构建棋类运动大数据平台　　引领棋类软件与棋手和谐发展

棋类运动在人工智能背景下的发展问题研究

胡贵平

（飞马智科信息技术股份有限公司　马鞍山 243000，中国）

摘要：本篇研究论文以建立棋类运动大数据平台为切入点，给从事棋类运动管理机构和部门提出了从构建棋类运动基础数据入手，以棋手选拔培养、训练提高到比赛分析、辅助改进、创新发展为主线 ， 以各种棋类软件为辅助工具，阐述了棋类运动大数据平台在棋类运动日常管理训练及比赛中发挥的重要作用，本篇研究论文最后为棋类运动大数据平台项目建设提出了指导性的意见。本篇研究论文对全球棋类运动管理机构和部门,在信息 2.0 时代下的管理思路和方法具有一定的参考价值，十分具有现实意义和可操作性。

关键词：棋类运动　大数据　棋类软件　人工智能

Constructing a Big Data Platform for Chess games. Leading the harmonious development of chess software and chess players. Research on the Development of Chess games in the Background of Artificial Intelligence.

Hu Guiping

(PHIMA Intelligence Technology CO.,Ltd.　Maanshan 243000,China)

Abstract

This research paper is based on the establishment of large data platform for chess games.This paper puts forward that the basic data of chess should be constructed for the management organizations and departments engaged in chess games.The main line is to improve the selection and training of chess players to match analysis, auxiliary improvement and innovative development. Using various chess software as assistant tools. This paper expounds the important role played by the big data platform of chess games in the daily management training and competition of chess games. In the end, the paper puts forward some guiding opinions for the construction of large data platform for chess games. This research paper has a certain reference value for the management ideas and methods of global chess management organizations and departments in the information 2.0 era. It has practical significance and operability.

Key words

Chess games; Big data; Chess software; Artificial intelligence

在本招标课题未被专家组审核通过前，版权归作者所有。

一、棋类运动大数据平台概要

1.1 引言

人类进入21世纪后，随着信息技术（IT）及人工智能（AI）技术的飞速发展，互联网+已经广泛渗透到各行各业，给人们传统的工作生活方式带来了无比深刻的变化。插上IT技术、AI技术翅膀的行业在与传统行业竞争中大获全胜，用IT技术、AI技术、互联网+思维改造传统行业是当今中国乃至世界范围内的一个热门且永恒不变的主题。

作为传统的智力运动产业，也面临同样的问题。近些年来，随着信息技术、云计算、大数据、人工智能技术飞速发展及各种棋类软件的问世，特别是当一些顶尖棋类对弈软件的水平超过世界上最优秀的专业棋手时，也带来了现实考量。人类究竟能不能战胜强软？在大数据及人工智能时代棋手如何生存？棋类运动在大数据及人工智能时代如何科学和谐发展等诸多疑问摆在了我们面前。

大数据是信息2.0时代一种新的数据体系及架构，最早提出大数据时代到来的是全球知名咨询公司麦肯锡，麦肯锡称：“数据，已经渗透到当今每一个行业和业务职能领域，成为重要的生产因素。人们对于海量数据的挖掘和运用，预示着新一波生产率增长和消费者盈余浪潮的到来。” 大数据在物理学、生物学、气象学等领域以及军事、金融、通讯等行业应用已有时日，却因为近些年来互联网和信息产业的迅猛发展而引起人们广泛关注。

构建棋类运动大数据平台对于研究棋类运动规律、科学指导棋手训练比赛、创新棋类运动发展方向、引领棋类软件与棋手在大数据及人工智能时代和谐发展具有十分重要的现实意义。

1.2 大数据的概念

大数据（big data），是指无法在一定时间范围内用常规软件工具进行捕捉、管理和处理的数据集合，是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产。

在维克托·迈尔-舍恩伯格及肯尼斯·库克耶编写的《大数据时代》中大数据指不用随机分析法（抽样调查）这样捷径，而采用所有数据进行分析处理。大数据具有5V特点（IBM提出）：Volume（大量）、Velocity（高速）、Variety（多样）、Value（低价值密度）、Veracity（真实性）。

麦肯锡全球研究所给出的大数据定义是：一种规模大到在获取、存储、管理、分析方面大大超出了传统数据库软件工具能力范围的数据集合，具有海量的数据规模、快速的数据流转、多样的数据类型和价值密度低四大特征。

大数据包括结构化、半结构化和非结构化数据，非结构化数据越来越成为数据的主要部分。据IDC的调查报告显示：企业中80%的数据都是非结构化数据，这些数据每年

在本招标课题未被专家组审核通过前，版权归作者所有。

都按指数增长60%。 大数据就是互联网发展到现今阶段的一种表象或特征而已，没有必要神话它或对它保持敬畏之心，在以云计算为代表的技术创新大幕的衬托下，这些原本看起来很难收集和使用的数据开始容易被利用起来了，通过各行各业的不断创新，大数据会逐步为人类创造更多的价值。

现代社会是一个高速发展的社会，科技发达，信息流通，人们之间的交流越来越密切，生活也越来越方便，大数据就是这个高科技时代的产物。阿里巴巴创始人马云早年在演讲中就提到，未来的时代将不是IT时代，而一定是DT时代，DT就是Data Technology数据科技，显示大数据对于未来、对于阿里巴巴集团举足轻重的地位。

有人把大数据比喻为蕴藏能量的煤矿，煤炭按照性质有焦煤、无烟煤、肥煤、贫煤等分类，而露天煤矿、深山煤矿的挖掘成本又不一样。与此类似，大数据并不在“大”，而在于“有用”。价值含量、挖掘成本比数量更为重要。对于很多行业而言，如何利用这些大规模数据是赢得竞争的关键。对于竞技棋类运动而言，棋类运动大数据就可以很方便的分析诸如以下有价值的信息。

各种布局的先手胜率分析；

各种布局的后手胜率分析；

任意一名棋手的先手布局喜好及胜率分析；

任意一名棋手的后手布局喜好及胜率分析；

任意一名棋手历史胜局及获胜原因分析；

任意一名棋手历史负局及失利原因分析；

任意一名棋手棋谱数据软件吻合度指标分析；

任意两名棋手之间历史交战数据分析；

……

1.3 棋类运动数据研究现状分析

作为传统智力竞技运动项目，棋类运动有着悠久的历史，棋类运动发展到现在比较普及的主要有象棋、围棋、国际象棋和国际跳棋，此外还有五子棋，军旗，飞行棋等大众游戏项目。

棋类运动理论上属于纯脑力运动，下棋可以锻炼我们的识记能力、思考能力和计算能力等，经常进行棋类运动可以对我们的大脑大有裨益，对于培养我们的优秀性格也有很大的作用。民间有“善弈者长寿”之说，可见下棋对于身心健康也有一定的作用。

棋类运动对培养人的耐力、养成强大的专注力、培养全局观念、锻炼思维的缜密性、加强记忆力、养成良好的个性都是十分有帮助的。同时还可以陶冶情操，易于形成积极乐观豁达的人生品质和处世哲学。

棋类运动发展到现今，特别是以信息技术、人工智能为代表的新技术飞速发展，人们对棋类运动数据的研究也从无到有，从弱到强。特别值得一提的是一些棋类软件研发

在本招标课题未被专家组审核通过前，版权归作者所有。

公司，他们在棋类运动数据收集整理方面已经悄然走在了前列，表现在他们对现有棋谱数据的收集度非常至高，目前顶尖棋类对弈软件的棋谱收藏可以说是海量的，几乎涵盖了所有专业及业余棋手的对局棋谱。拥有了这些数据，使得他们的软件如鱼得水、水涨船高。

譬如由谷歌（Google）旗下DeepMind公司戴密斯·哈萨比斯领衔的团队开发的阿尔法围棋（AlphaGo）就是世界上首个击败人类顶尖职业围棋选手、第一个战胜围棋世界冠军的人工智能机器人。2016年3月，阿尔法围棋与围棋世界冠军、职业九段棋手李世石进行围棋人机大战，以4比1的总比分获胜；2016年末2017年初，该程序在中国棋类网站上以“大师”（Master）为注册账号与中日韩数十位围棋高手进行快棋对决，连续60局无一败绩；2017年5月，在中国乌镇围棋峰会上，它与排名世界第一的世界围棋冠军柯洁对战，以3比0的总比分获胜。围棋界公认阿尔法围棋的棋力已经超过人类职业围棋顶尖水平，在GoRatings网站公布的世界职业围棋排名中，其等级分曾超过排名人类第一的棋手柯洁。

国内第一个象棋全局对弈软件是黄少龙大师1989年参与编写的NKW象棋软件，随后象棋软件如雨后春笋层出不穷。象棋软件的诞生是广大棋迷的福音，使象棋爱好者找到了学棋的好助手，通过利用机器帮助人来拆棋，研究布局、中局、残局着法，给人类提高棋力提供新的方式，经过多年的发展、加上电脑配置的不断提高，象棋软件可以说是百花齐放，百家争鸣。现在的顶尖大师几乎不能与象棋旋风、象棋名手等顶尖象棋软件相抗衡，特级大师全国冠军亦是难与之招架。

不管是哪一种棋类对弈软件，现阶段他们对棋类运动数据的分析研究大多数只是停留在棋谱数据单一层面，对于棋手个性数据（布局偏好、中局习惯、残局取向）等其他方面研究尚停留在初级阶段，原因是他们没有这方面的深挖数据作为支撑，因此软件也就存在一定的局限性。

本篇研究论文试图建立棋类运动大数据平台，以大数据为支撑，多角度探索优化棋类软件，为棋手、软件提供新的思路和设计理念。不论是对于棋手还是电脑软件，只有知己知彼，方能百战不殆，当然这只是棋类运动大数据平台的一个最简单的数据分析应用而已。

现阶段，就棋类运动而言，无论是国内还是国外目前还没有一个专门的数据分析公众平台能为各级棋类运动相关机构、个人提供数据分析服务。现有的各种棋类网站提供的也仅仅是一些没有经过任何加工分析的原汁原味的历史数据。随着人工智能时代的开启，特别是以云计算、大数据为代表的新兴技术的飞速发展，人们对棋类运动数据的研究需求也愈来愈迫切，这也从客观上为棋类运动数据研究快速迈入大数据时代提供了可能。

1.4 建立棋类运动大数据平台的重要意义

在本招标课题未被专家组审核通过前，版权归作者所有。

众所周知，大数据在带来巨大技术挑战的同时，也带来巨大的技术创新与商业机遇。不断积累的大数据包含着很多在小数据量时不具备的深度知识和价值，大数据分析挖掘将能为行业/企业带来巨大的商业价值，实现各种高附加值的增值服务，进一步提升行业/企业的经济效益和社会效益。大数据带来巨大的变革的同时，也改变我们的生活、工作和思维方式，改变我们的商业模式，影响我们的经济、政治、科技和社会等各个层面。据统计，大数据每年可为美国的医疗服务业节省3000亿美元，帮助美国零售业净利润增长60%。为欧洲的公共部门管理每年节省2500亿欧元，为全球个人位置数据服务提供商贡献1000亿美元，帮助制造业在产品开发、组装等环节节省50%的成本。

现阶段，建立棋类运动大数据平台，同样具有十分重要的现实意义，主要体现在以下几个方面。

1.4.1 建立棋类运动大数据平台是当今中国时代发展的需要。

2016年10月9日 下午，中共中央政治局就实施网络强国战略进行了第三十六次集体学习。中共中央政治局各位同志认真听取了讲解，并就有关问题进行了讨论。“建设全国一体化的国家大数据中心”是此次集体学习内容的重点之一。习近平总书记亲自参加并主持学习会，习总书记对大数据非常重视，在此次集体学习之前他曾就“大数据”进行过多次重点指示：

当今世界，科技进步日新月异，互联网、云计算、大数据等现代信息技术深刻改变着人类的思维、生产、生活、学习方式，深刻展示了世界发展的前景。

--致信祝贺国际教育信息化大会开幕(2015年05月23日)

我国大数据采集和应用刚刚起步，要加强研究、加大投入，力争走在世界前列。

--在贵州大数据产业园调研时强调(2015年06月18日)

加快制造大国向制造强国转变，推动移动互联网、云计算、大数据技术创新和应用。

--在二十国集团领导人第十次峰会第一阶段会议上关于世界经济形势的发言(2015年11月15日，安塔利亚)

“十三五”时期，中国将大力实施网络强国战略、国家大数据战略、“互联网+”行动计划，发展积极向上的网络文化，拓展网络经济空间，促进互联网和经济社会融合发展。我们的目标，就是要让互联网发展成果惠及13亿多中国人民，更好造福各国人民。

--在第二届世界互联网大会开幕式上的讲话(2015年12月16日，乌镇)

要依法加强对大数据的管理。一些涉及国家利益、国家安全的数据，很多掌握在互联网企业手里，企业要保证这些数据安全。企业要重视数据安全。如果企业在数据

在本招标课题未被专家组审核通过前，版权归作者所有。

保护和安全上出了问题，对自己的信誉也会产生不利影响。

--在网络安全和信息化工作座谈会上的讲话(2016年4月19日)

大数据、云计算、移动互联网等新一代信息技术同机器人和智能制造技术相互融合步伐加快，科技创新链条更加灵巧，技术更新和成果转化更加快捷，产业更新换代不断加快，使社会生产和消费从工业化向自动化、智能化转变，社会生产力将再次大提高，劳动生产率将再次大飞跃。

--在全国科技创新大会、两院院士大会、中国科协第九次全国代表大会上的讲话(2016年5月30日)

加强人文领域合作，深入开展教育、科技、文化、体育、旅游、卫生、考古等领域合作，建立大数据交流平台，共同打造“一带一路”智库合作网络。

--在乌兹别克斯坦最高会议立法院的演讲(2016年6月22日)

要完善人口健康信息服务体系建设，推进健康医疗大数据应用。

--在全国卫生与健康大会上强调(2016年8月19日至20日)

过去40年里，我国在工业化、信息化和网络化方面都取得了不俗的成绩。但是也要看到经济高速发展所带来的问题，包括基础设施建设、环境保护、食品安全等方面的挑战，都需要使用大数据这种创新方式提供更好的解决方案。巨大的需求客观上为中国大数据的研究提供了许多实践机会。

中国的人口和经济规模决定了中国大数据的规模为全球最大，可为大数据研究提供许多创新角度，同时也意味着中国大数据分析的解决方案将同样适用于其他国家，这将为从“中国制造”向“中国创造”转变提供难得的机遇。

大数据在推动中国经济转型方面也将发挥重要作用。其一，通过大数据的分析可以帮助解决中国城镇化发展中面临的住房、教育、交通等难题。例如，通过对交通流量数据的实时采集和分析，可以指导驾驶者选择最佳路线，改善城市交通状况。其二，通过大数据的研究有助于推动钢铁、零售等传统产业转型升级，向价值链高端发展迈进。其三，大数据的应用可以帮助中国在发展战略性新兴产业方面迅速站稳脚跟，巩固并提升竞争优势。

1.4.2 建立棋类运动大数据平台是研究棋类运动发展规律的需要。

棋类运动既属于智力运动，也是一种体育运动。因为它是一种比智力，比体力的综合竞技项目。棋类运动大数据平台建立对于分析棋类运动数据、研究棋手的个性特点、探索各种棋类运动项目内在规律、拓展棋类运动未来发展方向具有十分重要的意义。

棋类运动大数据平台也是全国一体化的大数据平台的组成部分，当我国各行各业的

大数据中心相继建成并建立共享机制之后，全国一体化的国家大数据中心就有了坚实的基础。

棋类运动大数据平台的海量数据是一笔宝贵的财富，它取之不尽用之不竭。棋类项目的创新发展离不开棋类大数据平台的支撑，棋类运动的研究离不开大数据平台的支撑，全国一体化的国家大数据中心建设离不开棋类运动大数据平台建设，从这个意义上说，建立棋类运动大数据平台已经迫在眉睫、跃然纸上。

1.4.3 建立棋类运动大数据平台是人工智能与棋手和谐发展的需要。

棋类运动大数据平台的建立也是人工智能与棋手和谐发展的需要。我们都知道，人类社会即将迈入或已经迈入人工智能时代，人工智能时代催生了很多智能化的产物，棋类对弈分析软件就是其中的一个代表，当前国际上成熟先进的棋类对弈软件战胜顶尖专业棋手已经是家常便饭不再是新闻，棋手在与人工智能的 PK 中完全处于逆势，人工智能的优势地位与人类棋手相比有多个方面，最明显的是人工智能有大的棋谱数据做支撑，而人类的大脑记忆力是有限的，单凭这一点，再加上人工智能棋类软件深层的分析计算能力，人类几乎没法与之抗衡。

建立棋类运动大数据平台可以很好地帮助棋手分析和训练，利用棋类大数据平台再借助棋类辅助软件，很容易从海量的数据中分析找到最佳应手，从而提高棋手的实战水平。当棋手在任一复杂局面下都能找到近似于软件或更优于软件的招法，那么人类棋手的竞技水平将会大大提高，人工智能与棋手和谐发展的目标就一定能够实现。

1.5 棋类运动大数据平台体系结构

在本招标课题未被专家组审核通过前，版权归作者所有。

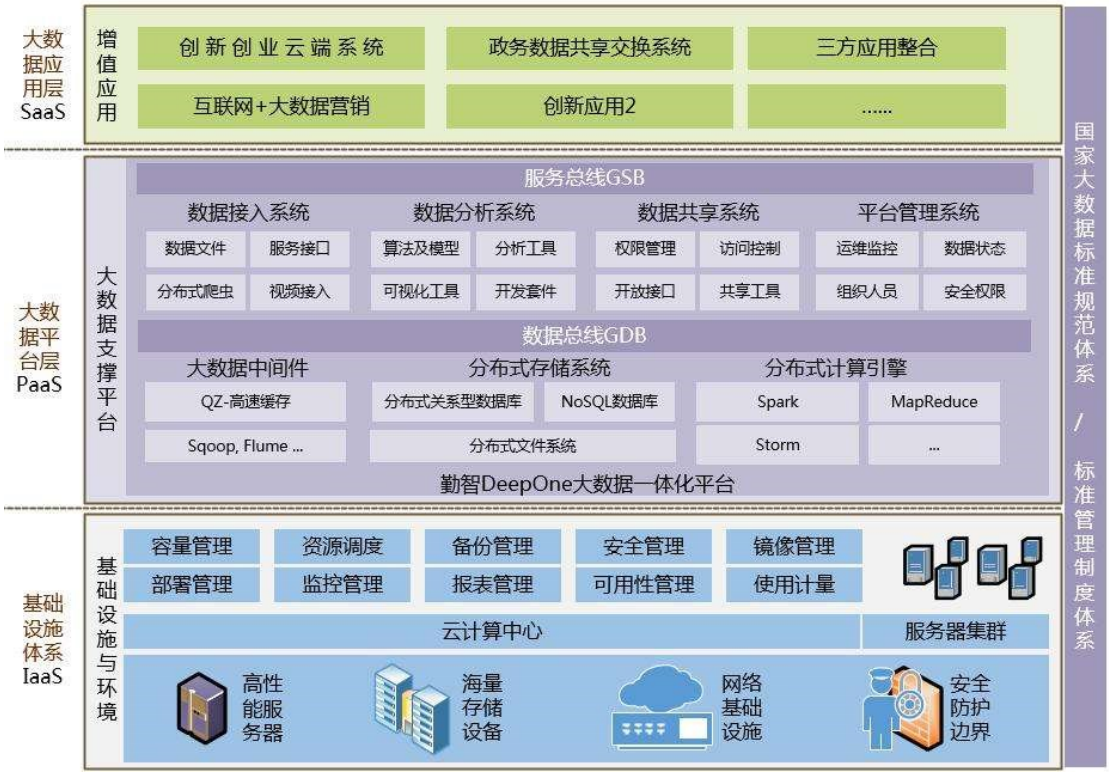

图 1：国家大数据标准规范体系 / 标准管理制度体系

图 2：棋类运动大数据平台系统架构

棋类运动大数据平台系统遵循国家大数据标准规范体系建设/标准管理制度体系标准来建设，系统采用四层架构。第一层（最底层）为基础软硬件层，这一层主要是指棋类运动大数据平台的软硬件环境，具体包括基础软硬件环境和大数据平台的相关软件支撑环境。第二层为信息资源库（基础数据层），该层主要存储棋类运动大数据有关基础

在本招标课题未被专家组审核通过前，版权归作者所有。

源数据、应用数据及共享数据。第三层为棋类运动大数据系统的支撑平台，该层主要完成基础数据后台的加工处理、挖掘分析计算及共享数据准备等工作，是棋类运动大数据平台的核心关键层，系统所有的基础数据都要经过本层加工产生目标数据提供给应用分析层使用。第四层即为智能服务区（大数据应用分析层），该层主要提供了一些分析工具和软件，为各类用户方便使用棋类运动大数据平台应用数据提供接口。

1.6 如何构建棋类运动大数据平台

构建棋类运动大数据平台是一项系统工程，需要统筹各方面力量，借助成熟先进的大数据技术来实现，棋类运动大数据平台以棋类数据资源整合为重点，以棋类运动大数据应用为建设核心，应当遵循“统筹规划、分步实施，整合资源、协同共享，突出重点、注重实效，深化应用、创新驱动”的建设原则，全面提升棋类运动信息化水平，引领棋类运动在人工智能时代科学健康和谐发展。

统筹规划、分步实施。结合棋类运动发展与人工智能应用现状，明确总体目标和阶段性任务，科学规划项目建设。先期完成大数据平台的底层架构建设，后期分步完成信息资源库、大数据支撑平台及智能服务区建设。

整合资源、协同共享。对棋类数据资源统一梳理，建立棋类运动标准资源库和数据规范，加快推进棋类数据资源整合，建设共享共用的棋类运动大数据中心，实现业务协同。

突出重点、注重实效。以用户为中心，以需求为导向，以服务为目的，突出重点，注重实效，加强平台可用性和易用性。

深化应用、创新驱动。深入了解各类用户需求，紧密跟踪信息技术发展趋势，不断深化应用、拓展新技术在应用中的广度和深度，促进跨界融合，丰富管理和服务手段。

二、棋类运动大数据平台数据组成

棋类运动大数据平台数据是由基础源数据、主题应用数据和计算挖掘数据三部分组成。期中基础源数据由以下六大块构成，分别是比赛大数据、棋手大数据、裁判（含教练）大数据、棋谱大数据、机构大数据和赞助商大数据。主题应用数据有等级分数据、技术等级数据等组成。计算挖掘数据是在基础源数据、主题应用数据基础上通过分析建模计算挖掘所产生的数据。

这三部分数据的关系如下，基础源数据是原汁原味的数据，比如棋谱大数据。主题应用数据是来源于基础源数据，是在此数据基础上根据给定的计算规则产生的数据，比如棋手等级分数据，它是根据棋手一定时间内比赛成绩计算得出的积分数据。计算挖掘数据是在基础源数据、主题应用数据基础上通过分析建模计算挖掘所产生的数据。比如棋手即时等级分数据，平台可以实时计算出一名棋手即时等级分，计算依据是根据棋手

在本招标课题未被专家组审核通过前，版权归作者所有。

上一积分周期积分加上计算日前比赛变动分值，这个计算过程平台是自动完成的。

棋类运动大数据组成图

2.1 基础源数据

2.1.1　比赛大数据。该部分数据主要记录棋类项目比赛信息，主要的源数据库记录表有比赛信息记录表、比赛成绩记录表。

2.1.2　棋手大数据。该部分数据主要按照棋别记录棋手个人档案及成长经历，主要的源数据库记录表有棋手档案表、成长经历表、取得成绩表。

2.1.3　棋谱大数据。该部分数据主要按照棋别记录棋类项目棋谱信息，主要的源数据库记录表为棋谱信息记录表。

2.1.4　裁判大数据。该部分数据主要按照棋别记录棋类运动裁判员及教练员信息，主要的源数据库记录表有裁判员信息表、执裁经历及评价表、教练员信息表、执教经历及评价记录表构成。

2.1.5　机构大数据。该部分数据主要按照棋别记录与棋类运动有关的机构信息，主要的源数据库记录表有棋类协会表、棋院表、棋类项目培训中心或机构表。

2.1.6　赞助商大数据。该部分数据主要按照棋别记录棋类比赛项目的赞助单位信息，主要的源数据记录表有赞助商信息表、历史赞助情况表。

在本招标课题未被专家组审核通过前，版权归作者所有。

2.2 主题应用数据

2.2.1　等级分数据。该部分数据主要记录专业及业余棋手等级分信息，主要的源数据库记录表有棋手周期等级分表。

2.2.2　技术等级数据。该部分数据主要记录专业或业余棋手技术等级历史信息，主要的源数据库记录表有棋手技术等级表、棋手技术等级成长表。

2.3 计算挖掘数据

挖掘数据或数据挖掘（*Data mining*）又译为资料探勘、数据采矿。它是数据库知识发现（英语：Knowledge-Discovery in Databases，简称：KDD)中的一个步骤。数据挖掘一般是指从大量的数据中通过算法搜索隐藏于其中信息的过程。数据挖掘通常与计算机科学有关，并通过统计、在线分析处理、情报检索、机器学习、专家系统和模式识别等诸多方法来实现上述目标。

计算挖掘数据是在棋类运动基础源数据、主题应用数据基础上通过建模分析计算挖掘所产生的数据。棋类运动基础源数据和主题应用数据，是棋类运动大数据平台的基础，基础源数据是海量的原汁原味的未经加工的数据，如何从这些海量的数据中挖掘分析整理计算，变成各种棋类运动从业人员共享利用的信息，这部分工作正是棋类运动大数据平台数据整合处理子平台及数据挖掘子平台来完成的，他们从这些枯燥无味的信息中分析、建模、提取、加工、计算后存贮有用的信息，提供给平台各类人员使用。

譬如棋类运动大数据平台通过计算存储棋谱数据与软件吻合度指标，可以方便查出围棋第一人柯洁、象棋第一人王天一最近 10 盘比赛棋谱与软件的吻合度平均数据，这个指标是判断一名棋手近期棋路是否“很专业”的一个重要指标，也是我们平时选拔棋手需要考量的一个重要指标。

再比如，一个企业要组队参加象棋甲级联赛，临时选调 4 名业余高手组队，我们可以方便地通过大数据平台找到业余棋手中近期与专业棋手对局胜率最高的几名棋手，再综合考虑其他因素优选其中 4 名，而不是一拍脑袋就直接聘请业余象棋界的四驾马车来征战。

计算挖掘数据是棋类运动大数据平台的精华，它们遵循一定的数学模型和算法分析计算后产生，提供给平台相关机构和人员使用，计算挖掘数据是棋类运动管理机构、棋手日常做决策的重要依据。

三、棋类运动大数据平台管理

棋类运动大数据平台面向全球棋类运动所有相关人员设计，服务各级别棋类运动管理机构、棋类运动教育培训机构、棋手、裁判员、教练员、媒体人员等。棋类运动大数据平台的用户身份的不同，决定了其能够访问利用平台资源的广度和深度，本篇概要地阐述了棋类运动大数据的平台管理方案和要点。

在本招标课题未被专家组审核通过前，版权归作者所有。

棋类运动大数据平台基础数据管理

3.1 用户管理

负责棋类运动大数据平台的用户管理，具体包括各类用户账户的设立管理、用户信息的维护及变更管理、用户账户密码的管理，用户帐号的锁定或删除管理等。

3.2 权限管理

负责棋类运动大数据平台的权限管理，具体包括授予各类用户账户平台数据访问的内容和范围，用户权限的变更管理等。

3.3 角色管理

负责棋类运动大数据平台的角色管理，具体包括授予各类用户账户平台的各种预定义角色身份（角色即一定权限范围的用户的统称），及用户角色的变更管理等。

3.4 数据管理

负责棋类运动大数据平台的基础数据维护管理，具体包括授权各类用户维护大数据平台基础数据的范围，基础数据的增加、修改、删除权限管理，基础数据维护的期限管理等。

3.5 安全管理

负责棋类运动大数据平台的基础数据安全管理，主要是指采用现代密码算法对

在本招标课题未被专家组审核通过前，版权归作者所有。

数据进行主动保护，如数据保密、数据完整性、双向身份认证等，二是采用现代信息存储手段对数据进行主动防护确保数据安全。

3.6 系统管理

负责棋类运动大数据平台的系统管理，具体包括大数据平台的各类基础数据读写保护管理、基础数据字典管理、基础数据备份管理等。

四、基于棋类运动大数据平台的棋手选拔

在棋类运动中，如何科学地选拔棋手是决定日后能否取得成功的关键。如何科学选拔棋手这个问题在大数据时代来临前，几乎所有的选拔方式都如出一辙。要么单独搞个比赛，按比赛成绩选拔棋手，没有时间搞比赛的，就直接拍脑袋，从历史成绩好的棋手中选拔。这已经成为棋界的定律，一种通行的做法。确实在没有更好的方法出现前，这也许是最公平合理的做法，在大数据时代，我们讲做任何决策都不能凭经验凭感情用事，在大数据时代最讲究的是决策的科学性，你的决策是否有数据做支持，这点至关重要。对于棋类运动而言，使用大数据做支撑选拔棋手具有十分明显的优势，主要表现在以下几个方面。

首先它可以防止感情用事、人情选人。就业余棋界来说，不可否认的是全国一些地方棋协在选拔棋手时，大多存在考虑人情要超过对棋手状态的考量，往往与棋协主要领导关系走的比较近的实力派棋手获得比赛的机会要更多。

其次它可以让决策更科学更合理。用大数据平台做支撑选出的棋手是最科学合理的，这是因为，大数据环境下选拔棋手是从多维度的数据做分析，用多维度多视角分析选择棋手，这样选拔出的棋手通常是经得起比赛检验的。

棋类运动大数据平台选拔棋手的主要指标有棋手年龄、历史战绩、最近身体状态、最近竞技状态、选拔赛成绩、最近棋谱软件吻合度等多项指标组成，综合考量后按每项K值（关键值）计算总得分按高低排名选择棋手。这样综合选出的棋手比从一个维度选择的要科学的多。

再举个简单的例子，某个专业棋院要选拔一批小队员进入棋院梯队。对于小棋手年龄这个指标来说，同等棋力年龄越大K值越小。从比赛成绩这个指标来看，同等名次全国少年赛的比赛成绩就要比其他比赛的成绩K值要高。从小棋手的教练（指导老师）来说，专业棋手做为他教练的就要比业余棋手做教练的K值要高。

最终平台是将各种指标转化成根据给定的数学模型并计算筛选出棋手的总得分排名，也就是说将这个选拔过程转化为一个数学问题，用量化的指标综合评价一名棋手，无疑这种选拔方式是最科学合理的。

五、基于棋类运动大数据平台的棋手训练

在棋类运动中，棋手训练是非常重要的一环，日常训练质量的高低直接影响棋手备

在本招标课题未被专家组审核通过前，版权归作者所有。

战比赛，如何科学地指导棋手训练，至今没有一个统一的标准体系，历来都是各家有各家的方法。随着人工智能时代的开启，棋手在平时训练中更多地喜欢用软件作为工具辅助训练，比如象棋第一人王天一，在成名前的相当长的时间里都是在用软件作为假想对手训练，这使他受益匪浅，23 岁便获得了全国象棋个人锦标赛冠军。在基于棋类运动大数据平台下，棋手可选的训练方法有很多，具体来说主要有以下几种。

5.1 人机实战训练法

人机实战训练法也称人机大战训练法，就是棋手与平台各类对弈软件直接 PK 对弈，棋手从低级别的软件到人工智能软件，逐级提升自己的水平，比如台湾软体公司早期开发的一款象棋人机对弈软件-“将族”游戏，这种训练方法大都棋手都曾采用过。

5.2 软件模拟训练法

软件模拟训练法也称大数据训练法，这种方法以大数据平台为依托，训练中完全以大数据的棋谱作拆棋数据研究基础，辅以模拟训练软件做工具，训练中能将历史棋谱中的优劣招拆解的非常透彻，为棋手针对性地研究目标对手提供了非常好的训练方法。

5.3 混合训练法

混合训练法为上述两种方法的有机结合，这种方法最适合创新招法的研究，它结合现有大数据和人工智能软件的超强分析计算能力，研究探索棋局中可能的，先前棋手在比赛中从未使用过的，新的变化或变招。

六、基于棋类运动大数据平台的比赛分析

棋类运动大数据平台的建立，为棋手提供了一个非常好的训练比赛辅助系统，使用棋类运动大数据平台集成的各种辅助软件，棋手除了可以赛前针对性的训练，还可以在比赛后利用平台对比赛棋谱进行系统的分析总结，比赛分析主要分为棋谱分析、辅助拆棋、创新招法研究等。

6.1 比赛棋谱分析

使用棋类运动大数据平台，棋手可以很方便将比赛棋谱导入各种智能分析软件，由软件输出详细的棋谱分析报告，该报告除了会详细为每步行棋质量打分，每步行棋后双方优劣势评分，还可以给出全盘棋谱综合评分数据、双方行棋软件吻合度指标等数据，报告的最后列出全盘关键步骤或招法。此外分析软件还可以搜索棋谱大数据将历史同谱资料一一列出，为棋手赛后综合复盘研究提供数据支持。

6.2 软件辅助拆棋

使用棋类运动大数据平台，棋手将比赛棋谱导入各种智能分析软件，使用平台提供的拆棋功能，系统就会将本局棋谱中关键招法一一列出，并列出每一步关键招法的可能

在本招标课题未被专家组审核通过前，版权归作者所有。

变化及该变化的优劣评分，使用此功能可以大大地缩短棋手赛后拆棋复盘时间，提高棋手赛后复盘的效率。

6.3 创新招法研究

使用棋类运动大数据平台，棋手将比赛棋谱导入各种智能分析软件，使用平台提供的创新招法研究功能，系统就会将本局棋谱中可以使用创新招法的关键步骤列出，并列出每一步关键步骤的创新招法及后续招法的优劣评分，此功能利用现有大数据和人工智能软件的超强分析计算能力，研究探索棋局中可能的，先前棋手在比赛中从未使用过的，新的变化或变招。

七、基于棋类运动大数据平台的综合信息服务

棋类运动大数据平台的建立，不仅提供了一个非常好的选拔、训练、比赛、分析和研究环境，同时它也可以通过接口面向公众，提供综合信息查询服务。这里的公众不仅指棋类运动管理培训机构、棋手、裁判员、教练员、媒体人员，所有想获取棋类运动数据及信息的自然人都可以成为棋类运动大数据平台的用户，平台在他们的用户角色权限范围内提供相关信息服务。

7.1 信息检索服务

信息检索服务也称信息查询服务，用户通过棋类运动大数据平台应用层信息检索模块，查询获取棋类运动大数据平台有关的信息。

譬如，在通用信息检索棋手一栏里，用户输入姓名“柯洁”，平台会自动生成柯杰的个人档案、成长经历、取得成绩、所有比赛棋谱等所有与之相关的维度信息表，你可以从中选择感兴趣的维度输出结果。

譬如，在通用信息检索裁判一栏里，用户输入姓名“栾晓峰”，平台会自动生成栾晓峰的个人档案、成长经历、执裁经历、执裁评价等所有与之相关的维度信息表，你可以从中选择感兴趣的维度输出结果。

譬如，在通用信息检索比赛一栏里输入“全国象棋个人锦标赛”，平台会自动生成全国象棋个人锦标赛历届举办时间地点、历届成绩排名、历届男女冠军、历届赞助单位、历届裁判员名单等与之相关的所有维度表，你可以从中选择需要的信息维度输出结果。

譬如，在通用信息检索布局一栏里输入“西西里防御 ”，平台会自动生成国际象棋西西里防御基本概况、发展情况、开局介绍、主要变例等多维度信息，你可以从中选择感兴趣的维度输出结果。

譬如，在通用信息检索机构一栏里输入“中国棋院杭州分院 ”，平台会自动生成中国棋院杭州分院基本概况、现任领导、组织机构、工商信息等多维度信息，你可以从中选择感兴趣的维度输出结果。

在本招标课题未被专家组审核通过前，版权归作者所有。

7.2 智能检索服务

智能检索服务也称组合查询服务，它是根据指定的输入组合关键信息，在棋类运动大数据平台中匹配有用的信息输出，也可简称多维信息检索。

譬如，在通用信息检索棋手一栏里输入“许银川”，同时在通用信息检索比赛一栏里输入“全国象棋个人锦标赛”，那么平台就会自动生成许银川的个人档案、成长经历、参加历届全国象棋个人赛的情况、取得冠军的届数与胜率、全国象棋个人赛棋谱等维度信息，你可以从中选择感兴趣的维度输出结果。

譬如，在通用信息检索棋别一栏里输入“围棋”，同时在通用信息检索赞助商一栏里输入“阿里巴巴”，那么平台就会自动生成阿里巴巴的工商注册信息、赞助围棋比赛的情况、赞助围棋比赛的总奖金、赞助围棋比赛的名称等多维度系信息，你可以从中选择感兴趣的维度输出结果。

譬如，在通用信息检索棋手一栏里输入“胡荣华”，同时在通用信息检索“布局”一栏里输入反宫马，那么平台就会自动生成胡荣华的个人档案、成长经历、取得成绩、反攻马的布局简介、发展变化、胡荣华使用反宫马布局的所有先后手对局胜率、胡荣华使用反攻马的所有对局棋谱等多维度信息，你可以从中选择感兴趣的维度输出结果。

八、棋类运动大数据平台项目实施建议

棋类运动大数据平台的项目建设是一项系统工程，需要统筹各方面的力量参与才能够完成，项目建设涉及技术、经费、人员、管理等诸多方面的因素，而且项目建设周期一般较长，短则 1-2 年，长则 3-5 年，如何统筹规划建设好棋类运动大数据平台项目，是棋类运动项目管理者首先要面临的一个问题。笔者以自身长期信息化项目实施经验给出如下建议，以供参考。

8.1 棋类运动大数据平台项目实施原则

棋类运动大数据平台的项目应坚持“统筹规划、分布实施、有序推进、注重实效”的项目建设原则。棋类运动大数据平台项目建设周期较长，在项目可行性研究方案通过，在平台需求及功能调研完成后，编制项目建设计划时，一定要坚持分阶段分步骤实施的原则，任何在短期内建成项目的思想都是不切实际的，项目建设要分阶段实施及分步骤验收，尤其注重项目实施的效果，在明确达到预期效果的基础上再稳步推进。

8.2 实施企业资质及经验要求

棋类运动大数据平台的项目实施企业的资质及经验要求如下：

实施企业必须具有优良的信誉及诚信的商业道德；

实施企业必须是在中国境内注册的并具有独立法人资格的企业；且具有软件企业的认定证书及全面的质量保证与控制体系；

在本招标课题未被专家组审核通过前，版权归作者所有。

实施企业必须投标产品原厂商认可的代理资质，并提供代理资质证明材料，同时拥有原厂商针对本项目的授权文件；

实施企业必须具有履行合同所必须的设备和专业技术能力；

实施企业必须在大数据行业内实施项目成功案例不少于三家，并提供成功案例实施证明材料；

实施企业必须保证在近三年内没有实施验收未通过的项目或由实施企业承担责任的重大用户投诉；

实施企业必须符合招标文件中关于投标人的其他规定；

8.3 实施采用招标采购法

棋类运动大数据平台的项目可采用国际上通行的招标采购法选择实施企业。在棋类运动大数据平台的项目需求调研及功能说明书编制完成后，可在国内招投标网上发布招标信息，同时聘请各方面技术专家组成评标小组，从优选的 3-5 家投标企业中综合评分，最后择优选择 1 家实施企业签订项目建设合同。

8.4 棋类运动大数据平台项目经费来源

棋类运动大数据平台的项目建设经费来源有两种方案。第一种是采用国家行政拨款，也就是国营资本独资的运作模式，第二种是引入社会民间资本参股的市场化运作模式。两种方案各有优缺点，笔者倾向于第二种。原因是市场化的运作体制更具有活力，也符合国内经济体制改革总体大方向。不论是采用国有独资还是引入社会民间资本参股的运作模式，棋类运动大数据平台的行业指导和管理单位（或控股单位）还是国家体育总局棋牌运动管理中心（或中国棋院杭州分院）。

九、棋类运动大数据平台的运营与增值服务

9.1 棋类运动大数据平台的运营

棋类运动大数据平台的运营应采用公司化独立法人的运作模式。我们都知道，公司化运作是多元股权结构企业、中小民营企业向现代企业制度迈进的必由之路，也是盈利性事业单位体制转型的方向。事业单位体制转型终极目标是独立核算自负盈亏，从而摆脱长期对政府财政资金补贴的依赖。

中国棋院杭州分院作为中国棋院的下属机构，自 2004 年 1 月 4 日经国家体育总局正式批准成立以来，在中共杭州市委、市人民政府和中国棋院的领导下，在浙江省和杭州市体育局的支持下，以普及围棋和其他棋牌运动项目为主业，立足国内、面向国际、努力打造成国内领先、世界一流的比赛、培训、国际交流“三个中心”。

笔者认为，中国棋院杭州分院的运作模式就是中国棋类运动产业化发展最好的实践形式，它行业上隶属中国棋院管理，承担部分政府的公共体育事业管理职能，近些年来，

在本招标课题未被专家组审核通过前，版权归作者所有。

他们在棋类事业市场化经营方面取得了不俗的成绩，特别是在棋类运动体制机制转换、棋类运动队伍建设、棋手选拔培养、棋类教学培训、棋类运动网络推广上积累了丰富的经验。棋类运动大数据平台项目交由中国棋院杭州分院来牵头建设运营无疑是最合适的。中国棋院杭州分院在棋类运动大数据平台项目建设运营中可适当引入社会民间资本参与进来，让有志于开拓棋类事业的文化公司、科技公司、投资公司参股进来，组成股份制联合企业，终极目标是以良好的经营业绩，将它推入资本市场，通过证券市场的再融资再扩大将棋类运动产业不断做大做强。

9.2 棋类运动大数据平台的增值服务

棋类运动大数据平台的上线运营必将引领棋类事业快速、健康、科学和谐地发展。棋类运动大数据平台的运营需要一个独立的专业的管理技术团队支持，棋类大数据平台在运营过程中会根据用户需求的变化不断更新应用。同时这种服务也是双向的，棋类运动大数据平台在给全球棋类运动相关机构和个人带来高效便捷服务的同时，也会酌情向用户收取一定的服务费用，这就是棋类运动大数据平台产生的增值服务费用。

棋类运动大数据平台的增值服务费用不单单是指向用户收取固定标准的服务费用，一切因棋类运动大数据平台应用带来的新的商机，比如基于棋类运动大数据平台衍生的软件产品销售费用、基于棋类运动大数据平台的棋类项目教学培训费用等都可以纳入增值服务费用范畴，具体来说增值服务主要有以下一些内容构成。

包年服务:棋类运动大数据平台的包年服务，按不同用户角色标准收费。
信息检索:棋类运动大数据平台的综合信息检索服务，按次或包月包年收费。
软件销售:棋类运动大数据平台衍生的软件产品销售费用，按软件产品价格销售。
教学视频:棋类运动大数据平台内的各种棋类教学视频，按次或包月包年收费。
在线课堂:棋类运动大数据平台内的各种在线课堂,按棋别按讲师层级按时间收费。
电子期刊:棋类运动大数据平台内的各种棋类电子期刊，按次或包月包年收费。
VIP 服务:棋类运动大数据平台一对一专享服务，按棋别按老师层级按时间收费。
网络赛场:棋类运动大数据平台网络比赛室，按棋别按次或包月包年收费。

参考文献

【1】“学习中国”推出《习近平与“十三五”十四大战略》系列文章之“国家大数据战略”篇[A].
【2】梅宏. 数据挖掘深度决定大数据应用价值 [N]. 《中国信息化周报》2017-2
【3】林晓立 陈恩红 任皖英. 高维数据的特征提取的算法研究与比较[G].《计算机科学》2003-4
【4】吴朝晖. 人工智能新发展:混合智能理论、方法及应用[R]. 全球人工智能杭州高峰论坛 2017-7
【5】刘知青 吴修竹. 解读 AlphaGo 背后的人工智能技术[J]. 《控制理论与应用》2016-12
【6】梅宏. 对云计算时代软件技术发展的若干思考和实践[R]. 第四届中国云计算大会报告
【7】黄宜华. 基于 Spark 的大规模文本 k-means 并行聚类算法[G]. 江苏，南京大学 2013
【8】袁冬. 基于海量文本的语义构建方法研究[D]. 山东，中国海洋大学 2012
【9】胡贵平. 运用 BW 数据仓库技术，构建企业决策分析系统[J]. 《冶金动力》2015-12

在本招标课题未被专家组审核通过前，版权归作者所有。

作者简介及联系方式

作者简介：胡贵平，男，汉族，1971 年 8 月生，安徽省马鞍山市人，现供职于飞马智科信息技术股份有限公司，主要从事软件研发、数据仓库及基于云平台的大数据系统的分析研究、基于企业 ERP 系统的决策分析系统设计，是一名高级工程师。

作者祖籍河南光山，出生地安徽马鞍山。幼时酷爱象棋，曾幻想当一名职业棋手，后在象棋与学业之间最终选择了后者，但始终保有一颗挚爱象棋之心，工作后下棋少了，看棋赛写棋评文章多了，现经常在广东象棋网（网名：起炮在中宫）撰写棋评文章。也常受邀观摩各地象棋比赛，印象最深的是受到碧桂园杯全国象棋冠军邀请赛组委会邀请，连续两届（第五届和第六届）作为观赛嘉宾赴广州增城碧桂园凤凰城酒店现场观赛。2018 年 10 月 20 日，作者以首届楚河汉界世界棋王赛经典时刻全球征集活动三等奖的获奖者身份，受到组委会邀请亲临荥阳鸿沟现场，实地观摩了第二届楚河汉界世界棋王赛决赛暨中国第二届象棋文化节相关活动。

联系方式：

工作单位：飞马智科信息技术股份有限公司

通讯地址：安徽省马鞍山市雨山区湖南西路 8 号

联系电话：13855578068　　微信号：hugp13855578068

电子信箱：huguiping@263.net

广东象棋网网名：起炮在中宫（诗城棋迷）

在本招标课题未被专家组审核通过前，版权归作者所有。