于显示以树的形式组织的分层数据。在树结构中, 有父节点和子节点。子节点可能进一步具有 n 级的后代。有许多方法可以使导航变得简单。其中一些是展开和折叠分支, 拆分树, 在单独的树中显示父节点, 缩放分支, 在不同方向滚动等。有效地处理大树仍然是一项困难的工作。这项工作仍在继续以更快的速度、更大的控制力、用户友好性和美观性来管理大量节点。本文阐述了从美国**专利**数据库中选择的树导航器上的五项发明。他们每个人都试图以不同的方式解决与树导航器相关的各种问题。每项发明也从 triz 的角度进行分析。少

2014 年 4 月 27 日提交;最初宣布 2014 年 4 月。

82. gui 中使用的对话框上的发明

作者:umakant mishra

摘要: 在特殊情况下显示警告、错误、确认等情况下, 这些对话框非常有用。一个典型的对话框显示在一个小窗口中, 其中包含一些文本消息以及一些供用户选择的选项。然而, 在编程和实现传统对话框中存在着某些困难, 如严格的编程工作、硬编码消息的刚性、模糊的屏幕空间等。有必要克服对话框中的这些困难, 使它们更有效率和更有用。对话框的模式也会产生一些限制。虽然模式对话框需要由用户显式关闭, 但无模式对话框的数量可能会增加, 并且难以控制。因此, 一个理想的对话框应该被剥夺上述所有缺点。该对话框不应模糊屏幕。用户应该能够打开多个对话框, 但不能模糊屏幕。本文分析了从美国专利数据库中选择对话框中的 5 项有趣的发明。每项发明都试图克服传统对话框的一些限制, 并提供一些创新功能。还从 triz 的角度分析每个解决方案。

少

2014 年 4 月 27 日提交;最初宣布 2014 年 4 月。

83. 关于图形用户界面中滚动和滚动条的 10 项发明

作者:umakant mishra

摘要: 滚动机制用于在有限的屏幕空间或窗口中显示和编辑大型文档。滚动机制可以包括垂直滚动条或水平滚动条,也可以同时移动文档的内容,或上下或向左和向右移动。屏幕上可能有表示键盘上导航键的导航按钮。用户可以单击这些按钮滚动屏幕。可能有非常不同的滚动方法,例如使用 "拇指",如 pdf 文档中普遍使用。滚动可以通过在免提环境中进行眼球跟踪来实现,在这种环境中,用户没有手或想用手进行其他活动。本文阐述了从美国专利数据库中选择十项有趣的滚动发明。少

2014 年 4 月 27 日提交;最初宣布 2014 年 4 月。

84. 图形用户界面中命令按钮上的 10 项发明

作者:umakant mishra

摘要: 命令按钮可能包含文本标签或图形图像,也可以同时包含这两个选项。它可以是静态的或动画的。可以有许多不同的功能,使命令按钮有吸引力和有效。由于命令按钮是典型的 gui 元素,因此通常对 gui 的大多数改进也适用于命令按钮。此外,还有一些发明,以改善各方面的命令按钮在具体。本文从美国专利数据库中展示了 10 项发明。每项发明之后都有一个基于 triz 的简要分析。少

2014 年 4 月 27 日提交;最初宣布 2014 年 4 月。

85. 关于改进工具栏的 10 项发明: 基于 triz 的分析

作者:umakant mishra

摘要: 工具栏是图形用户界面中最常用的元素之一。图形用户界面的其他流行元素是按钮、菜单、滚动条、对话框等, 所有这些都可以方便地访问 gui 系统的各种功能。工具栏通常执行与菜单类似的功能, 但存在某些差异。菜单的优点是可以容纳大量的项目, 而不需要任何额外的屏幕空间。相反, 工具栏上的每个按钮将永久占用屏幕上的一些空间。不可能通过工具栏实现大量的功能, 因为它们将占用越来越有价值的屏幕空间。但是, 工具栏的优势在于它提供了对任何功能的单击访问权限, 而不是菜单系统, 用户必须在子菜单中导航, 才能最终发现他要查找的项目。本文探讨了工具栏的理想功能和高级工具栏的理想功能。从 triz 的角度描述了这些矛盾, 并利用创新原则推导出了解决方案。此外, 本文还介绍了从美国专利数据库中选择的 10 项关于改进工具栏的发明。

少

2014 年 4 月 27 日提交;最初宣布 2014 年 4 月。

86. 适应菜单上的发明: 基于 triz 的分析

作者:umakant mishra

摘要: 菜单是图形用户界面中使用最广泛的元素之一。菜单系统的目的是以一种简单的方式向用户提供各种命令和功能, 以使用户只需从给定列表中选择所需的操作, 而无需在命令提示符中键入复杂的命令。在传统的菜单系统中, 菜单项或选项在计算机程序中进行硬编码。程序员或开发人员在开发时编写菜单项。开发人员尝试包括他认为将来用户可能需要的所有选项。虽然这些项目是根据 "需求分析" 和其他研究决定的, 但在未来的一段时间内, 很难知道用户的确切需求。这将导致在菜单中包含大量项目, 从而导致用户的困惑和沮丧。因此, 需要可根据用户要求更改可调整的菜单。根据程序上下文和用户选择的可能性, 可调整菜单中的项目应不时更改。本文定义了下拉菜单系统的理想最终结果, 定义了自适应菜单的理想功能, 发现并解决了在实现理想功能时面临的矛盾, 并在美国的适应性菜单上说明了六个选定的案例专利数据库。少

2014 年 4 月 27 日提交;最初宣布 2014 年 4 月。

87.使用单词关联向量的语法进行分析

作者:罗伯特·约翰·弗里曼

摘要: 本论文最初是在 2001 年起草的, 是美国美国专利 7, 392, 174 中描述的系统的正式化。它描述了一种基于上下文相似词的跨产品而不是向量的解析器实现系统。它现在是为了回应对语法

和语义的矢量组合模型的新生兴趣而出版的。该方法使用上下文相似的单词和单词组的积极替换,使产品向量能够保持在与它们的操作数相同的空间中,并使整个句子在句法上具有可比性,并有可能在语义上进行。生成的向量具有足够的表示强度,可以生成至少可与当代符号解析器相媲美的解析树。少

2014 年 3 月 10 日提交;最初宣布 2014 年 3 月。

88. 生物技术专利引文的决定因素:对整个工业和组织边界的专利影响分析

作者:[antonio messeni petrzzelli](#), [daniele rotolo](#), [vito albino](#)

文摘: 本文扩大了调查导致某些**专利**的关键驱动因素的文献,以便对随后的技术发展(发明)产生比其他技术发展(发明)更大的影响。我们调查了六个关键决定因素,如(一)科学知识的使用,(二)技术基础的广度,(三)专利开发方面合作的存在,(四)索赔的数量,(五)范围,(六)新颖性,以及如何当**专利影响**——以收到的**专利**的正向引用数量衡量——被区分为工业和组织内部和跨越范围时,这些决定因素的影响就会有所不同。我们对 1976 年至 2003 年授予 293 家美国生物技术公司的 5671 项**专利**样本进行了实证分析。结果表明,决定因素对**专利**影响的贡献因工业和组织边界的确定而异。例如,研究结果表明,科学知识的使用对**专利**在生物技术产业之外的**影响力**产生了负面影响,同时对**使专利与**

受让人的后续活动更加相关作出了积极贡献。技术发展。此外,专利范围越广,专利从随后的非生物技术专利中获得的引用数量就越多。这种关系是倒 u 形的,当考虑到专利对授予其他组织而不是专利受让人的发明的影响时。最后,专利的新颖性是与专利对随后跨越工业和组织边界的发明的影响有关的反 u。少

2014 年 3 月 9 日提交;最初宣布 2014 年 3 月。

89. 专利作为探索创新动态的工具: "光伏电池" 的地理和技术视角

作者 :loet leydesdorff, floortje alkemade, gaston heimeriks, rinke hoekstra

摘要: 创新的动态是非线性和复杂的: 地理、技术和经济选择环境可以预期是相互作用的。专利能否从不同的属性(如发明人地址、分类代码、后向和向前引用等)方面为这一过程提供一个分析镜头? 最近开发的两幅带有交互式叠加技术的专利地图——基于国际专利分类(ipc)引文关系的谷歌地图和地图——被详细地制作成动态版本,允许在线使用使用拆分屏幕进行动画和比较。探索各种形式的动画。美国专利和贸易局(uspto)和欧洲专利局(epo)最近开发的合作专利分类(cpc)为准确划分美国专利商标局数据和产品的样本提供了新的选择。epo 的全球专利统计数据库(patstat)。在 "减缓气候变化技术"(y02 类)中,我们放大了 9 种用于光伏电池的材料技术;并将其中一个(cuinse2)作为主要案例。基于 ipc 的地图中 rao-strarling 多样性的纵向发展为研究

本报告所述期间的技术世代研究提供了一个启发式方法 (1975–2012)。与 patstat 数据相比, 美国专利商标局的数据中的世代顺序更为普遍, 因为 patstat 汇总了处于不同技术发展阶段的国家的**专利信息**, 而人们可以预期美国专利商标局的**专利**将是在技术优势上的竞争力。少

2014 年 9 月 10 日提交;v1 于 2014 年 1 月 13 日提交;最初宣布 2014 年 1 月。

90. 视觉辅助导航的可观测性、可识别性和灵敏度

作者:[joshua hernandez](#), [konstantine tsotsos](#), [Stefano soatto](#)

摘要: 我们从视觉和惯性传感器的融合中分析了运动估计的可观测性。由于模型包含未知参数 (如传感器偏差), 因此问题通常被转换为混合标识/筛选, 而生成的可观测性分析为任何算法收敛到唯一点提供了必要条件估计。不幸的是, 大多数模型将传感器偏置率视为噪声, 独立于其他状态, 包括偏差本身, 这一假设在实践中**显然遭到了违反**。当这一假设被解除时, 生成的模型是不可观察的, 因此过去的分析不能用来得出结论, 与测量值无法区分的一组状态是一个单例。换句话说, 生成的模型是不可观察的。因此, 我们将分析重新转换为灵敏度分析: 我们不是试图证明无法区分的集合是一个单例, 而不是这样, 我们在其体积上推导出

边界, 作为输入特征及其充分激励的函数。这提供了可用于分析和验证目的无法区分的集合的显式描述。少

2015年4月26日提交;v1于2013年11月28日提交;**最初宣布** 2013年11月。

91. 观念传承、创新与集体创新

作者:[harris kyriakou](#), [jeffrey v. nikerson](#)

摘要: 为了创造新产品, 发明者搜索并结合以前的想法。很少有研究研究研究过导致新产品的搜索特征;大多数都集中在**专利**引文上, 这些引文往往是追溯性的, 可能不能反映发明的有用性。通过对在线虚拟社区中协作的分析, 检验了独创性对人气和实用性的影响。这些测试依次基于一种测量 3d 形状之间距离的方法。总之, 本文提出了一种衡量创新的新方法, 并提出了进一步了解技术在鼓励创新中的作用的的方法。从组织的角度来看, 这项工作提供了对创作过程的洞察, 特别是开放创新过程, 在这个过程中, 成千上万的个人共同进化设计, 而不属于相同的公司结构, 而不主张知识产权, 而不兑换金钱。少

2013年11月3日提交;**最初宣布** 2013年11月。

92. 模块化键盘上的 10 项发明: 基于 triz 的分析

作者:[umakant mishra](#)

摘要: 由于标准键盘相当宽敞,许多发明都试图利用键盘的空间进行各种活动。模块化键盘的设计方式使键盘的组件可以根据需要连接和分离。本文从美国**专利**数据库中阐述了 10 项模块化键盘上的发明。这些发明的目的是使键盘模块化,以便通过分离其某些组件使同一键盘变小,通过附加附加附加附件而使键盘变大。一些模块化键盘为外部附件(如鼠标、电话、扬声器、操纵杆和存储设备等)提供插槽。

2013 年 10 月 11 日提交;最初宣布 2013 年 10 月。

93. 10 笔记本电脑键盘上的发明基于美国专利的研究

作者:umakant mishra

摘要: 桌面键盘有几个部分,如字符键部分、导航键部分、数字键部分和功能键部分等,每个部分由多个键组成。但是,笔记本电脑没有那么大的空间来容纳键盘中的所有这些键。在设计笔记本电脑键盘时,有几个注意事项。本文说明了笔记本电脑和便携式计算机键盘上的 10 项发明。这些发明是从美国**专利**数据库中选择的。这些发明试图改善笔记本电脑键盘的各个方面,如缩小尺寸、折叠和隐藏、人体工学特征、提高质量和降低成本。少

2013 年 10 月 11 日提交;最初宣布 2013 年 10 月。

94. 10 可折叠键盘上的发明: 基于 triz 的分析

作者:umakant mishra

摘要: 虽然更大的键盘通常是舒适的工作，他们不能适应笔记本电脑和小型电脑盒。便携式和手持计算机需要小型键盘。因此，有必要减少键盘的大小，以适应笔记本电脑盒。有各种机制来减少键盘的大小;可折叠键盘就是其中之一。尽管不同的发明旨在实现减少键盘尺寸的相同目标，但它们的机制各不相同。例如，有些发明使用压缩机构，有些发明使用折叠机构，有些发明使用折叠机构等。本文从美国**专利**数据库中展示了可折叠键盘上的10项发明。每个案例都从 triz 的角度进行分析。少

2013 年 10 月 11 日提交;最初宣布 2013 年 10 月。

95. 对主要指南和键盘模板的 10 项发明

作者:umakant mishra

摘要: 键盘有许多功能键，每个功能键在与控制键、移位键和 alt 键一起使用时可以具有多种功能，用户很难记住功能键的功能。我们需要一种机制来指示为不同软件程序的每个功能键分配的操作。用于此目的的键盘指南或模板。本文阐述了从美国**专利**数据库中选择的键盘键指南和功能键模板上的 10 项发明。提出了键盘模板的各种机制，包括静态、动态、手动、机械、屏幕显示等。少

2013 年 10 月 11 日提交;最初宣布 2013 年 10 月。

96. 键盘按键开关机构的 10 项发明

作者: [umakant mishra](#)

摘要: 计算机系统键盘的按键开关一般由一个钥匙盖组成, 钥匙盖上有柱塞, 导电橡胶设置在膜电路上方, 并被柱塞压缩, 以触发膜电路, 使其产生电信号。计算机。一些按键开关使用弹簧。其他一些键盘使用橡胶圆顶或圆顶板, 它做弹簧的功能。当用户按下按键按钮时, 弹簧或圆顶就会塌陷。钥匙开关抑制钥匙杆, 按键启动膜电路上的按钮。当用户松开按钮时, 弹簧或橡胶圆顶将按钮按到静止位置。本文阐述了 10 项发明不同的关键开关机制。这些发明是从美国专利数据库中选择的。这些发明试图简化机构, 降低制造成本, 提高精度和可靠性, 提高钥匙的强度和鲁棒性。少

2013 年 10 月 11 日提交;最初宣布 2013 年 10 月。

97. 非单调的前缀点与学习

作者: [Stefano berardi](#), [ugo de ' liguoro](#)

摘要: 我们考虑了在任意知识空间上寻找交互式实现器的前缀点, 得到一个相对递归过程的问题。知识空间和交互式实现器是一个抽象的设置, 用来表示学习过程, 可以解释非建设性的证明。知识空间的原子信息被分层到一个层次上, 并根据知识状态被评估为真理值。然后, 实现器被用来定义通过添加和可能删除原子来扩

展给定状态的运算符: 在学习过程中, 知识状态不是单调变化。实现器的预定定点的存在相当于在某种知识状态下终止学习过程, 这种状态没有专利矛盾, 因此没有什么可补充的。本文将我们以前的结果推广到第 2 层知识空间和确定性算子的情况下, 对于欧米级知识空间和非确定性算子。少

2013 年 9 月 3 日提交;最初宣布 2013 年 9 月。

98. 键盘照明的发明–基于 triz 的分析

作者:umakant mishra

摘要: 传统的计算机不能在黑暗中使用, 因为用户看不到键盘正常运行。但在很多情况下, 用户可能喜欢在低光照条件下在电脑上工作, 比如, 在夜间, 在低光的飞行/火车上旅行, 晚上在花园里工作等。因此, 有必要为笔记本电脑提供一些光线, 以便在低光照情况下操作键盘。键盘的发光是便携式计算机的理想功能。可能有几种方法来照亮键盘, 最粗糙的可能是携带外部光和固定在一个点, 它可以照亮键盘。但这样的解决方案并不是很方便。发明者一直在寻找更好的解决方案。本文分析了 6 项**键盘照明专利**。一个好的照明系统不消耗更多的电池, 不增加笔记本电脑的大小和重量, 不增加笔记本电脑盒内的热量, 也不花费太多的制造成本。光线应均匀地照亮整个键盘, 而不应在其他不需要的区域上发光。少

2013 年 7 月 24 日提交;最初宣布 2013 年 7 月。

99. 软键盘上的发明-----基于 triz 的分析

作者:umakant mishra

摘要: 软键盘是具有字母数字字符和其他控件的物理键盘在屏幕上的表示形式。用户使用鼠标、手写笔或其他指针设备操作软键盘。软键没有任何机械部件。软键盘在许多公共场所用于信息目的、教育系统和金融交易系统。在某些情况下,硬键盘难以管理,则可方便使用软键盘。软键盘是物理键盘的替代品,并显示在屏幕上。它显示与实际键盘上的键相同类型的字母数字和控制键。软键盘上有许多发明,使软键盘更有效率和效力。本文从美国专利数据库中阐述了软键盘上的一些发明。少

2013 年 7 月 24 日提交;最初宣布 2013 年 7 月。

100.10 键盘附件上的发明-基于 triz 的分析

作者:umakant mishra

摘要: 尽管键盘将数据输入计算机的主要目的是,高级键盘还是会记住各种其他事项,例如,如何将同一键盘用于各种其他目的,或者如何使用其他各种键盘有效地使用同一键盘到键盘的附件。这一目标导致了键盘附件上的各种发明,其中一些在下面的本文中进行了说明。本文从美国专利数据库中说明了 10 项发明。每个

不同的附件都以某种方式提高了键盘的可用性。所说明的一些附件包括模板持有人、纸张持有人、指点设备附件、鼠标垫附件、手腕休息附件、存储设备附件和多媒体附件。少

2013 年 7 月 20 日提交;最初宣布 2013 年 7 月。

101. 10 提高键盘效率的发明: 基于 triz 的分析

作者:[umakant mishra](#)

摘要: 键盘是计算机最重要的输入设备。随着技术的发展, 基本键盘不希望保持在键盘的基本功能范围内, 而是要超越。有几种发明试图提高传统键盘的效率。本文阐述了美国专利数据库中的 10 项发明, 所有这些发明都提出了提高计算机键盘效率的非常有趣的方法。一些有趣的发明包括, 在键盘上添加工具栏按钮, 省电方法, 键盘冷却机制, 强大的键盘供公众使用, 键盘增强为儿童使用。这些发明中的每一个都从 triz 的角度进行了分析, 在许多情况下, 制定了理想的最终结果, 矛盾, 并突出了解决方案中使用的创新原则。少

2013 年 7 月 20 日提交;最初宣布 2013 年 7 月。

102. 键盘键布局上的 10 项发明: 基于 triz 的分析

作者:[umakant mishra](#)

摘要: 标准的 qwerty 键盘是在 19 世纪后期为打字机开发的。由于人们都知道, 同样的布局被保留在计算机键盘上。很多人觉得 qwerty 布局不是很有效的布局, 在不同的角色键布局上有很多发明。为了改进关键安排, 应解决两个主要问题。(i) 改进后的关键安排应能大大提高生产力。(二) 应尽量减少学习经改进的关键安排的培训时间。本文从美国**专利**数据库中分析了 10 项发明, 其中每项发明都提出了新的布局, 使我们具有一些具体的优势。发明试图实现以下一个或多个优势, 即减少打字过程中的手指动作, 实现数据输入速度, 减少打字错误, 使键盘易于学习, 使孩子们容易找到钥匙, 减少压力。手和手指, 适合特殊用途的计算机。少

2013 年 7 月 20 日提交;最初宣布 2013 年 7 月。

103. 一种非交换密钥交换协议的密码分析

作者: [贾科莫·米歇尔](#)

文摘: 在阿尔瓦雷斯等人和帕塔克、桑吉的论文中, 描述了一种非交换为基础的公钥交换。类似版本的它也获得了**专利**(7184551 美元)。本文提出了一种多项式时间攻击, 它打破了这两篇论文中提出的协议的变体。此外, 我们还表明, 打破**专利**密码系统 us7184551 可以很容易地简化为分解。我们还给出了一些例子来说明攻击的效率。少

2013 年 6 月 22 日提交;最初宣布 2013 年 6 月。

104. 股票市场预测中的人群智慧算法

作者: [marko velic](#), [toni grzinic](#), [ivan padavic](#)

摘要: 本文提出了一种在股票市场预测中协同过滤实现的数学模型。在大众文学中, 协同过滤, 也被称为人群智慧, 假设群体比个人有更多的知识, 而每个人可以通过其特定的信息输入来提高群体的表现。有商业上可用的工具, 用于合作股市预测和专利保护的基于 **web** 的软件解决方案。这些算法背后的数学并没有在文献中披露, 因此所提出的模型和算法实现是这项工作的主要贡献。

少

2013 年 6 月 21 日提交;最初宣布 2013 年 6 月。

105. 喷泉代码的入门和最新发展

作者: [jalaluddin qureshi](#), [chuan heng fohy](#), [蔡建飞](#)

摘要: 在本文中, 我们调查了最近提出并获得专利的各种擦除代码, 并在调查过程中提供了关于擦除和喷泉代码中许多基本概念和读数的介绍性教程。数据包擦除是数据存储和数据传输系统固有的一个基本特征。传统上采用了基于复制/再传输的技术来处理此类系统中的数据包擦除问题。虽然 reed-solomon (rs) 擦除码已经知道了相当长的时间, 以提高系统可靠性和减少数据冗余, rs 代码的高解码计算成本已经抵消了更广泛的应用 rs 代码。然

而，最近数据流量的指数增长和对更大数据存储容量的需求已经引起了人们对擦除代码的兴趣。最近的结果为解决擦除码固有的解码计算复杂度和冗余性权衡提供了很有希望的结果。少

2013 年 5 月 4 日提交;最初宣布 2013 年 5 月。

106. 新兴技术的战略情报: 科学计量覆盖图

作者 : [daniele roolo](#), [ismael raols](#), [michael hopkins](#), [loet leydesdorff](#)

摘要: 本文探讨了使用科学计量叠加映射作为 "战略情报" 工具, 以帮助新兴技术的治理。我们开发了一个综合综合不同的叠加映射技术和相关的观点, 在地理、社会 and 认知空间的技术出现。为此, 我们对医学领域新兴技术的三项案例研究进行了纵向分析 (有出版物和**专利**数据)。这些是 : ma 干扰 (mai)、人类状瘤病毒 (hpv) 宫颈癌检测技术和硫嘌呤甲基转移酶 (tpmt) 基因检测。鉴于叠加映射技术的灵活性 (即对不同数据源的适应性) 和粒度 (即跨多个数据聚合级别的适用性), 我们认为, 这些技术可以支持来自不同的背景和案例, 从而有可能作为一个平台, 为分析人员和决策者提供 "分布式" 战略情报。少

2015 年 9 月 7 日提交;v1 于 2013 年 3 月 19 日提交;最初宣布 2013 年 3 月。

107. 技术转移与贝约尔效应的终结: 专利作为产学研关系的分析镜头

作者: [loet leydesdorff](#), [martin meyer](#)

文摘: 自 1980 年《Bayh-Dole 法》以来, 美国专利和贸易局 (uspto) 的大学专利有三个时期: (1) 1995 年之前, 大学专利专利的第一个时期呈指数级增长 (申请日期) 或 1999 年 (签发日期); (2) 1999 年以来的相对下降期; (3) 近年来——自 2008 年以来——大学专利呈线性增长。我们认为, 这最后一个时期是由特定的非美国大学 (如东京大学和中国大学) 越来越多地在美国获得专利, 成为最具竞争力的高科技专利市场。少

2013 年 2 月 20 日提交; 最初宣布 2013 年 2 月。

108. stem 课程中知识产权课程的引入

作者: [madhur srivastava](#)

摘要: 本文阐述并强调了在 stem 学生的本科和研究生课程中纳入强制性知识产权 (ip) 课程的必要性。它规定了建议的知识产权内容, 涵盖本科水平, 并根据 stem 学生追求的学位毕业。更重要的是, 本文主张综合 phd/jd 学位, 以生产高技能的专利律师和代理人, 以克服专利质量差的问题, 以及专利相关的数量诉讼。建议建立一个框架, 使 stem 系和法学院能够成功地提供这一综

合学位, 并使学生能够以现实的方式成功地追求和完成这一学位课程;而不损害任何利益相关者的利益。少

2013 年 1 月 30 日提交;最初宣布 2013 年 1 月。

109. 中欧和东欧国家区域创新政策面临的挑战: 空间集中和外国对美国专利的控制

作者:[balázs lengyel](#), [tamás sebestyén](#), [loet leydesdorff](#)

摘要: 利用数据收集和映射技术作为谷歌地图的覆盖, 基于美国专利商标局 (uspto) 在线提供的**专利信息**——我们指出了两个主要的、相互关联的挑战: 中欧和东欧 (cee) 的决策者在对抗滞后的创新绩效时面临着现实。首先, 我们通过在城市一级的分布分析来解决空间集中问题。研究结果表明,**专利更多集中在后社会主义地区**, 而不是西方国家和地区。然而, 在中欧和东欧, 没有一个优秀的枢纽, 当我们比较美国专利商标局专利归一化的各自人口规模。其次, 我们认为, 外国对美国专利商标局**专利**的控制优势主要体现在个人层面的国际合作中, 很少出现在跨国公司子公司。我们认为, 在**专利方面**赶超中欧和东欧的可能性不大, 除非创新政策措施侧重于不断增长的枢纽, 并针对国内发明者和公司的国际关系。少

2013 年 8 月 2 日提交;^{v1} 于 2013 年 1 月 9 日提交;最初宣布 2013 年 1 月。

110. 突破创新的动态网络方法

作者:[russell j. funk](#), [jason owen-smith](#)

摘要: 本文概述了一个研究创新的框架, 该框架将发现视为不断发展的网络的补充。随着发明的进入, 它们通过影响连续地发现如何使用这些思想, 扩大或限制了它们所积累的思想的触角。这种方法的基础是对创新在多大程度上放大或破坏现状的新衡量。这些指标将发明对先前发现的后续用途的影响进行了索引。在这样做的过程中, 它们描述了创新的一个理论上重要但难以捉摸的特点。我们通过展示我们的方法来验证我们的方法: (1) 在分析美国**专利**时, 对类似影响的创新进行歧视;(2) 在选定的案例研究中发现放大和破坏技术流的发现;(3) 意味着破坏性**专利**在差异估计中减少其前身的使用 60%;(4) 在对 110 所美国大学的**专利**分析中得出了新的发现。少

2012 年 12 月 14 日提交;最初宣布 2012 年 12 月。

111. 基于国际专利分类 (ipc) 的美国专利 (uspto) 数据交互式叠加地图

作者:[loet leydesdorff](#), [duncan kushnir](#) , [ismael raols](#)

摘要: 我们报告了与美国**专利**商标局 (uspto) 接口的开发情况, 该接口允许将**专利**组合映射为重叠, 并根据所包含的所有**专利**之

间的引文关系构建的底图在这个数据库在期间在 1976–2011 期间。接口和数据都在公共领域;免费软件程序 vo 查看这家软件和/或 pajek 可用于可视化。这些底图和重叠可以在世界知识产权组织 (wipo) 国际**专利**分类 (ipc) 的 3 位数和 4 位数级别生成。底图可以为分析师提供一个稳定的心理框架, 以跟踪不同年份搜索的发展情况, 这可以被动画化。美国专利商标局高级搜索引擎的充分灵活性可用于生成一组**专利**和**专利**申请, 从而可以对其进行可视化和比较。该文书可解决技术距离、投资组合多样性等问题, 并随着时间的推移推动各组织技术和技术能力的发展。少

2012年11月18日提交;v1 于 2012 年 10 月 24 日提交;**最初宣布** 2012 年 10 月。

112. Économie des biens immatériels----无形商品经济学

作者:laurent fournier

摘要: 我们引入了一种适用于无形商品 (\backslash ssc) 的新的经济体系。我们认为, 这种系统现在可以在现实世界中使用分布式网络计算和密码学中的先进技术来实现。显示了所谓的 " \backslash net {}" 的规范。为了限制金融交易的数量, 系统被迫定义自己的货币, 有许多好处。新的 "杯" 货币, 扩展到世界各地, 是专门用于 \backslash sc ig}, 只提供给人與人之间的交易, 防止投机和适应税收回收没有额外的计算。这些关键功能使 \backslash net{} 成为一个新的民主工具, 解决 \backslash

sxig} 交易中的特定问题, 并恢复整个域活动。我们强调的事实是, 所有与本提案相关的语言的拟议文档、算法和程序都应该是开源的, 没有任何可能在系统或子系统上发布任何种类的**专利**。这种新的交易模式应该被认为是纯粹的智力结构, 就像数学的某些部分一样, 然后属于任何人或每个人, 比如 $1+1=2$ 。下一步将测试、验证各种实现详细信息的安全性, 并要求进行法律规则调整。论文初稿以法文编写, 并张贴在 arxiv.org 和 hal.archive-ouverte.fr。我们希望在圣诞节前提供英文翻译。少

2012年11月26日提交;v1于2012年10月15日提交;**最初宣布** 2012年10月。

113. 基于美国专利引文网络分析的新兴技术预测

作者: [péter erdi](#), [kingamakovi](#), [zoltán somogyvári](#), [katherine strandburg](#), [jan tobochnik](#), [péter volf](#), [lászló zalányi](#)

摘要: 引用连接的**专利**网络是一个不断发展的图形, 它提供了创新过程的表示。引用另一**项专利**意味着被引用的**专利**反映了引用的**专利**所积累的一项先前存在的知识。这里提出的一种方法 (一) 确定了实际的**专利**集群: 即技术分支, (二) 对集群结构的时间变化作出预测。定义了一个名为 {引文向量} 的预测器, 用于描述技术发展的特征, 以显示其他**专利**所引用的**专利**如何属于各种工业领域。所采用的聚类分析技术能够检测到新出现的重组, 并预测新出现的技术集群。以美国专利商标局第 11 子类别 "农业、食品、

纺织品" 为例, 说明了我们新方法的预测能力。根据 1991 年之前的引文数据确定了一组**专利**, 这表明 1997 年初形成的 442 类的专利有很大的重叠。这些新的预测分析工具可以支持科学和技术领域的决策过程, 并有助于制定行动建议。少

2013 年 4 月 4 日提交;v1 于 2012 年 6 月 18 日提交;**最初宣布** 2012 年 6 月。

114. 通过消除边界点删除量词

作者:[eugene goldberg](#), [panagiotis manolios](#)

摘要: 我们从布尔 cnf 公式中考虑了存在量词的消除问题。我们的做法基于以下观察。通过添加消除边界点的 f 析取子句, 可以消除对量化 cnf 公式 f 的一组变量的依赖。此方法类似于 [9] 中描述的量词消除方法。本文中描述的方法有两个区别: $\{\backslash \text{spr}\}$ 分支仅在量化变量上执行, $\{\backslash \text{sprt}\}$ 通过调用 sat 求解器来执行对边界点的显式搜索, 尽管我们之前发表了论文 [9]这一次, 从逻辑上讲, 本报告的方法是首先发展起来的。在 [10]、[11] 中对这种方法作了初步介绍。由于准备了**专利申请**, 我们推迟了这种方法的出版 [8]。少

2012 年 6 月 5 日提交;2012 年 4 月 8 日提交了 v1;**最初宣布** 2012 年 4 月。

115. 创新作为一个非线性过程, 科学计量视角, 以及 "创新机遇探索者" 的规范

作者:[loet leydesdorff](#), [daniele rotoolo](#), [wouter de noy](#)

摘要: 创新的过程遵循科学、技术和经济领域的非线性模式。新的文献计量学绘图技术可用于调查和代表关于创新过程（例如, "需求" 和 "供应"）以及这些观点之间的相互作用的独特但互补的观点。这些观点可以表示为与不同程度有关的数据的 "大陆"。例如, medline 数据库中的医疗主题标题 (mesh) 的不同分支提供了此类视角的来源（例如, "疾病" 与 "药物和化学品"）。多角度的方法使我们能够在选择机制形成可本地化的轨道方面重建创新动态的各个方面, 或导致更加全球化的制度。通过**扩展专利**和学术出版物的数据, 我们展示了这种多视角方法在 ma 干扰 (rnai) 中的应用。指定了开发 "创新机会资源管理器" 的可能性。少

2013年1月29日提交;v1 于 2012年2月28日提交;**最初宣布** 2012年2月。

116. 近似召回置信区间

作者:[威廉·韦伯](#)

文摘: 回忆, 相关文献检索的比例, 是衡量信息检索有效性的重要指标, 特别是在法律、**专利**和医疗领域。如果文件集太大, 无法

进行详尽的相关性评估, 则可以通过随机抽样的文件来估计召回情况;但也需要说明这一估计的可靠性。在本文中, 我们将研究几种估计双尾召回置信区间的方法。我们发现, 目前使用的正常近似在许多情况下提供的覆盖率很低, 即使经过调整以纠正其不适当的对称性。基于二项式比的解析法和贝叶斯法一般更准确, 但对小种群不准确。我们推荐的方法推导贝塔二项式后部在恢复和未检索屈服, 具有固定的超参数, 并对召回的后部分布进行蒙特卡罗估计。我们证明, 这种方法在几个场景中提供了在标称水平或接近名义水平的平均覆盖率, 同时保持了平衡和稳定。我们提供有关采样设计的建议, 包括将评估分配给检索到的段和未检索到的段, 并将建议的 β -二项式与官方报告的最近 trec 法律跟踪迭代的正常间隔进行比较。少

2012年10月23日提交;v1于2012年2月13日提交;**最初宣布** 2012年2月。

117. 无生命的路由方案

作者:万斯·费伯

摘要: 我们给出了任何允许的网络的无 livelock 路由算法。与此问题的其他一些解决方案不同: 1) 进入网络的数据包在到达目的地的时间上有绝对的上限;2) 在轻负载下, 数据包在几乎最佳的时间内传送到目的地;3) 距离拥塞区域较不远的需要路径的数据

包的路由时间将远远短于想要访问拥塞区域的数据包;4) 如果网络拥塞, 后来清除, 网络的运行就像它最初在一个轻负载时。这张纸条的主要思想在我 1994 年的专利 5,369,745 出现了不同的形式。本说明增加了这些结果, 使它们更具有数学意义。少

2012 年 1 月 21 日提交;最初宣布 2012 年 1 月。

118. 用于单极通信系统的 flip-ofdm

作者:nirmal fernando, yi hong, Emanuele viterbo

文摘 单极通信系统只能使用真实和正信号传输信息。这包括各种物理通道, 从光纤(光纤或自由空间)到使用非相干接收的幅度调制的射频无线, 再到基带单线通信。单极 ofdm 技术能够有效地补偿单极通信系统中的频率选择性失真。单极 ofdm 的主要例子之一是最初为光通信提出的非对称剪切光 ofdm (amco-ofdm)。flip-ofdm 是在专利中提出的一种替代方法, 但其性能和充分的潜力从未在文献中得到研究。本文首先比较了 flip-ofdm 和 aco-ofdm, 并证明了这两种技术具有相同的性能, 但具有不同的复杂性 (flip-ofdm 可节省 50%)。然后, 我们提出了一种新的检测方案, 使 flip-ofdm 接收机的噪声减少了近 3db。仿真结果支持了噪声滤波方案的分析性能。少

2011 年 12 月 13 日提交;v1 于 2011 年 11 月 30 日提交;最初宣布 2011 年 12 月。

119. 使用覆盖到谷歌地图的映射 (uspto) 专利数据

作者: [loet leydesdorff](#), [lutz bornmann](#)

摘要: 一种技术是利用在线 (美国**专利**和商标局) 提供的专利信息开发的,用于生成 google 地图。这些覆盖既表明了城市一级**专利**的数量, 也表明了**专利**的质量。这些信息与技术分析、创新研究和进化经济学以及经济地理学的研究问题有关。由此产生的地图也可能与技术创新政策和研发管理相关, 因为美国市场可以被视为**专利**和**专利**竞争的主要市场。除了地图, 这些例程还提供了有关**专利**的定量数据, 用于统计分析。地图上的城市根据意义测试的结果进行着色。荷兰的覆盖是一个 "国家创新体系", 并在两个新兴技术案例中进一步阐述: "rna 干扰" 和 "纳米技术"。少

2012 年 2 月 9 日提交;v1 于 2011 年 10 月 26 日提交;**最初宣布** 2011 年 10 月。

120. 一种利用投影空间格在超大规模集成电路应用中折叠式、流水线化体系结构的设计方法

作者: [hrishikesh sharma](#), [sachin patkar](#)

文摘: 在设计时需要节省硬件资源时, 只要需要节省硬件资源, 就会使用半并行或折叠的 vlsi 体系结构。最近基于投影几何 (pg) 的平衡二分图的应用也属于这一类。本文提出了一种基于 pg 二

部图的数据流图 (dfg) 的高级自上而下的设计方法来设计最佳的并行体系结构。例如, 在错误控制编码和矩阵计算中发现了此类应用。与许多其他折叠方案不同, 在此方法中, 物理元素之间的连接拓扑结构不会改变。另一个优点是易于实现。为了减少折叠造成的吞吐量损失, 我们还在设计方法中加入了多层流水线策略。通过在 C++ 中实现的综合工具, 验证了该设计方法, 并进行了验证。该工具是公开提供的。此外, 在合成工具设计之前, 对一个完整的解码器进行了手工原型化, 以验证本文中所有算法的发展, 实现了各种细化步骤。基于此方法的 ldpc 解码器的另一个特定高性能设计是在过去制定的, 并已获得**专利**。少

2012 年 8 月 4 日提交;v1 于 2011 年 8 月 19 日提交;**最初宣布** 2011 年 8 月。

121. 自动音量控制, 用于保持可读性

作者:富兰克林·费尔伯

摘要: 开发了一种新方法, 可在所有配备至少一个麦克风的音频设备上自动调整音量, 包括移动电话、个人媒体播放器、耳机和汽车收音机, 这些设备可能会在嘈杂的环境中使用, 如人群、汽车和户外。该方法使用一组**专利**的算法, 在此类设备中的芯片上实现, 以保持噪声环境中语音的恒定清晰度, 而不是恒定的信噪比。该算法对噪声背景进行实时分析, 只补偿频域和时域中干扰语音

清晰度的波动噪声。这种控制体积的方法的优点包括: 在不牺牲清晰度的情况下控制体积;仅针对持续的语音干扰噪声进行调整;平滑体积波动;并消除由噪声尖峰引起的静电样爆发。讨论了在手机中实现这些算法的实用人为因素方法。少

2013 年 7 月 17 日提交;v1 于 2011 年 3 月 29 日提交;**最初宣布** 2011 年 4 月。

122. 出版或专利: 国家融资战略中经验权衡的文献计量证据

作者:[robert d.shelton](#) , [loet leydesdorff](#)

摘要: 多元线性回归模型表明了国家研发投资分配的权衡。政府的资金和高等教育部门的支出似乎鼓励出版物, 而工业资金和工商部门支出等其他组成部分则鼓励**申请专利**。我们的研究结果有助于解释为什么美国在出版物上落后于欧盟, 因为它专注于工业融资——约占其研发总投资的 70%。相反, 这也有助于解释为什么欧盟在**专利方面**落后于美国。政府资金被认为是对高质量**专利的负面激励**。一旦知道了适当的输入指标, 这里的模型也可用于预测一个国家的产出指标。这通常是在数据集中完成的, 为期一年, 但如果能够对投入指标作出合理预测, 这一过程可以延长到未来几年预测产出。我们对美国、欧盟 27 国和中国在出版方面的进一步关系进行了新的预测。然而, 个别国家的模式可能比一系列国家的参数平均的回归模型更成功。少

2011 年 2 月 15 日提交;最初宣布 2011 年 2 月。

123. 交易与专利申请网络分析

作者:[hiroyasu inoue](#)

摘要: 如今,许多公司选择专业化而不是泛化,以此来保持竞争力。因此,它们不能仅仅依靠自己,而必须通过结合自己的优势进行合作。为了获得这种合作的实际情况,研究了基于两种不同类型数据的多层网络。第一类是日本公司的交易数据。根据这些数据建立的网络包括 961 363 家公司和 7 80760 个链接。第二类数据来自日本的联合**专利**申请。联合**专利**申请网络包括 54, 197 个节点和 154, 205 个链接。这两个网络合并为一个网络。第一次分析是基于投入产出表,并比较了三个不同的表。表之间的相关系数显示,与联合**专利**申请的联系比资金总额更紧密。货币和交易的总额几乎没有关系,这些关系可能与不同机制中的联合**专利**申请有关。第二项分析是在 p^* 模型的基础上进行的。对选择、多样性、互惠性、多重性和可传递性配置进行了评价。在所有被分析的行业中,倍数和互惠性的配置都很重要。多样性的结果意味着交易和联合**专利**申请链接密切相关。在一些被分析的行业中,多重互惠和可传递配置具有重要意义。很难在这些行业中找到任何共同的特点。第三部分分析采用了贝叶斯网络。学到的结构表明,如果知道两家公司之间的交易联系,企业产业的类别并不影响**专利**联系的存在。少

2010 年 9 月 21 日提交;最初宣布 2010 年 9 月。

124. 个人关注和知识贡献

作者:[lada a. adamic](#), [xiaowei](#), [姜阳](#), [sean gerrish](#), [kevin k. nam](#),
[gavin s.clarkson](#)

摘要: 在贡献新知识之前, 个人必须通过学校教育、培训、实践和经验获得必要的背景知识或技能。在时间有限的情况下, 个人往往选择要么专注于少数领域, 在这些领域建立深厚的专业知识, 要么不那么深入地钻研, 并将注意力和努力分配到几个领域。在本文中, 我们测量了在一系列传统和最近的知识共享媒体, 包括学术文章、**专利**、维基百科和在线问题中, 焦点的狭窄与贡献的质量之间的关系。回答论坛。在所有系统中, 我们观察到焦点和质量之间的微小但显著的正相关。少

2010 年 2 月 2 日提交;最初宣布 2010 年 2 月。

125. 沟通与知识: 一个经济体的知识库是如何构建的?

作者:[loet leydesdorff](#)

摘要: 知识经济中的竞争优势不能再归因于网络中的单个节点。以知识为基础的发展破坏了市场平衡和体制安排, 使政治经济日益重塑。该网络协调 (一) 财富生产、(二) 有组织的新颖生产和 (三) 私人拨款与公共控制的次级动态。交互术语产生了复杂的动态,

不能指望它包含中央协调。然而, 创新系统的知识基础设施可以用大学与工业-政府的关系等方式来衡量。这三个维度的相互信息表明了知识库的全球化。对这一指标的**专利统计和互联网数据**进行了比较。少

2009 年 12 月 7 日提交;最初宣布 2009 年 12 月。

126. 大学与行业的知识关系: 专利分析与技术的科学基础

作者:[loet leydesdorff](#)

摘要: 通过互联网, 信息科学家可以免费访问隐藏或深入网络中的大型数据库。这些数据库的结构往往远远超过互联网域名本身。本研究使用美国**专利和贸易局的专利数据库**, 根据这些**专利**中的文献参考来考察**专利**的科学基础。在将荷兰国民经济作为参考制度时, 将全球一级的**基于大学的专利**与结果进行比较。指定了访问联机数据库和显示结果的方法。结论是, "生物技术" 历史上产生了一个关于大学与产业关系的理论模型, 无法轻易推广到其他部门和学科。少

2009 年 12 月 7 日提交;最初宣布 2009 年 12 月。

127. 虚拟线程: 高级通用处理器体系结构

作者:[and 山甲毛](#)

文摘: 本文介绍了新的计算机架构, 其主要特点已在俄罗斯联邦专利 2312388 和美国专利申请 [11/991331](#) 中得到承认。这种架构旨在有效地支持通用并行计算 (gppc), 其实质是极其频繁地频繁地在活动状态和查看状态之间切换线程, 在本文中查看算法延迟。为了强调体系结构延迟和算法延迟对 gppc 的同样影响, 引入了广义延迟的新概念, 并定义了它的定量度量——广义延迟容差 (glt)。结果表明, 一种适合 gppc 实现体系结构的体系结构应该具有较高的 glt 级别, 并描述了这样的体系结构, 即所谓的虚拟线程机。这种体系结构是在活动虚拟化的方向上建立处理器虚拟化的, 它与众所周知的内存虚拟化方向是正交的。体系结构的关键要素是: 1) 体系结构寄存器文件的分布式细粒表示, 这些元素是通过微体系结构内存级别进行硬件交换的, 2) 优先的细粒度直接硬件多编程, 3) 访问控制的虚拟寻址和 4) 硬件驱动的信号量。这些功能的组合允许引入新的操作系统 (os) 编程样式, 这种编程没有中断, 使用操作系统服务的应用编程非常罕见。少

2009 年 10 月 21 日提交;最初宣布 2009 年 10 月。

128. 用于先验艺术搜索的多个检索模型和回归模型

作者: [patrice lopez](#), [laurent romary](#)

文摘: 本文介绍了为 2009 年线索网络实现的 ptatacars (**专利与文章跟踪、检索和分析**) 系统。我们的方法有三个主要特点: 1. 使

用多个检索模型 (kl、okapi) 和术语索引定义 (术语、短语、概念) 来介绍目前轨道中考虑的三种语言 (英语、法语、德语) 产生十种不同的信息排名结果集。2. 使用**从专利集合中创建的附加验证集**, 根据多元回归模型合并不同的结果。3. 利用**专利元数据和引文结构**来创建有限的初始工作专利集和生成最终重新排名回归模型。当我们仅在创建初始工作集时和在最后的排名后步骤中利用**专利文档的特定元数据和引文关系**时, 我们的体系结构仍然是通用的, 易于扩展。少

2009 年 8 月 30 日提交;最初宣布 2009 年 8 月。

129.orda: 利用 bpmn 对业务流程进行建模的方法

作者:[matteo buferli](#), [matteo magani](#), [danilo monthesi](#)

摘要: 本文提出了一种 bpmn 建模方法, 即业务流程表示的标准表示法。我们的方法简化了协作 bpmn 图的开发, 从而能够自动创建代表复杂编排的骨架流程图。为了评估和调整方法, 我们开发了一个支持它的工具, 我们将其应用于国际**专利**流程的建模作为一个工作示例。少

2009 年 7 月 7 日提交;最初宣布 2009 年 7 月。

130.灵活性的生存: 解释最近在快速发展的世界中自然启发优化的主导地位

作者:james m whitacre

摘要: 尽管研究人员经常评论自然启发的元启发式 (nim) 日益流行, 但很少有数据直接支持 nim 与其他优化技术相比越来越突出的说法。这项研究提供的证据表明, nim 的使用不仅在增长, 而且在与学术研究活动 (出版频率) 和商业相关的几个重要指标中似乎已经超过了数学优化技术 (mot) 活动 (**专利频率**)。在这些发现的推动下, 本文讨论了这种日益流行的一些可能的起源。我回顾了 nim 受欢迎程度的不同解释, 并讨论了为什么其中一些论点仍然不令人满意。我认为, 一个令人信服的和全面的解释应该直接解释大多数 nim 成功的实际方式, 例如通过杂交和定制不同的问题环境。本文从问题生命周期的角度出发, 重新审视了自然启发的元启发式从灵活性中获得的大部分效用的假设。我讨论了应用优化算法的业务环境中的全球趋势, 我推测高度灵活的算法框架在我们多样化和快速变化的世界中可能会越来越流行。少

2011 年 1 月 21 日提交;v1 于 2009 年 7 月 2 日提交;**最初宣布** 2009 年 7 月。

131. 论随机性的原因和原因

作者:soubhik chakraborty

摘要: 本文有几个目标。首先, 它将随机性与目无法纪区分开来, 并说明为什么即使是真正的随机性也不意味着目无法纪。其次,

它将问题分开——为什么要称之为随机现象？（并在第一部分中回答）从**专利**问题—什么是随机序列？—答案在于 kolmogorov 的复杂性（这在第二部分中解释）。在回答第一个问题时，该说明指出，为什么应该有四个激励因素来调用随机现象：无菌、认知、伪和望远镜，前两个描述真正的随机性，最后两个描述错误。第三，无菌和认知随机性与主动和认知概率不同。第四，它鼓励学生成为应用统计学家，并建议不要成为扶手椅理论家，但有趣的是，这是通过直接应用伸缩随机性来实现的。总体而言，它告诉（老师）不要跳到概率而不正确地解释随机性第一，并类似地建议学生阅读（和理解）随机性，然后再采取概率。少

2009 年 2 月 7 日提交;最初宣布 2009 年 2 月。

132. 将隐写数据插入 voip 呼叫的程序分析

作者:[wojciech mazurczyk](#), [jozef lukcz](#)

文摘: 本文讨论了一种主要针对 voip 的隐写方法的性能分析，该方法最近以 lack 的名义申请**专利**。该方法的性能取决于将隐蔽数据插入音频数据包流的过程。在简要介绍了 back 方法之后，重点分析了插入过程对 voip 调用持续时间概率分布的依赖关系。少

2008 年 8 月 25 日提交;v1 于 2008 年 6 月 5 日提交;最初宣布 2008 年 6 月。

133.ntcir-4 研讨会专利对专利检索和专利地图生成的测试集

作者:[atsushi fujii](#), [makoto iwayama](#), [noriko kando](#)

文摘: 本文介绍了第四届 ntcir 研讨会中的**专利检索**任务, 以及在这一任务中产生的测试集合。我们执行无效搜索任务, 在该任务中, 每个参与者组搜索专利集合的**专利**, 这些专利可能会使现有权利要求无效。我们还执行**自动专利地图生成**任务, 在该任务中, 与特定主题相关的**专利**被组织在多维矩阵中。少

2004 年 4 月 10 日提交;最初宣布 2004 年 4 月。

134. 引文网络分析的有效算法

作者:[弗拉基米尔·巴塔格利](#)

文摘: 本文提出了在引文网络中确定洪蒙和多瑞安弧线权权重的算法, 并给出了这些权重的一些理论性质。讨论了引文网络中的非无环性问题。提出了一种基于电弧权重的重要小子网的识别方法, 并在 som (自组织地图) 文献和美国**专利**的引文网络上进行了说明。少

2003 年 9 月 13 日提交;最初宣布 2003 年 9 月。

135. 一个多语言专利检索系统

作者:shigeto Shigeto, masatoshi fukui, atsushi fujii, Ishikawa
tetsuya

摘要: 鉴于在多个国家/地区申请的**专利**数量不断增加, 用户对跨语言检索**专利**感兴趣。我们提出了一个多语言专利检索系统, 它将用户查询转换为目标语言, 搜索多语言数据库中与查询相关的**专利**, 并通过机器提高浏览效率翻译和集群。我们的系统还从由可比专利组成的**专利**家族中提取新的翻译, 以增强翻译词典。少

2002 年 6 月 24 日提交;最初宣布 2002 年 6 月。

136. 一种混合查询翻译方法在日本语/英语跨语言专利检索中的应用

作者:fuui masatoshi, shigeto Shigeto, youichi nakatani, masao
tanaka, atsushi fujii, shishikawa Tetsuya

文摘: 本文将现有的查询翻译方法应用于跨语言**专利**检索。在我们的方法中, 使用多个字典来派生输入查询的所有可能的翻译, 并使用搭配统计信息来解决翻译歧义问题。我们使用日本/英语并行**专利**摘要进行比较实验, 我们的方法比简单的基于字典的查询翻译方法, 并实现了 76% 的单语检索在平均精度方面。少

2002 年 6 月 24 日提交;最初宣布 2002 年 6 月。

137. 专利法技术是否有特殊性?

作者:dan l. burk, mark a. lemley

摘要: 尽管**专利法**旨在涵盖所有技术, 但我们注意到最近在生物技术和软件**专利方面**的标准存在差异: 联邦巡回法院采用了非常宽松的标准。在生物技术的明显性, 但一个高度限制性的披露要求。软件**专利**的情况正好相反, 在我们看来, 这与这两个行业的健全政策完全相反。这些**专利**标准的基础是 "在艺术方面有普通技能的人" 或 *phoita* 的法律虚构。我们讨论 *phosita* 标准的适当性, 得出的结论是, 它适当地为**专利**制度提供了灵活性。然后, 我们讨论了在不同行业应用这一标准的困难, 并就如何修改该标准以避免生物技术和软件**专利**中出现的问题提出了建议。少

2001年9月26日提交;2001年9月25日提交的 v1;最初宣布 2001年9月。

138. 全球通信时代的标准与知识产权——第三代移动系统的国际标准化述评

作者:比约恩·海尔姆

摘要: 当欧洲电信标准协会 (etsi) 选择了一种基于宽带码分多址 (wcdma) 的无线电接入技术时, 由欧洲电信设备制造商爱立信和诺基亚赞助。第三代无线通信系统是 etsi 和 qualcomm inc. 之间的激烈纠纷, 高通威胁说, 除非欧洲人同意制造无线电接入技术, 否则将保留其在 cdma 技术上的知识产权与高通青睐的 cdma 版本 cdmaone 向后兼容。爱立信和高通之间还在关键

cdma 技术上就知识产权问题爆发了纠纷，两家公司都向美国法院提交了**专利侵权**诉讼。这一争端停止了标准活动，给全世界的运营商以及国际电信联盟（国际电联）带来了麻烦。少

2001年10月30日提交;v1于2001年9月24日提交;**最初宣布**2001年9月。