ABSTRACT: Many science archive centres publish very large volumes of image, simulation, and experiment data. In order to integrate and analyse the available data, scientists need to be able to (i) identify and locate all the data relevant to their work; (ii) understand the multiple heterogeneous data models in which the data is published; and (iii) interpret and process the data they retrieve. RDF has been shown to be a generally successful framework within which to perform such data integration work. It can be equally successful in the context of scientific data, if it is demonstrably practical to expose that data as RDF. In this paper we investigate the capabilities of RDF to enable the integration of scientific data sources. Specifically, we discuss the suitability of SPARQL for expressing scientific queries, and the performance of several triple stores and RDB2RDF tools for executing queries over a moderately sized sample of a large astronomical data set. We found that more research and improvements are required into SPARQL and RDB2RDF tools to efficiently expose existing science archives for data integration.