# 协作式商业智能虚拟助手参考模型

奥尔加 切列德尼琴科 [1]，法哈德 穆罕默德 [1]，热罗姆 达尔蒙 [1] 和塞西尔 法弗尔 [1]

[1] *Univ Lyon, Univ_Lyon 2, UR ERIC* – 法国 *Bron Cedex 69676*，*Mendès France* 大道 *5* 号

**摘要**

协作式业务分析（CBA）是一种方法论，旨在汇集不同的利益相关者，包括业务用户、分析师和技术专家，共同分析数据并深入了解业务运营。CBA 的主要目标是鼓励参与业务分析的不同群体之间进行知识共享与协作，因为这有助于更全面地理解数据，并做出更好的决策。CBA 通常涉及一系列活动，包括数据收集与分析、头脑风暴、问题解决、决策和知识共享。这些活动可以通过多种渠道开展，例如面对面会议、虚拟协作工具或在线论坛。本文探讨了作为商业智能（BI）平台重要组成部分的虚拟协作工具。协作式商业智能（CBI）工具正变得更加用户友好、更易于访问且更加灵活，使用户能够定制自身的使用体验，并根据具体需求进行调整。虚拟助手的目标是让更广泛的用户群体能够更轻松地进行数据探索，并减少数据分析所需的时间和精力。本文介绍了统一商业智能语义模型，以及与数据仓库和协作单元相结合以应用数据挖掘技术的方案。此外，我们提出了一种用于 CBI 的虚拟助手，以及一个用于 CBI 的虚拟工具参考模型，该模型由三个组件构成：对话智能体、数据探索智能体和推荐智能体。我们认为，将这三项功能任务进行分配，有助于构建 CBI 问题的结构，并同时应用相关且高效的模型，以实现类人对话、文本到命令转换和推荐。基于这三个方面的综合方法为协作虚拟工具奠定了基础。CBI 鼓励通过人员、流程和技术，使每个人都能够共享并利用集体专长、知识和数据，从而获得有价值的见解并做出更好的决策。这有助于更迅速、更有效地应对市场或内部运营的变化，并推动工作进展。

**关键词**

人工智能、协作式商业智能、虚拟辅助、对话代理

## 1. 引言

数据分析帮助我们根据从数据中发现的洞察和模式做出更好的决策。通过分析数据，我们可以识别趋势、关系和模式，而这些趋势、关系和模式在其他情况下可能难以甚至无法察觉。利用数据为决策提供依据，我们能够改善结果并避免代价高昂的错误。此外，在当今大数据时代，数据分析的重要性日益凸显，因为我们可以获取海量信息，但要理解和梳理所有这些信息可能会令人不知所措。数据分析工具和技术有助于处理和理解这些海量信息，

[1]COLINS-2023：第七届计算语言学与智能系统国际会议，2023年4月20日至21日，乌克兰哈尔科夫电子邮箱：olga.cherednichenko@univ-lyon2.fr（O. Cherednichenko）；fahad.muhammad@eric.univ-lyon2.fr（F. Muhammad）；jerome.darmont@univ-lyon2.fr（J. Darmont）；cecile.favre@univ-lyon2.fr（C. Favre）ORCID：0000-0002-9391-5220（O. Cherednichenko）；0000-0002-7258-9884（F. Muhammad）；0000-0003-1491-384X（J. Darmont）；-8658-7564（C. Favre）© 2023 本文作者版权所有。允许依据知识共享署名4.0国际许可协议（CC BY 4.0）使用。 CEUR研讨会论文集（ org）

使其更易于管理和采取行动。协作式业务分析（CBA）是一个协作分析数据并做出业务决策的过程。根据分析，数据研究的需求不仅限于追求利润的企业或应对国家问题的政府，也包括个人和整个社会。这些群体需要通过数据分析做出对社会具有重要意义的决策或个人决策。然而，要促进潜在用户、决策者和数据分析技术专家之间的协作可能颇具挑战。BI4people 项目 [1] 旨在利用在线分析处理（OLAP），通过软件即服务模式提供交互式分析和数据可视化，以解决这些问题。这使更广泛的受众能够使用数据仓库，包括多源异构数据集成。

协作式商业智能（CBI）涉及利用社交网络、测验、头脑风暴会议，甚至团队中两三位同事之间的简单聊天，共同解决问题 [2]。CBI 的核心理念是将人们聚集在虚拟或实体空间中，鼓励他们分享自己的观点和评论，以实现共同利益。此外，重复利用其他协作者的评论或成果也被视为 CBI 的一部分，这最终有助于形成一种更加全面的商业智能（BI）方法。

CBI 和集体智能（CI）是相关概念，两者都涉及将人们聚集在一起开展协作并分享知识。不过，两者之间存在一些重要差异。CI 是一个概念，指群体或网络共同解决问题、做出决策并形成洞见的能力，而这些成果往往超出单独工作的个人所能达到的水平。其理念是，通过利用群体的集体知识和专业技能，我们能够获得更好的解决方案，做出更好的决策，而不是各自为战。CBI 特指利用协作工具和方法来支持商业智能流程，其中可能包括社交网络、头脑风暴会议、聊天应用程序以及其他能够促进业务用户、分析师和技术专家之间知识共享与协作的技术。CBI 的目标是充分发挥群体的集体知识和洞见，以改进决策、推动业务成果并获得竞争优势。

由于数据探索是商业智能的核心，因此设计强大且高效的工具来支持商业智能流程（如数据分析或可视化）非常重要。另一方面，对于缺乏经验的用户而言，管理数据分析可能极其困难。分析表明，人工智能（AI）的一些解决方案可以帮助解决这一问题。机器学习（ML）和自然语言处理（NLP）可以帮助组织实现数据分析流程的自动化与优化，并基于从数据中获得的洞察，做出更准确、更明智的决策。

聊天机器人是一种计算机程序，可以通过与用户进行文本或语音交互来模拟人类对话 [3]。聊天机器人通常用于执行特定任务或向用户提供信息。它们通常是基于规则的，这意味着它们会遵循预先定义的脚本或决策树来响应用户输入 [4,5]。一些聊天机器人和对话代理（CAs）允许用户通过嵌套对话组合命令，以完成开放式数据分析任务。

1. IBM Watson Assistant [6] 允许用户创建复杂的对话流程，并组合各种命令以完成数据科学任务。

2. Dialogflow [7] 提供了一个自然语言理解系统，使用户能够创建对话流程和嵌套对话，以处理复杂的数据科学任务。

3. Rasa [8] 是一个开源的对话式人工智能框架，使开发人员能够构建具备先进自然语言处理能力和灵活对话管理功能的人工智能助手。

4. Botpress 平台 [9] 提供了一个工作流构建器，使用户能够设计具有嵌套对话、自定义脚本以及与第三方服务集成功能的聊天机器人。

5. Wit.ai 平台 [10] 提供自然语言处理系统，使开发人员能够创建具有复杂对话流程的智能聊天机器人、语音助手和其他对话式人工智能应用。

这里只列举了几个示例：这些聊天机器人和会话代理允许用户通过嵌套对话组合命令，以完成开放式数据科学任务。

顺应前沿发展趋势，本文重点探讨对话式聊天机器人如何改进 CBI 工具，并为企业、政府和公民带来新的机遇。我们提出了一种面向 CBI 的协作式虚拟助手参考模型，该模型由三个组件构成：对话智能体、数据探索智能体和推荐智能体。我们认为，将这三项功能任务进行分配，有助于梳理 CBI 问题，并应用相关且高效的模型。

同时实现类人对话、文本到命令转换和推荐功能。基于这三个方面的复杂方法为协作虚拟工具奠定了基础。

本文其余部分安排如下。下一节将阐述 CBI 的概念，以及使用聊天机器人的原因，并重点介绍研究问题。随后将综述最新研究进展。第四节简要概述所用方法，第五节介绍 CBI虚拟助手参考模型。接下来一节将展开讨论，最后总结研究结果。

# 2. 背景

## 2.1. CBI 概念

CBA 是优化商业分析流程、加强协作并推动业务取得更佳成果的有效方式。CBA 通常借助现代协作工具，例如社交网络、头脑风暴会议和聊天应用程序，使利益相关者能够不受地理位置限制、实时协同工作。通过这些工具，利益相关者可以以更具互动性且更及时的方式分享想法、观点和见解，从而让分析流程更加高效。CBI 旨在让用户彼此分享数据、见解和观点，以改进决策并推动业务取得更佳成果 [2, 11, 12]。CBI 概念的要点可概括如下。

- 协作：CBI 以团队成员之间的协作和知识共享为核心理念。通过促进沟通与协作，组织可以营造数据驱动决策的文化。• 数据共享：CBI 涉及跨团队和部门共享数据。这意味着打破数据孤岛，并确保每个人都能访问相同的数据，无论数据存储在哪里。

- 可视化：CBI 利用数据可视化工具，使数据更易于访问和理解。这使团队成员能够快速识别趋势和模式，并做出明智的决策。

- 敏捷性：CBI 具有敏捷性，这意味着它能够帮助团队快速适应不断变化的业务需求，并根据最新数据做出决策。
- 持续改进：CBI 是一个持续改进的过程，团队成员共同努力，识别改进机会，并相应地调整其工作流程。

因此，CBI 是一种强大的数据驱动决策方法，使组织能够汇聚团队的集体知识与专业能力，从而取得更好的业务成果 [12, 13]。

## 2.2. 聊天机器人

聊天机器人是利用人工智能与人类用户模拟对话的计算机程序 [14]。聊天机器人在协作工具中越来越受欢迎，原因有很多。
- 便利性：聊天机器人为用户提供便捷的方式，让他们无需离开通信平台即可获取信息并完成任务。这让团队成员可以轻松协作和共享信息，无需在多个工具之间切换。• 效率：聊天机器人可以自动处理常规任务和流程，让团队成员腾出时间专注于更高价值的工作。这有助于提高生产力，让团队更高效地工作。• 个性化：聊天机器人可以根据每位用户的具体需求进行定制，提供个性化体验，从而提升用户参与度和采用率。• 可扩展性：聊天机器人可以同时处理大量请求和交互，因此非常适合服务大量用户的协作工具。

- 可用性：聊天机器人全天候（每周 7 天、每天 24 小时）提供服务，让用户随时获取信息和支持。这对于跨不同时区工作的全球团队尤其有价值。

聊天机器人为用户提供了一种便捷高效的方式，使其能够在协作工具中访问信息并执行任务。通过自动化常规任务并提供个性化体验，聊天机器人可以帮助团队更有效地协作，更高效地实现目标。

本研究聚焦于从对话式助理概念（即聊天机器人）的角度开发CBI框架。本研究的主要目标是调查并改进基于虚拟助手的CBI工具。为此，我们提出以下研究问题。

1. 虚拟平台上的协作工具主要应用于什么场景?
2. 哪种类型的会话代理最适合作为CBI辅助工具?
3. CBI虚拟助手应采用哪些人工智能模型和自然语言理解技术?

## 3. 相关工作

CBI [11] 有多种实现方式，包括社交商业智能、移动商业智能、基于云的商业智能、自助式商业智能和嵌入式商业智能 [12]。

社交商业智能利用社交媒体平台收集见解，并与更广泛的受众共享信息 [15]。这种方法可用于促进不同团队和部门之间的协作与信息共享。它涉及从 Twitter、Facebook、LinkedIn 等社交媒体渠道收集和分析数据，以深入了解客户行为、市场趋势和品牌声誉。社交商业智能还涉及团队成员之间的协作，例如共享和讨论数据、可视化和报告，以便做出更明智的决策。

移动商业智能（MBI）[16] 通过智能手机和平板电脑等移动设备提供对商业智能工具和信息的访问。这种方式适用于远程团队或经常需要外出的团队。

采用云端商业智能（CBBI）时，[17], 数据存储在云端的集中位置，而非本地服务器或硬盘上。这使组织无论身处何处或使用何种设备，都能更轻松地访问和共享数据。CBBI 工具通常可通过网页浏览器或移动应用程序访问，为用户提供了一种简单便捷的数据访问和分析方式。CBBI 能够带来多项优势，例如降低成本、可扩展性和灵活性。组织可以避免构建和维护本地部署的商业智能基础设施所需的高昂成本，并根据需要扩大或缩减商业智能能力。此外，云端商业智能允许用户在任何有互联网连接的地方开展工作，使远程团队能够更轻松地协作和共享见解。

自助式商业智能（SSBI）是指一种赋能业务用户自行创建报告、仪表板和分析的方法，而无需依赖信息技术部门或数据分析师 [18]。它使用户无需具备技术技能或获得 IT人员的协助，即可访问和分析数据。SSBI 工具通常具有用户友好的界面，使用户能够轻松访问和操作数据，包括拖放功能、可视化工具和自然语言处理能力。这种方法允许业务用户自行回答问题、探索数据并生成洞察，从而有助于更快地进行决策并取得更好的业务成果。

最终，嵌入式商业智能（EBI）将商业智能工具和洞察直接集成到其他业务应用程序和工作流中 [19]。这可以让用户在工作环境中轻松访问相关数据和洞察，从而促进协作。EBI 的主要优势在于，用户可以直接在已经使用的应用程序中访问分析和洞察，无需在不同工具或界面之间切换。这可以提高生产力，因为用户能够更快速、更轻松地做出决策；同时，由于商业智能能力更易于访问，也可以提高其采用率。

此外，聊天机器人可以成为促进 CBI 的强大工具，通过为用户提供快速、便捷的数据和洞察获取途径，改善协作与反馈，并帮助在整个组织内实现数据和洞察的民主化。例如，在一篇研究文章 [20], 中，作者提出，聊天机器人可用于促进团队成员之间的协作与知识共享，尤其适用于信息检索和决策等任务。该领域的其他专家也倡导了这一观点。

用于在 CBI 中采用聊天机器人，以改善协作、提高可访问性并简化工作流 [21, 22]。

聊天机器人可根据其功能以及所支持的协作类型划分为不同类别 [4, 14, 22]。聊天机器人可用于根据现有数据提供见解、建议或预测。协作型聊天机器人可用于分配任务、安排会议，或在团队成员之间共享信息。例如，聊天机器人可用于收集用户对仪表板或报告的反馈，或促进团队成员围绕关键见解或指标展开讨论。聊天机器人还可根据预定义的触发条件（例如数据变化或关键指标异常）向用户发送通知和警报。这有助于团队及时了解相关信息，并根据实时信息及时做出决策。

聊天机器人是会话代理的一种，但并非所有会话代理都是聊天机器人。会话代理是一个更广泛的术语，涵盖任何能够与用户进行自然语言交互的计算机程序或系统 [21, 23]。会话代理可以基于规则，也可以依靠机器学习和自然语言处理技术来理解并回应用户输入。

Iris 会话代理是由斯坦福大学自然语言处理组开发的对话系统 [24]。它旨在使用自然语言与用户进行基于文本的对话。Iris 旨在帮助数据科学家和其他用户通过自然语言对话与数据和机器学习模型进行交互 [24]。Iris 可以集成到各种平台和工具中，例如 Slack、Jupyter Notebooks 和应用程序接口。用户可以向 Iris 提出有关数据的问题，例如“我所在公司的员工平均工资是多少？”或“我的数据集中年龄的分布情况如何？”Iris 还可以协助用户创建可视化内容，并利用机器学习模型分析数据。Iris 使用自然语言处理和机器学习技术来理解用户的意图并生成相关回复。该系统旨在从用户交互中学习，并随着时间的推移不断改进。Iris 还能够处理复杂的查询以及多个数据源。

一些类似的会话代理包括 Apple 的 Siri [25], Amazon 的 Alexa [26] 和 Google Assistant [27]。这些虚拟助手也使用自然语言处理来帮助用户完成任务、回答问题和执行操作。此外，还有一些可用的聊天机器人平台，例如 Dialogflow [28] 和 Microsoft Bot Framework [29], ，这些平台允许开发人员创建自己的会话代理。

此外，还有多种常用于设计会话代理的神经网络架构。其中，应用最广泛的是循环神经网络（RNNs）、长短期记忆（LSTM）、生成对抗网络（GANs）、序列到序列模型以及 Transformer [4, 24, 30, 31]。循环神经网络是一种特别适合处理序列数据（例如文本）的神经网络 [32]。它们通常用于语言建模，即根据序列中前面的词语预测下一个词语。在会话代理的应用场景中，循环神经网络可用于根据用户输入生成回复。长短期记忆网络是一种循环神经网络，旨在解决梯度消失问题，该问题可能出现在深度神经网络的训练过程中 [33]。长短期记忆网络在对序列数据中的长期依赖关系进行建模方面尤其有效，因此非常适合需要理解对话上下文的会话代理。生成对抗网络是一种可用于生成逼真合成数据（例如文本）的神经网络 [34]。在会话代理的应用场景中，与传统的基于规则或基于模板的方法相比，生成对抗网络能够生成更多样化、更具创造性的用户输入回复。

序列到序列模型是一种常用于机器翻译和文本摘要等自然语言处理任务的神经网络架构 [24, 35]。例如，在 Iris [24]，中，序列到序列模型用于处理用户查询并生成回复。注意力机制有助于模型在生成输出序列时关注输入序列中相关的部分。除了序列到序列模型，Iris 还使用强化学习（RL）来逐步提升性能 [24]。强化学习是一种机器学习方法，智能体通过在环境中采取行动来学习，以最大化累积奖励信号 [35]。在传统的强化学习中，奖励信号由环境提供；而在基于人类反馈的强化学习中，奖励信号则由人类专家提供。因此，强化学习是构建会话代理的常用方法。在 Iris 中，强化学习用于提高序列到序列模型生成回复的准确性和相关性。最后，Transformer 是一种神经网络架构，旨在并行处理输入数据序列，而不是像循环神经网络那样按顺序处理 [36]。因此，Transformer 的速度和效率远高于循环神经网络，非常适合用于构建大规模会话代理。

这些只是用于会话代理（CAs）的神经网络架构中的几个示例。具体架构的选择取决于当前任务和数据集的具体情况，以及可用于训练和部署的资源。

总体而言，会话代理开发领域正在迅速发展，人们采用各种方法来创建有效的对话式界面。其中一个最突出的趋势是利用神经网络和深度学习技术来提高会话代理的准确性和自然度。这包括使用预训练语言模型来生成更加真实、更加接近人类表达方式的回复。此外，人们对会话代理在各种应用领域中的使用越来越感兴趣，包括客户服务、医疗保健、教育和娱乐。这推动了面向特定领域的专用会话代理的开发。

我们可以得出结论，聊天机器人和会话代理在 CBI 中的应用日益普遍，因为它们为检索和共享数据提供了便捷易用的界面。当前 CBI 的发展趋势包括采用基于云的解决方案、移动商业智能，以及整合自然语言处理和机器学习等先进技术。自助式商业智能也越来越受欢迎，它赋予用户创建自己的报告和分析的能力。

## 4. 方法与材料

在探索如何界定业务问题的过程中，有多种技术可用于确定需要提出的关键问题。以下是一些适用的技术。

1. SWOT分析：这是一种结构化框架，有助于识别与业务问题相关的优势、劣势、机遇和威胁 [37]。它有助于确定与该业务问题相关的重要问题。

2. 根本原因分析：这项技术用于识别问题的深层根本原因 [38]。它有助于确定与根本原因相关、需要在探索过程中进一步探究的关键问题。

3. 鱼骨图：这是一种用于直观呈现问题可能成因的图形工具 [39]。它可用于确定与需要探究的成因相关的关键问题。

4. 五个为什么法：这是一种简单的提问技巧，通过连续五次询问“为什么”来找出问题的根本原因 [40]。它有助于确定与根本原因相关、需要进一步探究的关键问题。

5. 最小化d 图：这是一种视觉技巧，有助于围绕相关想法和概念进行头脑风暴。 e 商业智能 p 业务问题 [41]。它可用于识别与该业务问题相关的关键问题。

数据探索是 BI 流程中的一个重要组成部分，涉及收集、识别和分析数据，以发现有意义的洞察和模式。数据探索的主要目标是识别能够推动决策并提升业务绩效的关键业务机遇与挑战。BI 数据探索的主要步骤可概括如下。

1. 定义业务问题涉及确定数据探索可能发挥作用的业务领域、明确需要解决的问题，以及确定在探索过程中需要提出的关键问题。

2. 数据收集涉及从各种来源收集原始数据，例如数据库系统、数据仓库、电子表格、网页、社交媒体或其他相关来源。

3. 数据准备涉及对数据进行清理和格式化，以确保其准确性、完整性、一致性和质量。此步骤还包括将数据转换为更易于使用的格式，例如表格、图形、图表和其他可视化。

4. 探索数据是指通过分析数据，识别不同变量之间的模式、趋势、相关性和关系。此步骤可能涉及探索性分析技术，例如聚类分析、回归分析、相关性分析、时间序列分析以及其他统计方法。

5. 沟通洞察涉及以清晰、简洁的方式呈现研究结果，使用可视化、仪表板、报告或其他易于利益相关者理解的沟通形式。

6. 行动规划包括利用从数据探索中获得的洞察来推动行动并作出明智决策，从而提升业务绩效。此步骤可能涉及制定策略、设定目标和具体目标，以及根据从数据探索中获得的洞察实施行动计划。

此外，会话代理开发中还常用几种架构。

1. 基于规则的架构：会话代理使用一组预定义规则和决策树来理解用户输入并生成回复 [42]。这种方法简单且易于实现，但可能无法处理复杂交互，也难以适应用户行为的变化。

2. 有限状态机（FSM）架构：会话代理使用一组状态和状态转换来建模对话流程 [43]。与基于规则的架构相比，这种方法具有更高的灵活性，能够处理更复杂的交互，但设计和维护状态机可能需要投入更多的开发工作。

3. 自然语言理解（NLU）架构：会话代理使用机器学习算法来理解用户输入并生成响应 [44]。这种方法更为先进，能够适应用户行为随时间发生的变化，但需要大量训练数据和计算资源。

4. 混合架构：许多会话代理结合使用上述架构，以平衡简洁性、灵活性和准确性 [3, 31, 44, 45]。例如，会话代理可以使用基于规则的技术处理简单交互，并使用自然语言理解（NLU）处理更复杂的交互。

总体而言，架构的选择取决于会话代理的具体需求，以及可用于开发和训练的资源。

图 1 展示了典型的面向任务的对话智能体由三个模块组成。这些模块分别是：(1) 自然语言理解（NLU），用于识别用户意图并提取相关信息；(2) 对话管理（Dialogue Management），用于跟踪对话状态，以捕获对话中的所有关键信息，并根据当前状态选择下一步操作；(3) 自然语言生成（NLG），用于将智能体的操作转换为自然语言响应。近年来，越来越多的研究倾向于通过深度神经网络将这些模块结合起来，从而构建完全由数据驱动的系统。

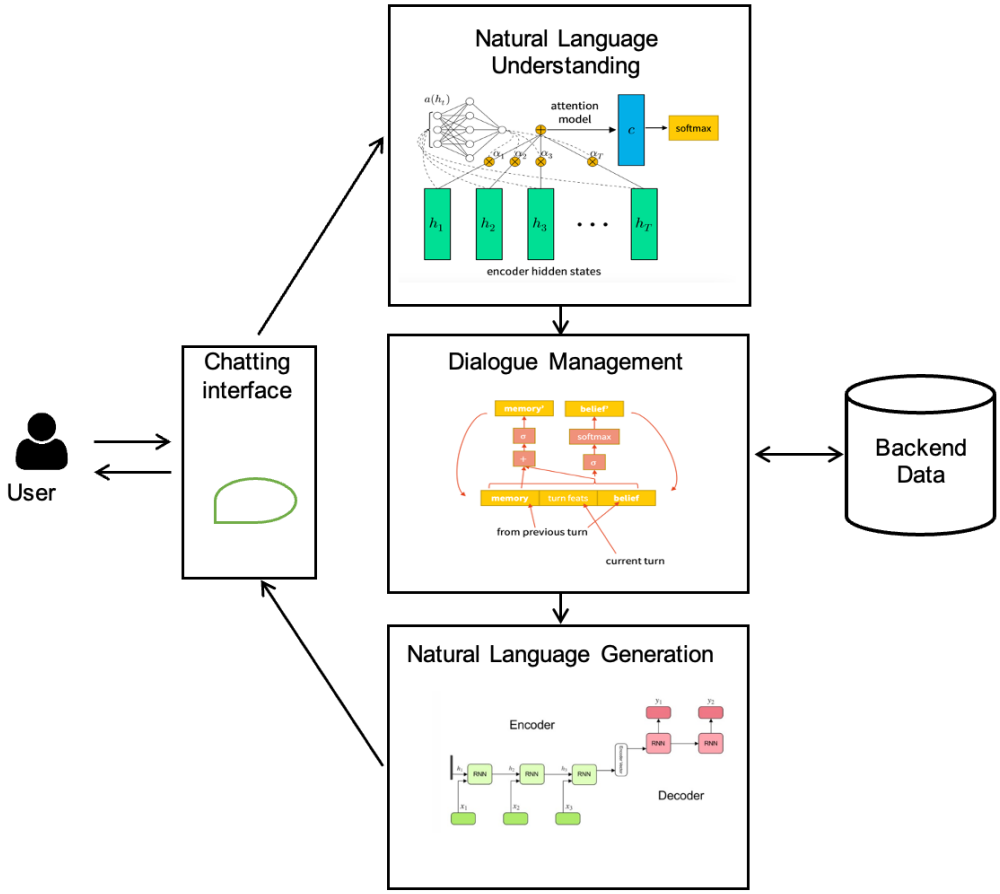


**图 1**：典型对话智能体

因此，为了为 CBI 虚拟助手选择合适的参考模型，我们需要明确协作模型、数据探索的步骤以及基本的神经网络架构。

## 5. 结果

让我们更详细地探讨组织联合数据分析任务的特点。实现用于 CBI 的交互式智能体时，需要考虑三个主要方面（图 2）。首先，会话代理必须能够使用自然语言进行交流，避免使用高度专业化的词汇，并开展类似人类之间的对话。此类任务通常通过采用 Transformer 等神经网络架构来解决，例如 ChatGPT。其次，虚拟助手必须能够将用户指令转换为数据处理命令（如筛选、中位数计算、特征分组等）。Iris 对话智能体 [24] 就是成功解决此类问题的一个例子。然而，Iris 的使用受到可用命令集合的限制，而且用户需要具备一定的经验才能解决数据分析问题。第三，虚拟助手必须支持协作过程，并根据对其他用户行为的分析提供推荐。在电子商务等领域，已有许多成功的推荐代理解决方案。然而，对于 CBI 任务，这些方案仍需进行适配。因此，CBI 虚拟助手必须融合这三个主要的会话代理方向。

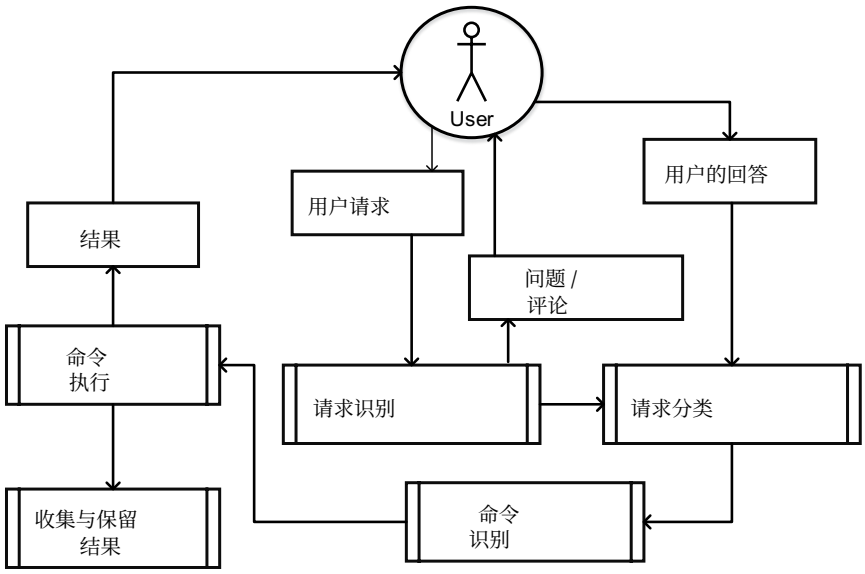


**图 2**：CBI 对话示例

本研究的核心思想是，通过智能体之间的交互，对分布式虚拟团队中的 CBI 流程进行建模。本文将智能体定义为一种置于外部环境中的计算机系统，能够与环境交互，并自主、理性地采取行动以实现特定目标。智能体与一般系统的主要区别在于其主动性，即能够独立执行各种操作 [46]。此外，智能体通常被视为一个整体，而非一组组成部分；例如，在研究系统属性时，采用的主要就是后一种视角。智能体的另一个特点是，它既可以是实体，也可以是独立运行的程序。然而，即使不影响物质世界（仅运行于计算机或计算机系统内部），这种程序也能执行有用的操作。

图 3 展示了会话代理（CA），它由三个组件构成：类人交互、数据探索和推荐代理。交互代理基于迁移学习，直接与用户沟通，并根据推荐提出数据研究方案。推荐代理收集并处理用户行为，并建议用户探索数据，以实现协作目标。数据探索代理则将自然语言表达转换为数据处理命令。

最终，为了构建用于 CBI 的虚拟助手，必须整合三个主要组件：会话代理、数据探索和推荐。建议将这些模块开发为多智能体系统 [46], ，其中包含三种不同类型的智能体，分别适用于各自的任务。这些智能体能够支持与用户的对话，捕捉数据探索和决策的上下文，并提供建议。我们的方法涉及以下智能体之间的交互：

这三种类型的智能体各自负责在对话过程中执行特定任务。

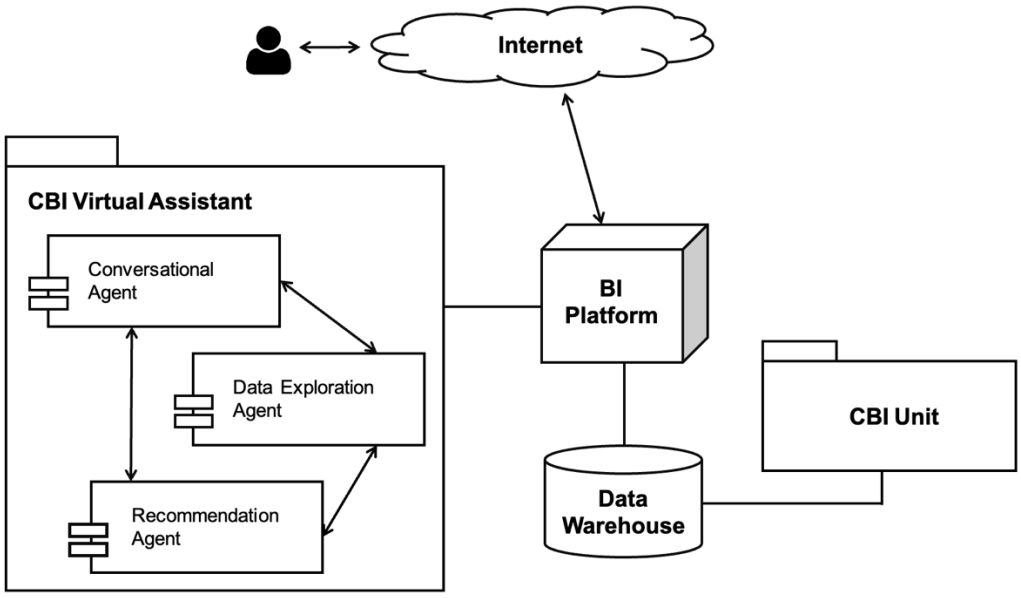


**图 3**：CBI虚拟助手参考模型

## 6. 讨论

CBI 的最新发展可描述为一个不断壮大的领域，越来越多的工具和技术正在开发之中，以促进决策过程中的团队有效协作。CBI 工具正变得更加用户友好、易于使用且灵活，使用户能够自定义使用体验并适应自身的具体需求。CBI 最重要的进展包括社交媒体功能、移动端可访问性和基于云的解决方案的集成。这些发展使用户能够开展协作，并随时随地通过任何设备访问数据。此外，自然语言处理和机器学习的应用使用户能够更轻松地与数据交互并提取洞见，从而使决策过程更加高效和有效。

近年来，会话代理开发领域的最新发展一直在快速演进。深度学习和自然语言处理的出现，使得创建更加先进、复杂的会话代理成为可能；这些会话代理能够更准确地理解和回应人类语言。自然语言查询（NLQ）也能让非技术用户更轻松地访问和分析数据。这有助于在组织内部普及数据和洞见，使更广泛的受众能够更方便地获取这些内容 [4, 14, 47]。虚拟助手的目标是让更广泛的用户群体能够更便捷地进行数据探索，并减少数据分析所需的时间和精力。它可以应用于医疗保健、金融、旅游和电子商务等多个领域。创建创新型会话代理的理念，是改变用户与数据及机器学习模型交互的方式，并让更广泛的用户能够更容易地接触数据科学。

Iris 会话代理基于深度学习和强化学习技术的强大组合，能够理解用户查询并生成相关响应 [24]。这使 Iris 成为数据科学家及其他希望通过自然语言对话与数据和机器学习模型进行交互的用户的强大工具。然而，Iris 的局限性在于它只能转换和执行一组特定的命令。它可以帮助那些不太熟悉 Python库和 Jupyter Notebook 的数据分析人员，但像 Iris 这样的会话代理无法为协作决策提供支持。

聊天机器人和会话代理可以作为自助式商业智能的一部分加以实现，让用户无需信息技术部门或数据分析师的协助即可轻松访问和分析数据。这样，用户便能用自然语言提问并即时获得答复，使数据分析更加便捷、高效。我们的参考模型结合了自然语言处理和机器学习技术，包括深度学习和强化学习这一前沿趋势。

因此，我们可以得出结论：虚拟助手能够通过运用深度强化学习和自然语言处理来改进协作工具。实现预训练模型（如 Iris 会话代理），并针对数据探索和协作进行微调，可为强大的 CBI 提供完整工具包。

## 7. 结论

CBI 与 CI 是相关概念，都涉及人们聚集起来协同工作并分享知识。CBI 主要指利用协作工具和方法支持 BI 流程。这可能包括社交网络、头脑风暴会议、聊天应用程序以及其他促进业务用户、分析师和技术专家之间知识共享与协作的技术。CBI 的目标是利用群体的集体知识和见解，改进决策、推动业务成果并获得竞争优势。CI 是一个更为广泛的概念，适用于任何一群人协力解决问题或做出决策的情境。这可能包括商业领域之外的场景，例如社会或政治运动、科学研究或开源软件开发。CI 的目标是发挥群体的集体智慧，从而得出更好的解决方案或做出比单靠个人努力更好的决策。

聊天机器人可有效用于 CBI。它们可以与协作平台或工具集成，例如聊天应用程序或社交网络，以促进业务用户、分析师和技术专家之间的知识共享与协作。会话代理在许多不同的应用场景中都能发挥作用。例如，人们正在研究如何利用会话代理帮助人们获取法律信息。也有研究探讨会话代理在车辆中的作用，尤其是在自动驾驶汽车领域。游戏也是会话代理的一个相关应用场景。研究人员还探讨了在平台上让社群成员与会话代理交谈，能否帮助游戏社群发展壮大并增进凝聚力。最后，人们经常提议在不同领域采用会话代理充当推荐系统。会话代理很可能会成为一种能够带来诸多益处的技术。这项技术或将彻底改变人们与机器交互的方式，并让所有人都能平等地使用在线和数字服务。

我们的 CBI 解决方案涉及使用虚拟助手。虚拟助手的目标是让更多用户能够更轻松地进行数据探索，并减少数据分析所需的时间和精力。为实现这一目标，我们回答了研究问题，并总结如下。

1. 基于文本通信的虚拟助手是协作工具的主要场景，它涉及以清晰简洁的方式呈现研究结果，并使用用户易于理解的可视化。它包括用户与聊天机器人之间的对话，以便探索数据、选择操作并创建可视化。聊天机器人需要提供建议，并协助执行数据处理操作。

2. 我们选择由三个单元组成的三组件对话代理，分别是对话、数据探索和推荐智能体。它们将彼此交互，并可作为多智能体模块实现，以支持 BI4People 平台上的协作单元。

3. 由于可选方案众多，人工智能模型和自然语言理解应在未来工作中通过实验研究加以确定。

## 8. 致谢

本文所述研究由法国国家研究署（ANR）资助，项目为 ANR-19-CE23-0005 BI4people（面向大众的商业智能）。

## 9. 参考文献

[1] 面向大众的商业智能。网址：https://eric.univ-lyon2.fr/bi4people/index-en.html。[2] D. Power、R. Sharda，《商业智能与分析》，Wiley，2015。doi:10.1002/9781118785317.weom070011。

[3] E. Adamopoulou、L. Moussiades，《聊天机器人：历史、技术与应用》，《机器学习应用》2（2020）100006。doi:10.1016/j.mlwa.2020.100006。[4] J. Masche、N.-T. Le，《对话系统技术综述》，载于：《计算机科学、应用数学与应用国际会议》，Springer，2017，第212–225页。[5] D. Avalverde，《聊天机器人的简史》。感知、控制、认知，2019。网址：https://pcc.cs.byu.edu/2018/03/26/a-brief-history-of-chatbots/。[6] Watson Assistant。网址：https://www.ibm.com/products/watson-assistant。[7] Dialogflow。网址：https://cloud.google.com/dialogflow。[8] 对话式人工智能平台。网址：https://rasa.com。[9] Botpress。网址：https://botpress.com。[10] Wit。网址：https://wit.ai [11] S. Rizzi, S，《协作式商业智能》，载于：M.A. Aufaure、E. Zimányi（编）《商业智能》（eBISS 2011），《商业信息处理讲义》，第96卷，Springer，柏林、海德堡，2012，第186–205页。doi:10.1007/978-3-642-27358-2_9。[12] J. Kaufmann、P. Chamoni、Peter，《协作式商业智能的结构化：文献综述》，载于：《夏威夷国际系统科学年会论文集》，2014。doi:10.1109/HICSS.2014.465。[13] F. Muhammad、J. Darmont、C. Favre，《协作式商业智能本体（CBIOnt）》，载于：第18届商业智能与大数据研讨会（EDA 2022），《信息科学与技术评论》，B-18卷，2022。[14] N. Svenningsson、M. Faraon，《对话代理中的人工智能：聊天机器人感知人性化相关因素研究》，载于：2019年第2届人工智能与云计算会议论文集（AICCC 2019），计算机协会，纽约，美国，2020，第151–161页。doi:10.1145/3375959.3375973。[15] B. Dinter、A. Lorenz，《社会化商业智能：文献综述与研究议程》，载于：信息系统国际会议（ICIS 2012），2012。[16] L. K. Chan 等，《移动商业智能：最新研究进展与关键成功因素综述》，载于：第20届国际商业信息管理协会会议（IBIMA 2013）论文集，第1卷，IBIMA，2013，第708–717页。[17] J. Ereth、B. Henning，《基于云的商业智能与分析应用——商业价值与可行性》，载于：亚太信息系统会议（PACIS 2015）论文集，2015。[18] C. Lennerholt、J. Van Laere、E. Söderström，《自助式商业智能面临的用户相关挑战》，《信息系统管理》38(4)（2021）309–323。[19] T. T. Parbat，《嵌入式商业智能市场的分析与趋势研究》，载于：《应用科学与工程技术国际研究期刊》9(12)（2021）1025–1031。[20] M. H. Jarrahi、D. Askay、A. Eshraghi、P. Smith，《人工智能与知识管理：人与人工智能的伙伴关系》，《商业视野》66(1)（2023）87–99。doi:10.1016/j.bushor.2022.03.002。[21] M. Akhtar、J. Neidhardt、H. Werthner，《聊天机器人的潜力：聊天机器人对话分析》，载于：2019年IEEE第21届商业信息学会议（CBI），IEEE，2019，第397–404页。doi:10.1109/CBI.2019.00052。[22] G. Caldarini、S. Jaf、K. McGarry，《聊天机器人最新进展的文献调查》，《信息》13(1)（2022）41。doi:10.3390/info13010041。[23] R. Bavaresco、D. Silveira、E. Reis、J. Barbosa、R. Righi、C. Costa、R. Antunes、M. Gomes、C. Gatti、M. Vanzin、S. C. Junior、E. Silva、C. Moreira，《商业领域的对话代理：系统性文献综述与未来研究方向》，《计算机科学评论》36（2020）100239。[24] E. Fast、B. Chen、J. Mendelsohn、J. Bassen、M. Bernstein，《Iris：用于复杂任务的对话代理》，载于：2018年CHI计算系统中的人为因素会议论文集，2018，第1–12页。doi:10.1145/3173574.3174047。[25] Siri – Apple。网址：https://www.apple.com/siri/ [26] Amazon Alexa语音人工智能。网址：https://developer.amazon.com/en-US/alexa。[27] Google Assistant。网址：https://assistant.google.com/。

[28] Dialogflow。网址：https://cloud.google.com/dialogflow。 [29] Microsoft Bot Framework。网址：https://dev.botframework.com/。 [30] J. Schmidhuber，神经网络中的深度学习：概述，《神经网络》61（2015）85–117。[31] D. W. Otter，J. R. Medina，J. K. Kalita，自然语言处理中深度学习用途综述，载于：《IEEE神经网络与学习系统汇刊》，2020年，第604–624页。doi:10.1109/TNNLS.2020.2979670。[32] X. Zhu，基于RNN语言处理模型的口语对话系统建模方法，《计算智能与神经科学》（2022）6993515。doi:10.1155/2022/6993515。 [33] M. H. Toding Bunga，S. Suyanto，使用长短期记忆开发完整的对话系统，载于：2019年信息技术与智能系统研究国际研讨会（ISRITI），印度尼西亚日惹，2019年，第326–329页。doi:10.1109/ISRITI48646.2019.9034567。[34] M. Moradshahi，C. Utkarsh，使用生成对抗网络进行语言建模，arXiv，2018年。网址：https://arxiv.org/abs/1804.02617。 [35] Y. L. Hsueh，T. L. Chou，基于长短期记忆和强化学习的面向任务的聊天机器人，《ACM亚洲及低资源语言信息处理汇刊》22（1）（2022）1–27。doi:10.1145/3529649。[36] T. Wolf，L. Debut，V. Sanh，J. Chaumond，C. Delangue，A. Moi，P. Cistac，T. Rault，R. Louf，M. Funtowicz，J. Brew，Transformer：自然语言处理的最新研究进展，载于：自然语言处理实证方法会议，2019年，第38–45页。doi:10.18653/v1/2020.emnlp-demos.6。 [37] S. Gürel，M. Tat，SWOT分析：理论综述，《国际社会研究杂志》10（2017）994–1006。 [38] A. M. Doggett，根本原因分析：工具选择框架，载于：《质量管理期刊》12（2005）34–45。[39] M. Coccia，用于技术分析和前瞻的鱼骨图，《国际前瞻与创新政策期刊》14（2020）225–247。doi:10.1504/IJFIP.2020.10033239。 [40] N. James，开展定性分析的五种方法，《英国教育研究期刊》38（2012）883–884。[41] B. Camburn，R. M. Arlitt，D. Anderson，P. Sanaei，S. Raviselam，D. Jensen，K. L. Wood，通过众包和机器学习生成计算机辅助思维导图，《工程设计研究》31（2020）1–27。doi:10.1007/s00163-020-00341-w。 [42] M. B. Islam，G. Governatori，RuleRS：用于决策支持系统的基于规则的架构，《人工智能与法律》26（2018）315–344。[43] C. Chakrabarti，G. F. Luger，用于客户服务聊天机器人的人工对话：架构、算法和评估指标，《应用专家系统》42(20)（2015）
6878 6897 10 1016 2015 04 067 –。doi: 。/j.eswa. 。 。A. Radford，K. [44]
Narasimhan，T. Salimans，I. Sutskever，通过生成式预训练改进语言理解，2018年。网址：https://paperswithcode.com/paper/improving-language-understanding-by。 [45] R. Matić，M. Kabiljo，M. Zivkovic，M. D. Čabarkapa，使用自然语言理解元模型的可扩展聊天机器人架构，《电子学》10(18)（2021）2300。doi:10.3390/electronics10182300。[46] C. Clarke，J. Peper，K. Krishnamurthy，W. J. Talamonti，K. Leach，W. S. Lasecki，Y. Kang，L. Tang，J. Mars，一个代理统领一切：迈向多智能体对话式人工智能，载于：计算语言学协会论文集：ACL 2022，爱尔兰都柏林，2022年，第3258–3267页。doi:10.18653/v1/2022.findings-acl.257。 [47] D. Adiwardana，M.-T. Luong，D. R. So，J. Hall，N. Fiedel，R. Thoppilan，Z. Yang，A. Kulshreshtha，G. Nemade，Y. Lu，V. Le Quoc，迈向类人开放领域聊天机器人，arXiv，2020年。网址：https://arxiv.org/abs/2001.09977。