

16. Pilipenko A.M., Biryukov V.N. Gibridnye metody resheniya obyknovennykh differentsial'nykh uravneniy zhestkikh i/ili kolebatel'nykh tsepey [Hybrid methods for solving ordinary differential equations of rigid and/or oscillatory circuits], *Radiotekhnika* [Radio Engineering], 2011, No. 1, pp. 11-15.
17. Pilipenko A.M. Gibridnye metody vysokogo poriyadka tochnosti dlya chislennogo analiza vo vremennoy oblasti zhestkikh i kolebatel'nykh tsepey [Hybrid methods of high order of accuracy for numerical analysis in the time domain of rigid and oscillatory circuits], *Modelirovanie, optimizatsiya i informatsionnye tekhnologii* [Modeling, Optimization and Information Technology], 2017, No. 3 (18). Available at: <https://moitvvt.ru/ru/journal/pdf?id=367> (accessed 24 August 2025).
18. Khayrer E., Vanner G. Reshenie obyknovennykh differentsial'nykh uravneniy: Zhestkie i differentsial'no-algebraicheskie zadachi [Solution of ordinary differential equations: Stiff and differential-algebraic problems], transl. from the second Engl. ed. by E.L. Starostina i dr., ed. by S.S. Filippova. Moscow: Mir, 1999, 685 p.
19. Rakitskiy Yu.V., Ustinov S.M., Chernorutskiy I.G. Chislennye metody resheniya zhestkikh system [Numerical methods for solving stiff systems]. Moscow: Nauka, 1979, 208 p.
20. Maffezzoni P., Codecasa L., D'Amore D. Time-domain simulation of nonlinear circuits through implicit Runge-Kutta methods, *IEEE Transactions. Circuits and Systems*, 2007, Vol. 54, No. 2, pp. 391-400.

Пилипенко Александр Михайлович – Южный федеральный университет; e-mail: ampilipenko@sfedu.ru; г. Таганрог, Россия; тел.: +78634371632; кафедра теоретических основ радиотехники; зав. кафедрой; к.т.н.; доцент.

Pilipenko Alexandr Mikhailovich – Southern Federal University; e-mail: ampilipenko@sfedu.ru; Taganrog, Russia; phone: +78634371632; the Department of Fundamentals of Radio Engineering; head of the department; cand. of eng. sc.; associate professor.

УДК 004.89 + 004.042

DOI 10.18522/2311-3103-2026-1-122-147

В.В. Гапочка, Е.Е. Полупанова

ОБЗОР ТЕХНОЛОГИЙ СЕМАНТИЧЕСКОГО АНАЛИЗА И ИНТЕГРАЦИИ ГЕТЕРОГЕННЫХ ИНФОРМАЦИОННЫХ ПОТОКОВ В СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Современные системы поддержки принятия решений (СППР) работают с потоками данных из IoT-сенсоров, баз данных, текстовых сообщений и социальных медиа, представленных в различных форматах и обладающих неодинаковыми смысловыми моделями и качеством. Отсутствие семантически согласованной интеграции приводит к неоднозначной интерпретации сущностей, дублированию и потере контекста, что снижает качество и оперативность решений. Цель работы – систематизировать методы семантического анализа и интеграции гетерогенных информационных потоков для СППР, выявить их преимущества, ограничения и области применения. Исследование выполнено в форме аналитического обзора публикаций 2018–2025 гг. по семантической интероперабельности, онтологиям и знаниевым графам, многоисточниковому слиянию данных, федеративному доступу и потоковой обработке. Показано, что обеспечение семантической совместимости достигается за счёт онтологий и знаниевых графов, задающих общие сущности и идентификаторы и обеспечивающих гибкость схемы интеграции. Для режимов реального времени наиболее результативны гибридные решения, сочетающие семантический слой с алгоритмами слияния и учётом доверия к источникам; в прикладных кейсах отмечаются улучшения точности порядка 15–20% и сокращение времени отклика до 70–80% при использовании многоисточниковых моделей. Для неструктурированных потоков ключевую роль играют NLP и машинное обучение, обеспечивающие извлечение сущностей и отношений и семантическое обогащение данных. Полученные результаты могут быть использованы при проектировании СППР в доменах умного города, промышленности и здравоохранения. Кроме того, работа подчеркивает важность стандартов, таких как SHACL для валидации и SPARQL для запросов, которые усиливают практическую применимость семантических подходов. Перспективы включают автоматизацию ontology alignment для снижения трудозатрат и интеграцию с ИИ для динамической адаптации к новым источникам данных.

Семантическая интеграция данных; гетерогенные информационные потоки; система поддержки принятия решений; онтология; знаниевый граф; семантическая интероперабельность; слияние данных; федерация данных; обработка потоков; интернет вещей.

V.V. Gapochka, E.E. Polupanova

**SEMANTIC ANALYSIS AND INTEGRATION OF HETEROGENEOUS
INFORMATION STREAMS IN DECISION SUPPORT SYSTEMS:
A TECHNOLOGY REVIEW**

Modern decision support systems (DSS) increasingly rely on heterogeneous data streams from IoT sensors, databases, text messages, and social media, represented in different formats and characterized by diverse semantic models and quality levels. The lack of semantically aligned integration results in inconsistent entity interpretation, duplication, and loss of context, which reduces the quality and timeliness of decisions. The aim of this paper is to systematize methods for semantic analysis and integration of heterogeneous information streams in DSS and to identify their benefits, limitations, and application domains. The study is conducted as an analytical review of publications from 2018–2025 focusing on semantic interoperability, ontologies and knowledge graphs, multi-source data fusion, data federation, and real-time stream processing. The review shows that semantic compatibility is primarily achieved through ontologies and knowledge graphs that define shared entities and identifiers and provide a flexible integration schema. For real-time decision-making, hybrid solutions combining a semantic layer with data fusion algorithms and source trust assessment are the most effective; published case studies report accuracy gains of about 15–20% and response-time reductions of up to 70–80% in multi-source settings. For unstructured streams, NLP and machine learning play a key role by extracting entities and relations and enabling semantic enrichment. The results can be used to design DSS for smart city, industrial, and healthcare domains. Furthermore, the paper highlights the role of standards like SHACL for validation and SPARQL for querying, enhancing the practical applicability of semantic approaches. Future directions include automating ontology alignment to reduce labor costs and integrating with AI for dynamic adaptation to new data sources.

Semantic data integration; heterogeneous information streams; decision support system; ontology; knowledge graph; semantic interoperability; data fusion; data federation; stream processing; Internet of Things.

Введение. В эпоху больших данных и интернета вещей организации и города сталкиваются с огромными потоками разнородной информации, которые необходимо учитывать при принятии решений. Гетерогенные информационные потоки включают структурированные данные (таблицы БД), полуструктурированные (XML/JSON документы, лог-файлы) и неструктурированные данные (текстовые отчеты, изображения, датчики, посты соцсетей). Каждый тип данных обладает своими форматами, временными характеристиками и семантическими представлениями – словарями и контекстом, в рамках которых данные имеют смысл. Традиционные СППР, как правило, разрабатываются под один источник или тип данных и не обладают встроенными средствами унификации информации из разных источников. В результате актуальной становится проблема семантической интероперабельности: данные из разных систем и датчиков могут иметь разный смысл или структуру, что приводит к ошибкам интерпретации, дублированию и пробелам в знании.

Актуальность проблемы. В научной и прикладной сфере назрел запрос на решения, обеспечивающие целостное представление ситуации на основе мультиформатных данных. Например, в умном городе для управления дорожным движением нужны данные от транспортных датчиков, погодных служб, социальных сетей (сообщения о ДТП) и городских реестров. Эти сведения разнородны по природе – числовые показания, текстовые сообщения, ГИС-данные – и поступают с различной периодичностью (потоково и дискретно). Без их интеграции городские службы не смогут получить полную картину и своевременно реагировать на события. Аналогичные вызовы существуют в медицине (совмещение электронных медкарточек, результатов анализов и научных публикаций), в промышленности (данные с разных участков производства), в финансах (агрегирование новостей, котировок, экономических индикаторов) и т.д. Везде правильность решений зависит от способности системы обработать весь спектр доступной информации. Однако сегодня многие информационные системы остаются изолированными и разработаны под узкие задачи, что затрудняет их совместное использование. Даже при наличии технической интеграции на уровне форматов (синтаксическая совместимость) остается проблема семантической несовместимости – разных терминологий, онтологий и контекстов данных в разных источниках.

Анализ текущего состояния. В последние пять лет наблюдается интенсивное развитие методов и инструментов для интеграции гетерогенных данных [1–4]. Согласно обзору 2024 г. [3], большинство исследований сфокусировано на семантических аспектах объединения данных, особенно на проблеме интеграции неструктурированных данных и разрешении различий в терминологии. Это объясняется тем, что именно семантические несоответствия (разные названия одних сущностей, разные классификационные схемы, единицы измерения и т.д.) являются главным препятствием к полноценной интеграции [3]. Появились специализированные решения: разработаны онтологии для различных доменов (например, онтологии сенсорных сетей, медицинские онтологии), создаются знаниевые графы, охватывающие множество источников, и разрабатываются алгоритмы выравнивания схем данных (schema matching) на основе правил и машинного обучения. С другой стороны, эволюция технологий больших данных привела к тому, что интеграция осуществляется на двух уровнях: физическом и виртуальном [3]. Физический подход (например, хранилища данных, data warehouse) обеспечивает более эффективные запросы, но дорог в реализации и обновлении; виртуальный подход (посредники, вытягивающие данные по запросу из источников) набирает популярность благодаря гибкости и снижению дублирования данных. Многие современные работы применяют сочетание этих подходов – создание единого семантического слоя, который виртуально объединяет данные из разных хранилищ.

Однако остаются нерешенные вопросы. Во-первых, обеспечение масштабируемости: с ростом числа источников и объема данных возрастает нагрузка на системы интеграции, требуя распределенных архитектур и оптимизированных алгоритмов. Во-вторых, автоматизация семантического согласования: ручная разработка онтологий и правил сопоставления трудоемка, поэтому актуально применение методов ИИ, способных автоматически находить соответствия между понятиями разных датасетов. В-третьих, вопрос качества и доверия к данным: например, данные из социальных сетей ценны для быстрого обнаружения событий, но содержат шум и могут быть недостоверными. Нужно внедрение механизмов верификации, очистки и оценки доверия к источникам (например, учет рейтинга надежности источника в алгоритмах слияния). Наконец, проблема стандартизации: отсутствие единых стандартов описания и передачи данных затрудняет интероперабельность, хотя в отдельных отраслях наблюдается прогресс (например, стандарты HL7 FHIR в здравоохранении, ISO 15926 в промышленности) [4–6].

Цель исследования. С учетом изложенного, целью данной работы является обобщение и систематизация существующих методов семантического анализа и интеграции разнородных информационных потоков для СППР, а также оценка их эффективности. Иными словами, работа отвечает на вопрос: *какие технологии позволяют объединить гетерогенные данные разных типов в единую информационную основу для принятия решений, и какой эффект дает их применение?*

Объект и предмет исследования. Объектом исследования являются процессы и системы интеграции данных в рамках СППР, оперирующие множественными источниками информации. Предмет исследования – методы семантического анализа и интеграции, позволяющие преодолеть структурные и семантические различия данных (включая онтологии, алгоритмы слияния, архитектурные решения). Гипотеза исследования заключается в том, что использование семантических моделей знаний (онтологий, знаниевых графов) в сочетании с автоматизированными методами обработки данных (машинное обучение, потоковая аналитика) значительно повышает полноту и точность информации, доступной СППР, что в итоге повышает качество решений.

В дальнейшем изложении статьи сначала рассматриваются типы гетерогенных данных и связанные с ними проблемы интеграции, затем описываются основные подходы и методы решения поставленной задачи. Далее приводятся примеры реализации интеграции в различных доменах и обсуждаются достигнутые результаты, а в заключении сформулированы основные выводы, новизна и направления дальнейших исследований.

Типы гетерогенных информационных потоков и проблемы их интеграции.

Для проектирования интеграции в СППР целесообразно диагностировать гетерогенность по нескольким осям:

- 1) структура представления данных;
- 2) смысловая (семантическая) модель;
- 3) качество и динамика (задержка, актуальность, шум);
- 4) способы доступа и протоколы.

На практике эти аспекты накладываются: один и тот же объект может одновременно присутствовать в табличных данных, потоковых JSON-событиях и текстовых сообщениях, отличаясь терминологией и уровнем доверия к источнику (табл. 1).

Таблица 1

Матрица типов потоков

	Batch	Near-real-time	Streaming
Structured	DWH/OLAP выгрузки, таблицы БД → ETL/ELT, CDC	Оперативные таблицы, частые инкременты → CDC, upsert	Телеметрия как time-series, агрегаты → window/CEP, merge
Semi-structured	XML/JSON файлы, логи → парсинг, канонизация	API-ответы JSON, schema drift → schema registry, нормализация	JSON events (Kafka), IoT payload → stream ETL, JSON-LD [7]/RDF
Unstructured	Документы/отчёты/PDF → NLP (NER/RE), извлечение фактов	Новости/соцмедиа пачками → NLP	Соцмедиа/чат/аудио/видео → online NLP/CV, KG enrichment

Примечание: CDC – Change Data Capture; CEP – Complex Event Processing; NER/RE – извлечение сущностей/отношений; KG – Knowledge Graph (граф знаний); CV – computer vision.

Далее гетерогенность рассмотрим по ключевым аспектам.

Структурная гетерогенность. Данные различаются по структуре и модели представления. Одни источники могут хранить информацию в виде реляционных таблиц со строгой схемой, другие – в виде иерархических документов (XML/JSON), третьи – как неструктурированный текст или мультимедиа. Например, система мониторинга города содержит табличные данные о транспортных потоках, геопривязанные временные ряды с датчиков (формат JSON) и текстовые сообщения о происшествиях. При объединении таких данных возникает проблема согласования разных схем и форматов. Требуется привести их к единой представленной форме, не потеряв смысл. Часто используют преобразование всех данных к формату RDF (Resource Description Framework) или другим промежуточным моделям, позволяющим описать разную структуру единым способом [8].

Семантическая гетерогенность. Разные источники описывают схожие сущности различными терминами и моделями. Классический пример – различия в онтологиях: в одной базе поле названо “*CitizenID*”, в другой “*PersonCode*”, и, хотя по смыслу это одно и то же, без явного соответствия система воспримет их как разные понятия (рис. 1). Аналогично, единицы измерения (например, градусы Цельсия и Фаренгейта) или классификаторы (разные рубрики товаров в двух магазинах) – все это создает барьеры для автоматического объединения. Семантическая несовместимость ведет к тому, что агрегированные данные могут быть неправильно интерпретированы или не сопоставлены друг с другом. Решение требует введения слоя знаний – словарей, тезаурусов или онтологий, которые задают единую систему понятий для всех источников [9]. Составление такого словаря – сложная методическая задача, особенно если данные охватывают несколько предметных областей сразу.

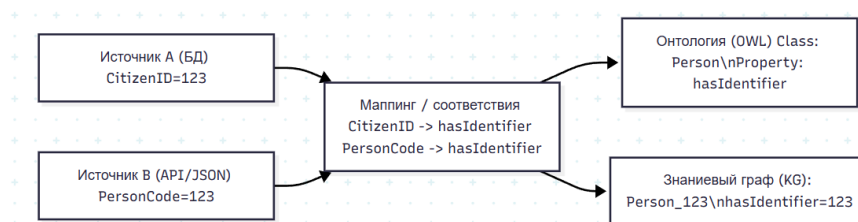


Рис. 1. Пример мэппинга данных из разных источников

Различие в качестве и динамике данных. Информационные потоки различаются по достоверности, точности, частоте обновления. Например, данные от официальных датчиков обычно надежны, но обновляются с интервалом в минуты, тогда как сообщения граждан в соцсетях появляются мгновенно, но могут содержать ошибки или дезинформацию [10]. Интегрируя такие источники, СППР должна учитывать вес или доверие к каждому. Возникает задача валидации и фильтрации: выявлять аномалии, пропуски, противоречия между источниками. Необходимо обеспечить, чтобы слияние данных повышало информационную ценность, а не вносило дополнительный шум. Здесь помогают методы фильтрации на основе правил качества данных и статистические модели, оценивающие достоверность (к примеру, кросс-проверка сообщений соцсетей с датчиками или другими независимыми источниками). Современные решения включают также *семантическую валидацию* – проверку соответствия входящего потока заданным онтологиям и ограничениям (например, в умном городе – сравнение показаний датчиков с допустимыми диапазонами и контекстом) [11] (рис. 2).

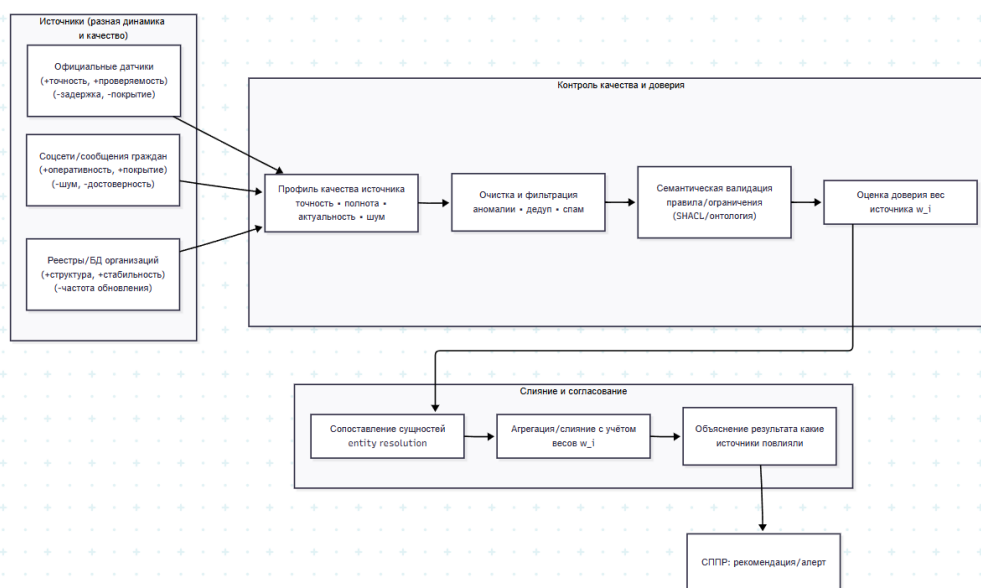


Рис. 2. Контроль качества и учёт доверия к источникам при интеграции потоков данных

Существует большое разнообразие протоколов и форматов доступа. Одни данные доступны через API в реальном времени, другие – как статические файлы, третьи – через базы данных. Интеграционная система должна уметь *подключаться* к разным источникам: читать стримы событий (например, через MQTT или Kafka), выполнять SQL-запросы к БД, парсить веб-страницы и т.д. Это больше технический аспект, но без его решения невозможно собрать все данные воедино. В рамках современных платформ все

чаще применяется микросервисная архитектура, где специализированные коннекторы собирают данные с каждого типа источника и передают их в единый центр хранения или обработки [12] (рис. 3).

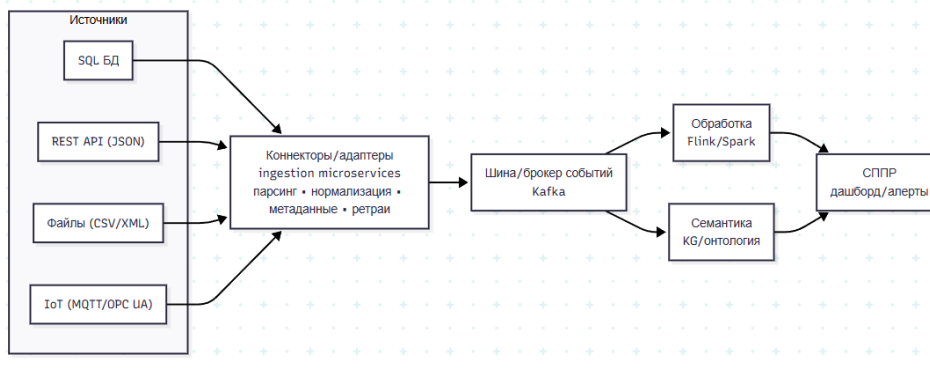


Рис. 3. Слой коннекторов доступа. Унификация протоколов и форматов

Понимание этих видов гетерогенности позволяет сформировать требования к методам интеграции. В частности, необходимы:

- 1) средства трансформации данных между форматами;
- 2) семантические шлюзы – словари или сервисы сопоставления понятий;
- 3) механизмы оценки и обеспечения качества данных;
- 4) модульная архитектура, поддерживающая подключение новых источников.

Методы и подходы к семантической интеграции данных. После описания видов гетерогенности и связанных с ними проблем (структурной, семантической, качественной и гетерогенности доступа) возникает практический вопрос: какими методами и технологиями эти проблемы решаются в современных системах поддержки принятия решений. В отличие от “классической” интеграции данных (ETL/хранилища), семантическая интеграция требует не только приведения форматов, но и согласования смысла, разрешения неоднозначностей сущностей, контроля доверия к источникам и поддержки запросов поверх множества разнородных потоков. Поэтому в современной литературе и практике наблюдается переход к многоуровневым решениям, где семантический слой (онтологии/графы знаний) работает совместно с алгоритмами слияния, средствами потоковой обработки и архитектурными механизмами виртуализации/федерации доступа [3, 12].

Классификация подходов и критерии выбора. Современные подходы к семантической интеграции гетерогенных информационных потоков в СППР целесообразно сгруппировать по доминирующему принципу решения задачи на три класса:

- 1) знаниевые (knowledge-driven);
- 2) алгоритмические (data-/compute-driven);
- 3) архитектурно-гибридные.

Знаниевые подходы ориентированы на согласование смысла данных через онтологии, словари и знаниевые графы, обеспечивая единый “язык” предметной области и сквозную идентификацию сущностей. Алгоритмические подходы концентрируются на вычислительной стороне интеграции: сопоставлении сущностей, слиянии данных, обработке неопределенности и учёте доверия к источникам, включая методы машинного обучения для entity matching и извлечения фактов из неструктурированных потоков. Архитектурно-гибридные решения обеспечивают внедрение этих методов в условиях больших объемов и скоростей: потоковая обработка, микросервисные коннекторы, семантический слой поверх lake/warehouse, а также виртуальная интеграция и федеративные запросы без полного копирования данных.

Выбор конкретного набора методов определяется эксплуатационными требованиями и характеристиками данных. При строгих требованиях к задержке (latency) и реакции в реальном времени приоритет получают потоковые и гибридные решения, где семантический слой

дополняет stream processing и механизмы оценки качества. При высокой доле неструктурированных данных (тексты, изображения, сообщения) критичны NLP/CV и извлечение сущностей/отношений с последующим связыванием с онтологиями и графом знаний. Если для лица, принимающего решение, важна объяснимость (почему система выдала рекомендацию), предпочтительны модели с явным знанием (онтологии, KG) и трассируемыми правилами валидации/качества. При большом числе источников и частых изменениях схем возрастает роль гибких архитектур и механизмов виртуализации/федерации, снижающих стоимость поддержки интеграции. Наконец, ограничения на хранение и копирование данных (организационные, юридические, технические) напрямую влияют на выбор между материализацией интегрированных данных и их виртуальным объединением.

Сводное сравнение подходов приведено в табл. 2.

Таблица 2

Классификация подходов и критерии выбора

Подход	Сильные стороны	Ограничения	Когда применять	Примеры технологий/практик
Знаниевый (онтологии, KG)	Единая смысловая модель; устранение терминологических конфликтов; сквозная идентификация сущностей; поддержка семантических запросов и объяснимости	Трудоёмкость разработки/поддержки онтологий; требования к качеству маппингов; масштабируемость больших графов	Нужно “единое понимание” данных, сильная объяснимость, много разнородных источников и терминологий	RDF/OWL [13]; Knowledge Graph; SPARQL; триплетные хранилища; ontology alignment
Алгоритмический (matching/fusion/ML)	Хорошо работает с шумом и неопределённостью; автоматизация сопоставления записей; повышение точности при мульти-источниковости; применим к потокам	Может быть менее объясним; требует данных для обучения/калибровки; чувствителен к дрейфу данных	Есть конфликтующие источники, нужна оценка доверия, много событий/сенсоров, требуется online-обработка	Data fusion (Bayes/Kalman/DS); entity resolution; ML-matching; NLP-извлечение фактов
Архитектурно-гибридный (stream + семантика + федерация)	Низкая задержка; масштабируемость; гибкое подключение новых источников; возможность не копировать данные (виртуализация/федерация)	Сложность внедрения и эксплуатации; зависимость от зрелости инфраструктуры; возможный рост латентности при федерации	Реальное время, много коннекторов, быстро меняющиеся источники, ограничения на копирование данных	Kafka/Flink/Spark; микросервисы ingestion; Trino/федерация; семантический слой поверх lake/warehouse
Виртуальная интеграция (семантическая федерация)	Актуальные данные “из источника”; минимизация дублирования; снижение затрат на хранение	Латентность запросов; сложность оптимизации; зависимость от доступности источников	Нужны “самые свежие” данные и/или нельзя реплицировать источники	Ontop (R2RML); Trino; query federation; wrappers/mediators
Материализация (ETL/ELT + семантический слой)	Высокая скорость аналитики; устойчивость и воспроизводимость результатов; удобство качества/аудита	Высокая скорость аналитики; устойчивость и воспроизводимость результатов; удобство качества/аудита	Требуются тяжёлые аналитические запросы и историчность, допустима задержка обновления	DWH/Lakehouse + семантическая витрина; периодические загрузки; индексирование/кэш

Знаниевые подходы: онтологии как «единый язык» семантической интеграции. Онтологический подход основан на том, что корректная интеграция гетерогенных источников невозможна без предварительного согласования понятийной модели предметной области. Если на уровне синтаксиса данные можно привести к общим форматам (таблицы, JSON, события), то на уровне смысла остаются различия в терминологии, классификаторах, единицах измерения и контекстах. Онтология задаёт формальное описание классов, свойств и отношений домена, а также ограничения, которые фиксируют “как правильно понимать данные”. В результате источники получают согласованный словарь, который снижает риск ложных сопоставлений и дублирования сущностей.

В типовом сценарии онтологически-ориентированная интеграция включает следующие шаги:

- 1) *выбор или разработка базовой онтологии*, покрывающей ключевые сущности и отношения всех источников (например, *Sensor, Event, Location, Organization*);
- 2) *семантическая аннотация источников*, при которой элементы схемы источника (поля, сущности, сообщения) сопоставляются понятиям онтологии;
- 3) *приведение данных к семантически совместимому представлению* – либо материализация фактов в форме RDF-триплетов («субъект–предикат–объект»), либо формирование экземпляров классов онтологии;
- 4) *хранение и обработка* интегрированных данных в средствах, поддерживающих семантические запросы (например, SPARQL [14]) и, при необходимости, логический вывод.

Ключевая практическая ценность онтологии в СППР состоит в том, что правила интеграции и запросы формулируются в терминах предметной области, а не в терминах отдельных источников. Например, в сценарии умного города отдельные потоки могут представлять данные о трафике, погоде и инцидентах в разных форматах и с разными атрибутами. При наличии онтологии, где “инцидент” связан с “локацией”, а “погодные условия” – с той же локацией и временем наблюдения, становится возможным единообразно задавать запросы и правила вида: “обнаружить аварии рядом с перегруженными дорогами при температуре ниже 0°C”, не привязываясь к структурам конкретных API или БД.

Важно разграничивать онтологию как схему смысла и данные как факты/наблюдения: онтология фиксирует классы и отношения, а факты (экземпляры) постоянно поступают из потоков и пополняют интегрированное представление. Это разграничение удобно визуализировать (рис. 4): слева располагается слой знаний (классы и свойства), справа – слой экземпляров (конкретные датчики, события, измерения), связанный с онтологией через типизацию и свойства.

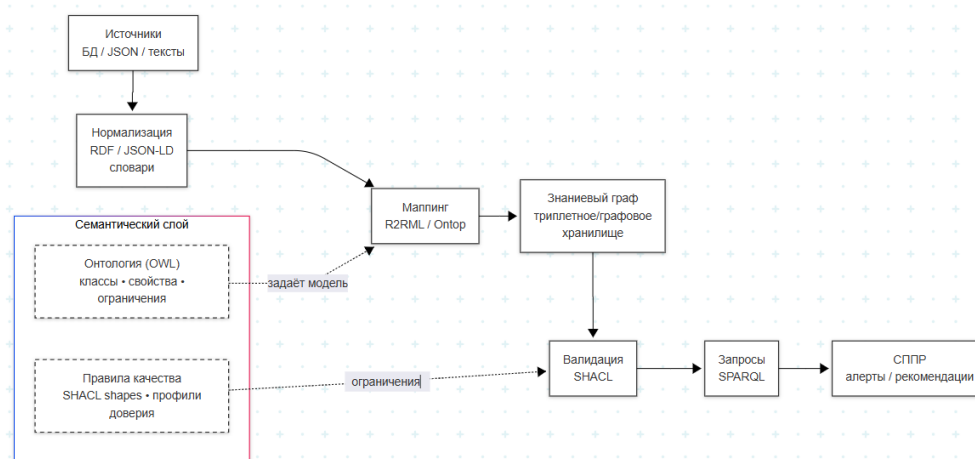


Рис. 4. Онтология (классы/свойства) и экземпляры данных (факты) в семантической интеграции

Ограничениями онтологического подхода являются трудоёмкость разработки и поддержки доменной модели, необходимость сопровождения маппингов источников и потенциальные проблемы масштабируемости при росте объема фактов. На практике эти ограничения снижаются:

- 1) модульным построением онтологии (ядро + доменные расширения);
- 2) автоматизацией части аннотации и сопоставлений;
- 3) сочетанием материализации и виртуализации данных;
- 4) использованием гибридной архитектуры, где онтология задаёт смысловой слой, а потоковые и федеративные механизмы обеспечивают производительность и актуальность данных.

Знаниевые графы и разрешение сущностей. Знаниевые графы (Knowledge Graphs, KG) развивают онтологический подход, делая акцент не только на формализации понятий, но и на сквозной идентификации объектов реального мира и явном представлении связей между ними. В KG узлы соответствуют сущностям (объектам, событиям, организациям, людям), а ребра – отношениям. Ключевая практическая роль KG в семантической интеграции заключается в том, что разные источники могут описывать одну и ту же сущность фрагментарно и разными словами, а граф позволяет собрать эти описания в единый “профиль” сущности, сохранив происхождение фактов и контекст.

Центральным механизмом при построении KG является разрешение сущностей (entity resolution) и связывание сущностей (entity linking). Под entity resolution понимают задачу выявления, относятся ли записи из разных источников к одному и тому же объекту (например, “ул. Ленина, 10” в реестре и “Lenina 10” в сообщении гражданина). Под entity linking – привязку упоминаний из неструктурированного текста к конкретным узлам графа (например, “ДТП на улице X” → узел *Accident_123* с геокоординатами). На практике эти задачи решаются комбинацией методов:

- 1) правила и нормализация (приведение адресов/имен/единиц к каноническому виду);
- 2) метрики сходства строк и атрибутов;
- 3) вероятностные модели и классификаторы соответствий;
- 4) эмбединги и трансформеры для сравнения текстовых описаний;
- 5) использование контекста графа (соседние связи повышают уверенность в совпадении).

Результат работы entity resolution – операция “слияния узлов” (merge): если два узла признаны описаниями одной сущности, они объединяются, а атрибуты агрегируются с учетом доверия к источникам и правил качества. Важно, что KG обычно хранит не только итоговое значение, но и происхождение: из какого источника пришел факт, когда он получен и какова его надежность. Это повышает объяснимость СППР: система может показать, какие источники подтвердили событие и какие признаки повлияли на решение. Типовой пример для умного города: сообщение в соцсетях “авария на улице X” связывается с узлом события, который также получает подтверждение от дорожного датчика (аномальное падение скорости) и от камеры (распознавание инцидента). После слияния СППР получает единое событие с геопривязкой, временной меткой, уровнем доверия и ссылками на источники.

Рис. 5 иллюстрирует принцип “слияния узлов”: до разрешения сущностей граф содержит два фрагмента описания события из разных источников, после – один объединенный узел с агрегированными атрибутами и указанием происхождения фактов.

Согласование знаний и маппинги источников. Даже при наличии онтологии задача семантической интеграции не решается автоматически: реальные информационные системы часто используют разные словари, классификаторы и модели данных, сформированные независимо друг от друга. В результате возникает необходимость согласования онтологий (ontology alignment) и построения маппингов между локальными схемами источников и глобальной (целевой) моделью. С практической точки зрения, alignment отвечает на вопрос “какие понятия и свойства разных систем означают одно и то же”, а маппинг – “как именно преобразовать данные источника, чтобы они соответствовали целевым понятиям”.

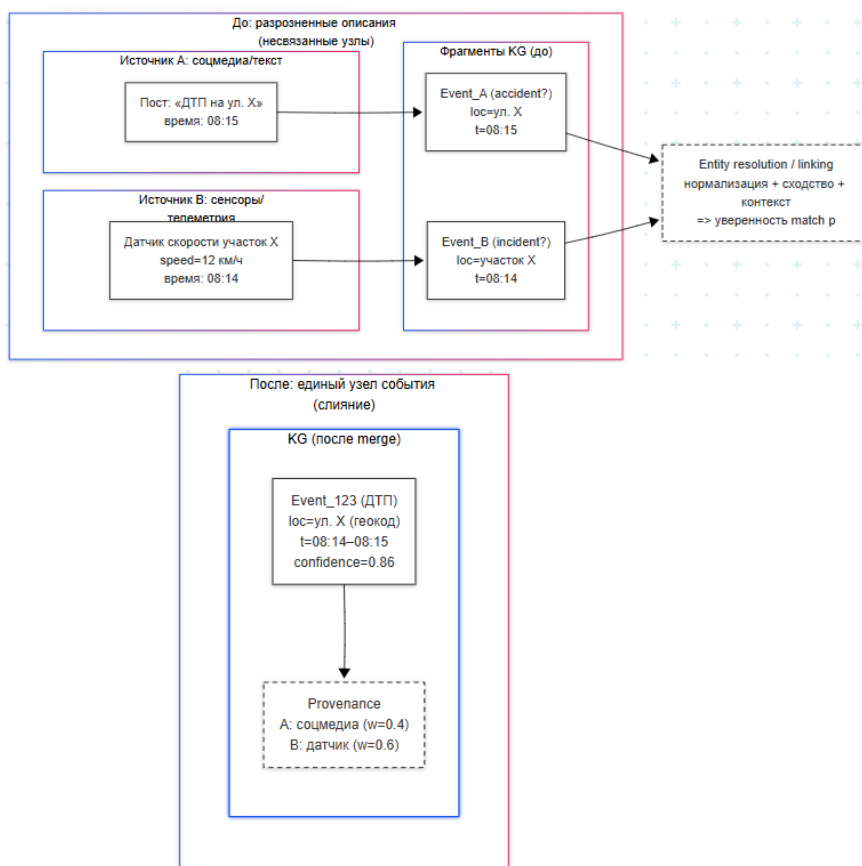


Рис. 5. Разрешение сущностей и «слияние узлов» в знаниевом графе

Методы ontology alignment могут быть автоматическими и полуавтоматическими. Наиболее распространены:

- ◆ лексические и структурные эвристики (сходство названий, синонимы, близость по структуре графа/иерархии);
- ◆ статистические методы (сходство распределений значений, совпадение доменов/типов полей);
- ◆ машинное обучение (включая эмбединги и трансформеры для сопоставления описаний и метаданных);
- ◆ экспертная верификация, необходимая в критичных доменах и при высокой цене ошибки.

В литературе встречается архитектурный шаблон “глобальная онтология + локальные онтологии источников + слой соответствий”. Например, в рамках подхода DIF (Data Integration Framework) формируется глобальная онтология, для каждого источника используется локальная модель, а механизм маппинга соединяет локальные и глобальные элементы, позволяя консолидировать данные [3]. Такой подход снижает связность системы: изменение одного источника требует обновления только его маппингов, а не всей интеграционной схемы.

Для структурированных источников (прежде всего реляционных БД) ключевым инструментом являются маппинги “RDB→RDF”, которые формально описывают, как таблицы, строки и столбцы отображаются в RDF-триплеты и классы/свойства онтологии. На практике широко применяется спецификация R2RML [15], а также системы, которые используют совместимые принципы отображения.

При проектировании интеграции важно выбрать стратегию работы с данными: материализацию либо виртуализацию (рис. 6).

Материализация предполагает извлечение (ETL/ELT) и преобразование данных в RDF/граф (или в унифицированное хранилище) с последующей загрузкой в триплетное/графовое хранилище. Плюсы – быстрые семантические запросы, удобство контроля качества и историчности; минусы – стоимость сопровождения загрузок и риск устаревания копий.

Виртуализация предполагает, что данные остаются в исходных системах, а маппинг (R2RML/Ontop-подход) позволяет выполнять запросы к “виртуальному RDF”: поступающий SPARQL запрос транслируется в SQL/подзапросы и выполняется в источнике, после чего результаты собираются и приводятся к семантическому представлению. Плюсы – актуальность данных и отсутствие копирования; минусы – более сложная оптимизация и зависимость латентности от источников.

Таким образом, ontology alignment и формальные маппинги выступают “мостом” между разнородными системами и семантической моделью. Они обеспечивают масштабируемость внедрения: добавление новых источников и изменение существующих сводится к обновлению соответствий и маппингов, а не к перепроектированию всей СППР.

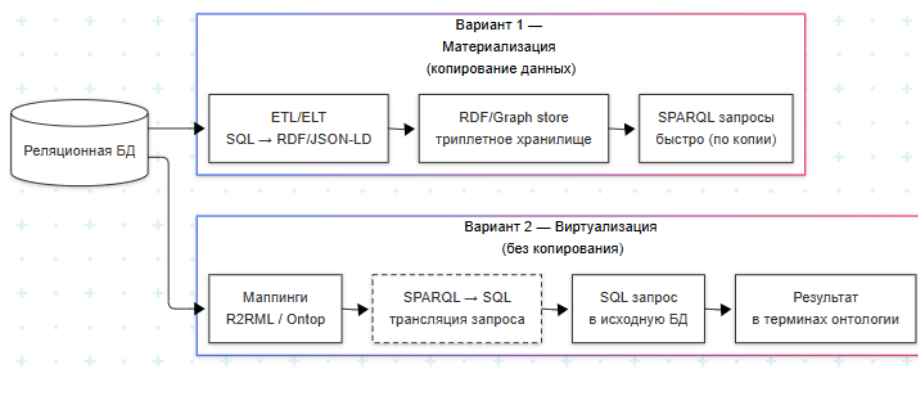


Рис. 6. Два сценария интеграции: материализация данных и виртуализация через маппинги

Семантическое обогащение неструктурированных потоков (NLP/CV). Для СППР критически важны не только структурированные источники (БД, реестры, телеметрия), но и неструктурированные потоки: текстовые сообщения, отчёты, обращения граждан, новости и социальные медиа, а также изображения и видео. Именно такие источники часто дают ранние сигналы о событиях (инцидентах, сбоях, изменениях ситуации), однако в “сыром” виде они плохо пригодны для автоматической интеграции: содержат неоднозначные упоминания, шум, жаргон, неполные описания и контекстные ссылки. Поэтому в современных системах интеграции применяется семантическое обогащение (semantic enrichment) – преобразование неструктурированного контента в формализованные факты, согласованные с онтологией и знаниевым графом (рис. 5).

Семантическое обогащение в текстовых потоках обычно включает следующие этапы:

- 1) *предобработка* (язык, токенизация, нормализация, удаление спама/дубликатов);
- 2) *извлечение сущностей (NER)* – выделение объектов (персон, организаций, локаций, событий, показателей);
- 3) *извлечение отношений (RE)* – выявление связей между сущностями (например, “инцидент произошёл в локации”, “организация заключила контракт”, “датчик измерил параметр”);
- 4) *нормализация и канонизация* – приведение сущностей к стандартному виду (адреса, единицы измерения, временные выражения);

5) *entity linking* – привязка упоминаний к конкретным узлам графа знаний (или создание новых узлов при отсутствии);

б) *формирование фактов* в виде триплетов/графовых утверждений с метаданными (время, источник, доверие).

Для мультимодальных потоков (изображения/видео) аналогичную роль выполняют методы компьютерного зрения: детекция объектов и сцен (например, “ДТП/скопление людей/затопление”), распознавание текстов на изображении (OCR) и последующая привязка результата к сущностям графа (локация, время, тип происшествия). Важный практический момент – хранение provenance: любая извлечённая сущность и связь должны сопровождаться указанием источника, времени извлечения и оценкой уверенности модели. Это позволяет СППР управлять качеством: понижать вес сомнительных наблюдений, подтверждать события по нескольким независимым каналам и объяснять пользователю, на каких данных основана рекомендация.

Семантическое обогащение существенно расширяет область применения онтологий и знаниевых графов: они перестают быть “слоем для БД” и начинают работать как единый контур интеграции, включающий тексты и мультимедиа. В результате СППР получает не просто поток сообщений, а поток структурированных событий и фактов, которые можно объединять с телеметрией и реестрами, валидировать правилами качества и использовать для оперативных решений.

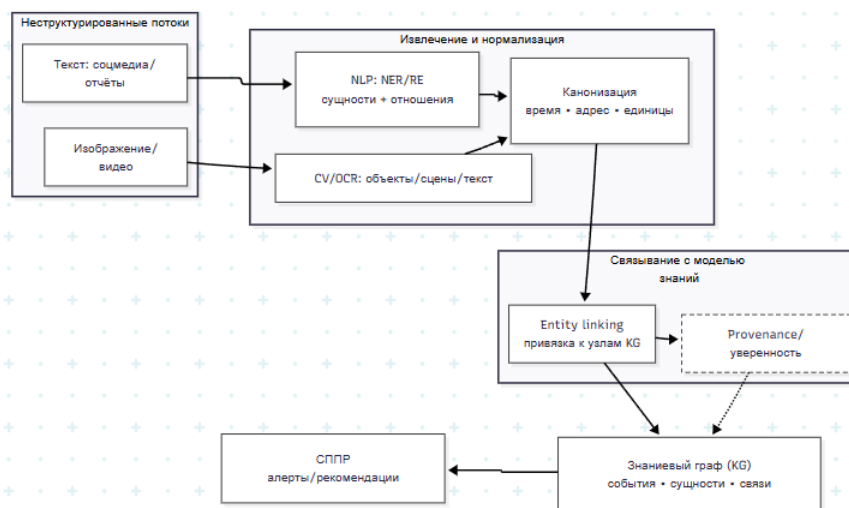


Рис. 7. Семантическое обогащение неструктурированных потоков:
NLP/CV → linking → KG

Алгоритмические подходы: data fusion и управление неопределённостью. Даже при наличии единой семантической модели (онтологии/графа знаний) интеграция в СППР требует алгоритмических решений, которые отвечают на вопрос: как объединять наблюдения из разных источников, если они поступают с разной частотой, имеют разные погрешности и могут противоречить друг другу. Эта группа методов относится к классу data fusion (слияние данных) и включает агрегирование измерений, обновление гипотез о событиях, оценку состояния системы, а также детектирование инцидентов в режиме реального времени. В мульти-источниковых сценариях цель fusion – получить более точную и устойчивую оценку, чем даёт любой отдельный источник, при этом сохранив информацию о доверии и происхождении фактов.

Неопределённость в потоке данных проявляется в нескольких формах:

- 1) шум и случайные ошибки измерений (типично для сенсоров);
- 2) латентность и пропуски (часть источников обновляется редко или теряет сообщения);

- 3) семантические неоднозначности (одно и то же событие описано разными словами);
- 4) конфликтующие свидетельства (соцмедиа сообщает об аварии, а датчики её не подтверждают).

Поэтому алгоритмы fusion обычно включают:

- 1) нормализацию (приведение к общим единицам);
- 2) взвешивание источников по надежности;
- 3) механизм разрешения конфликтов;
- 4) обновление оценки по мере поступления новых свидетельств.

Классический набор методов, применяемых в data fusion, включает вероятностные и эвристические подходы. Для числовых временных рядов и задач оценивания состояния часто используют *фильтр Калмана* и его расширения, позволяющие сглаживать шум и оценивать скрытые переменные системы. Для событийных сценариев, где поступают неоднородные “свидетельства”, применяются *байесовские методы* (обновление апостериорных вероятностей гипотез) и *теория Демпстера–Шафера*, удобная для объединения частичных и неопределённых свидетельств и явного представления “массы доверия” к гипотезам. В задачах, требующих интерпретируемости при неоднозначных правилах, используются *нечёткие модели* (fuzzy logic), где экспертные знания формализуются через лингвистические правила (“высокая температура” и “резкое падение скорости” равно “вероятность инцидента высокая”).

Для современных СППР важным является *связующее звено* между контуром качества и fusion: результаты профилирования качества и валидации формируют *веса доверия источникам* (или confidence к наблюдениям), которые затем участвуют в объединении. Практически это реализуется как взвешенные агрегаты для числовых данных, ранжирование и пороговые решения для инцидентов, а также гибридные схемы, где разные методы используются последовательно (например, Калман для сенсорной части + Демпстер–Шафер для событийных свидетельств + байесовское обновление итоговой уверенности). Такой подход особенно эффективен в real-time СППР, где нужно быстро реагировать, но при этом снижать риск ложных срабатываний за счёт мульти-источникового подтверждения (рис. 8).

Таблица 3

Методы data fusion и управление неопределённостью в СППР

Метод	Для каких данных	Что даёт	Типовые задачи в СППР
Взвешенная агрегация / voting	однотипные сенсоры, показатели	быстрое объединение с весами доверия	агрегирование телеметрии, первичный скоринг
Фильтр Калмана (EKF/UKF)	временные ряды	оценка состояния, сглаживание, прогноз	трафик, производство, предиктивное обслуживание
Байесовское обновление	гипотезы/события, вероятн. признаки	обновление вероятности события	диагностика, оценка риска/инцидента
Демпстер–Шафер	неоднородные свидетельства	объединение частичных/конфликтных данных	подтверждение инцидентов (сенсоры+текст+CV)
Нечёткая логика	экспертные признаки	интерпретируемые правила “если–то”	триггеры реакции, риск-скоринг
ML-модели (ансамбли/NN)	смешанные признаки (в т. ч. NLP/CV)	извлечение сложных паттернов	классификация событий, прогноз, приоритизация
KG-контекст (graph-aware)	факты/связи в KG	учет контекста и связности	консолидация событий, объяснение решений

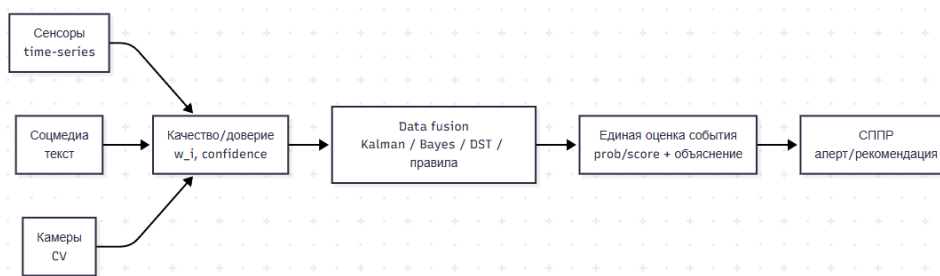


Рис. 8. “Fusion-узел”: как доверие и потоки сходятся в оценку события

Контроль качества и семантическая валидация (SHACL/правила). Контроль качества данных является обязательным компонентом семантической интеграции: даже идеально согласованная онтология не компенсирует шум, пропуски, задержки и противоречия, характерные для многопоточных источников. В СППР качество влияет непосредственно на надежность рекомендаций, поэтому контур качества обычно строится как непрерывный процесс, работающий как на входе (до попадания фактов в интегрированное представление), так и на уровне уже сформированного графа знаний.

Практически контроль качества включает четыре взаимосвязанных слоя.

На слое *профилирования и мониторинга качества* происходит измерение полноты, частоты обновления, задержки, доли пропусков и аномалий для каждого источника и типа данных. Эти метрики используются для расчёта веса доверия источнику и выбора политики обработки (например, “для источника X применять более строгие фильтры и подтверждение по второму каналу”).

На слое *очистки и фильтрации* осуществляется дедупликация, подавление выбросов, удаление спама и сервисных сообщений, нормализация единиц измерения и форматов времени/географии. Для потоков важны online-проверки (окна, пороги, gate-limit), позволяющие отбрасывать явные ошибки до того, как они повлияют на решения.

На слое *семантической валидации* реализован контроль соответствия фактов доменной модели (типам, обязательным свойствам, диапазонам, допустимым связям). В отличие от “обычной” валидации схемы, семантическая проверяет корректность в контексте предметной области: например, что событие “ДТП” обязательно имеет время и локацию, а скорость измерения не может быть отрицательной; что датчик, измеряющий скорость, связан с дорожным участком; что временная метка находится в разумном диапазоне относительно текущего времени и т.п.

Последний слой – *управление неопределенностью и конфликтами*. При обнаружении противоречий между источниками система может:

- 1) понижать уверенность/вес спорных фактов;
- 2) переводить событие в статус “требуется подтверждения”;
- 3) инициировать повторную проверку или запрос к дополнительному источнику;
- 4) сохранять несколько альтернативных значений с указанием происхождения и уверенности.

Для RDF/knowledge graph-представлений широко применяется язык ограничений SHACL (Shapes Constraint Language) [16], позволяющий описывать правила валидации как “шаблоны” (shapes), которые должны выполняться для узлов определенного типа. SHACL удобен тем, что делает правила качества явными и проверяемыми, а также позволяет унифицировать валидацию для данных, поступающих из разных источников. В потоковых СППР SHACL-валидация используется либо на этапе пакетной проверки графа, либо выборочно “на входе” для критичных сущностей (событий и ключевых измерений), чтобы снизить риск ложных срабатываний.

На рис. 9 приведен пример SHACL-ограничений для события типа *Accident*: обязательны локация и время наблюдения, требуется хотя бы одно свидетельство (*evidence*), а также задается допустимый диапазон уверенности события. Дополнительно введено доменное правило: временная метка не должна существенно уходить в будущее (допускается небольшой “люфт” из-за задержек доставки и различий системного времени).

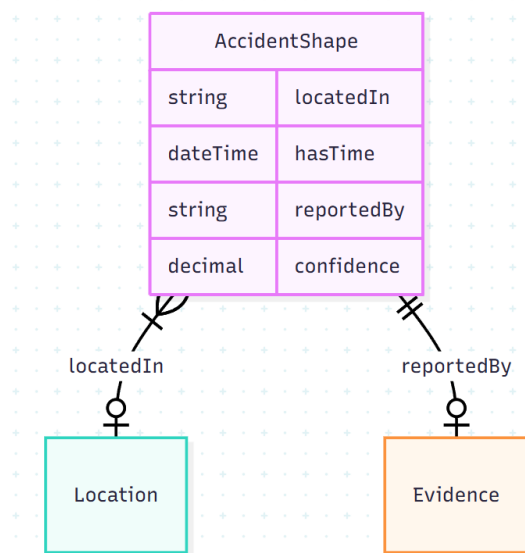


Рис. 9. Пример SHACL-ограничений (NodeShape) для события *Accident*: обязательные свойства и связи с классами *Location* и *Evidence*

Практическая ценность подобных shapes заключается в том, что проверка качества становится формализованной и воспроизводимой: правила можно версионировать, применять к разным источникам и запускать как в пакетном режиме (аудит графа), так и в потоковом режиме на критичных сущностях. При обнаружении нарушений система, как правило, не ограничивается “ошибкой”, а применяет политику обработки (*quality policy*). На практике используются три базовые стратегии:

- 1) *reject* – отклонение факта/события (например, отрицательная скорость или координаты вне диапазона);
- 2) *quarantine* – помещение в карантин/очередь на уточнение (например, событие без локации или времени);
- 3) *soft-fail* – пропуск с понижением доверия (например, неполные данные из источника с низкой надежностью), когда важно не терять сигнал, но необходимо отразить его неопределенность.

Выбор стратегии зависит от критичности атрибута и домена СППР: в оперативных сценариях часто допустим *soft-fail* для раннего предупреждения, но с обязательным последующим подтверждением (кросс-проверкой) и обновлением *confidence*. Таким образом, контур качества выполняет две ключевые функции:

- 1) защищает СППР от “плохих” данных и снижает риск ошибочных рекомендаций;
- 2) формирует метрики доверия и уверенности, которые далее используются в алгоритмах слияния и принятия решений.

Визуально взаимодействие этапов контроля качества удобно представить как замкнутый контур “профиль качества → фильтрация → семантическая валидация → веса доверия → fusion/решение”.

Архитектурно-гибридные решения: потоковая обработка, федерация и семантический слой. Практическая реализация семантической интеграции в СППР почти всегда требует архитектурно-гибридного подхода: семантический слой (онтологии, граф

знаний, правила валидации) должен быть встроен в инфраструктуру сбора и обработки данных, способную работать с высокой скоростью поступления событий и большим числом источников. Поэтому современные решения сочетают:

- 1) потоковую обработку для низкой задержки;
- 2) семантическое моделирование и KG для согласования смысла и сквозной идентификации сущностей;
- 3) контур качества/валидации для надежности результатов;
- 4) федеративный доступ для случаев, когда данные нельзя или нецелесообразно копировать в единое хранилище.

В потоковом слое выполняются операции, которые критичны по времени:

- ◆ прием событий;
- ◆ оконные агрегации;
- ◆ корреляция;
- ◆ первичный детект инцидентов;
- ◆ расчёт признаков.

На этом же этапе обычно производится нормализация форматов (например, приведение JSON-событий к каноническим полям), а также первичная фильтрация и присвоение меток качества. Типовой стек включает брокер сообщений и потоковый движок, позволяющий обрабатывать события по “event-time”, выдерживать задержки и переупорядочивание сообщений.

Семантический слой может быть реализован двумя способами:

- 1) материализованно, когда факты загружаются в KG/триплетное хранилище и далее доступны для SPARQL-запросов и вывода;
- 2) виртуально, когда KG выступает как “семантическая витрина” поверх источников, а запросы транслируются к исходным системам через маппинги и федерацию.

На практике часто применяется компромисс: материализуются наиболее важные, “оперативные” сущности (ключевые события, справочники, идентификаторы), а тяжелые детальные данные остаются в источниках и подключаются по мере необходимости.

Федеративный доступ становится критичным в двух ситуациях:

- 1) данные живут в разных хранилищах (SQL/NoSQL/объектные хранилища/API) и их копирование дорого или запрещено;
- 2) требуется максимально актуальный ответ в момент запроса.

В этом случае СППР использует единый слой запросов, который распределяет подзапросы по источникам и собирает результат, а семантические описания источников (метаданные, маппинги, словари) помогают корректно “сшивать” данные. Ограничением федерации является рост латентности и сложность оптимизации — поэтому в архитектуре обычно применяются кэширование, материализация горячих данных и предвычисление признаков.

В итоге архитектурно-гибридная схема позволяет построить СППР, которая одновременно реагирует в реальном времени, сохраняет семантическую согласованность данных, умеет контролировать качество и доверие, и не требует обязательного полного копирования всех данных в единое хранилище. Обобщённая конфигурация таких решений показана на рис. 10.

Типовой конвейер семантической интеграции потоков и соответствие технологий. Рассмотренные выше методы (онтологии и знаниевые графы, entity linking, семантическое обогащение, data fusion, контроль качества и архитектурные решения) на практике реализуются не в виде одного “универсального” модуля, а как конвейер взаимосвязанных этапов, через которые проходят данные от момента поступления из источников до формирования рекомендаций СППР. Такой конвейер позволяет:

- 1) разделить ответственность компонентов;
- 2) масштабировать систему по узким местам (например, отдельно масштабировать обработку потоков или извлечение сущностей из текстов);
- 3) внедрять контроль качества как непрерывный процесс;
- 4) обеспечить трассируемость фактов и их происхождения.

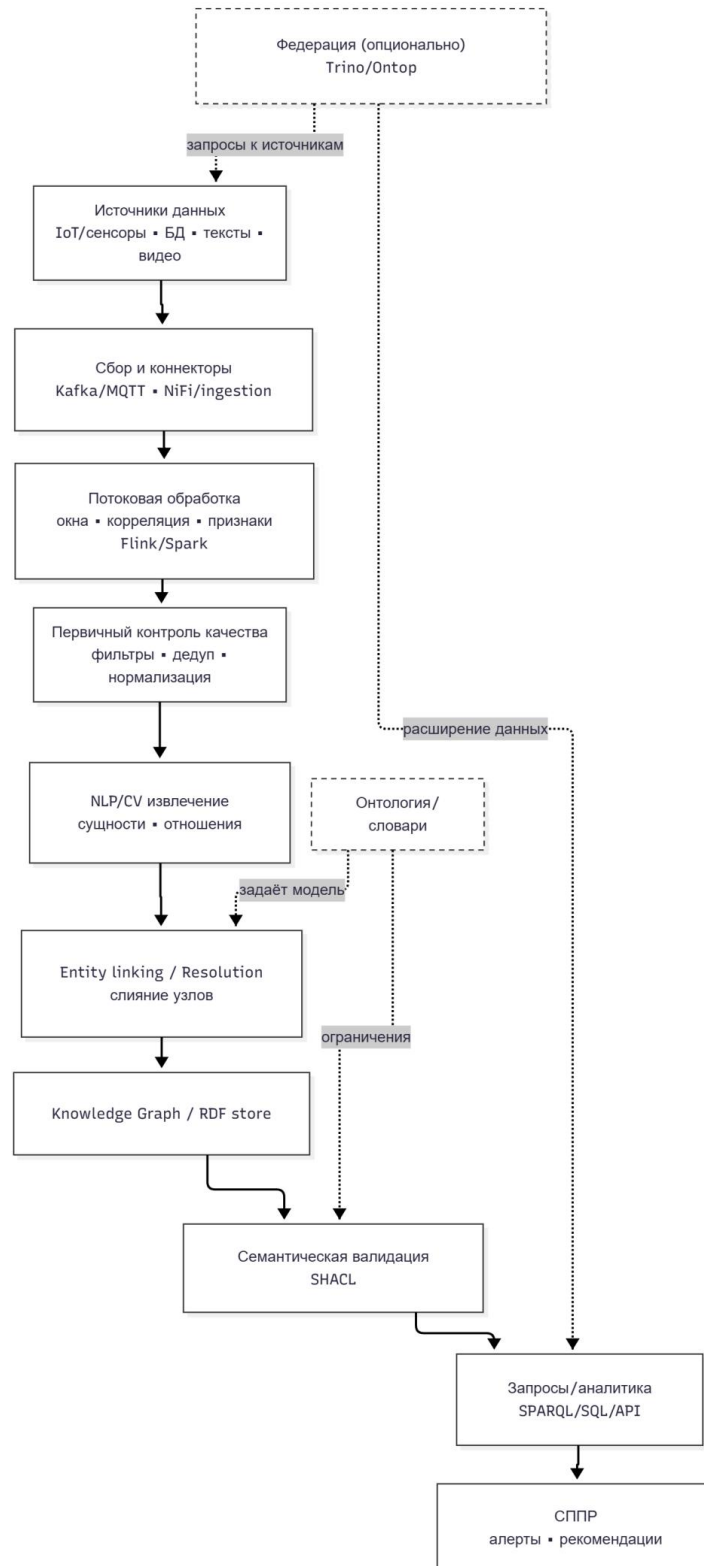


Рис. 10. Архитектура семантической интеграции для СППР: потоковый слой + KG/валидация + федерация источников

Типовая последовательность шагов включает:

- 1) сбор данных (стрим/файлы/запросы);
- 2) потоковую обработку (окна, корреляция, детект событий);
- 3) нормализацию/гармонизацию форматов и единиц;
- 4) семантическое моделирование домена;
- 5) маппинг структурированных источников (включая RDB→RDF);
- 6) валидацию качества (в том числе через SHACL);
- 7) семантические запросы к интегрированному представлению.

На каждом этапе могут применяться как знаниевые, так и алгоритмические методы: например, извлечение сущностей и связывание с KG выполняется на этапе обогащения, а веса доверия и слияние данных – на стыке контроля качества и принятия решения.

Сопоставление этапов интеграционного конвейера с типовыми стандартами и примерами технологий приведено в табл. 2. Эта таблица может использоваться как практическая “карта” при проектировании СППР: она помогает определить, какие компоненты являются обязательными для заданного класса задач (например, real-time мониторинг), а какие – опциональными и зависят от ограничений (например, федерация при запрете на копирование данных).

Таким образом, семантическая интеграция в СППР представляет собой комбинацию семантического слоя (онтологии/KG) и инженерной инфраструктуры (потоки, коннекторы, федерация), дополненную алгоритмами слияния и контуром качества. В следующем разделе рассматриваются примеры внедрения этих подходов в прикладных доменах и обсуждаются достигнутые эффекты (точность, скорость реакции, устойчивость к неоднозначностям и качеству данных).

Практические эффекты применения семантической интеграции в СППР. Эффективность семантической интеграции в СППР оценивается не только по тому, насколько корректно объединяются данные из разных источников. Важно также, как меняются прикладные показатели работы системы. Обычно смотрят, стала ли ситуационная картина более полной, повысилась ли точность рекомендаций и уменьшилось ли время реакции. Отдельное внимание уделяют устойчивости к шуму и противоречиям между источниками. Анализ публикаций и практических внедрений показывает, что наиболее заметный эффект появляется там, где система опирается на несколько каналов с разной надежностью и скоростью обновления. Существенную роль играет наличие неструктурированной информации, которая поступает в виде текстов и мультимедиа и требует извлечения смысловых фактов. Кроме того, во многих прикладных областях важно, чтобы результаты были объяснимыми и трассируемыми. В этом разделе обобщаются прикладные эффекты внедрения по типовым доменам и рассматриваются факторы, которые сильнее всего влияют на результат. К ним относятся выбранная архитектура, качество маппингов между источниками и моделью знаний, зрелость контура валидации, а также правила учета доверия к источникам.

Повышение полноты и точности ситуационной осведомлённости. Одним из ключевых эффектов семантической интеграции в СППР становится рост ситуационной осведомлённости. Под этим понимается способность системы формировать целостное и непротиворечивое представление о происходящем на основе данных из разных источников. Полнота здесь означает, что в интегрированном представлении присутствуют все значимые объекты и события, а также их важные атрибуты. Точность означает, что эти объекты корректно интерпретированы, не задублированы и правильно связаны друг с другом. Например, событие должно быть привязано к корректной локации и времени, а показатели – к нужным сущностям и единицам измерения.

Рост полноты достигается прежде всего за счёт расширения охвата источников. Структурированные системы обычно дают устойчивые данные, но описывают ситуацию “частично”. Текстовые потоки и сообщения людей часто дают ранние сигналы, а также контекст, который отсутствует в телеметрии. Когда применяется семантическое обогащение на основе NLP и компьютерного зрения, неструктурированный контент преобразуется в формализованные факты. После этого такие факты можно сопостав-

лять с показаниями датчиков и записями реестров. В результате система реже пропускает важные события только потому, что они появились в канале, который изначально не был “табличным”.

Повышение точности обеспечивается несколькими механизмами, которые работают совместно, но выполняют разные задачи. Онтологии и знаниевые графы задают общую смысловую модель. За счёт этого уменьшается количество ошибок, связанных с разной терминологией и разными правилами интерпретации. Механизмы entity resolution и entity linking снижают дублирование. Они позволяют “сшивать” записи из разных источников, если речь идёт об одном и том же объекте или событии. При этом система может сохранять происхождение фактов и уверенность в них, что важно для объяснимости.

Отдельную роль играет контур качества. Он отсекает явно некорректные данные и снижает влияние сомнительных сообщений. Семантическая валидация помогает проверять факты в контексте предметной области. Она выявляет ситуации, когда событию не хватает обязательных атрибутов, когда значения физически невозможны или когда связи между сущностями противоречат модели. Если данные всё же потенциально полезны, система может не удалять их полностью. Вместо этого она может снизить доверие и запросить подтверждение по другим каналам. Алгоритмы слияния данных дополняют этот контур. Они позволяют повышать уверенность события, когда разные независимые источники дают согласующиеся признаки.

В прикладных кейсах это приводит к тому, что СППР формирует не набор разрозненных сигналов, а единую “карту событий”. Для каждого события становится понятнее, что произошло, где и когда это случилось, какими источниками это подтверждается и насколько этому можно доверять. В городских сценариях интеграция сообщений граждан с дорожной телеметрией помогает быстрее обнаруживать инциденты и одновременно снижать число ложных тревог за счёт подтверждения по нескольким каналам. В здравоохранении графовые модели упрощают согласование терминов и объединение клинических данных с внешними знаниями, что снижает неоднозначность и повышает связность фактов. В бизнес-сценариях объединение разнородных потоков повышает точность прогнозов и рекомендаций по сравнению с решениями, построенными на одном источнике [4–6].

Чтобы измерять эффект полноты и точности, в исследованиях и внедрениях обычно фиксируют несколько показателей. В первую очередь оценивают долю найденных событий относительно эталона и точность обнаружения. Затем измеряют качество “сшивки” сущностей между источниками. Дополнительно считают количество дублей до и после entity resolution. Часто отдельно проверяют долю фактов, прошедших правила качества и семантическую валидацию. Эти метрики позволяют показать, что улучшилось не только количество данных, но и их пригодность для принятия решений.

Сокращение времени реакции и автоматизация решений. Второй важный практический эффект семантической интеграции связан со скоростью реакции СППР. В задачах оперативного управления ценность данных снижается по мере роста задержки: ситуация может измениться быстрее, чем оператор успеет собрать информацию из разных систем и согласовать её смысл. Семантическая интеграция уменьшает задержку за счёт того, что данные заранее проходят единый конвейер нормализации, связывания сущностей и проверки качества. В результате система быстрее формирует понятное событие и выдаёт рекомендацию, а не поток разрозненных сигналов.

Сокращение времени реакции обычно обеспечивается сочетанием трёх механизмов (рис. 11):

1. *Потоковая обработка и корреляция.* События обрабатываются “на лету”, агрегируются по окнам времени и связываются по пространственному и временному контексту. Это сокращает время на ручную “сборку” картины из фрагментов.
2. *Семантическое связывание и устранение дублей.* Entity linking/resolution позволяет быстро объединять признаки одного события из разных источников в единый объект (инцидент/событие), что уменьшает задержки на сопоставление и снижает число ложных уведомлений.

3. *Контур качества и подтверждение по нескольким каналам.* SHACL/правила качества отсекают явно некорректные факты, а механизмы доверия к источникам позволяют быстро повышать уверенность события при наличии независимых подтверждений. Это ускоряет принятие решения без роста ложных тревог.

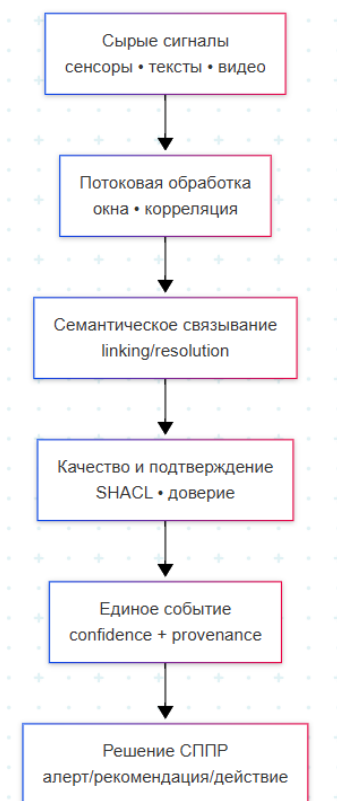


Рис. 11. Конвейер сокращения времени реакции в СППР: от сырого сигнала к подтверждённому событию и действию

Автоматизация решений в СППР обычно развивается поэтапно и зависит от домена и требований к рискам:

1. *Поддержка оператора:* система формирует алерт и объяснение, а решение принимает человек.
2. *Полуавтоматический режим:* часть действий выполняется автоматически, но требует подтверждения или может быть отменена.
3. *Автоматический запуск процедур:* система выполняет заранее определённые действия, если выполнены требования качества и уверенность события достаточна.

Чтобы ускорение не приводило к росту ошибок, в реальных внедрениях вводят политику обработки событий при нарушении качества:

- 1) *reject* – отклонение факта или события (физически невозможные значения, критические ошибки);
- 2) *quarantine* – карантин до уточнения (не хватает обязательных атрибутов);
- 3) *soft-fail* – пропуск с понижением уверенности (сохранение слабого сигнала с пометкой доверия).

Эффект сокращения времени реакции измеряют несколькими метриками:

- 1) задержка от поступления первичного сигнала до формирования события в интегрированном представлении;
- 2) время до выдачи рекомендации оператору;

- 3) в автоматизированных сценариях – время полного цикла от события до действия;
- 4) доля ложных тревог и доля событий, подтвержденных несколькими источниками.

Эффективность и экономический эффект. Помимо роста точности и сокращения времени реакции, семантическая интеграция даёт прикладной эффект в виде повышения эффективности процессов и снижения затрат. Это связано с тем, что решения начинают приниматься на более полной и согласованной картине. Система быстрее выявляет отклонения и причины, а также точнее определяет приоритеты действий. В результате уменьшаются простои, снижаются потери ресурсов и улучшаются целевые показатели работы организаций и городских служб [17].

Наиболее наглядно экономический эффект проявляется в сценариях, где есть измеримые операционные показатели. В промышленности это связано с предиктивным обслуживанием и управлением производством. Интеграция телеметрии оборудования, журналов ремонтов и данных ERP позволяет точнее определять состояние активов и планировать обслуживание до наступления отказа. При этом знаниевый граф помогает связать симптомы с конкретными узлами оборудования и историями неисправностей, а контур качества снижает риск ошибочных срабатываний из-за шумных датчиков. В городских системах эффект выражается в более точном распределении ресурсов. Это касается транспорта, коммунальной инфраструктуры и реагирования на инциденты. Когда события подтверждаются несколькими источниками, диспетчерские службы тратят меньше времени на проверку и меньше ресурсов на ложные выезды.

Экономическая результативность усиливается, когда система не просто фиксирует события, а поддерживает полный цикл управления. Сначала происходит обнаружение и подтверждение инцидента. Затем система предлагает оптимальное действие или автоматически запускает процедуру. После этого результаты фиксируются и используются для уточнения правил и веса доверия к источникам. Такая обратная связь постепенно снижает количество лишних действий и повышает точность планирования (рис. 12).



Рис. 12. Цикл повышения эффективности

Эффект внедрения можно представить через набор измеримых ключевых показателей эффективности (KPI - Key Performance Indicators). Чаще всего используют показатели, связанные с простоем и затратами, а также с качеством сервиса. Для промышленности это длительность простоев и расходы на обслуживание. Для городских сценариев это

время реагирования, нагрузка на службы и доля подтверждённых инцидентов. Для бизнес-сценариев это точность прогнозов и эффективность решений, влияющих на выручку и потери (табл. 4).

Таблица 4

Примеры KPI для оценки экономического эффекта семантической интеграции

Домен	KPI “до/после”	Как связан с интеграцией
Промышленность	простои, MTTR, затраты на ТО	точнее диагностика и планирование по мульти-источниковым данным
Умный город [18]	время реагирования, ложные выезды	подтверждение событий по нескольким каналам + качество данных
Бизнес/туризм	точность прогноза спроса, скорость пересчёта	fusion + потоковые признаки + обогащение текстов [19]
Здравоохранение	согласованность данных, время поиска контекста	KG объединяет ЭМК и знания, снижает ручные проверки

Улучшение качества данных и накопление знаний. Помимо прямых эффектов в виде ускорения реакции и повышения точности, семантическая интеграция даёт важный “вторичный” результат. Когда данные из разных систем приводятся к общей модели и связываются на уровне сущностей, становятся заметны ошибки, пробелы и противоречия, которые раньше оставались скрытыми внутри отдельных источников. Это приводит к росту качества данных в целом и формированию устойчивой базы знаний, которая может использоваться повторно в разных задачах СППР (табл. 5).

Таблица 5

Типовые проблемы качества и результат семантической интеграции

Проблема	Как выявляется	Что меняется после интеграции
Дубли сущностей	entity resolution	единые идентификаторы, меньше дублей
Пропуски атрибутов	SHACL minCount	карантин/дозаполнение
Некорректные значения	диапазоны/правила	отбрасывание/понижение доверия
Конфликты источников	provenance + веса	подтверждение по нескольким каналам

Улучшение качества происходит за счёт нескольких механизмов.

1. *Выявление несоответствий между источниками.* При сведении данных в единый граф знаний обнаруживаются расхождения в атрибутах, датах, единицах измерения и классификаторах. Например, один реестр может содержать устаревшую информацию, а фактические данные сенсоров или отчётов показывают иное состояние объекта.

2. *Дедупликация и уточнение идентичностей.* Entity resolution снижает количество дублей и “размноженных” сущностей. В результате справочники и реестры становятся более консистентными, а отчёты – более сопоставимыми.

3. *Формализация правил корректности.* SHACL и доменные правила переводят требования к данным из неявных “ожиданий” в проверяемые ограничения. Со временем это повышает дисциплину формирования данных в источниках, потому что ошибки начинают фиксироваться и возвращаться ответственным системам.

4. *Обратная связь и исправление источников.* Практика внедрений показывает, что интеграционная платформа становится не только “потребителем”, но и инструментом контроля качества. Нарушения правил качества могут автоматически формировать отчёты для владельцев источников или запускать процедуры корректировки.

Накопление знаний проявляется в том, что граф знаний постепенно превращается в “память” системы. Он хранит не только текущие состояния объектов, но и связи между ними, историю событий, подтверждения и происхождение фактов. Это облегчает объяснимость решений: система может показать, какие факты и из каких источников привели к рекомендации. Кроме того, знания становятся переносимыми между задачами. Одна и та же семантическая модель может использоваться для мониторинга, анализа причин, прогнозирования и планирования действий.

Практически это означает, что организация получает единое семантически согласованное пространство данных, которое со временем становится более чистым и более полезным. Даже если первоначальная цель внедрения была связана с оперативным реагированием, побочным результатом часто становится рост качества мастер-данных, повышение доверия к аналитике и снижение затрат на ручную сверку информации.

Ограничения и риски внедрения. Несмотря на выраженные преимущества, внедрение семантической интеграции в СППР связано с рядом ограничений и рисков, которые необходимо учитывать при проектировании и эксплуатации. В отличие от “классической” интеграции на уровне форматов и ETL, семантический подход затрагивает смысловые модели, правила качества и процессы принятия решений. Поэтому ошибки на этапе моделирования или эксплуатации могут приводить не только к техническим сбоям, но и к неправильным управленческим выводам.

Ключевые риски можно сгруппировать следующим образом.

1. *Трудоёмкость разработки и сопровождения онтологий и маппингов.* Построение доменной модели, согласование терминов и настройка отображений источников на общую схему требуют времени и участия экспертов. При расширении числа источников объём работ растёт, а качество маппингов становится критичным фактором для точности интеграции. Дополнительная сложность возникает при изменениях в источниках, когда схемы “дрейфуют” и маппинги нужно регулярно обновлять.

2. *Масштабируемость и производительность.* Знаниевые графы и семантические запросы дают гибкость и выразительность, но могут становиться узким местом при больших объёмах событий и высокой частоте обновлений. В реальном времени особенно чувствительны операции связывания сущностей, валидации и сложные запросы, которые требуют оптимизации, кэширования и иногда частичной материализации данных.

3. *Риск накопления “шума” в графе знаний.* Если потоковые данные поступают без строгого контроля качества или с низкой надёжностью источников, граф может быстро насыщаться ложными и дублирующимися фактами. Это ухудшает качество рекомендаций и усложняет последующую очистку. Поэтому необходимы политики обработки нарушений качества, а также стратегии жизненного цикла фактов, включая устаревание, архивирование и удаление.

4. *Сложности с конфликтами источников и неопределённостью.* В мульти-источниковой среде противоречия неизбежны. Если веса доверия заданы неверно или подтверждение по нескольким каналам реализовано формально, система может либо “не замечать” реальные события, либо генерировать ложные тревоги. Для снижения риска требуется аккуратная настройка правил слияния, хранение provenance и периодическая калибровка доверия к источникам.

5. *Эксплуатационные риски и дефицит компетенций.* Поддержка семантических компонентов требует специалистов, которые понимают одновременно предметную область, данные, графовые модели и потоковую инженерию. При недостатке компетенций растёт вероятность ошибок конфигурации, деградации качества и “зарастания” системы техническим долгом.

6. *Безопасность, конфиденциальность и соответствие требованиям.* Интеграция объединяет данные из разных контуров, что повышает риск несанкционированного доступа и повторной идентификации персональных данных. Особенно чувствительны медицинские и финансовые сценарии. Помимо технических мер защиты, требуется управление доступом на уровне сущностей и атрибутов, ведение аудита, а также корректная работа с правами и политиками использования данных.

7. *Риск потери прозрачности при использовании ML.* Когда интеграция включает модели машинного обучения для извлечения сущностей и принятия решений, возрастает вероятность скрытых ошибок из-за дрейфа данных и изменения контекста. Это требует мониторинга качества моделей, контроля версий и обязательного сохранения объяснимых следов в графе знаний.

С практической точки зрения минимизация указанных рисков достигается за счёт организационных и технических мер. Важно заранее определить границы доменной модели, выбрать реалистичный уровень детализации онтологии и внедрять интеграцию поэтапно, начиная с наиболее критичных сущностей и сценариев. Требуется формализовать правила качества и политики обработки нарушений, а также обеспечить хранение provenance и механизмов пересмотра доверия к источникам. Для производительности обычно применяется гибридная стратегия: материализация “горячих” данных, кэширование, ограничение тяжёлых запросов в реальном времени и перенос части аналитики в пакетный контур.

Заключение. В работе выполнен обзор технологий семантического анализа и интеграции гетерогенных информационных потоков в системах поддержки принятия решений. На основе анализа публикаций 2018–2025 гг. обобщены подходы к обеспечению семантической совместимости данных и показано, как они встраиваются в потоковые и гибридные архитектуры СППР.

Основные результаты можно сформулировать следующим образом.

1. Показано, что ключевым препятствием для интеграции в СППР остаётся не синтаксическая, а семантическая неоднородность данных. Даже при совпадении форматов различия в терминологии, единицах измерения, классификаторах и контексте приводят к ошибкам интерпретации и дублированию сущностей.

2. Систематизированы методы семантической интеграции и их роль в типовом конвейере: нормализация и гармонизация, семантическое моделирование, маппинг источников, связывание сущностей, контроль качества и семантическая валидация, семантические запросы и федерация. Показано, что максимальная практическая ценность достигается при комбинировании знаниевого слоя (онтологии/граф знаний) с алгоритмами слияния данных и механизмами доверия к источникам.

3. Отмечено, что для неструктурированных потоков решающую роль играют методы NLP и компьютерного зрения, которые обеспечивают извлечение сущностей и отношений и превращают текст/мультимедиа в формализованные факты, пригодные для включения в граф знаний и последующего анализа.

4. На основании рассмотренных кейсов показано, что применение семантической интеграции в СППР повышает полноту и согласованность ситуационной картины, сокращает время реакции за счёт потокового конвейера и снижает долю ложных срабатываний за счёт валидации и подтверждения по нескольким каналам. В ряде прикладных работ фиксируются улучшения точности порядка 15–20% и существенное снижение времени отклика при многоисточниковом объединении данных [5, 20].

5. Выделены ограничения внедрения: трудоёмкость создания и сопровождения онтологий и маппингов, проблемы масштабируемости графов и запросов в реальном времени, риск накопления “шума” при слабом контуре качества, а также требования к безопасности и управлению доступом при интеграции чувствительных данных.

Практическая значимость работы состоит в том, что предложенная структура методов и конвейерная логика интеграции могут служить ориентиром при проектировании СППР для умного города, промышленности и здравоохранения. Перспективы дальнейших исследований связаны с автоматизацией семантического согласования (alignment/mapping), динамическим обновлением моделей при дрейфе схем источников, а также с развитием механизмов объяснимости на основе provenance и графовых связей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Cook M. Semantic Integration // *Diffbot Knowledge Graph Glossary*. – 2025. – URL: <https://blog.diffbot.com/knowledge-graph-glossary/semantic-integration/>.
2. Zuiev P., Kuchansky A., Biloshchytskyi A. et al. Development of Complex Methodology of Processing Heterogeneous Data in Intelligent Decision Support Systems // *Eastern-European Journal of Enterprise Technologies*. – 2020. – Vol. 4. – P. 14-23. – DOI: 10.15587/1729-4061.2020.209154.
3. Putrama I.M., Martinek P. Heterogeneous data integration: Challenges and opportunities // *Data in Brief*. – 2024. – Vol. 56. – Art. 110853. – DOI: 10.1016/j.dib.2024.110853.
4. Balaha F., Albinali H., Alrabiah H., Ali M., Bahroun Z. An analytical review of data integration for decision support in smart manufacturing // *Decision Analytics Journal*. – 2025. – Vol. 17. – Art. 100647. – DOI: 10.1016/j.dajour.2025.100647.
5. Rajabi E., Kafaie S. Knowledge Graphs and Explainable AI in Healthcare // *Information*. – 2022. – Vol. 13, No. 10. – Art. 459. – DOI: 10.3390/info13100459.
6. Rathore M.M., Ahmad A., Paul A., Rho S. Towards smart transportation: data integration and mining for large scale data // 2016 IEEE Region 10 Conference (TENCON Spring). – 2016. – DOI: 10.1109/TENCONSpring.2016.7519392.
7. R2RML: RDB to RDF Mapping Language. – W3C Recommendation, 27.09.2012. – URL: <https://www.w3.org/TR/r2rml/>.
8. JSON-LD 1.1. – W3C Recommendation, 16.07.2020. – URL: <https://www.w3.org/TR/json-ld11/>.
9. Gao F., Ali M. I., Mileo A. Semantic Discovery and Integration of Urban Data Streams // *CEUR Workshop Proceedings*. – 2014. – Vol. 1280. – P. 15-30. – URL: <https://ceur-ws.org/Vol-1280/>.
10. Hashem I.A.T., Chang V., Anuar N. B. et al. The role of big data in smart city // *International Journal of Information Management*. – 2016. – Vol. 36, No. 5. – P. 748-758. – DOI: 10.1016/j.ijinfomgt.2016.05.002.
11. Bamgboye O., Liu X., Cruickshank P. Semantic-Driven Approach for Validation of IoT Streaming Data in Trustable Smart City Decision-Making and Monitoring Systems // *Big Data and Cognitive Computing*. – 2025. – Vol. 9, No. 4. – Art. 108. – DOI: 10.3390/bdcc9040108.
12. Ribeiro M.B., Braghetto K.R.A Data Integration Architecture for Smart Cities // *Anais do XXXVI Simpósio Brasileiro de Bancos de Dados (SBBDD)*. – 2021. – DOI: 10.5753/sbbd.2021.18030.
13. OWL 2 Web Ontology Language Document Overview (Second Edition). – W3C Recommendation, 11.12.2012. – URL: <https://www.w3.org/TR/owl2-overview/>.
14. SPARQL 1.1 Query Language. – W3C Recommendation, 21.03.2013. – URL: <https://www.w3.org/TR/sparql11-query/> (дата обращения: 07.12.2025).
15. SPARQL 1.1 Query Language. – W3C Recommendation, 21.03.2013. – URL: <https://www.w3.org/TR/sparql11-query/> (дата обращения: 07.12.2025).
16. Shapes Constraint Language (SHACL). – W3C Recommendation, 20.07.2017. – URL: <https://www.w3.org/TR/shacl/>.
17. Михайлова С.С., Будаев Е.С. и др. Интеграция разнородных данных пространственного развития региона в СППР // *Нелинейный мир*. – 2025. – Т. 23, № 4. – С. 107-120.
18. Psyllidis A., Bozzon A., Vocconi S., Titos Bolivar C. A Platform for Urban Analytics and Semantic Data Integration in City Planning // *Computer-Aided Architectural Design Futures. The Next City – New Technologies and the Future of the Built Environment* / eds. G. Celani, D.M. Sperling, J.M.S. Franco. – Cham: Springer, 2015. – (Communications in Computer and Information Science; Vol. 527). – P. 21-36. – DOI: 10.1007/978-3-662-47386-3_2.
19. Song M. Research on intelligent decision support platform for tourism enterprises based on multi-source heterogeneous data fusion // *Scientific Reports*. – 2025. – Vol. 15, No. 1. – Art. 39810. – DOI: 10.1038/s41598-025-23486-x.
20. Осипов В.П., Сивакова Т.В., Судаков В.А., Нечаев Ю.И. Интеллектуальное ядро системы поддержки принятия решений // *Препринты ИПМ им. М.В. Келдыша*. – 2018. – № 205. – 23 с. – URL: <https://library.keldysh.ru/preprint.asp?id=2018-205>.

REFERENCES

1. Cook M. Semantic Integration, *Diffbot Knowledge Graph Glossary*, 2025. Available at: <https://blog.diffbot.com/knowledge-graph-glossary/semantic-integration/>.
2. Zuiev P., Kuchansky A., Biloshchytskyi A. et al. Development of Complex Methodology of Processing Heterogeneous Data in Intelligent Decision Support Systems, *Eastern-European Journal of Enterprise Technologies*, 2020, Vol. 4, pp. 14-23. DOI: 10.15587/1729-4061.2020.209154.
3. Putrama I.M., Martinek P. Heterogeneous data integration: Challenges and opportunities, *Data in Brief*, 2024, Vol. 56, Art. 110853. DOI: 10.1016/j.dib.2024.110853.

4. Balaha F., Albinali H., Alrabiah H., Ali M., Bahroun Z. An analytical review of data integration for decision support in smart manufacturing, *Decision Analytics Journal*, 2025, Vol. 17, Art. 100647. DOI: 10.1016/j.dajour.2025.100647.
5. Rajabi E., Kafaie S. Knowledge Graphs and Explainable AI in Healthcare, *Information*, 2022, Vol. 13, No. 10, Art. 459. DOI: 10.3390/info13100459.
6. Rathore M.M., Ahmad A., Paul A., Rho S. Towards smart transportation: data integration and mining for large scale data, *2016 IEEE Region 10 Conference (TENCON Spring)*, 2016. DOI: 10.1109/TENCONSpring.2016.7519392.
7. R2RML: RDB to RDF Mapping Language. W3C Recommendation, 27.09.2012. Available at: <https://www.w3.org/TR/r2rml/>.
8. JSON-LD 1.1. W3C Recommendation, 16.07.2020. Available at: <https://www.w3.org/TR/json-ld11/>.
9. Gao F., Ali M. I., Mileo A. Semantic Discovery and Integration of Urban Data Streams, *CEUR Workshop Proceedings*, 2014, Vol. 1280, pp. 15-30. Available at: <https://ceur-ws.org/Vol-1280/>.
10. Hashem I.A.T., Chang V., Anuar N. B. et al. The role of big data in smart city, *International Journal of Information Management*, 2016, Vol. 36, No. 5, pp. 748-758. DOI: 10.1016/j.ijinfomgt.2016.05.002.
11. Bamgboye O., Liu X., Cruickshank P. Semantic-Driven Approach for Validation of IoT Streaming Data in Trustable Smart City Decision-Making and Monitoring Systems, *Big Data and Cognitive Computing*, 2025, Vol. 9, No. 4, Art. 108. DOI: 10.3390/bdcc9040108.
12. Ribeiro M.B., Braghetto K.R.A. Data Integration Architecture for Smart Cities, *Anais do XXXVI Simpósio Brasileiro de Bancos de Dados (SBB D)*, 2021. DOI: 10.5753/sbbd.2021.18030.
13. OWL 2 Web Ontology Language Document Overview (Second Edition). W3C Recommendation, 11.12.2012. Available at: <https://www.w3.org/TR/owl2-overview/>.
14. SPARQL 1.1 Query Language. W3C Recommendation, 21.03.2013. Available at: <https://www.w3.org/TR/sparql11-query/> (accessed 07 December 2025).
15. SPARQL 1.1 Query Language. W3C Recommendation, 21.03.2013. Available at: <https://www.w3.org/TR/sparql11-query/> (accessed 07 December 2025).
16. Shapes Constraint Language (SHA CL). W3C Recommendation, 20.07.2017. Available at: <https://www.w3.org/TR/shacl/>.
17. Mikhaylova S.S., Budaev E.S. i dr. Integratsiya raznorodnykh dannykh prostranstvennogo razvitiya regiona v SPPR [Integration of heterogeneous data on the spatial development of a region in decision-making systems], *Nelineyny mir* [Nonlinear World], 2025, Vol. 23, No. 4, pp. 107-120.
18. Psyllidis A., Bozzon A., Bocconi S., Titos Bolivar C. A Platform for Urban Analytics and Semantic Data Integration in City Planning, *Computer-Aided Architectural Design Futures. The Next City – New Technologies and the Future of the Built Environment*, eds. G. Celani, D.M. Sperling, J.M.S. Franco. Cham: Springer, 2015. – (Communications in Computer and Information Science; Vol. 527), pp. 21-36. DOI: 10.1007/978-3-662-47386-3_2.
19. Song M. Research on intelligent decision support platform for tourism enterprises based on multi-source heterogeneous data fusion, *Scientific Reports*, 2025, Vol. 15, No. 1, Art. 39810. DOI: 10.1038/s41598-025-23486-x.
20. Osipov V.P., Sivakova T.V., Sudakov V.A., Nechaev Yu.I. Intellektual'noe yadro sistemy podderzhki prinyatiya resheniy [Intelligent core of the decision support system], *Preprinty IPM im. M.V. Keldysha* [Preprints of the Keldysh Institute of Problems of Applied Mathematics], 2018, No. 205, 23 p. Available at: <https://library.keldysh.ru/preprint.asp?id=2018-205>.

Гапочка Владислав Валерьевич – Кубанский государственный университет; e-mail: gapochka2000@mail.ru; тел.: +89883667270; кафедра вычислительных технологий; аспирант.

Полупанова Елена Евгеньевна – Кубанский государственный университет; e-mail: jienka@mail.ru; г. Краснодар, Россия; тел.: +79284013301.; кафедра вычислительных технологий; к.т.н.; доцент; ORCID: <https://orcid.org/0000-0002-0364-1132>.

Gapochka Vladislav Valerievich – Kuban State University; e-mail: gapochka2000@mail.ru; Krasnodar, Russia; phone: +89883667270, the Department of Computational Technologies; graduate student.

Polupanova Elena Evgenievna – Kuban State University; e-mail jienka@mail.ru; Krasnodar, Russia; phone: +79284013301; the Department of Computational Technologies; cand. of eng. sc.; associate professor; ORCID: <https://orcid.org/0000-0002-0364-1132>.