

As tendências das infraestruturas europeias de dados marinhos

P. Nunes (1), A. Saramago (1), C. Borges (1), L. Veiga (1) e S. Almeida (1)

(1) Instituto Hidrográfico, Rua das Trinas, nº 49 Lisboa. antunes.nunes@hidrografico.pt.

Resumo: As infraestruturas de dados marinhos *Seadatanet* e *EMODnet* têm sido edificadas ao longo de vários anos sobre a égide da Comissão Europeia por consórcios de organizações e centros de dados marinhos europeus. O principal objetivo destas iniciativas é tornar acessíveis ao cidadão e às organizações (públicas e privadas) os dados marinhos que se encontram dispersos por múltiplas organizações em formatos diversos. Através da partilha de dados e informação pretende-se potenciar o valor agregado da informação seguindo o princípio de – “observar uma vez, utilizar várias vezes”. Através da disponibilização de uma infraestrutura tecnológica de dados marinhos harmonizados e interoperáveis, pretende-se incrementar o potencial económico e ecológico do recurso informação. O presente artigo apresenta as principais tendências das infraestruturas europeias de dados marinhos (*Seadatanet* e *EMODnet*) na procura de uma maior partilha de informação marinha e na implementação dos princípios de dados abertos.

Palavras-chave: Infraestruturas de dados espaciais marinhos, Sistemas de informação geográfica, EMODnet, *Seadatacloud*, Partilha de informação.

1. INTRODUÇÃO

As infraestruturas Europeias de informação marinha são iniciativas suportadas pela Comissão Europeia no âmbito da Economia Azul.

Em 2007, a Comissão Europeia estabeleceu o objetivo de criar uma infraestrutura sustentável e interoperável de dados marinhos que substituísse os múltiplos repositórios de dados, desarticulados, não homogêneos e inacessíveis (Comissão Europeia, 2014).

A rede *European Marine Observation and Data Network* (EMODnet) tem o objetivo explícito de melhorar o acesso aos dados marinhos e, dessa forma, reduzir os custos para os utilizadores, potenciando a inovação e o conhecimento disponível sobre os mares e oceanos (Comissão Europeia, 2012).

O projeto EMODnet apareceu cerca do ano 2008 através do estabelecimento de parcerias entre várias organizações no âmbito da iniciativa *marine knowledge* alinhada com *Blue Book on the Integrated Maritime Policy (IMP)* da Comissão Europeia (Comissão Europeia, 2012, Europeia, 2016). Atualmente é considerado um projeto líder que visa a geração de um mapa digital dos espaços marinhos europeus, até 2020 (Directorate-General for Maritime Affairs and Fisheries, 2018). Os parceiros da rede EMODnet desenvolveram de uma forma colaborativa uma plataforma tecnológica constituída por vários portais temáticos interoperáveis, acessíveis de uma forma comum pela comunidade de utilizadores (EMODnet Consortium, s.d.).

O princípio da rede EMODnet é a manutenção dos repositórios de dados nas organizações criando-se as

condições para que o acesso aos dados seja garantido de uma forma harmonizada, num princípio de acesso sem custos e livre de restrições (Comissão Europeia, 2014).

Como é possível observar na Fig. 1 a rede EMODnet apresenta um conjunto de objetivos concorrentes para a Estratégia Europeia de Crescimento Azul.

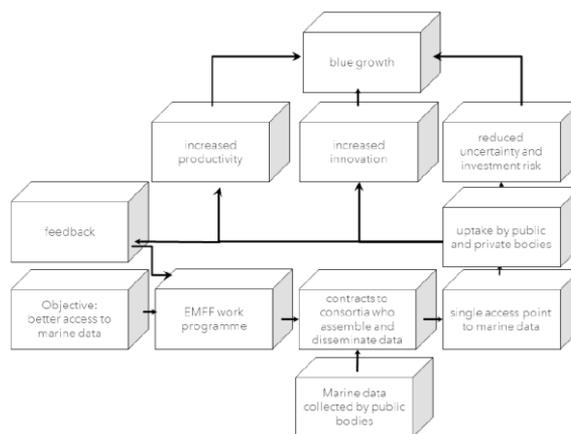


Fig. 1. Esquema conceitual da articulação dos objetivos da rede EMODnet e o seu contributo para o crescimento azul. Extraído do relatório de avaliação de rede EMODnet (Directorate-General for Maritime Affairs and Fisheries, 2018).

O projeto *SeaDataCloud* (SDC) é a última geração de uma linha de desenvolvimento que começou nos anos 90 e culminou na infraestrutura pan-europeia de dados marinhos, com os esforços e a colaboração de Centros de Dados de 35 países.

Os primórdios da infraestrutura foram os arquivos de metadados MEDAR/MedAtlas que depois evoluíram para o projeto *Sea-Search* (2002-2005), a *SeaDataNet* (2006-2011), *SeaDataNet II* (2011-2015) e a *SeaDataCloud* (2016-2020) (Schaap e Fichaut, 2017b).

SDN é uma infraestrutura que apresenta uma arquitetura distribuída para gerir as diversas tipologias de dados marinhos (observações *in situ*). Sendo constituída maioritariamente por centros de dados marinhos de referência, apresenta dados de elevada qualidade integrados em bases de dados.

O projeto apresenta um portal único de acesso e disponibiliza as ferramentas necessárias para garantir a interoperabilidade e harmonização dos dados, com serviços de pesquisa e acesso *on-line* a dados, metadados e dicionários comuns de vocabulário (SeaDataNet Consortium, s.d.).

A infraestrutura SDN tem procurado seguir as boas práticas na partilha de dados e informação, contribuindo de forma ativa para o desenvolvimento de *standards* em estreita colaboração com as organizações de normalização, e.g. ISO (19115, 19139), e OGC (WMS, WFS, CS-W e SWE). O projeto tem garantido o desenvolvimento, a manutenção e redistribuição de ferramentas de transformação de dados e criação de metadados. Dessa forma, torna possível garantir a acessibilidade aos dados marinhos Europeus de uma forma coerente e interoperável com um impacto significativo num universo alargado de utilizadores com diferentes níveis de *expertise* na área. A SDN está a estender as capacidades de computação através de parcerias com os centros europeus de computação de alto desempenho do consórcio EUDAT. O projeto SDC tem trabalhado no desenvolvimento e adoção de metodologias que fazem uso da *cloud* e da computação de alto desempenho. Estas novas capacidades colocam à disposição dos investigadores uma coleção de serviços e ferramentas em ambientes virtuais de desenvolvimento na *cloud* orientados aos requisitos específicos dos utilizadores. A estratégia de desenvolvimento da SDN/SDC é feito em estreita colaboração com redes, comunidades e outras infraestruturas: *Science*, *EMODnet*, *Copernicus Marine Environmental Monitoring Service* (CMEMS) e *EuroGOOS* (Schaap e Fichaut, 2017b).

2. EVOLUÇÃO ESPERADA COM OS PROJETOS EMODNET E SDC

As divisões técnicas do IH participam nos diferentes grupos temáticos da rede EMODnet desde as fases preliminares de edificação da rede, contribuindo com o *know-how* dos investigadores e com dados para os portais temáticos.

Os principais desenvolvimentos que são atualmente perseguidos são (EMODnet Secretariat, 2017):

- Produzir o mapa das bacias marinhas europeias com uma maior resolução recorrendo à assimilação de novos dados batimétricos que permitem um adensamento das malhas já

disponibilizadas no portal EMODnet *Bathymetry*;

- Desenvolver a assimilação de novas tipologias de dados, como os fluxos de água doce dos rios europeus;
- Assimilar dados que permitam melhorar o conhecimento da distribuição espacial dos plásticos.
- Criar uma ferramenta de *crowdsourcing* que permita a assimilação de dados classificados como dados abertos marinhos. Existe uma forte convicção que existem muitos dados com interesse para a comunidade, fora das redes e infraestruturas europeias, que importa assimilar através um processo de facilitação do trabalho de quem contribui.

A rede EMODnet tem apresentado uma estratégia de desenvolvimento baseada no incremento dos nós da rede através da cativação de centros de dados europeus familiarizados com as cadeias de gestão de informação marinha, desde a aquisição dos dados, passando pelos processos de processamento e controlo de qualidade, finalizando no arquivo dos dados e informação, e, por fim, a gestão da disponibilização da informação e conhecimento à comunidade científica. A grande novidade nos desenvolvimentos da rede EMODnet está a ser a implementação de um novo portal que pretende dotar a rede da capacidade de *crowdsourcing*. Esta inovação oferece alguns desafios, e.g. a promoção do processo e as ações de *marketing* orientada a uma comunidade não familiarizada com as questões técnicas da gestão de informação.

O projeto *EMODnet Ingestion and safe-keeping of marine data* tem como objetivo a edificação da componente de *crowdsourcing*. O *Data Ingestion Portal*, a portal de entrada do projeto, está acessível no endereço: <https://emodnet-ingestion.eu/>.



Fig. 2. Vista do EMODnet Data Ingestion Portal (EMODnet, 2017).

A submissão de dados no portal é um processo simples, começa com a criação de um *MarineId* que identifica o utilizador na rede EMODnet. Após este passo inicial, a submissão de dados consiste no

preenchimento de um formulário de metadados que acompanha o conjunto de dados submetido no portal numa pasta comprimida (*Zip folder*) (Fig. 3).

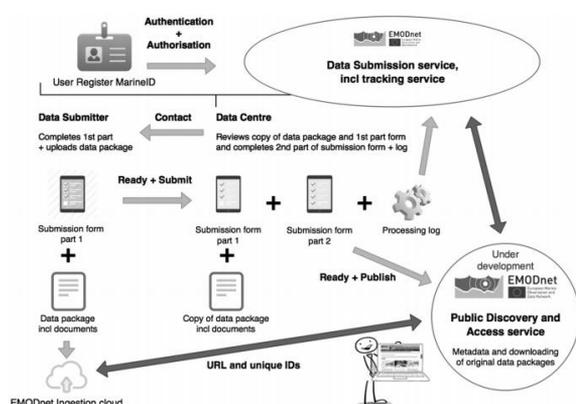


Fig. 3. Perspetiva do processo de submissão de dados no Data Ingestion Portal. Extraído de Schaap (2017).

Depois de concluída a fase I do processo de submissão, dá-se início à segunda fase durante a qual os centros de dados nacionais da rede EMODnet realizam a análise da completude do formulário preenchido na fase I e procedem a uma análise da coerência dos formatos de dados. Após este estágio, realizam o processo de harmonização dos dados por forma a ficarem compatíveis com os diversos catálogos e acessíveis nos diferentes portais temáticos da rede EMODnet e infraestruturas europeias (Fig. 4).

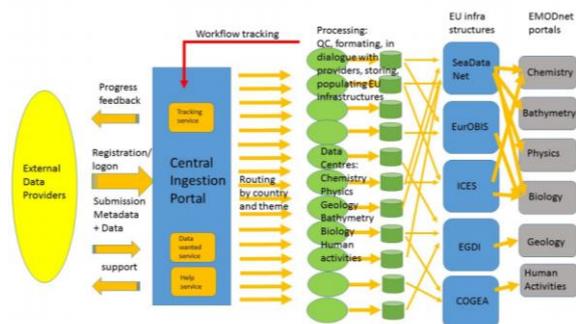


Fig. 4. Workflow do processo de submissão e assimilação de dados. Extraído de Schaap (2017).

Através do processo descrito atrás é possível combinar o *know-how* técnico dos centros de dados com o cidadão para maximizar o valor dos dados e informação, gerando um benefício comum para toda a comunidade de utilizadores.

O *Data ingestion portal* apresenta também um serviço facilitador do processo de procura de dados – o *Data Wanted service* (<https://www.emodnet-ingestion.eu/data-wanted>). Com este serviço os utilizadores identificados com um *marineID* em busca de conjuntos de dados específicos podem dar a conhecer à comunidade EMODnet as suas necessidades através da publicação de *posts* descrevendo os dados ou informação procurada. A comunidade EMODnet tem-se esforçado por criar

sinergias entre os utilizadores através de ferramentas facilitadoras e inovadoras como o *Data wanted Service*. A Fig. 5 apresenta uma visão geral sobre a organização da rede EMODnet.

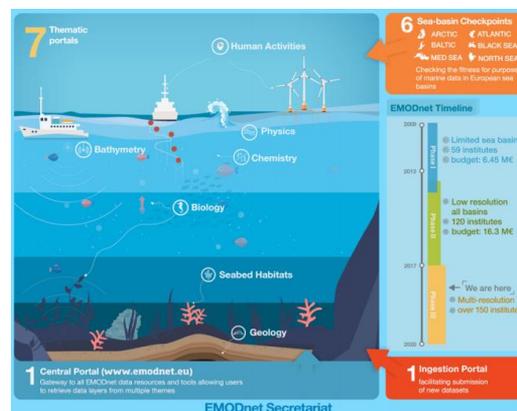


Fig. 5. “EMODnet at a Glance”. Extraído de EMODnet Secretariat (2017).

O projeto SDC é a edição atual de desenvolvimento da infraestrutura SDN (Fig. 6) estando integrado no programa H2020. Com o SDC o consórcio está a trabalhar no desenvolvimento tecnológico, na integração de novas capacidades de processamento e de processos de geração automático de conhecimento. Os dois campos principais de inovação são: a integração da capacidade de *semantic web* e *machine-to-machine services and interfaces*, e o desenvolvimento de um *Virtual Research Environment* (VRE) que coloca à disposição dos investigadores as capacidades de processamento de alto desempenho através de servidores virtuais na *cloud*. Isto é conseguido com parcerias com as infraestruturas como a EUDAT *Collaborative Data Infrastructure* (CDI) (Fig. 7).

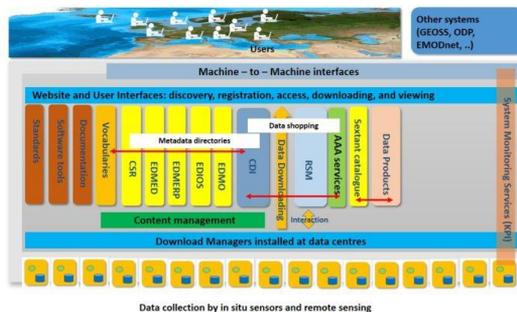


Fig. 6. Arquitetura atual da infraestrutura SDN. Extraído de Schaap e Fichaut (2017a).

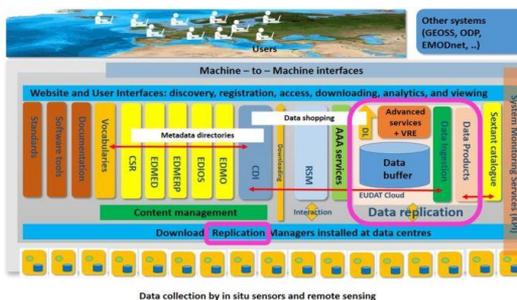


Fig. 7. Arquitetura futura SDN/SDC. Extraído de Schaap e Fichaut (2017a).

O resultado esperado com os desenvolvimentos é tornar mais eficientes os acessos aos serviços e ferramentas disponíveis na infraestrutura e continuar a tornar acessíveis os dados marinhos nos catálogos da infraestrutura (Fig.8).

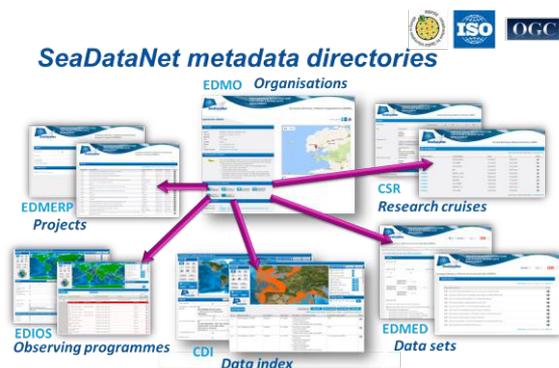


Fig. 8. Catálogos da infraestrutura SDN. Extraído de Schaap e Fichaut (2017a).

Com os desenvolvimentos operados existe a ambição que a infraestrutura SDN/SDC e a rede EMODnet continuem a agregar cada vez mais recursos informacionais do meio marinho e que continuem a ser referências no contexto europeu.

3. CONCLUSÃO

As infraestruturas de informação geográfica assumiram um papel preponderante no contexto regional. A partilha de dados de qualidade é hoje reconhecida como vetor fundamental para entender o mundo que nos rodeia.

Os dois exemplos apresentados no artigo demonstram que as sinergias criadas em torno de projetos europeus com instituições de uma ampla distribuição espacial, contextos e realidades muito diferentes, são importantes para criar comunidades de investigação e atingir uma agregação de informação que seria impossível de outra forma. Todos estes dados contribuem para um conhecimento global dos processos transfronteiriços. A dimensão destas redes obriga a um investimento de manutenção que dificilmente seria suportável sem o apoio de fundos comunitários.

Os projetos nestas áreas estão alinhados com as estratégias europeias na área da economia azul e com as principais tendências de investigação e desenvolvimento: *crowdsourcing*, *semantic web*, *machine-to-machine interfaces* e *cloud computing*.

Os projetos apresentados continuam o seu ciclo de desenvolvimento com esforços que decorrem desde a década de noventa acumulando dados de vários países e o esforço e conhecimento de investigadores com o objetivo último de contribuir para um melhor ambiente e para o bem comum. É percepção geral que os dados e informação são um recurso dispendioso de adquirir que é muito importante preservar, explorar e partilhar. Só assim os dados guardados

em vários centros europeus ao longo de várias décadas podem contribuir para um conhecimento cada vez mais exato dos processos ambientais e possibilitar melhores decisões regionais para gestão do ambiente.

Agradecimentos

O projeto SDC é financiado pelo programa H2020 através do *grant agreement* 730960. O projeto *EMODnet Ingestion and Safe-keeping of marine data* é objeto de financiamento pelo *European Maritime and Fisheries Fund* (EMFF) através do contrato: EASME/EMFF/2015/1.3.1.3/SI2.727770.

REFERÊNCIAS

- Comissão Europeia 2012. COM(2012) 473 final - Green Paper Marine Knowledge 2020: from seabed mapping to ocean forecasting.
- Comissão Europeia 2014. COM(2014) 254 final - Marine Knowledge 2020: roadmap.
- Directorate-General for Maritime Affairs and Fisheries 2018. Ares(2018)316404 - Evaluation of the European Marine Observation and Data Network (EMODnet).
- EMODnet. 2017. *EMODnet Data Ingestion Portal* [Em linha]. Disponível em: <https://emodnet-ingestion.eu/> [Consult. maio 2018].
- EMODnet Consortium. s.d. *EMODNet - Central Portal* [Em linha]. Disponível em: <http://www.emodnet.eu/> [Consult. maio 2018].
- EMODnet Secretariat 2017. *EMODnet Annual report 2016*.
- Europeia, C. 2016. *Marine data infrastructure study* [Em linha]. Disponível em: https://webgate.ec.europa.eu/maritimeforum/en/node/3942#_ftnref3.
- Schaap, D. 2017. *EMODnet Ingestion and safe-keeping of marine data - 1st Annual Interim Report*.
- Schaap, D.M.A. e Fichaut, M. 2017a. *SeaDataCloud - Further developing the pan-European SeaDataNet Infrastructure for marine and ocean data management*. *Em EGU2017 ed*.
- Schaap, D.M.A. e Fichaut, M. 2017b. *SeaDataCloud - Further developing the pan-European SeaDataNet infrastructure for marine and ocean data management*. *EGU General Assembly 2017*.
- SeaDataNet Consortium. s.d. *SeaDataNet - PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT* [Em linha]. Disponível em: www.seadatanet.org/ [Consult. maio 2018].