© Запотылько В. С., Свиридов С. А., 2026, Океанологические исследования, 2026, Том 54, № 3, С. 173–198

УДК 551.46.071:004.9:528 DOI: 10.29006/1564-2291.JOR-2026.54(3).9

РАЗРАБОТКА ПРОФИЛЯ МЕТАДАННЫХ ОКЕАНОЛОГИЧЕСКОЙ ЭКСПЕДИЦИИ НА ОСНОВЕ ГОСТ Р 57668-2017 И ЕГО ПРОГРАММНАЯ РЕАЛИЗАЦИЯ В METADATAEXP

В. С. Запотылько, С. А. Свиридов

Институт океанологии им. П. П. Ширшова РАН,
Россия, 117997, Москва, Нахимовский проспект, д. 36,
e-mail: esimo@ocean.ru

В статье представлен специализированный профиль метаданных для океанологических экспедиционных данных, разработанный в соответствии с требованиями ГОСТ Р 57668-2017 и международного стандарта ISO 19115. Центральным элементом работы выступает *UML*[1]-модель, формализующая структуру метаданных и обеспечивающая семантическую совместимость с международными информационными инфраструктурами (*EMODnet*, *OBIS*, *SeaDataNet*) и соответствие принципам *FAIR*[2] (*Findable*, *Accessible*, *Interoperable*, *Reusable*). Модель построена вокруг класса *MD_Metadata*, выступающего контейнером для всей описательной информации, и включает четыре ключевых блока: идентификационные сведения (*MD_Identification*), данные о способе получения информации (*MI_AcquisitionInformation*), линию происхождения данных (*LI_Lineage*) и условия распространения (*MD_Distribution*). Такая структура позволяет не только зафиксировать статические характеристики массива (пространственные, временные, тематические границы, назначение, статус), но и описать динамические аспекты его жизненного цикла – от используемых приборов и платформ до последовательности шагов обработки и источников исходных материалов. Отличительной особенностью предложенного профиля является акцент на «цифровой истории жизни» каждого массива наблюдений: в модель включены элементы для формализованного описания временного разрешения измерений, пространственного охвата, методологии сбора и протоколов калибровки оборудования – параметров, критически важных для верификации результатов и вторичного анализа. Реализация профиля в виде *UML*-диаграммы (с выделением классов *MD_Metadata*, *MD_Identification*, *EX_Extent*, *TM_Duration* и др.) обеспечивает строгую семантическую структуру, пригодную для автоматической валидации, поиска и интеграции данных. Практическая значимость работы заключается в создании инструмента, позволяющего преодолеть фрагментацию океанологических данных и повысить их пригодность для междисциплинарных и международных исследований. Новизна работы заключается в создании первого специализированного профиля метаданных для океанологических экспедиций, полностью соответствующего ГОСТ Р 57668-2017, и его программной реализации, обеспечивающей практическое применение. Предложенный профиль может быть использован при проектировании баз данных, разработке сервисов каталогизации и построении распределенных инфраструктур управления научными данными.

Ключевые слова: информационная система, экспедиционные исследования, база данных, *MetadataExp*, *ODBASE*, *UML*, профиль метаданных, оценка качества данных, стандарты метаданных, ГОСТ Р 57668-2017, ISO

[1] UML (Unified Modeling Language) – унифицированный язык моделирования, который используется для графического описания и визуализации различных систем и процессов.

[2] FAIR-принципы (Findable, Accessible, Interoperable, Reusable) – это набор стандартов для научных данных, которые делают их: находимыми, доступными, совместимыми, повторно используемыми.

Введение

Современное развитие океанологических исследований характеризуется экспоненциальным ростом объемов получаемых данных, обусловленным расширением парка измерительных приборов и переходом к автоматизированным методам наблюдений. В этих условиях особую значимость приобретают процессы систематизации и семантической унификации данных, обеспечивающие их долговременное хранение, верификацию и многократное использование в рамках междисциплинарных и международных научных проектов.

В отечественной океанологической практике накоплен значительный опыт построения информационных систем, ориентированных на сбор, хранение и предоставление доступа к экспедиционным данным. Так, в работах (Голик, Фищенко и др., 2004; Фищенко, Голик и др., 2007) представлена концепция корпоративной океанографической геоинформационной системы ТОИ ДВО РАН, обеспечивающей интеграцию спутниковых и подспутниковых наблюдений по северо-западной части Тихого океана; при этом архитектурные решения были направлены преимущественно на объединение разнородных источников данных и их визуализацию в рамках единой ГИС-платформы. Аналогичные задачи решались в монографии (Дмитриева, Ростов, 2004), где основное внимание уделено организации баз данных океанографических наблюдений в северной части Тихого океана, процедурам обработки первичных материалов и формированию информационной продукции для конечных пользователей.

Важный вклад в развитие технологий хранения и структурирования океанологических данных внесли работы специалистов ИО РАН, заложившие основы отечественной практики управления научными данными. В частности, в начале 2000-х годов был реализован ряд инициатив, направленных на создание специализированных хранилищ и форматов данных, адаптированных к специфике океанологических наблюдений. Так, в работе (Осипенко, Свиридов и др., 2001) обосновано применение технологий хранилищ данных для обработки океанологической информации, что позволило систематизировать разнородные массивы наблюдений и обеспечить их долговременное хранение. Дальнейшее развитие этого направления представлено в публикациях А. А. Метальникова с соавторами, где описана архитектура хранилища *OceanDB* (Метальников, Осипенко и др., 2005), предложен специализированный формат данных *OceanML* (Метальников, Осипенко и др., 2003). Эти разработки обеспечили структурированное хранение параметров станций, методов измерений и условий наблюдений, заложив технологическую основу для последующей стандартизации описаний данных. Несмотря на то, что упомянутые разработки не использовали международные стандарты описания метаданных (ISO 19115) и не обеспечивали прямой интеграции в глобальные информационные системы, они доказали практическую значимость формализованного подхода к описанию океанологических данных, заложили основу для создания стандартизированных профилей метаданных, создали предпосылки для перехода к современным требованиям *FAIR* и подготовили базу для соответствия национальному стандарту ГОСТ Р 57668-2017.

В более поздних работах (Еремеев, Халиулин и др., 2014; Коваленко, Мельников и др., 2015) фокус смещается на программно-математическое обеспечение банков данных, механизмы контроля качества и оперативного доступа, а также на построение специализированных баз метаданных для отдельных акваторий (в частности, Черного моря).

Отдельного упоминания заслуживает опыт создания отечественных систем метаданных, в частности Единой государственной системы информации об обстановке в Мировом океане (ЕСИМО) (Вязилов, 2005; Вязилов, Михайлов, Кобелев, 2008). В рамках этой системы еще в 2000-х годах была разработана база метаданных на основе стандарта ISO 19115:2003, что позволило описать более двух тысяч информационных ресурсов (Вязилов, Михайлов, 2005). Однако между редакцией ISO 19115:2003, использованной в ЕСИМО, и редакцией ISO 19115-1:2014, на основе которой разработан ГОСТ Р 57668-2017, существуют принципиальные структурные различия (Danko, 2014). В частности, в версии 2014 г. было добавлено 194 новых подэлемента и исключено 66, а также изменена структура ряда ключевых классов, включая *MD_Metadata*, *CI_Responsibility* и *MD_Identifier*. Например, *fileIdentifier* (*CharacterString*) заменен на *metadataIdentifier* (*MD_Identifier*), *language* – на *defaultLocale* (*PT_Locale*), а *parentIdentifier* – на *parentMetadata* (*CI_Citation*) (ISO 19115-1:2014. Annex G). Эти изменения носят не поверхностный, а семантический характер: они затрагивают правила идентификации, локализации и наследования метаданных. Поэтому прямое использование структур ЕСИМО, созданных по ISO 19115:2003, без их предварительной адаптации к модели ISO 19115-1:2014 невозможно. Вместе с тем, такие элементы ЕСИМО, как Единый словарь параметров, справочники организаций и форматов, могут быть рассмотрены как источник для гармонизации создаваемого профиля с отечественной практикой – при условии их доработки. Анализ показывает, что справочники ЕСИМО в текущем виде не применимы для Института океанологии или слабо применимы для формирования профилей метаданных по ГОСТ Р 57668-2017: во-первых, отсутствует описание форматов и структуры самих справочников в открытых источниках; во-вторых, нет возможности их загрузки, что делает неясным их использование в программных продуктах; в-третьих, записи не сопровождаются параметрами, аналогичными справочнику BODC (уникальные идентификаторы, расширенные описания, ссылки на актуальные версии, история изменений). Например, объект «Организации» в ЕСИМО содержит около десяти полей (точнее 11 или 12 – по разным источникам), из которых лишь пять относятся к контактной информации (Вязилов, 2005; Сведения…: apicrafter.ru), тогда как ГОСТ Р 57668-2017 (п. 6.6.2, рис. 20) для раздела «Информация о стороне и ее роли» требует значительно большего набора атрибутов, включая роль, телефон и его тип, структурированный адрес, тип контакта, период доступности, дополнительные инструкции и сведения об онлайн-ресурсах. Таким образом, справочники ЕСИМО могут быть использованы как отправная точка для гармонизации, но требуют существенной доработки: добавления машиночитаемых описаний, уникальных идентификаторов, версионирования и обеспечения доступа к данным в стандартизованных форматах (ОТС ЕСИМО, 2011).

Международный опыт также демонстрирует широкий спектр подходов к управлению морскими данными, при этом эволюция решений прослеживается от локальных систем к панъевропейским инфраструктурным проектам. Характерным примером является инициатива *SeaDataNet* (Schaap, Lowry, 2010), в рамках которой решалась задача преодоления фрагментации морских данных за счет создания распределенной инфраструктуры с централизованными механизмами гармонизации метаданных, контроля качества и стандартизации форматов обмена. Проект наглядно показал, что ключевым барьером для интеграции данных является не столько техническая доступность массивов, сколько отсутствие унифицированной семантики описания ресурсов, позволяющей обеспечить сквозную интероперабельность между центрами данных. Вместе с тем, несмотря на высокий уровень проработки вопросов каталогизации и валидации по основным метаданным в сервисах – *CSR*[3], *EDMED*[4], *EDMERP*[5], *EDIOS*[6] и *EDMO*[7] (Pecci et al., 2020), в архитектуре *SeaDataNet* основной акцент делался на инфраструктурных и технологических решениях, а не на формализации профиля метаданных в соответствии с ISO 19115 и принципами *FAIR*.

На уровне региональных научных задач аналогичные вызовы решались при создании специализированных океанографических баз данных. В частности, в работе (Smirnov, Korablev, 2010) описан опыт формирования региональной БД по Норвежскому морю – от первичных наблюдений до климатического массива. Авторы подчеркивают, что переход к интегрированному продукту потребовал тщательной проработки схем описания параметров станций, методов измерений и показателей качества данных, поскольку именно эти элементы определяли сопоставимость и пригодность данных для вторичного анализа. Однако и в этом случае задача стандартизации метаданных решалась в рамках внутренних форматов и предметно-ориентированных моделей, без опоры на единые профили, совместимые с международными системами обмена.

Параллельно в ряде исследований разрабатывались технологические решения для структурирования морских данных на основе *XML*-форматов (Jiang et al., 2009; Wang Yan et al., 2012; Ikegami et al., 2003), которые являлись частью более широкого движения *MarineXML*, инициированного ЮНЕСКО/IOC и *OceanML*, развиваемого в ИО РАН (Метальников, Осипенко и др., 2003). Это позволяло обеспечить структурную совместимость, но не гарантировало семантическую интероперабельность в соответствии с международными профилями метаданных, поскольку семантические определения элементов данных

[3] *CSR* (Cruise Summary Reports) – Отчет о результатах морских экспедиций, который служит для документирования измерений и проб, собранных во время морских исследований.

[4] *EDMED* (European Directory of Marine Environmental Data) – Европейский каталог морских экологических данных, созданный для систематизации и поиска информации о морских исследованиях.

[5] *EDMERP* (European Directory of Marine Environmental Research Projects) – Европейский каталог морских исследовательских проектов, который охватывает широкий спектр научных дисциплин.

[6] *EDIOS* (European Directory of the Initial Ocean-observing System) – Европейский каталог океанографических систем наблюдения, созданный в рамках инициативы EuroGOOS.

[7] *EDMO* (European Directory of Marine Organisations) – это европейский каталог морских организаций, который содержит актуальную информацию о научно-исследовательских институтах, центрах хранения данных, агентствах мониторинга, правительственных и частных организациях, участвующих в океанографических исследованиях.

XML находятся в форме документа, не представлены явно и не используются совместно между системами, что приводит к возможности ошибок в семантических интерпретациях. В других работах основное внимание уделялось интеграции метаданных с предметными моделями пространственных данных – так, (Isenor, Spears, 2013) комбинировали *Arc Marine Framework*[8] с географическими метаданными для поддержки акустического моделирования океана, фокусируясь на прикладной совместимости, а не на унификации описания ресурсов. Отдельные исследования посвящены управлению данными в специализированных инфраструктурных проектах: например, (Xiu Li et al., 2013) разработали систему управления данными для сети донных обсерваторий, где приоритет был отдан потоковой передаче и хранению, а не формализованному описанию жизненного цикла данных. Кроме того, в публикациях, посвященных автоматизации пространственных метаданных (Olfat et al., 2012), рассматриваются методы генерации и валидации метаданных, но без привязки к конкретным профилям, совместимым с *FAIR*-принципами. Наконец, в ранних работах по стратегии управления данными для программ океанических исследований (Mineart, Klein, 2002) закладывались концептуальные основы каталогизации и доступа, однако технологические и нормативные рамки того времени не позволяли реализовать полноценную семантическую модель описания данных.

Современные исследования усиливают акцент на *FAIR*-принципах как ключевом условии устойчивого развития морских данных. В частности, (Tanhua et al., 2019) формулируют концепцию *FAIR*-сервисов для океанических данных, подчеркивая необходимость машиночитаемых метаданных, прозрачных идентификаторов и стандартизованных протоколов доступа. Дальнейшее развитие этой концепции раскрывается в работе тех же авторов (Tanhua et al., 2021), где предложена целостная модель *FAIR*-продуктов для морской науки, включающая требования к описанию происхождения данных, методов измерений и условий их повторного использования. Эти работы задают методологическую основу для проектирования профилей метаданных, ориентированных не только на хранение, но и на активное повторное использование данных в автоматизированных аналитических цепочках.

Важным направлением является также модульная каталогизация морских данных, позволяющая гибко адаптировать метаданные под специфику наблюдений и исследовательских задач. Adam Marc Leadbetter с соавторами в 2020 г. предлагают модульный подход к каталогизации, в котором отдельные блоки метаданных (пространство, время, метод, качество) могут комбинироваться в зависимости от типа ресурса, что повышает масштабируемость и удобство интеграции в разнородные инфраструктуры (Leadbetter et al., 2020). Такой подход согласуется с практикой стандартизации метаданных на базе *ISO* и *OGC*[9], в том числе в смежных областях, и Marcos Jiménez с соавторами еще в 2014 г.

[8] Arc Marine Framework (или ArcGIS Marine Data Model) – это специализированная модель базы данных, разработанная для морской ГИС-среды, создана в 2001 г. исследователями из Орегонского университета, Университета Дьюка, NOAA, Датского гидрологического института и ESRI. Представляет собой шаблон базы геоданных для морских исследований.

[9] Стандарты OGC – Open Geospatial Consortium, набор открытых спецификаций для работы с геопространственными данными.

реализуют систему метаданных, формализующую параметры измерений, калибровки и условий съемки (Jiménez et. al., 2014). Эти решения подтверждают универсальность и масштабируемость стандартов ISO 19115 и *OGC* для описания разнородных научных данных, включая океанологические наблюдения.

В контексте биогеографических и экологических данных значимый опыт накоплен в рамках инфраструктуры *OBIS*[10], которая демонстрирует возможности интеграции разнородных массивов в единую систему обмена при условии строгого следования стандартам метаданных и семантической унификации описаний (Klein et. al., 2019). Опыт, приобретенный при создании и развитии *OBIS*, подчеркивает важность не только технической совместимости, но и согласованной терминологии, контролируемых словарей и прослеживаемости данных – элементов, которые должны быть заложены на уровне профиля метаданных.

Вместе с тем, в указанных отечественных и зарубежных исследованиях не ставилась задача формирования профиля метаданных в соответствии с требованиями ГОСТ Р 57668-2017 («Пространственные данные. Метаданные») или его международного прототипа ISO 19115:2014 (“Geographic information – Metadata”), регламентирующих унифицированную модель описания пространственных ресурсов. Это обусловлено как хронологическими рамками публикаций (большинство работ относится к 2002–2013 гг., тогда как ГОСТ Р 57668-2017 утвержден в 2017 г.), так и методологическими приоритетами: акцент делался на локальных форматах хранения, *XML*-сериализации[11], интеграции с ГИС-платформами или на предметно-ориентированных моделях, а не на обеспечении семантической совместимости и интеграции в международные информационные инфраструктуры (*EMODnet*, *OBIS*, *SeaDataNet*) в соответствии с принципами *FAIR* (*Findable*, *Accessible*, *Interoperable*, *Reusable*). В частности, в них не реализованы механизмы формализованного описания таких критически важных для вторичного анализа параметров, как временное разрешение наблюдений (*temporalResolution*), пространственный охват (*extent* в терминах ISO 19115), а также детализация методологии измерений и протоколов калибровки оборудования, необходимых для верификации результатов. Подробный обзор развития и эволюции международных стандартов метаданных географической информации (GI-metadata) с акцентом на деятельность Технического комитета ISO/TC 211 «географическая информация/Геоматика» представлен в работе (Brodeur et. al., 2019).

Таким образом, несмотря на высокий уровень проработки вопросов хранения, доступа и предварительной обработки океанологических данных, существующие отечественные и зарубежные разработки не обеспечивают полного соответствия современным требованиям к интероперабельности и машиночитаемости метаданных.

---

[10] *OBIS – Ocean Biodiversity Information System*, ранее *Ocean Biogeographic Information System*, – глобальная инфраструктура для сбора, интеграции и предоставления открытых данных о морском биоразнообразии: о том, где, когда и в каком количестве встречаются морские виды.

[11] *XML*-сериализация – процесс преобразования структурированных данных (чаще всего объектов в программе) в *XML*-формат, чтобы их можно было хранить, передавать по сети или обмениваться между разными системами.

Данное обстоятельство обусловливает необходимость разработки специализированного профиля метаданных, адаптированного к специфике экспедиционных океанологических наблюдений и базирующегося на стандартизованной модели описания ресурсов.

В настоящей работе предложен профиль метаданных для океанологических экспедиционных данных, разработанный на основе ГОСТ Р 57668-2017, который представляет собой адаптированную версию международного стандарта ISO 19115 с учетом национальных особенностей. Данный профиль метаданных соответствует требованиям ISO 19115 и современным принципам FAIR. Его отличительной особенностью является формализация не только статических характеристик наборов данных, но и динамических аспектов их получения и обработки, что позволяет сформировать «цифровую историю жизни» каждого массива наблюдений и обеспечить соответствие принципам *FAIR*. Реализация профиля в рамках *UML*-модели (с выделением классов *MD_Metadata, MD_Identification, EX_Extent, TM_Duration* и др.) обеспечивает строгую семантическую структуру, пригодную для автоматической валидации, поиска и интеграции данных.

Методы

Выбор и обоснование нормативной базы

В качестве нормативной основы для формирования профиля метаданных океанологической экспедиции принят ГОСТ Р 576682017 «Пространственные данные. Метаданные», представляющий собой национальную адаптацию международного стандарта ISO 191151:2014. Выбор стандарта обусловлен тремя ключевыми факторами, релевантными для океанологических исследований:

– Совместимость с международными системами обмена данными. Профиль, построенный на базе ISO 19115, обеспечивает интероперабельность с крупными инфраструктурными проектами (*EMODnet*, *OBIS*, *Argo*), что критично для интеграции российских экспедиционных данных в глобальные научные сети.

– Полнота описания жизненного цикла данных. Стандарт включает пакет *LI_Lineage*, позволяющий фиксировать происхождение, обработку и преобразование данных – это напрямую отвечает требованиям воспроизводимости научных результатов, особенно в условиях многоэтапного анализа океанологических измерений.

– Гибкость структуры. За счет механизма расширений (*MD_MetadataExtensionInformation*) и возможности выборочного использования пакетов стандарт допускает специализацию под конкретные научные задачи, не теряя формальной строгости.

В настоящем разделе описана методика адаптации стандарта к специфике морских экспедиций ИО РАН, а также алгоритм формирования минимально достаточного профиля.

Методика формирования профиля метаданных для морской экспедиции

Профиль метаданных в данном контексте понимается как специализация стандарта для конкретного научного сообщества: в него включаются только те элементы, которые необходимы для однозначной интерпретации и повторного использования данных, а избыточные поля исключаются либо переводятся в опциональные.

Для формирования профиля был проведен анализ 12 экспедиционных отчетов ИО РАН за 2018–2023 гг. с целью выявления типичных метаданных, реально используемых при поиске, цитировании и повторном анализе данных. На основе этого анализа были определены три критерия отбора полей:

– Критерий воспроизводимости. Поле считается обязательным, если его отсутствие делает невозможным повторение эксперимента или проверку расчетов. Например, указание типа CTD-зонда и методики расчета солености критично для сопоставления данных разных экспедиций.

– Критерий однозначной идентификации. Поле необходимо, если оно позволяет однозначно найти и процитировать ресурс: название экспедиции, даты, район работ, статус, уникальный идентификатор.

– Критерий практической реализуемости. Поле включается в профиль, только если его заполнение выполнимо силами экипажа и научных сотрудников непосредственно в рейсе без привлечения внешних специалистов.

На основании этих критериев сформирован минимально достаточный профиль, включающий следующие пакеты (таблица 1):

Табл. 1 – Состав пакетов минимального профиля метаданных морской экспедиции

Пакет	Назначение	Обязательность	Ключевые поля
MD_Metadata	Корневой пакет, агрегирующий остальные компоненты	Обязателен	metadataIdentifier, contact, dateInfo, defaultLocale
MD_Identification	Уникальная идентификация экспедиции	Обязателен	citation (название), abstract (краткое описание), purpose (цель), extent (пространственно временной охват), topicCategory, status
MI_ AcquisitionInformation (ISO 191152)	Описание методов и средств сбора данных	Обязателен (через расширение)	platform (судно), instrument (оборудование), operation (станции, профили)
LI_Lineage	Происхождение и обработка данных	Обязателен	statement (общее описание), processStep (шаги обработки), source (исходные данные)
MD_Distribution	Условия доступа и распространения	Обязателен	distributor (поставщик), distributionFormat (формат), description (способы доступа)

Остальные пакеты стандарта (например, *MD_Constraints*, *MD_MaintenanceInformation*, *MD_SpatialRepresentation*) в базовый профиль не включены, но могут быть добавлены на последующих этапах при необходимости. Такой подход обеспечивает баланс между полнотой описания и удобством заполнения в экспедиционных условиях.

Детализация обязательных пакетов и правила валидации полей

Пакет ***MD_Metadata*** *я*вляется корневым и агрегирует все остальные пакеты. Обязательные поля:

– *metadataIdentifier* – уникальный идентификатор метаданных (рекомендуется использовать *UUID*[12], либо внутренний регистрационный номер экспедиции ИО РАН).

– *contact* – ответственное лицо за метаданные (организация, ФИО, должность, контактные данные; роль указывается через классификатор, например, *custodian* или *pointOfContact*).

– *dateInfo* – дата создания метаданных (в формате ISO 8601).

– *defaultLocale* – локальный язык описания (по умолчанию – «ru»).

Пакет ***MD_Identification*** обеспечивает однозначную идентификацию экспедиции. Обязательные поля и правила заполнения:

– *citation* – название экспедиции (в свободной форме, но с указанием номера рейса и судна, например: «Экспедиция НИС «Академик Мстислав Келдыш», рейс 85»).

– *abstract* – краткое описание целей и задач экспедиции (не более 250 слов).

– *purpose* – научное обоснование экспедиции (цель, гипотезы, связь с программами исследований).

– *extent* – пространственновременной охват: географические границы (четыре точки широты/долготы), даты начала и окончания работ.

– *topicCategory* – тематическая категория (используются коды: *oceans*, *geoscientificInformation*; при необходимости добавляются *climatologyMeteorologyAtmosphere*).

– *status* – статус ресурса (единственное допустимое значение для завершенной экспедиции – ***completed***).

Пакет ***MI_AcquisitionInformation*** **(**через *MD_MetadataExtensionInformation)* – позволяет создать структурированное описание методов и средств измерений непосредственно в метаданных. Для каждого типа измерений возможно формирование отдельной записи. Обязательные элементы:

– *platform* – платформа (судно): название, позывной, основные характеристики.

– *instrument* – оборудование: тип прибора, модель, серийный номер, дата последней калибровки.

– *operation* – операции: перечень типов работ (гидрологические станции, профили, буксировки и т. п.) с указанием количества и методики.

[12] *UUID* (*Universally Unique Identifier*) – универсальный уникальный идентификатор, который представляет собой стандартизированный код для однозначной идентификации объектов, данных или сущностей в компьютерных системах.

Здесь важно отметить, что «пакет» предусматривает возможность создания отдельных записей для различных типов измерений, таких как: зондирование, гидрохимические исследования, геологические измерения и для других видов работ.

Такая структура позволяет детально описывать каждый компонент измерительной системы и методику проведения работ.

Пакет ***LI_Lineage*** – «Сведения о происхождении и обработке данных». Пакет ***LI_Lineage*** фиксирует сведения о происхождении и обработке данных: этапы измерений и расчетов, исполнителей, версии методик и ПО, а также условия, при которых данные были получены и преобразованы, что важно для воспроизводимости океанологических данных. В предлагаемом профиле реализованы следующие элементы:

– *statement* обеспечивает первичную идентификацию источника и базового метода получения данных (оборудование, тип измерений, первичные пересчеты), например: «данные по электропроводности получены с помощью зонда CTD Sea-Bird и затем пересчитаны в соленость»;

– *processStep* позволяет документировать каждый шаг обработки с атрибуцией (автор, дата, методика, версии ПО), что дает возможность другим исследователям повторить расчеты;

– *source* фиксирует исходные материалы, на базе которых сформирован набор, исключая неоднозначность трактовки. Информация об исходных данных, использованных при создании данных, определенных областью применения);

– *scope* – опциональное поле, которое повышает гибкость описания, указывает тип ресурса и протяженность, к какому уровню детализации относится *lineage*[13] – ко всей экспедиции или станции.

Важно отметить, что минимально достаточный профиль, реализованный в предварительном экспедиционном отчете, сознательно ограничивает детализацию ***processStep***. Заполнение полей охватывает только базовые сведения, критически важные для первичной идентификации данных. Полная детализация этапов обработки – с указанием методик, версий ПО, ответственных организаций и исполнителей – выполняется на этапе полного экспедиционного отчета. Такой двухэтапный подход соответствует критерию практической реализуемости и не противоречит требованиям ГОСТ Р 57668-2017, поскольку *LI_ProcessStep* допускает как минимальное, так и расширенное заполнение. Более того, профиль изначально проектировался с учетом того, что экспедиционные данные могут включать не только гидрофизические, но и гидрохимические, биологические и геологические материалы – для каждого из этих типов данных может быть сформирован отдельный *LI_ProcessStep*, отражающий специфику обработки.

Для комплексных отчетов, объединяющих данные нескольких отрядов, предусмотрено использование нескольких пакетов *LI_Lineage*, каждый из которых описывает свой блок данных. Такой подход гарантирует, что *lineage* остается прозрачным и проверяемым на любом уровне агрегации данных.

[13] *Lineage* (происхождение данных) – это информация о происхождении и истории обработки набора данных.

Пакет ***MD_Distribution*** обеспечивает доступность данных. Обязательные элементы:

– *distributor* – поставщик данных: организация, контактное лицо, способы связи;

– *distributionFormat* – формат передачи данных (опционально, но рекомендуется указывать основные форматы, например, CSV[14], *NetCDF*[15], *SHP*[16]);

– *description* – краткое описание способов доступа (в свободной форме).

Для всех полей, представленных объектами других классов (например, *CI_Citation*, *CI_Responsibility*), действует правило: если объект заполняется, то все его обязательные атрибуты также должны быть заполнены. Это снижает риск частичного и противоречивого описания.

Процедура заполнения: две формы экспедиционного отчета

С учетом практики экспедиционных работ в ИО РАН предусмотрены две формы заполнения метаданных:

1. **«Предварительный экспедиционный отчет»** – формируется после завершения экспедиции в течение нескольких недель. Включает минимально необходимые сведения: *MD_Metadata*, *MD_Identification, MI_AcquisitionInformation*. Цель – зафиксировать базовые параметры экспедиции, методы и оборудование, чтобы избежать утраты критически важной информации сразу после окончания работ. Объем заполнения минимален: только обязательные поля; опциональные оставляются на этап полного отчета.

2. **«Полный экспедиционный отчет»** – формируется после экспедиции. Включает дополнительные пакеты (*LI_Lineage*, *MD_Distribution*) и детализированные описания. При необходимости может быть создано несколько экземпляров пакетов (например, несколько *MD_Identification* или *LI_Lineage*) для описания отдельных отрядов или типов данных.

Такой двухэтапный подход снижает нагрузку на исполнителей в рейсе и одновременно обеспечивает полноту описания после завершения работ.

Программная реализация и автоматизация валидации

Формирование метаданных реализовано в приложении *MetadataExp* (*C# WPF*[17], ОС Windows), интегрированном в ИС *ODBASE* (Запотылько и др., 2024). Интерфейс приложения структурирован по 13 вкладкам, каждая из которых соответствует одному из пакетов стандарта. Это позволяет пользователю последовательно заполнять профиль, не нарушая логической структуры ГОСТ.

---

[14] *CSV* (*Comma-Separated Values*) – текстовый формат для хранения табличных данных, в котором значения разделяются запятыми;

[15] *NetCDF* (*Network Common Data Form*) – бинарный формат для хранения научных данных, самоописывающийся формат с метаданными;

[16] *SHP* (*Shapefile*) – векторный формат для хранения географических данных.

[17] *C# WPF* (*Windows Presentation Foundation*) – это технология разработки графических интерфейсов для *Windows*-приложений, которая является частью платформы *.NET.*

Встроенная валидация обеспечивает базовый контроль качества данных при заполнении:

– Проверка обязательных полей. При попытке сохранить метаданные система проверяет наличие всех обязательных атрибутов в каждом пакете.

– Контроль форматов. Система проверяет корректность структуры вводимых данных, включая формат дат согласно ISO 8601.

– Справочники и классификаторы. Для полей с ограниченным набором значений (например, *topicCategory, status*) используются выпадающие списки с кодами, стандартизированными в соответствии с международными классификаторами и отраслевыми стандартами.

– Валидация вложенных объектов. Если пользователь заполняет объект (например, *CI_Responsibility*), система автоматически проверяет наличие всех его обязательных атрибутов.

Окончательная проверка уникальности идентификаторов и целостности данных осуществляется на этапе загрузки в базу данных. Процедура заполнения опирается на заранее подготовленные справочники (коды типов измерений, шаблоны описаний, списки ролей и статусов), что сокращает ручной ввод и повышает единообразие метаданных.

Валидация методики и оценка качества профиля

Для оценки качества формируемого профиля использовались следующие метрики:

– Полнота покрытия обязательных полей. Доля заполненных обязательных полей в каждом пакете (целевой показатель – 100 %).

– Согласованность данных. Проверка соответствия между полями разных пакетов (например, совпадение дат в *MD_Identification* и *LI_Lineage*).

– Соответствие международным стандартам. Автоматическая проверка на совместимость с форматами *EMODnet* и *OBIS* (на уровне кодов категорий, форматов дат и идентификаторов).

Валидация проводилась на тестовых наборах данных, подготовленных разработчиками системы с учетом типичных сценариев заполнения метаданных для экспедиционных исследований. В процессе тестирования были проанализированы основные типы океанологических измерений и соответствующие им структуры метаданных. На основе анализа типовых ошибок и требований к описанию данных были:

– скорректированы формулировки шаблонов ввода;

– уточнены справочники допустимых значений;

– доработаны правила валидации полей;

– оптимизированы алгоритмы проверки целостности данных.

Дальнейшая валидация планируется с привлечением научных сотрудников после внедрения пользовательского интерфейса и подготовки обучающих материалов.

Результаты

ГОСТ Р 57668-2017 требует минимального заполнения двух пакетов *MD_Metadata* и *MD_Identification*. Специфика экспедиционной деятельности требует более расширенного описания (Коротаев и др., 2005; Кузьмин, Ипатов, 2011). В настоящей работе предлагается более расширенный, но минимальный профиль, который подходит для океанологической экспедиции. В перспективе будет разработан полный профиль экспедиции, включающий отчеты начальников отрядов. На текущем этапе создан минимальный профиль, включающий информацию из предварительного отчета начальника экспедиции, в процессе создания которого были выполнены:

– анализ элементов экспедиционной отчетности, на основе которого разработаны требования к метаданным экспедиции;

– на основе требований, разработанных для метаданных экспедиции, были выбраны необходимые пакеты ГОСТа;

– сгруппированы избранные пакеты в минимально необходимый профиль;

– частично описан инструментарий для создания метаданных минимального профиля.

Таким образом, предложенный профиль метаданных океанологической экспедиции содержит минимально необходимый и достаточный набор пакетов и полей для однозначной идентификации, описания происхождения и условий доступа к данным. На рисунке 1 представлена *UML*-диаграмма разработанного профиля метаданных океанологической экспедиции, отражающая структуру и связи ключевых пакетов. Остальные пакеты стандарта ГОСТ Р 57668-2017 могут быть заполнены или актуализированы на последующих этапах без потери совместимости.

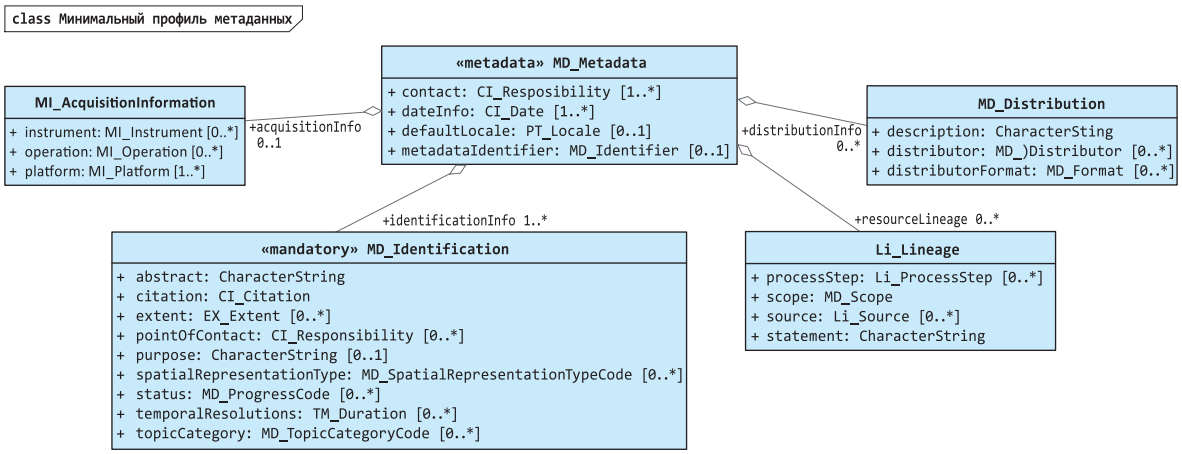


Рис. 1 – UML-диаграмма профиля океанологической экспедиции

На изображении приведена *UML*-диаграмма классов, описывающая логическую структуру профиля метаданных для океанологических экспедиционных данных. Модель построена в соответствии с канонами ISO 19115 и адаптирована под специфику экспедиционных наблюдений; она служит основой для формализации «цифровой истории жизни» массива данных и обеспечения *FAIR*-принципов.

Центральным элементом модели выступает класс *MD_Metadata* – он задает контейнер для всего набора метаданных и содержит базовые атрибуты:

– ***contact***: *CI_Responsibility* [1..*] – обязательные сведения об ответственных лицах/организациях;

– ***dateInfo***: *CI_Date* [1..*] – набор дат (создание, публикация, обновление);

– ***defaultLocale***: *PT_Locale* [0..1] – язык и региональные настройки по умолчанию;

– ***metadataIdentifier***: *MD_Identifier* [0..1] – уникальный идентификатор самого метаописания.

От ***MD_Metadata*** отходят четыре ключевые агрегации, раскрывающие содержательную часть профиля:

– +***identificationInfo*** [1..*] – обязательная идентификационная информация (кто, что, где, когда, зачем);

– +***acquisitionInfo*** [0..1] – сведения о способе получения данных (инструменты, операции, платформы);

– +***resourceLineage*** [0..*] – линия происхождения данных (источники, шаги обработки, утверждения);

– +***distributionInfo*** [0..*] – информация о распространении (описание, дистрибьюторы, форматы).

Класс ***MD_Identification*** реализует блок идентификации и включает широкий спектр характеристик ресурса:

– ***abstract***, ***citation***, ***description*** – текстовые описания и ссылки на публикации;

– ***extent***: *EX_Extent* [0..*] – пространственные, временные и тематические границы данных;

– ***pointOfContact***: *CI_Responsibility* [0..*] – контакты для запросов и поддержки;

– ***purpose***, ***status***, ***topicCategory*** – назначение, статус и тематические категории;

– ***spatialRepresentationType***, ***temporalResolutions*** – тип пространственного представления и временное разрешение наблюдений.

Класс ***MI_AcquisitionInformation*** фиксирует технические детали сбора данных:

– ***instrument***: *MI_Instrument* [0..*], ***platform***: *MI_Platform* [1..*] – используемые приборы и платформы (судно, буй, дрон);

– ***operation***: *MI_Operation* [0..*] – выполненные операции (зондирование, отбор проб).

Класс ***LI_Lineage*** отвечает за прослеживаемость данных:

– ***processStep***: *LI_ProcessStep* [0..*] – последовательность шагов обработки;

– ***source***: *LI_Source* [0..*] – исходные массивы и их метаданные;

– ***statement*** – общее описание линии происхождения.

Класс ***MD_Distribution*** описывает условия и способы распространения данных:

– ***description***, ***distributor***: *MD_Distributor* [0..*] – описание и список дистрибьюторов;

– ***distributorFormat***: *MD_Format* [0..*] – поддерживаемые форматы передачи.

Такая структура позволяет однозначно описать основные аспекты экспедиции, включая не только статические характеристики (пространство, время, тему), но и динамические аспекты (как, кем, на чем и в какой последовательности были получены

данные). Это обеспечивает семантическую совместимость с международными системами (*EMODnet*, *OBIS*, *SeaDataNet*) и соответствие принципам *FAIR* за счет машиночитаемой и стандартизованной схемы.

Программная реализация формирования экспедиционных метаданных MetadataExp

Ручное формирование *XML*-документов по стандарту является трудоемкой и подверженной ошибкам задачей, для заполнения профиля разработан программный инструмент *MetadataExp,* используемый при работе с информационной системой «База данных ИО РАН». Приложение реализовано на стеке средств разработки *C# WPF* под ОС *Windows*. Программа MetadatExp реализована как универсальная платформа для создания метаданных по ГОСТ Р 576682017 и формирует файл XML для загрузки в БД ИС «ODBASE». Этот же файл может использоваться и другими БД и программами, которые поддерживают ГОСТ Р 576682017 и ISO 19115, например, “OceanDB”, “SeaDataNet” и др. Интерфейс *MetadataExp* структурирован по тематическим вкладкам: каждая из 13 вкладок соответствует одному из пакетов стандарта, что позволяет пользователю последовательно заполнять метаданные. Такой подход снижает риск пропуска обязательных полей и повышает полноту описания. На рисунках 2 и 3 отображены интерфейсы для заполнения пакетов «*MD_Metadata*» и «*MD_Indentification*» программы *MetadataExp*.

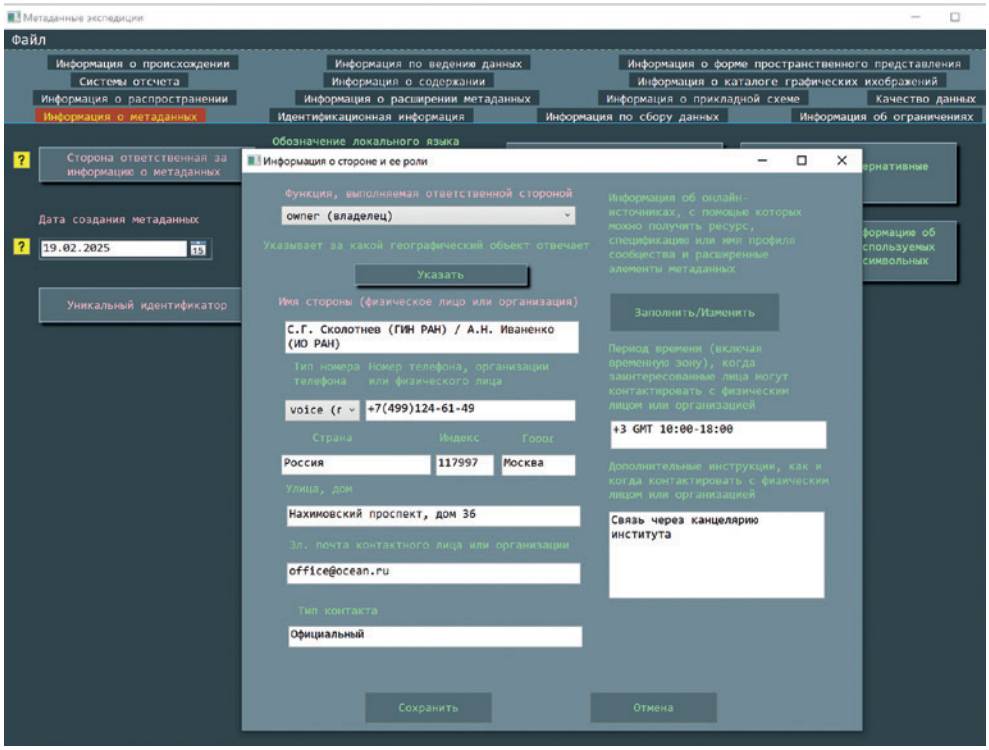


Рис. 2 – Общий вид интерфейса пакета информации о метаданных «*MD_Metadata*» в программе «*MetadataExp*»

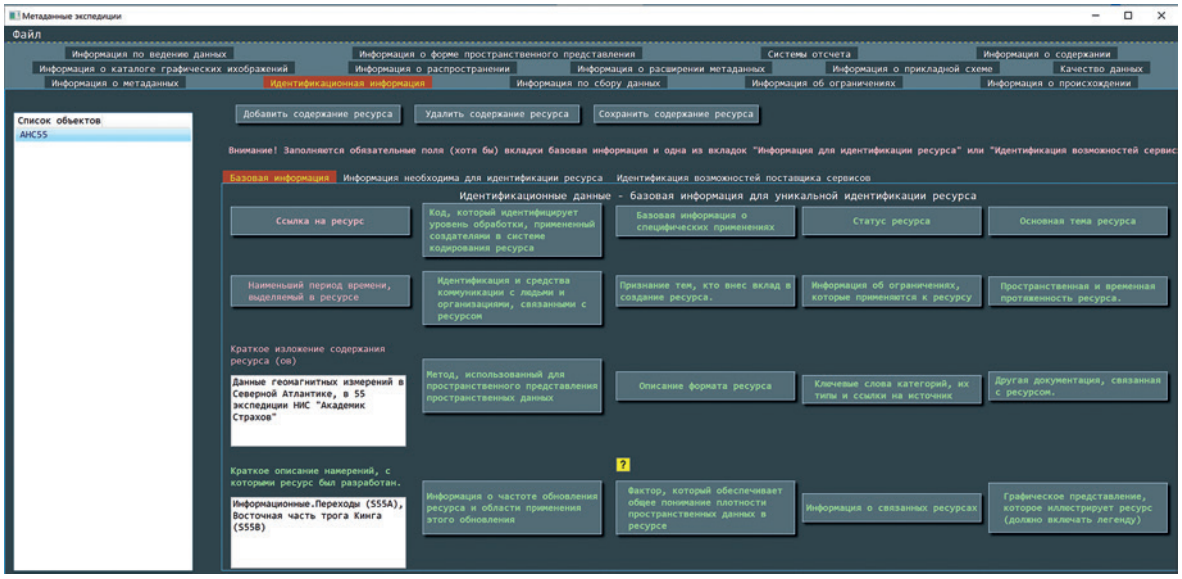


Рис. 3 – Общий вид интерфейса пакета идентификационной информации «MD_Indentification» в программе «MetadataExp»

Дискуссия

Предложенный профиль метаданных и инструмент *MetadataExp* направлены на преодоление ключевых барьеров в управлении экспедиционными данными в океанологии – информационных разрывов, фрагментации хранилищ и утраты сведений о происхождении данных (*lineage*). Эти проблемы особенно критичны в условиях роста объемов измерений и повышения требований к воспроизводимости результатов: без однозначной фиксации условий сбора и обработки данные теряют научную ценность уже спустя несколько лет после экспедиции.

Перспективным направлением развития предложенного профиля *MetadataExp* является его согласование с рекомендациями междоменной интероперабельности (*CDIF*[18]) (Gregory et al., 2024), координируемой *CODATA*[19]. Трехлетний проект *CDIF4EOSC*, стартовавший 1 июня 2026 г., направлен на создание практического руководства по *FAIR*-интеграции данных в Европейском открытом научном облаке (*EOSC*[20]) и за его пределами. Примечательно, что один из трех ключевых сценариев использования этого проекта – науки об океане. Это делает интеграцию с *CDIF* не просто желательной, а стратегически необходимой для обеспечения совместимости разрабатываемого

[18] *CDIF* (*Cross-Domain Interoperability Framework*) – рамочная модель междоменной интероперабельности.
[19] *CODATA* (*Committee on Data for Science and Technology*) – Комитет по данным для науки и техники – международная научная организация, продвигающая работу с данными в науке и технике в составе Международного научного совета (ISC). Раньше была частью Международного совета по науке (ICSU). Основана в 1966 г.
[20] *EOSC* (*European Open Science Cloud*) – Европейское открытое научное облако – инфраструктура для хранения, обмена и повторного использования научных данных в Европе – экосистема правил, стандартов и подключенных ресурсов, которая объединяет национальные и дисциплинарные хранилища данных.

профиля с европейскими инфраструктурами и для расширения международной видимости российских экспедиционных данных.

В связи с этим определяются следующие приоритетные направления интеграции с *CDIF*. Во-первых, предполагается согласование идентификационных и описательных полей с профилем обнаружения ресурсов *CDIF* (*Discovery*) (Richard et al., 2023), который задает минимальный набор метаданных для обеспечения обнаружения ресурсов и опирается, среди прочего, на стандарт ISO 19115-1. Во-вторых, будет рассмотрена реализация рекомендаций *CDIF Discovery* по контролируемым словарям и «универсальным» элементам (время, география, единицы измерения), что напрямую перекликается с задачей приведения к единому виду справочников ЕСИМО. В-третьих, планируется отслеживать развитие *CDIF* v1.1 и последующих версий, чтобы обеспечить совместимость MetadataExp и генерируемых *XML*-документов с *EOSC Federation*. Эти шаги закладывают основу для практической реализации принципов *FAIR* на междисциплинарном уровне и создают предпосылки для включения российских океанологических данных в международный обмен.

Опыт интеграции специализированных наборов в ИС *ODBASE* подтверждает остроту проблемы. Так, при архивации морских геомагнитных данных были созданы новые таблицы и новый тип данных для нанесения на карту, потребовалась постэкспедиционная нормализация разнородных материалов, сопровождавшаяся риском утраты сведений о калибровках и условиях измерений (Запотылько и др., 2025). Предложенный подход смещает акцент с трудоемкой постобработки на стандартизацию в момент сбора данных: минимально достаточный профиль фиксирует только те атрибуты, которые реально необходимы для интерпретации и повторного использования результатов.

Отбор пакетов и полей профиля основан на балансе между полнотой описания и практической применимостью при работе с данными. В отличие от полного стандарта ГОСТ Р 576682017, охватывающего широкий спектр пространственных данных, в профиль включены только критически важные компоненты:

– ***MD_Identification*** и ***MD_Metadata*** – для однозначной идентификации экспедиции; ***MI_AcquisitionInformation,*** интегрированный через механизм расширений ***MD_MetadataExtensionInformation*** (в перспективе будет добавлен ISO 19115-2) – для описания методов и оборудования;

– ***LI_Lineage*** – для документирования этапов обработки;

– ***MD_Distribution*** – для указания условий доступа.

Такой состав позволяет обеспечить воспроизводимость и совместимость с международными системами (*EMODnet*, *OBIS*, Argo) без избыточной нагрузки на исполнителей в экспедиции.

Программная реализация в виде приложения *MetadataExp* (*C# WPF*) делает профиль практически применимым. Автоматизация заполнения и встроенная валидация обязательных полей существенно снижают риски ошибок и пропусков, типичных при ручном вводе. Совместное использование *MetadataExp* и ИС *ODBASE* закладывает основу для систематизации экспедиционных данных на уровне института и обеспечивает их долговременную сохранность в соответствии с принципами *FAIR*.

В настоящее время приложение *MetadataExp* находится на этапе тестирования. Существующая структура интерфейса, ориентированная на прямое соответствие пакетам стандарта, обеспечивает гибкость: пользователь может заполнять любые пакеты ГОСТ Р 57668-2017, в том числе выходящие за рамки минимального профиля. Вместе с тем, такая полнота может усложнять работу для пользователей без специальной подготовки. Для устранения этого барьера разрабатывается специализированный мастер для пошагового заполнения, который будет направлять пользователя по ключевым этапам ввода данных, акцентируя внимание на обязательных полях и предоставляя контекстные подсказки. Это позволит сократить время подготовки метаданных и снизить количество ошибок при заполнении.

В процессе дальнейшего развития планируется:

– создание подробных методических рекомендаций по заполнению метаданных для каждого вида исследований;

– организация системы обучения пользователей работе с инструментом.

Реализация этих подходов будет способствовать формированию единого информационного пространства для экспедиционных данных институтов РАН, что в перспективе повысит прозрачность исследований, обеспечит их воспроизводимость и усилит международную видимость российских морских научных работ.

Выводы

1. Разработан профиль метаданных океанологической экспедиции на базе ГОСТ Р 57668-2017: сформирован минимально достаточный набор пакетов и полей, обеспечивающий однозначную идентификацию экспедиции, фиксацию методов, оборудования, протоколов калибровки и жизненного цикла данных.

2. Определены критерии отбора и семантика полей с учетом специфики экспедиционных океанологических данных: выделены обязательные атрибуты, критичные для воспроизводимости результатов и повторного использования данных.

3. Реализован программный инструмент *MetadataExp* для заполнения профиля и генерации машиночитаемых метаданных (*XML*) с функциями валидации обязательных полей и контроля корректности заполнения.

4. Подтверждена применимость профиля на типовых сценариях океанологических экспедиций (Запотылько и др., 2025): продемонстрированы примеры заполнения, показана совместимость с требованиями ИС *ODBASE* и международных систем (*EMODnet*, *OBIS*, *Argo*), а также валидность генерируемых *XML*-документов.

5. Профиль и *MetadataExp* апробированы на практике в системе ODBASE при составлении описаний экспедиций и могут быть использованы институтами РАН для повышения прозрачности, воспроизводимости и обмена океанологической информацией в соответствии с принципами *FAIR*.

Благодарности. Работа выполнена в рамках Государственного задания по теме № FMWE-2024-0024. Авторы выражают признательность рецензентам за ценные замечания и рекомендации, а также научному руководителю технического сектора Института океанологии Н. А. Римскому-Корсакову за прочтение работы, рекомендации и критические замечания, сделанные в процессе написания статьи, которые существенно улучшили эту работу.

Список литературы

1. *Вязилов Е. Д.* Консолидация метаданных в области наук об окружающей среде // Вычислительные технологии. 2005. № S2. EDN: NUOJBH
2. *Вязилов Е. Д., Михайлов Н. Н.* Интеграция гетерогенных информационных ресурсов в области морской деятельности // Вычислительные технологии. 2005. Т. 10. Ч. 1. Спец. Выпуск. С. 21–29. EDN: NUOJAX
3. *Вязилов Е. Д., Михайлов Н. Н., Кобелев А. Е.* Единая государственная система информации об обстановке в Мировом океане: интеграция информационных ресурсов и метаданные // Современные проблемы дистанционного зондирования Земли из космоса. 2008. Т. 5. № 2. С. 69–75. EDN: NDPVOD
4. *Голик А. В., Фищенко В. К., Дубина В. А., Митник Л. М.* Корпоративная океанографическая ГИС ДВО РАН: интеграция спутниковых и подспутниковых данных по северо-западной части Тихого океана // Исследование Земли из космоса. 2004. № 6. С. 73–80. EDN: OVYEKL
5. ГОСТ Р 57668-2017 (ИСО 19115-1:2014). Пространственные данные. Метаданные. Часть 1. Основные положения: национальный стандарт Российской Федерации. М., 2017.
6. *Дмитриева Е. В., Ростов И. Д.* Разработка и реализация баз океанографических данных по северной части Тихого океана. Владивосток: Дальнаука, 2004. ISBN 5-8044-0403-2. EDN: QKEKDL
7. *Еремеев В. Н., Халиулин А. Х., Ингеров А. В., Жук Е. В., Годин Е. А., Пластун Т. В.* Современное состояние банка океанографических данных МГИ НАН Украины: программно-математическое обеспечение // Морской гидрофизический журнал. 2014. № 2. С. 54–66. EDN: TEYRAP
8. *Запотылько В. С., Попов А. П., Свиридов С. А., Римский-Корсаков Н. А.* ODBASE – система хранения и отображения данных морских исследований Института океанологии РАН // Океанологические исследования. 2024. Т. 52. № 2. С. 206–233. EDN: GTSOTT. https://doi.org/10.29006/1564–2291.JOR-2024.52(2).11.5
9. *Запотылько В. С., Пальшин Н. А., Иваненко А. Н., Свиридов С. А.* Цифровой архив морских геомагнитных данных и его интеграция в ИС ODBASE // Океанологические исследования. 2025. Т. 53. № 1. С. 205–221. EDN: NEDWCH. https://doi.ocean.ru/10.29006/1564-2291.JOR-2025.53(1).10
10. *Коваленко М. В., Мельников В. В., Мельник Л. А., Гребнева Е. А.* База метаданных ИПТС по экспедиционным исследованиям в Черном море // Системы контроля окружающей среды. 2015. № 1 (21). С. 28–32. EDN: VRDAOT
11. *Коротаев Г. К., Смирнов В. Г., Агеев М. Д., Еремеев В. Н., Ястребов В. С.* Океанология. Средства и методы океанологических исследований: Монография. М.: Наука, 2005. 795 с.
12. *Кузьмин С. Б., Ипатов А. Ю.* Современные приборы и технологии наблюдения за гидрологическими условиями в Северном Ледовитом океане / Глава из книги «Океанография и морской лед». М.: Paulsen, 2011. С. 7–22. ISBN 978-5-98797-065-2.

13. *Метальников А. А., Осипенко М. В., Свиридов С. А., Соловьев В. А., Утяков Л. Л., Филипчук Ю. Б.* Структура хранилища и формат исследовательских данных по океанологии. В кн.: Сб. трудов VIII Международной конференции МСОИ. М.: ИО РАН, 2003. С. 273–293. EDN: EJIZIK
14. *Метальников А. А., Осипенко М. В., Свиридов С. А., Соловьев В. А., Филипчук Ю. Б.* Хранилище комплексных океанологических данных // Океанология. 2005. Т. 45. № 4. С. 622–628. EDN: HRWTAJ
15. Общие технические спецификации ЕСИМО // Полнофункциональная версия. (версия v.1.0 от 15.06.2011) / Под Ред.: Михайлов Н. Н. ФГБУ «ВНИИГМИ-МЦД» Росгидромета. г. Обнинск, 2011. http://portal.esimo.ru/esimo_static/documents/ts_esimo_2011.pdf
16. *Осипенко М. В., Свиридов С. А., Соловьев В. А., Утяков Л. Л., Филипчук Ю. Б.* Об использовании технологий хранилищ данных для обработки океанологической информации // Современные методы и средства океанологических исследований: Сб. трудов VII Международной науч.-техн. конференции МСОИ-2001. М.: ИО РАН, 2001. С. 87–88. EDN: YZGUOL
17. Сведения об организациях России и мира. https://data.apicrafter.ru/tables/esimo/wdc_2877/docs.
18. *Фищенко В. К., Голик А. В., Антушев С. Г., Кисленок Е. Г., Суботэ А. Е.* Корпоративная океанологическая геоинформационная система ТОИ ДВО РАН: концепция, архитектура, реализация // Дальневосточные моря России. Т. 1. М.: Наука, 2007. С. 558–606. EDN: TTPUHT
19. *Brodeur J., Coetzee S., Danko D., Garcia S., Hjelmager J.* Geographic Information Metadata–An Outlook from the International Standardization Perspective // ISPRS International Journal of Geo-Information. 2019. Vol. 8. No. 6. P. 280. EDN: FTBQMD. https://doi.org/10.3390/ijgi8060280
20. *Dave Danko*. Overview of ISO 19115-1 / Esri, Berlin Workshop 2014. https://www.iso.org/files/live/users/fh/aj/aj/tc211contributor%40iso.org/files/Presentations/2014-06%20Berlin/Danko_Overview_of_ISO_19115-1.pdf
21. *Gregory A., Bell D., Brickley D., Buttigieg P. L., Cox S., Edwards M., Fils D., Gonzalez Morales L. G., Heus P., Hodson S. (Work package leader), Kanjala C., Le Franc Y., Maxwell L., Molloy L., Richard S., Rizzolo F., Winstanley P., Wyborn L., Burton A.* WorldFAIR (D2.3) Cross-Domain Interoperability Framework (CDIF) // (Report Synthesising Recommendations for Disciplines and Cross-Disciplinary Research Areas). 2024. Zenodo. https://doi.org/10.5281/zenodo.11236871
22. *Ikegami M., Kameyama Y., Furukawa K., Sakakibara T.* Development of coastal and estuarine environmental data sharing using XML // Proceedings of the 3rd International Workshop on Scientific Use of Submarine Cables and Related Technologies. IEEE, 2003. P. 293–297. https://doi.org/10.1109/ssc.2003.1224166
23. *Isenor A. W., Spears T. W.* Combining the Arc Marine Framework with Geographic Metadata to Support Ocean Acoustic Modeling // Transactions in GIS. 2013. Vol. 18. No. 2. P. 183–200. https://doi.org/10.1111/tgis.12029
24. ISO 19115-1:2014. Geographic information. Metadata. Part 1: Fundamentals.
25. ISO 19115-1:2014. Annex G (Revisions). P. 157–166. https://www.iso.org/standard/53798.html
26. *Jiang Y., Lu L., Guo* Z. Design of Marine Information Metadata and Directory Service System Based on XML // 2009 First International Workshop on Database Technology and Applications. IEEE, 2009. P. 574–577. https://doi.org/10.1109/dbta.2009.58
27. *Jiménez M., Fernández-Díaz M. C., González J. A., Martínez M., Sánchez A., Rodríguez E., García F.* Field Spectroscopy Metadata System Based on ISO and OGC Standards // ISPRS International Journal of Geo-Information. 2014. Vol. 3. No. 3. P. 1003–1022. https://doi.org/10.3390/ijgi3031003

28. *Klein E., Appeltans W., Provoost P., Saeedi H., Benson A., Bajona L., Peralta A. C., Bristol R. S.* OBIS Infrastructure, Lessons Learned, and Vision for the Future // Frontiers in Marine Science. 2019. Vol. 6. P. 588. EDN: MOUHSN. https://doi.org/10.3389/fmars.2019.00588
29. *Leadbetter A., Meaney W., Tray E., Conway A., Flynn S., Keena T., Kelly C., Thomas R.* A modular approach to cataloguing marine science data // Earth Science Informatics. 2020. Vol. 13. No. 2. P. 537–553. EDN: WZZBSA. https://doi.org/10.1007/s12145-020-00445-w
30. *Mineart G. M., Klein F. C.* A data management strategy for the ocean exploration program // Oceans '02 MTS/IEEE. IEEE. 2002. P. 1642–1647. https://doi.org/10.1109/oceans.2002.1191880
31. *Olfat H., Kalantari M., Rajabifard A., Williamson I. P.* Towards a foundation for spatial metadata automation // Journal of Spatial Science. 2012. Vol. 57. No 1. P. 65–81. https://doi.org/10.1080/14498596.2012.686361
32. *Pecci L., Fichaut M., Schaap D.* SeaDataNet, an enhanced ocean data infrastructure giving services to scientists and society // IOP Conference Series: Earth and Environmental Science. 2020. Vol. 509. No. 1. P. 012042. EDN: HECZOV. https://doi.org/10.1088/1755-1315/509/1/012042
33. *Richard S., Gregory A., Hodson S., Fils D., Kanjala C., Bell D., Winstanley P., Edwards M., Heus P., Brickley D., Rizzolo F., Maxwell L., Luis G., Buttigieg P. L., & Le Franc Y.* Cross-Domain Interoperability Framework (CDIF) Working Group. Cross Domain Interoperability Framework (CDIF): Discovery Module (v01 draft for public consultation) (Version 01). 2023. Zenodo. https://doi.org/10.5281/zenodo.10252564
34. *Schaap D. M. A., Lowry R. K.* SeaDataNet – Pan-European infrastructure for marine and ocean data management: unified access to distributed data sets // International Journal of Digital Earth. 2010. Vol. 3. Suppl. 1. P. 50–69. https://doi.org/10.1080/17538941003660974.
35. *Smirnov A. V., Korablev A. A.* A regional oceanographic database for the Nordic Seas: from observations to climatic dataset // Proc. International Marine Data and Information Systems Conference IMDIS 2010 (29–31 March 2010) / Eds. M. Fichaut, V. Tosello. IFREMER/SISMER: UNESCO, 2010. P. 81–82.
36. *Tanhua T., Pouliquen S., Hausman J., O'Brien K., Bricher P., de Bruin T., Buck J. J. H., Burger E., Carval T., Casey K. S., Diggs S., Giorgetti A., Glaves H., Harscoat V., Kinkade D., Muelbert J. H., Novellino A., Pfeil B., Pulsifer P. L., Van de Putte A., Robinson E., Schaap D., Smirnov A., Smith N., Snowden D., Spears T., Stall S., Tacoma M., Thijsse P., Tronstad S., Vandenberghe T., Wengren M., Wyborn L., Zhao Z.* Ocean FAIR Data Services // Frontiers in Marine Science. 2019. Vol. 6. P. 440. EDN: MTEMOA. https://doi.org/10.3389/fmars.2019.00440
37. *Tanhua T., Key R. M., González-Dávila M., Pfeil B., Olsen A., Lauvset S. K., Bakker D. C. E., Hoppema M., Johannessen T., Steinfeldt R., Gruber N., Wallace D. W. R., Wanninkhof R., Landschützer P.* A vision for FAIR ocean data products // Communications Earth & Environment. 2021. Vol. 2. No. 1. EDN: ZTERHG. https://doi.org/10.1038/s43247-021-00209-4.
38. *Wang Yan, Le Jiajin, Huang Dongmei* A metadata management framework for marine information based on XML // IET International Conference on Information Science and Control Engineering 2012 (ICISCE 2012). Institution of Engineering and Technology. 2012. P. 1.01–1.01. https://doi.org/10.1049/cp.2012.2258
39. *Xiu L., Yan T., Gao F., Zhou L., Yu J., Guo Z.* Design of Data Management System for Seafloor Observatory Network // 2013 International Conference on Service Sciences (ICSS). IEEE. 2013. P. 147–150. https://doi.org/10.1109/icss.2013.50

Статья поступила в редакцию 08.08.2026, одобрена к печати 29.09.2026.
Для цитирования: *Запотылько В. С., Свиридов С. А.* Разработка профиля метаданных океанологической экспедиции на основе ГОСТ Р 57668-2017 и его программная реализация в MetaDataExp // Океанологические исследования. 2026. Т. 54. № 3. С. 173–198. https://doi.ocean.ru/10.29006/1564-2291.JOR-2026.54(3).9

DEVELOPMENT OF A METADATA PROFILE FOR AN OCEANOGRAPHIC EXPEDITION BASED ON GOST R 57668-2017 AND ITS SOFTWARE IMPLEMENTATION IN METADATAEXP

V. S. Zapotylyko, S. A. Sviridov

Shirshov Institute of Oceanology, Russian Academy of Sciences,
36, Nakhimovskiy prospekt, Moscow, 117997, Russia,
esimo@ocean.ru

This article presents a specialized metadata profile for oceanographic expedition data, developed in accordance with the requirements of GOST R 57668-2017 and the international standard ISO 19115. The central element of the work is a UML model that formalizes the metadata structure and ensures semantic compatibility with international information infrastructures (EMODnet, OBIS, SeaDataNet) and compliance with the FAIR (Findable, Accessible, Interoperable, Reusable) principles. The model is built around the MD_Metadata class, which serves as a container for all descriptive information, and includes four key blocks: identification information (MD_Identification), information on the method of data acquisition (MI_AcquisitionInformation), data lineage (LI_Lineage), and distribution conditions (MD_Distribution). This structure allows capturing not only the static characteristics of the dataset (spatial, temporal, and thematic boundaries, purpose, and status) but also the dynamic aspects of its life cycle—from the instruments and platforms used to the sequence of processing steps and source materials. A distinctive feature of the proposed profile is its emphasis on the "digital life history" of each observation dataset: the model includes elements for a formalized description of the temporal resolution of measurements, spatial coverage, collection methodology, and equipment calibration protocols—parameters critical for the verification of results and secondary analysis. Implementing the profile as a UML diagram (with the MD_Metadata, MD_Identification, EX_Extent, TM_Duration, and other classes) provides a rigorous semantic structure suitable for automated data validation, retrieval, and integration. The practical significance of this work lies in creating a tool for overcoming the fragmentation of oceanographic data and improving its suitability for interdisciplinary and international research. The novelty of the work lies in the development of the first specialized metadata profile for oceanographic expeditions—fully compliant with GOST R 57668-2017—and its software implementation, which enables practical application. The proposed profile can be used in database design, cataloging service development, and the construction of distributed scientific data management infrastructures.

Keywords: information system, expeditionary research, database, MetadataExp, ODBASE, UML, metadata profile, data quality assessment, metadata standards, GOST R 57668-2017, ISO

Acknowledgments: The work was carried out within the framework of the State Assignment, Topic No. FMWE-2024-0024. The authors are grateful to N. A. Rimsky-Korsakov, Scientific Supervisor of the Technical Sector, for reviewing the paper, providing recommendations, and offering critical comments during the writing process — these significantly improved the work.

References

1. Brodeur, J., S. Coetzee, D. Danko, S. Garcia, and J. Hjelmager, 2019: Geographic Information Metadata–An Outlook from the International Standardization Perspective. *ISPRS International Journal of Geo-Information*, **8** (6), 280, https://doi.org/10.3390/ijgi8060280
2. Dave Danko, 2014: *Overview of ISO 19115-1* / Esri, Berlin Workshop 2014. https://www.iso.org/files/live/users/fh/aj/aj/tc211contributor%40iso.org/files/Presentations/2014-06%20Berlin/Danko_Overview_of_ISO_19115-1.pdf
3. Dmitrieva, E. V. and I. D. Rostov, 2004: *Development and Implementation of Oceanographic Databases for the North Pacific*. Vladivostok, Dal'nauka, ISBN 5-8044-0403-2, EDN: QKEKDL
4. Eremeev, V. N., A. Kh. Khaliulin, A. V. Ingerov, E. V. Zhuk, E. A. Godin, and T. V. Plastun, 2014: Current state of the oceanographic data bank of the Marine Hydrophysical Institute, National Academy of Sciences of Ukraine: software and mathematical support. *Physical Oceanography*, **2**, 54–66, EDN: TEYRAP
5. Fishchenko, V. K., A. V. Golik, S. G. Antushev, E. G. Kislenok, and A. E. Subote, 2007: Corporate oceanological geoinformation system of the Pacific Oceanological Institute, Far Eastern Branch of the Russian Academy of Sciences: concept, architecture, implementation. In: *Far Eastern Seas of Russia*, **1**, Moscow, Nauka, 558–606, EDN: TTPUHT
6. *General Technical Specifications for ESIMO*. Full-function version (version v.1.0, June 15, 2011), Edited by N. N. Mikhailov, FSBI "VNIIGMI-WDC" (Roshydromet). Obninsk, 2011, http://portal.esimo.ru/esimo_static/documents/ts_esimo_2011.pdf
7. Gregory, A., D. Bell, D. Brickley, P. L. Buttigieg, S. Cox, M. Edwards, D. Fils, L. G. Gonzalez Morales, P. Heus, S. Hodson, (Work package leader), C. Kanjala, Y. Le Franc, L. Maxwell, L. Molloy, S. Richard, F. Rizzolo, P. Winstanley, L. Wyborn, and A. Burton, 2024: *WorldFAIR (D2.3) Cross-Domain Interoperability Framework* (CDIF) // (Report Synthesising Recommendations for Disciplines and Cross-Disciplinary Research Areas). Zenodo. https://doi.org/10.5281/zenodo.11236871
8. Golik, A. V., V. K. Fishchenko, V. A. Dubina, and L. M. Mitnik, 2004: Corporate oceanographic GIS of the Far Eastern Branch of the Russian Academy of Sciences: integration of satellite and subsatellite data for the northwestern Pacific. *Earth Research from Space*, **6**, 73–80, EDN: OVYEKL
9. GOST R 57668-2017 (ISO 19115-1:2014), 2017: Geographic information – Metadata – Part 1: Fundamentals. National Standard of the Russian Federation, Moscow.
10. Ikegami, M., Y. Kameyama, K. Furukawa, and T. Sakakibara, 2003: Development of coastal and estuarine environmental data sharing using XML. *Proceedings of the 3rd International Workshop on Scientific Use of Submarine Cables and Related Technologies, IEEE*, 293–297, https://doi.org/10.1109/ssc.2003.1224166
11. *Information on organizations in Russia and worldwide*. https://data.apicrafter.ru/tables/esimo/wdc_2877/docs
12. Isenor, A. W. and T. W. Spears, 2013: Combining the Arc Marine Framework with Geographic Metadata to Support Ocean Acoustic Modeling. *Transactions in GIS*, **18** (2), 183–200, https://doi.org/10.1111/tgis.12029
13. *ISO 19115-1:2014*, 2014: Geographic information – Metadata – Part 1: Fundamentals.
14. *ISO 19115-1:2014*. Annex G (Revisions), 157–166, https://www.iso.org/standard/53798.html
15. Jiang, Y., L. Lu, and Z. Guo, 2009: Design of Marine Information Metadata and Directory Service System Based on XML. *Proceedings of the 2009 First International Workshop on Database Technology and Applications, IEEE*, 574–577, https://doi.org/10.1109/dbta.2009.58

16. Jiménez, M., M. C. Fernández-Díaz, J. A. González, M. Martínez, A. Sánchez, E. Rodríguez, and F. García, 2014: Field Spectroscopy Metadata System Based on ISO and OGC Standards. *ISPRS International Journal of Geo-Information*, **3** (3), 1003–1022, https://doi.org/10.3390/ijgi3031003
17. Klein, E., W. Appeltans, P. Provoost, H. Saeedi, A. Benson, L. Bajona, A. C. Peralta, and R. S. Bristol, 2019: OBIS Infrastructure, Lessons Learned, and Vision for the Future. *Frontiers in Marine Science*, **6**, 588, https://doi.org/10.3389/fmars.2019.00588
18. Kovalenko, M. V., V. V. Melnikov, L. A. Melnik, and E. A. Grebneva, 2015: Metadata database of the IPTS on expeditionary research in the Black Sea. *Environmental Control Systems*, (1/21), 28–32, EDN: VRDAOT
19. Korotaev, G. K., V. G. Smirnov, M. D. Ageev, V. N. Eremeev, and V. S. Yastrebov, 2005: *Oceanology. Instruments and Methods of Oceanological Research*: Monograph, Moscow, Nauka, 2005, 795 p.
20. Kuzmin, S. B. and A. Yu. Ipatov, 2011: *Modern Instruments and Technologies for Monitoring Hydrological Conditions in the Arctic Ocean*. Chapter from the book "Oceanography and Sea Ice". Moscow, Paulsen, 2011, 7–22, ISBN 978-5-98797-065-2.
21. Leadbetter, A., W. Meaney, E. Tray, A. Conway, S. Flynn, T. Keena, C. Kelly, and R. Thomas, 2020: A modular approach to cataloguing marine science data. *Earth Science Informatics*, **13** (2), 537–553, https://doi.org/10.1007/s12145-020-00445-w
22. Metalnikov, A. A., M. V. Osipenko, S. A. Sviridov, V. A. Solovyev, L. L. Utyakov, and Yu. B. Filipchuk, 2003: Structure of the repository and format of research data in oceanology. *Proceedings of the VIII International Conference MSOI*, P. P. Shirshov Institute of Oceanology RAS, Moscow, 273–293, EDN: EJIZIK
23. Metalnikov, A. A., M. V. Osipenko, S. A. Sviridov, V. A. Solovyev, and Yu. B. Filipchuk, 2005: Repository of complex oceanological data. *Oceanology*, **45** (4), 622–628, EDN: HRWTAJ
24. Mineart, G. M. and F. C. Klein, 2002: A data management strategy for the ocean exploration program. *Proceedings of Oceans '02 MTS/IEEE, IEEE*, 1642–1647, https://doi.org/10.1109/oceans.2002.1191880.
25. Olfat, H., M. Kalantari, A. Rajabifard, and I. P. Williamson, 2012: Towards a foundation for spatial metadata automation. *Journal of Spatial Science*, **57** (1), 65–81, https://doi.org/10.1080/14498596.2012.686361
26. Osipenko, M. V., S. A. Sviridov, V. A. Solovyev, L. L. Utyakov, and Yu. B. Filipchuk, 2001: On the use of data warehouse technologies for processing oceanological information. In: *Modern Methods and Means of Oceanological Research*. Proceedings of the VII International Scientific-Technical Conference MSOI-2001, Shirshov Institute of Oceanology RAS, Moscow, 87–88, EDN: YZGUOL
27. Pecci, L., M. Fichaut, and D. Schaap, 2020: SeaDataNet, an enhanced ocean data infrastructure giving services to scientists and society. *IOP Conference Series: Earth and Environmental Science*, **509** (1), 012042, https://doi.org/10.1088/1755-1315/509/1/012042
28. Richard, S., A. Gregory, S. Hodson, D. Fils, C. Kanjala, D. Bell, P. Winstanley, M. Edwards, P. Heus, D. Brickley, F. Rizzolo, L. Maxwell, G. Luis, P. L. Buttigieg, & Y. Le Franc, 2023: *Cross-Do main Interoperability Framework (CDIF) Working Group.* Cross Domain Interoperability Framework (CDIF): Discovery Module (v01 draft for public consultation) (Version 01). Zenodo. https://doi.org/10.5281/zenodo.10252564
29. Schaap, D. M. A. and R. K. Lowry, 2010: SeaDataNet – Pan-European infrastructure for marine and ocean data management: unified access to distributed data sets. *International Journal of Digital Earth*, **3** (Suppl. 1), 50–69, https://doi.org/10.1080/17538941003660974
30. Smirnov, A. V. and A. A. Korablev, 2010: A regional oceanographic database for the Nordic Seas: from observations to climatic dataset. In: *Proceedings of the International Marine Data*

and Information Systems Conference IMDIS 2010, Eds. M. Fichaut and V. Tosello, IFREMER/SISMER, UNESCO, 81–82.

31. Tanhua, T., S. Pouliquen, J. Hausman, K. O'Brien, P. Bricher, T. de Bruin, J. J. H. Buck, E. Burger, T. Carval, K. S. Casey, S. Diggs, A. Giorgetti, H. Glaves, V. Harscoat, D. Kinkade, J. H. Muelbert, A. Novellino, B. Pfeil, P. L. Pulsifer, A. Van de Putte, E. Robinson, D. Schaap, A. Smirnov, N. Smith, D. Snowden, T. Spears, S. Stall, M. Tacoma, P. Thijsse, S. Tronstad, T. Vandenberghe, M. Wengren, L. Wyborn, and Z. Zhao, 2019: Ocean FAIR Data Services. *Frontiers in Marine Science*, **6**, 440, https://doi.org/10.3389/fmars.2019.00440
32. Tanhua, T., R. M. Key, M. González-Dávila, B. Pfeil, A. Olsen, S. K. Lauvset, D. C. E. Bakker, M. Hoppema, T. Johannessen, R. Steinfeldt, N. Gruber, D. W. R. Wallace, R. Wanninkhof, and P. Landschützer, 2021: A vision for FAIR ocean data products. *Communications Earth & Environment*, **2** (1), https://doi.org/10.1038/s43247-021-00209-4.
33. Vyazilov, E. D., 2005: Metadata consolidation in the environmental sciences. *Computational Technologies*, No. S2, EDN: NUOJBH
34. Vyazilov, E. D. and N. N. Mikhailov, 2005: Integration of heterogeneous information resources in the field of marine activities. *Computational Technologies*, **10** (1), Special Issue, 21–29, EDN: NUOJAX
35. Vyazilov, E. D., N. N. Mikhailov, and A. E. Kobelev, 2008: Unified State System of Information on the World Ocean Environment: integration of information resources and metadata. *Current Problems in Remote Sensing of the Earth from Space*, **5** (2), 69–75, EDN: NDPVOD
36. Wang, Y., J. Le, and D. Huang, 2012: A metadata management framework for marine information based on XML. *Proceedings of the IET International Conference on Information Science and Control Engineering 2012* (ICISCE 2012), Institution of Engineering and Technology, pp. 1.01–1.01, https://doi.org/10.1049/cp.2012.2258
37. Xiu, L., T. Yan, F. Gao, L. Zhou, J. Yu, and Z. Guo, 2013: Design of Data Management System for Seafloor Observatory Network. *Proceedings of the 2013 International Conference on Service Sciences (ICSS), IEEE*, 147–150, https://doi.org/10.1109/icss.2013.50
38. Zapotylyko, V. S., A. P. Popov, S. A. Sviridov, and N. A. Rimsky-Korsakov, 2024: ODBASE – storage and visualization system for marine research data of the Institute of Oceanology, Russian Academy of Sciences. *Journal of Oceanological Research*, **52** (2), 206–233, https://doi.org/10.29006/1564–2291.JOR-2024.52(2).11.5
39. Zapotylyko, V. S., N. A. Palshin, A. N. Ivanenko, and S. A. Sviridov, 2025: Digital archive of marine geomagnetic data and its integration into the ODBASE information system. *Journal of Oceanological Research*, **53** (1), 205–221, https://doi.ocean.ru/10.29006/1564-2291.JOR-2025.53(1).10

Submitted 08.08.2026, accepted 29.09.2026.
For citation: Zapotylyko, V. S. and S. A. Sviridov, 2026: Development of a metadata profile for an oceanographic expedition based on GOST R 57668-2017 and its software implementation in MetadataExp. *Journal of Oceanological Research*, 2026, **54** (3), 173–198. https://doi.ocean.ru/10.29006/1564-2291.JOR-2026.54(3).9

基于 GOST R 57668-2017 的海洋学考察元数据配置文件开发及其在 METADATAEXP 中的软件实现

V. S. Zapotylyko, S. A. Sviridov

Shirshov Institute of Oceanology, Russian Academy of Sciences,
36, Nakhimovskiy prospekt, Moscow, 邮编 : *117997, Russia,*
电子邮件 : *esimo@ocean.ru*

本文介绍了一套专用于海洋学考察数据的元数据配置文件，该文件依据 GOST R 57668-2017 和国际标准 ISO 19115 的要求开发。工作的核心是一个 UML 模型，该模型形式化了元数据结构，确保与国际信息基础设施 (EMODnet、OBIS、SeaDataNet) 的语义兼容性，并符合 FAIR 原则 (可发现、可访问、可互操作、可重用)。该模型围绕 MD_Metadata 类构建，该类作为所有描述性信息的容器，包含四个关键模块 : 标识信息 (MD_Identification)、信息获取方式数据 (MI_AcquisitionInformation)、数据溯源 (LI_Lineage) 和分发条件 (MD_Distribution)。这一结构不仅能够记录数据集的静态特征 (空间、时间、主题范围、用途、状态)，还能描述其生命周期的动态方面 – 从所使用的仪器和平台到处理步骤序列及原始材料来源。所提出配置文件的显著特点在于强调每个观测数据集的“数字生命史”: 模型中包含了对测量时间分辨率、空间覆盖范围、采集方法和设备校准协议进行形式化描述的元素 – 这些参数对于结果验证和二次分析至关重要。该配置文件以 UML 图形式实现 (区分出 MD_Metadata、MD_Identification、EX_Extent、TM_Duration 等类)，提供了严格的语义结构，适用于数据的自动验证、检索和集成。本工作的实际意义在于创建了一种工具，能够克服海洋学数据的碎片化问题，提高其对于跨学科和国际研究的适用性。本工作的创新之处在于开发了首个完全符合 GOST R 57668-2017 的海洋学考察专用元数据配置文件及其软件实现，确保了实际应用。所提出的配置文件可用于数据库设计、编目服务开发以及科学数据分布式管理基础设施的构建。

关键词 : 信息系统，考察研究，数据库，MetadataExp，ODBASE，UML，元数据配置文件，数据质量评估，元数据标准，GOST R 57668-2017，ISO

致谢 : 本研究在俄罗斯联邦科学与高等教育部下达给 Shirshov Institute of Oceanology, Russian Academy of Sciences 的国家任务 (课题编号 : FMWE-2024-0024) 框架内完成。作者感谢审稿人提出的宝贵意见和建议，同时感谢俄罗斯科学院海洋研究所技术部门学术负责人 N.A. Rimsky-Korsakov 对本文的审阅、建议以及在文章撰写过程中提出的批评意见，这些意见显著提升了本工作的质量。

脚注 :

[1]: UML (Unified Modeling Language) – 统一建模语言，用于对各类系统和过程进行图形化描述与可视化。
[2]: FAIR原则 (Findable, Accessible, Interoperable, Reusable) – 一套科学数据标准，旨在使数据可发现、可访问、可互操作、可重用。