

关于期货交易所统计信息共享平台的需求研究

作者：陈建平 王柏松

统计信息是期货交易所的重要信息资源，期货交易所的统计信息不仅包括交易所内部的交易、结算、交割、会员、客户管理等主要业务信息，还包括从不同渠道、以不同方式收集的与周边市场、交易品种相关的各类辅助信息。其特点是：内容庞杂，数据量大，可开发利用价值高。但目前，期货交易所统计信息处理特别是日常统计工作自动化程度较低，统计信息共享性较差，既不能满足所内外用户对统计信息日益增长的需求，更不能对统计数据进行深度分析和挖掘。信息开发利用的未来发展趋势是自动化、集成化、产品化，建设统计信息共享平台，是优化信息资源利用、提高信息使用效率的重要手段。为此，从期货交易所信息工作实际出发，结合信息开发利用的未来发展趋势，我们提出建设统计信息共享平台（Statistics Data Information Share Platform）项目的设想，希望在第一阶段实现下述目标：一方面能为期货交易所各业务部门收集、整理、存储、统计、分析各种数据信息提供自动化程度较高的技术支持，另一方面能为期货交易所员工和会员检索所需统计信息提供一个统一的系统平台，最终为逐步建成所级数据仓库探索思路并积累经验。

统计信息共享平台（以下简称 SDISP）本质上是建立基于统计业务的数据集市并采用门户技术来实现信息发布管理的服务系统，以达到信息有序组织和有效共享的目的。即技术系统管理人员利用信息门

户工具建立信息网络平台和数据集市,信息采编人员利用信息分析工具对数据集市进行统计分析以及使用各种 WEB 工具进行信息汇总,信息管理人员利用信息门户工具对各种不同类型信息进行认证发布和权限管理,内外部终端客户使用 WEB 工具按照不同权限访问系统平台中存储的各类信息。

一、总体目标和阶段目标

期货交易所综合信息集成系统最终的目标是以企业数据仓库为基础,综合运用商业智能、人工智能、Web 服务等技术,实现企业运营中有效的信息管理,提供经营分析决策的有力支持手段。

但在第一阶段实施的 SDISP 项目中,我们建议主要应实现以下几方面的功能:

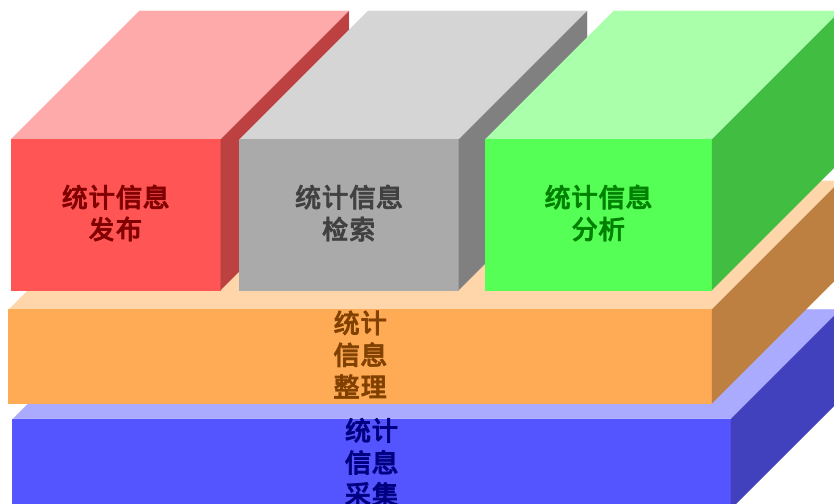
- 1、自动采集部分统计源数据;
- 2、统计信息集中存储与管理;
- 3、自动生成并分发日常统计报表;
- 4、自定义统计图表。基于统计源数据,用户可自定义生成部分统计报表或图形;
- 5、统计报表的条件检索。

二、业务需求分析

(一) 业务模块划分

SDISP 系统从统计业务上可以实现以下五个方面的功能,如图一所示:

图一 SDISP 系统的业务需求模块



1、 统计信息采集

统计信息采集的主要任务是实现与信息管理部门现行系统的接口，从信息管理部门的各种数据源中获取相关统计数据，这些数据源包括：

业务系统统计数据 SDISP 系统每日收盘后从业务系统中获得，以高度结构化的数据表的形式存放。包括交易、结算、交割、会员、客户等按日汇总的统计数据。

信息终端和网站统计数据 包括国内外知名信息服务公司终端和国内外知名交易所网站公布的相关上市品种行情与成交统计数据等。

通过电子文档、电子邮件等方式采集到的国际国内宏观经济指标、期货交易所会员、客户统计数据以及相关上市品种现货市场统计数据。

信息机自动采集的统计数据，这些数据主要来自有相互交换协议的周边其它金融市场或国内外其它交易所的相关数据信

息。

采集到的这些统计信息将存放到临时数据库中，等待信息整理步骤进行整理分类。

2、统计信息整理

信息整理的主要任务是对信息采集步骤中获取到的有关统计数据进行处理、清洗、转换、分类，并最终加载到 SDISP 系统数据库中。SDISP 系统建成后该部分工作要达到结构化数据自动清洗、转换和加载的要求。

3、统计信息发布

统计信息发布的主要任务是将 SDISP 中的统计数据以多种数据格式并通过多种途径如：交易所网站、用户 WEB 终端、电子邮件等对外发布。发布的对象包括交易所内部各部门，政府主管机关，会员公司，其他相关金融机构和统计信息用户等。

4、统计信息检索

信息检索主要完成对 SDISP 系统的查询功能，允许 WEB 用户按照特定的条件查询需要的统计信息。对于结构化的数据，要求能够根据业务需要制定固定查询功能，并能够根据信息管理员的临时需要进行自定义的简单查询，最终以报表或图形的方式进行展现。

5、统计信息分析

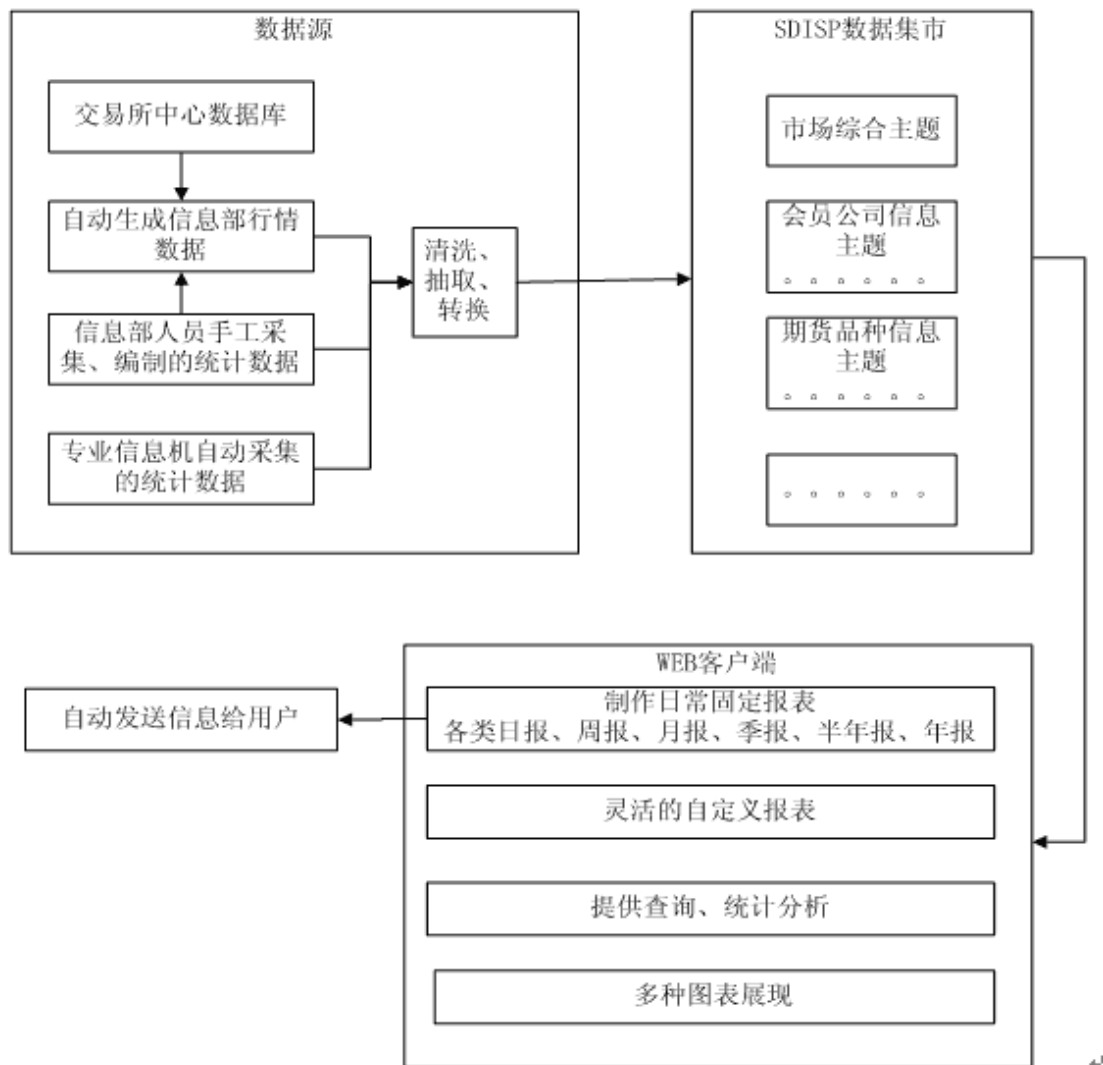
上述四个步骤完成后，接下来要做的就是更进一步的统计分析工作，也就是通常所说的数据挖掘。通过使用各种数据挖掘的工具，建立不同的统计业务分析模型，从而实现对统计信息的进一步开发利用。

用。

（二）日常统计业务流程

SDISP 系统的日常统计业务流程如下图所示：

图二 SDISP 系统基本业务流程



（三）系统用户类型

SDISP 系统的日常使用中，包含了以下一些类型的用户：

系统管理员：使用 ETL 工具或手工编程方式对信息的自动收

集系统进行维护和管理，负责日常的从业务系统中采集、清洗数据的任务；对数据集市的数据结构进行日常维护；维护内部站点系统结构及监控系统运行。

信息分析员：使用分析工具生成 OLAP 分析报表，并将其发布到合适的信息采编人员或用户。

信息采编员：使用内容管理工具，对不同的信息内容，如统计报表，分析图表、外部文件等，进行逻辑组织，分类和编辑，生成信息平台所需要的内容素材。

信息管理员：使用内容管理工具，建立用户权限，信息频道组织以及设定工作流程，负责内容的最终审定和发布。

终端用户：使用 WEB 方式访问各种统计信息资源。

三、技术需求分析

作为期货交易所数据仓库的概念验证，SDISP 系统应该采用数据仓库的理念进行系统构架设计。SDISP 系统必须针对数据仓库的一些关键特性进行测试，为今后期货交易所建设所级数据仓库打通道路、积累经验。这些关键特性包括：共享、权威、高效、增值、易用、开放、安全、统一、扩展、智能，它们是一个成功的数据仓库所必须具备的共同特征。作为数据仓库的前期准备，考虑到投资和工期的限制，SDISP 系统进行技术构架时目前可主要侧重于其中的共享、权威、高效、增值、安全、扩展等六个关键特性的规划，而其他特性应进行必要的技术论证。总体而言，SDISP 系统建设的技術目标至少应满足：

1、系统具有完整性、准确性、及时性

SDISP 系统是数据处理十分频繁，数据交换吞吐量很大的集成系统，数据处理的速​​度直接关系到信息共享平台的正常使用。

因此，在客观条件下，SDISP 系统必须在大数据量的情况下同时保持快速的实时响应能力，以保证系统的通畅运行。为了将来建立所级数据仓库，必须保证系统数据源的完整、准确、及时性。

2、系统既有稳定性又有健壮性

SDISP 系统必须能够做到连续、安全、无故障、不间断运行作业。在充分考虑资金投入效益的基础上，为了防止系统某一环节出现故障导致崩溃，保证系统在最短的时间内恢复正常，将损失降低到最低限度，在条件许可情况下应该考虑采用成熟的热切换备份恢复技术。

另外，SDISP 系统数据存贮功能必须强大，因为每天都有大量的业务数据、分析数据产生，所以，SDISP 系统和相应的中央数据库管理系统必须具备极高的数据存储与管理能力。

3、系统具有开放性和可扩展性

随着交易所业务的迅速发展，交易品种、数据来源、服务方式和其它各种创新业务必将不断产生与演化，这要求整个系统必须具备良好的开放性与可扩展性。为此 SDISP 系统采用结构化、模块化、层次化的设计原则，设计灵活的外部接口和内部接口。考虑到建立更全面的更高层次的其他外接系统的需要，本系统应具有良好的开发接口。

4、系统具有先进性和实用性

先进性和实用性的辩证关系要求 SDISP 系统在采用成熟技术、

稳定架构的基础上，充分考虑技术的先进性，使整个系统既符合信息技术发展潮流，同时又具有很高的实用性，确保系统在可预见的将来的可用性。

5、系统具有安全性

SDISP 系统未来将扩展为一个大型的网络应用系统，并将提供包括所有固定用户在内的网络接入和访问能力，因此，这将对安全性提出很高的要求，保护系统和系统信息数据不被非法用户侵入是系统设计和建设所要考虑的重要因素。因此在系统分析与设计的各个阶段必须从各个角度、各个方面充分考虑整个系统的安全性，另外系统中所有的操作必须绝对留痕并规范管理。

6、系统具有易用性

SDISP 系统作为一个人机交互界面密集型的 IT 应用系统，在操作流程与界面设计方面应参考目前已有的网络应用系统，总体上应达到以下要求：界面友好、美观、实用，流程清晰合理、符合现有工作流程。