

Representação do Conhecimento Gerado pelo Projeto AgroHidro: Estrutura Conceitual e Aplicação Computacional

Rodrigo Bonacin^{1,2}
Olga Fernanda Nabuco¹
Ivo Pierozzi Júnior³

Um dos objetivos da organização e da representação do conhecimento é oferecer uma estrutura conceitual, formalmente construída, viabilizando uma visão compartilhada sobre uma realidade e, ao mesmo tempo, viabilizando aplicações computacionais sobre tal conhecimento como, por exemplo, recuperação da informação. No contexto do trabalho colaborativo, executado em rede, como o caso do Projeto AgroHidro, recursos de organização e representação do conhecimento minimizam dificuldades nos processos de geração, tratamento, compartilhamento, disseminação e reuso do conhecimento. As ontologias de domínio têm-se apresentado como alternativa para explicitar de maneira formal conceitos e relações compartilhadas dentro de uma determinada área de assunto. Nessa perspectiva, a ontologia OntoAgroHidro está sendo construída e descrita em OWL 2.0 usando o software Protégé v. 4.0 e, em seu atual estado de desenvolvimento, reúne mais de 8500 conceitos do domínio, ou seja, os impactos da agricultura e das mudanças climáticas nos recursos hídricos dos biomas brasileiros. O processo geral de construção da OntoAgroHidro envolveu o trabalho de engenheiros do conhecimento e especialistas de domínio e foi iniciado com a proposta de uma estrutura conceitual de alto nível, envolvendo os seguintes conceitos: transformação ambiental; objeto da transformação ambiental; agente da transformação ambiental; definições; dimensões; evento; agente de informação. Na sequência, esses conceitos gerais foram granulados em subconceitos mais específicos, onde puderam ser categorizados e interrelacionados a maioria dos fenômenos naturais ou antrópicos envolvidos no domínio modelado. Dessa forma, foram representados conceitos como: definição biológica, química e física; definição espacial e temporal; dimensão ambiental, socioeconômica, espacial e temporal; agricultura; mudanças climáticas; sistemas de produção; uso e cobertura das terras; biomas; qualidade e quantidade da água; propriedades físicas, químicas, biológicas da água, assim como toda uma gama de subcategorias relacionadas a essa conceitualização mais geral, permitindo chegar às instâncias, tais como nomes científicos de animais, plantas ou microrganismos indicadores de qualidade biológica da água; nomes químicos de pesticidas; nome de bacias hidrográficas e rios componentes, etc. Nesse processo de granulação e categorização conceitual foram reusados e agregados padrões de intercâmbio de dados e ontologias já consolidados ou validados em outros contextos como, por exemplo, o Dublin Core (produção e documentação técnico-científica); o Perfil de Metadados Geoespaciais do Brasil (dados geoespaciais); as ontologia Cuahsi (desenvolvida por um consórcio de universidades americanas para monitoramento da qualidade da água nos Estados Unidos) e, em parte, o conjunto de ontologias SWEET (desenvolvido pela NASA para representação do Planeta Terra). Um aspecto metodológico importante e que facilitou o desenvolvimento da OntoAgroHidro foi a interação forte entre os engenheiros do conhecimento e os especialistas de domínio, processo no qual foram utilizados mapas conceituais como ferramentas para facilitação do entendimento conceitual do domínio. Paralelamente à estruturação conceitual do domínio, um extenso exercício terminológico está sendo desenvolvido. Considerando que a unidade elementar do conhecimento é o conceito e que todo conceito é nominado por um termo, pretende-se extrair da OntoAgroHidro uma terminologia, por si igualmente representativa do domínio modelado. Essa terminologia, trabalhada nos moldes de um vocabulário controlado e nos padrões de um tesouro, pode originar mecanismos mais efetivos e eficazes de recuperação da informação, com base em linguagem natural. A versão atual da OntoAgroHidro já constitui uma conceitualização interessante do domínio ao qual pretendeu representar. Algum esforço adicional ainda é necessário para ultrapassar limitações de compreensibilidade e escopo. Um volume enorme de conceitos ainda é necessário ser incorporado no modelo atual para garantia de mais precisão e efetividade de representação. O mapeamento da estrutura conceitual, assim como da terminologia, com outros sistemas de organização do conhecimento como os tesouros é uma continuidade do trabalho. Outras etapas envolvem questões relacionadas às aplicações computacionais da ontologia, como problemas de interoperabilidade, para permitir sua utilização como ferramenta de recuperação de informação em bases de dados distribuídas, por exemplo.

¹ Centro de Tecnologia da Informação Renato Archer, Rod. Dom Pedro, Km 143,6, 13069-901 Campinas, SP

² FACCAMP, Rua Guatemala, 167, 13231-230 Campo Limpo Paulista, SP

³ Embrapa Informática Agropecuária, ivo.pierozzi@embra.br