

Ю.М.Лихачев, Д.Н.Тогобицкая, А.Ф.Хамхотько

СИСТЕМНЫЕ СРЕДСТВА ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ НАУКИ О РАСПЛАВАХ.

Представлен ретроспективный анализ информационного обеспечения металлургии расплавов. Обосновано применение систем управления базами данных для решения задач прогнозирования свойств расплавов. Намечены пути расширения сферы применения баз данных за счет адаптации программного обеспечения к требованиям исследователя.

Современное состояние вопроса.

Фундаментальная и прикладная наука о металлургических расплавах накопила огромный объем числовых и описательных данных характеризующих их поведение во всех металлургических переделах. Технологические параметры производства металлопродукции накоплены в различных средствах хранения (от бумажных до баз данных на ЭВМ) и продолжают возрастать с каждым днем. Вычислительная техника имеет средства, позволяющие не только рационально хранить такие данные, но и удобно оперировать ими, – это базы данных.

Базы данных разделяются по цели их использования : числовая информация которая однородна по глубине, объединяется в таблицы и применяемый поисковый алгоритм учитывает именно это свойство и дает адекватный результат на выходе; документальная или описательная часть хранится в текстовом виде, что требует для доступа к ней выполнения индексации групп текста, а это при современном развитии информатики не является тривиальной задачей. Однако при правильном структурировании информационного фонда и учете предметной области выделение смысловых групп текста упрощается, а его формализация ложится на подготовленного интерпретатора–предметника.

Постановка задачи. При создании информационного фонда в 90–х годах прошлого столетия при ведущем участии ИЧМ для систематизации межотраслевого банка данных «Металлургия» были заложены принципы и методология создания сложно–структурированных мультибазовых документально–фактографических данных [1–6]. Информационную среду исследований восстановления металла составили базы фундаментальных данных о свойствах железорудных материалов «Шихта», шлаковых расплавов «Шлак», термодинамических данных «Шлак–Металл–Газ», железорудных материалов «Ironmat», железорудного сырья «Iron», термодинамических свойств окислов «Termet». Эти базы строились по документально–фактографическому принципу, что совместно с индексацией текстовых полей позволяет использовать их неподготовленному пользователю.

Анализ последних достижений.

Начало создания баз данных пришлось на первый этап развития персональной вычислительной техники, что обусловило использование ранних версий операционных систем. Такой подход, хотя и позволял решать стоящие вопросы по информационным проблемам обработки баз данных (хранение информации, достаточно быстрый доступ к данным, построение приемлемого интерфейса), однако с развитием технических и системных средств вычислительной техники обусловил дальнейшее развитие новых программных средств ведения баз данных с учетом огромного потенциала последнего поколения разработок как технических (с существенно возросшим быстродействием, огромным объемом хранимой информации, расширенными возможностями обработки графики), так и с системным программным обеспечением (существенно возросшая мощность операционных систем, улучшенный программный сервис по созданию интерфейса пользователя, возможность обработки графической информации).

Для долговременного хранения и оперативной обработки больших объемов данных используются магнитные накопители, которые осуществляют последовательную обработку данных. Это стимулирует использование методов оптимизации для записи и считывания необходимых порций информации на магнитные накопители. Существующие системы обработки физической информации работают как с последовательной структурой (реляционные СУБД /Системы Управления Базами Данных/), так и с кусочно–последовательной /страничной/ (SQL СУБД). В обоих случаях используется процедура индексации для ускорения поиска необходимого фрагмента информации.

Последовательный способ хранения довольно просто реализуется, однако требует увеличенного объема физической копии, что существенно усложняет реализацию корректировки фрагментов данных на носителях. Страничный способ хранения более сложный в реализации, однако позволяет оптимально размещать физические порции данных, не критичен к корректировке фрагментов и имеет существенно большую скорость обработки по сравнению с последовательным. Использование страничного способа хранения позволяет не только легко выполнять корректировку фрагментов, но и создавать мульти базовые структуры со сложным (разветвленным) внутренним строением. Страничный способ позволяет легко хранить графические образы в одном файле с другой информацией, что существенно при построении файл–серверов.

Для расширения диапазона пользователей баз данных необходимо предоставлять им удобный сервис для поиска и обработки хранимых данных. Человеческий фактор требует интеллектуализации программного обеспечения для слияния возможностей пользователя с работой СУБД. Одно из главных направлений такого слияния является разносторонняя индексация информационного фонда. В случае документальных баз дан-

ных индексация является органическим элементом структуры фонда, что позволяет для больших объемов информации обеспечить произвольный поиск за короткое время.

Главной целью индексации в документальных базах является адекватность результатов запроса поисковому предписанию. Использование многоиндексного поиска для получения релевантных результатов не только увеличивает скорость поиска, но и снижает уровень шума в полученных выборках. Правильный выбор набора индексов позволяет для документальных баз выполнять полное покрытие выборкой исследуемый информационный материал.

Индексация в фактографических базах не столь актуальна как в документальных, однако и здесь сбрасывание ее со счетов ведет к увеличению реакции системы на запрос, кажущемуся зависанию запроса при большом объеме физического фрагмента исследуемой информации, снижению интенсивности охвата покрытия информационного фонда для множественного критерия выборки в случае сложных запросов (человеческий фактор).

Эффективность обработка и интерпретация числовой информации увеличивается при визуальном представлении. Анализ визуального образа информации на рис.1 позволяет сделать заключение о характере рассматриваемой области. Фрагментация исследуемой области на части может повысить достоверность прогнозной модели и обосновать вид аппроксимирующей поверхности.

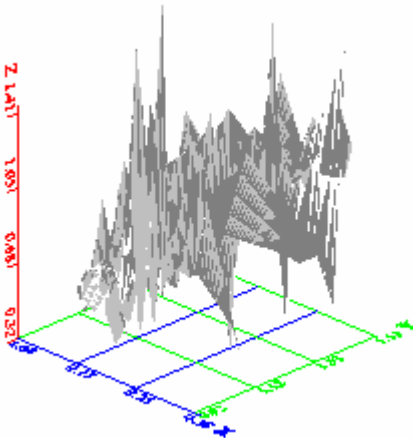


Рис.1. Выборка из технологической базы данных о работе доменной печи. (X – Al_2O_3/MgO , Y – CaO/SiO_2 , Z – [Si])

Выделение характерных областей легче осуществить при графической визуализации объекта, чем при абстрактном его представлении. В этом случае присутствие шума в некоторых фрагментах данных проще обосновать когнитологическим

опытом исследователя. Удалив фрагменты, искажающие поведение исследуемого объекта, исследователь снова может выполнить анализ и таким последовательным промыванием «золотого» песка информационного фонда получить ту нужную картину, которая удовлетворяет выбранным критериям качества.

Изложение основных материалов исследования.

СУБД предназначены для широкого спектра работ с базами данных (первоначальное накопление, корректировка существующего фонда, построение и удаление индексов, организация запросов, выгрузка информации из базы). Информационно-поисковые системы (ИПС) обладают меньшими возможностями, однако в них структурно включаются средства поиска документальных фрагментов, что делает их очень удобным средством работы с документальными базами. ИПС обычно разрабатываются по некоторому целевому признаку, который учитывает структуру хранимого информационного фонда и поэтому существенно приближает пользователя к решению поставленных задач.

Металлургическая наука (подобно другим наукам) проводит эксперименты, публикует результаты и формирует выводы о свойствах или явлениях материалов и процессов. Такой подход в основном представляется в виде документально-фактографической информации. Накопление такой информации в базах данных изначально ассоциируется с ИПС.

ИПС для науки о расплавах должна обеспечивать функционирование взаимосвязанных частей, призванных удовлетворить металлурга-исследователя информацией о необходимых составах и свойствах различных расплавов.

Создание ИПС [1] для функционирования баз данных включает формализацию предметной области [2]. Здесь необходимо выделить существенные классы описываемых параметров с целью минимизации типов структур данных. Такая минимизация позволит четко представить структуру хранимого фонда и дифференцировать различия при составлении запросов.

Второй этап включает генерацию паспорта данных [3–4]. Происходит уточнение понятий этапа формализации, выработки подходов для машинного хранения информации (построение удобных для машинной обработки названий, создание полей для индексации, формирование вида предметной области для машинного хранения).

Третий этап включает ввод информации в базу данных, выявление ошибок подготовки данных и уточнение адекватности введенных данных предметной области [5–6].

Четвертый этап позволяет осуществлять проведение запросов к базе данных, получение на них ответов, вывод результатов запросов для дальнейшей обработки.

Ниже представлен паспорт экспериментальных данных для базы данных «МЕТАЛЛ – ШЛАК – ГАЗ».

ОСНОВНЫЕ СВЕДЕНИЯ.

СИСТЕМА ФАЗ: Металл, шлак.

ОБЪЕКТ: [S], неметаллические включения, [O], [N].

СОСТАВ ФАЗ: Металл: $Fe, C, Mn, Si, Cr, S, O, N$; шлак: $CaO - Al_2O_3 - SiO_2$; элементный: $Ca - Al - Si - O$.

АВТОРЫ: Ершов Г.С., Михайликов С.В.

БИБЛИОГРАФИЯ: О рафинирующей способности шлаковых распла-
вов системы $CaO - Al_2O_3 - SiO_2$ //Изв.АН СССР. Металлы. 1965 – N2 –
С.36–41.

РЕФЕРАТ: Изучена рафинирующая способность синтетических шла-
ков системы $CaO - Al_2O_3 - SiO_2$ при обработке стали в ковше. Установле-
но, что между степенью десульфурации стали и содержанием кремнезема
в синтетических шлаках (в интервале от 0,4 до 20% SiO_2) связи не существ-
вует, поэтому можно использовать материалы с повышенным содержи-
мом SiO_2 для выплавки синтетических шлаков. Обработка стали синтетиче-
скими шлаками в ковше привела к значительному снижению загрязненно-
сти ее оксидными и силикатными неметаллическими включениями. Кон-
центрация кислорода в стали при рафинировании ее в ковше синтетиче-
скими шлаками заметно снижается, а азота – повышается, так как шлаки
выплавляли в восстановительных условиях, что способствовало растворе-
нию в них азота в больших количествах.

РАСШИРИТЕЛЬ: Приведены составы шлаков и составы стали до и
после обработки их шлаками в ковше. Показано изменение содержания
серы, неметаллических включений кислорода и азота в сталях.

РЕФЕРЕНТ: Хамхотько А.Ф.

ХАРАКТЕРИСТИКА ЭКСПЕРИМЕНТА.

ХАРАКТЕРИСТИКА УСТАНОВКИ: Нагреватель – индукционная
печь – для расплавления металла; печь сопротивления с угольным нагре-
вателем – для выплавки шлаков; рабочее пространство – нет данных; ат-
мосфера – нет данных; тигель – кислый – для выплавки сталей плавки NN
1–17, магнетитовый – для выплавки сталей плавки NN 18–25; термометр –
вольфрам – ренийевая термопара.

ХАРАКТЕРИСТИКА ОБРАЗЦА: Состав металла $C - 0,1-1,42; Mn - 0,07-1,91; Si - 0,09-0,83; Cr - 0-1,73; S - 0,008-0,04$; состав шлака – $CaO - 28,5-57,6; Al_2O_3 - 2-55; SiO_2 - 0,4-43,4$; состав газа – нет данных; мате-
риалы – металл – армко–железо с присадками ферросплавов; шлак – хи-
мически чистые реагенты; газ – нет данных; подготовка образца – нет
данных; изменение температуры – 1700 град.С перед выпуском шлака.

ХАРАКТЕРИСТИКА МЕТОДИКИ: Метод – выплавка стали, скачи-
вание шлака с ее поверхности перед выпуском плавки, слив с высоты 2
метра в ковш с жидким синтетическим шлаком в количестве 6% от веса
металла; методика отбора проб – отбор литых проб металла для определе-
ния серы, газов и неметаллических включений до и после обработки стали
в ковше; анализ образца – по химическому анализу металла и шлака и
содержанию газов – нет данных; по количеству и составу неметалличе-
ских включений – методом электролитического растворения.

Востребованность базы данных зависит от потребительских качеств: достоверности хранимой информации, полноты охвата предметной области, удобства доступа к данным, дружественного интерфейса пользователя, – определяющих свойства баз и обеспечивающих их успешное функционирование. Мощный и дружественный интерфейс является существенной частью разработки программного обеспечения базы данных. Дружественность интерфейса включает: настройку параметров для каждого пользователя, подсказку при работе с базой, понятное представление действий при работе, отсутствие непонятных пользователю реакций системы на запрос, предупреждение о необходимости дальнейших действий.

Рассмотренные вопросы по информационному обеспечению дают представление о целях и возможностям обработки больших объемов данных. Привлечение информации из баз данных для построения прогнозных моделей, выполнение на их основе оценки свойств материалов – несомненно мощный инструмент в руках исследователя. Однако боязнь программной обработки у рядового пользователя и частичная его неподготовленность сдерживает широкое распространение создаваемых информационных фондов. В этом плане пропаганда идей компьютерной обработки накопленных числовых данных, создание максимально приближенных к пользователю программных средств является одной из приоритетных задач информационного обеспечения металлургической науки.

На рисунках 2–4 представлены фрагменты работы с базой.

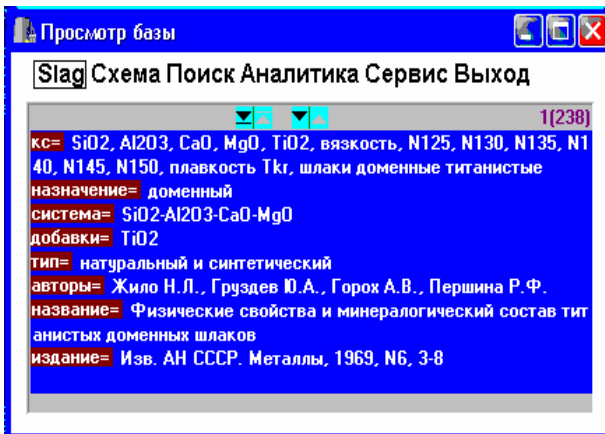


Рис.2.Видеокадр просмотра базы данных.

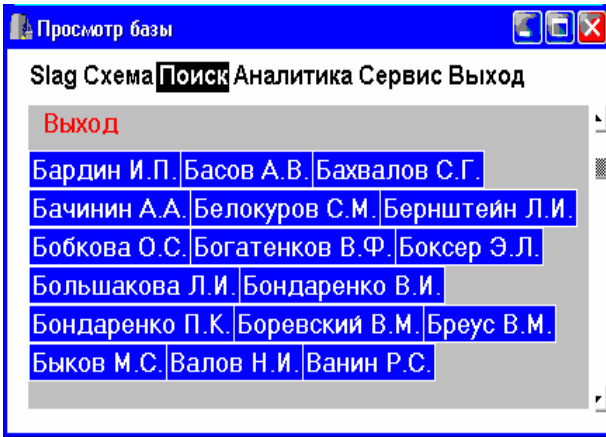


Рис.3.Фрагмент за-проса к базе по авто-грам.

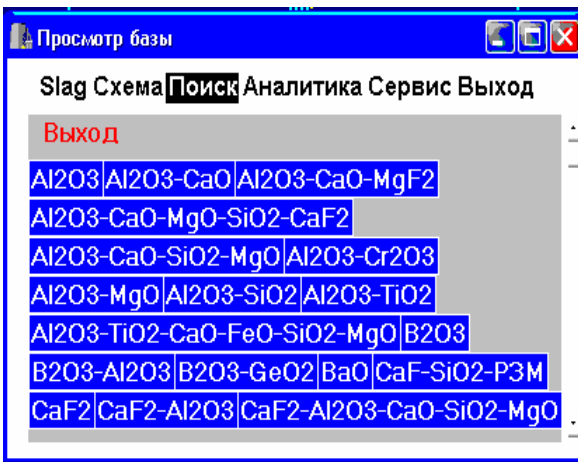


Рис.4.Фрагмент запроса к базе по системе (хи-мический состав).

Выводы. Таким образом, развитие информационного обеспечения для металлургической науки является перспективным как для интеллектуализации программных средств, так и адаптации к нуждам пользователя–металлурга. Охват всей предметной области базами данных будет способствовать ускорению получения прогнозных свойств там, где они были недоступны из-за зашумленности, разреженности, огромного объема исходных данных. Разработка дружественного программного обеспечения позволит существенно расширить сферы применения информационных средств в металлургии, а актуализация накопленного информационного фонда о расплавах выведет на новый уровень решение вопроса по уточнению существующих и разработки новых прогнозных моделей о свойствах исходной продукции, металлургических расплавов, технологий производства.

1. *Базы физико–химических и технологических для оптимизации металлургических технологий.* Всесоюзное совещание. Днепропетровск: ИЧМ. Минмет СССР. –1988. –282с.; –Курган: КМИ, 1990. – 232 с.
2. *Жмойдин Г.И., Тогобицкая Д.Н.* Проблемы информационного обеспечения теоретической и прикладной металлургии // Изв. АН СССР. Металлы. – 1991. –№ 3. –С.218–223.
3. *О паспортизации* экспериментальных материалов для банка данных «Металлургия» / Г.И.Жмойдин, Э.В.Приходько, Д.Н.Тогобицкая и др. // Изв.ВУЗов. Черная металлургия.–1988.–С.136–139.
4. *Тогобицкая Д.Н., Хамхотько А.Ф., Лихачев Ю.М.* Оптимизация металлургических технологий и концепция создания информационно–интеллектуальных систем: // Сб.н.т. Фундаментальные и прикладные проблемы черной металлургии. – Киев: Наукова Думка, 1995. – С.242–249.
5. *Приходько Э.В., Тогобицкая Д.Н.* Методология создания базы знаний о свойствах сталей и сплавов. // Металознавство та обробка металів.–Киев.–1996.–№3.– С.50–56.
6. *Приходько Э.В., Тогобицкая Д.Н.* Базы физико–химических и технологических данных для создания информационных технологий в металлургии. // Металлургия и горнорудная промышленность.–1999.–№3.– С.17–21.

Статья рекомендована к печати д.т.н., проф. Э.В.Приходько