

An Information Integration Appliance

Todd C. Hughes
Defense Advanced Projects Agency
Information Exploitation Office

Commercial, government, and military organizations complain bitterly that their information is not adequately integrated to meet operational objectives. However, research and development in information integration does not enjoy support commensurate with the severity of the problem. Part of the explanation of this disparity is surely that the very concept of “information integration” connotes an abstract, even ethereal, capability that is difficult for stakeholders to make concrete. Even worse, success of some information integration technologies is predicated on a global adoption of expensive, unproven, and equally abstruse systems: service oriented architectures and the Semantic Web being the most notable examples. Information integration research would fare better under a more tangible and affordable model.

One approach is to embed information integration capabilities in a dedicated physical object: a black box of sorts. This could be viewed best as a network appliance, similar in many respects to a router or intrusion detection system. In general terms, network appliances intercept network communication, perform operations on it, and send the information on its way to its intended destination. Similarly, an information integration appliance would intercept content (e.g. queries, reports) from an information source, transform it into syntactically well-formed and semantically equivalent messages, and send them to other sources. Such transformation would be transparent to the resources themselves.

Situating information integration into a physical box doesn't by itself solve any technology problems, but it does provided a new perspective on use cases and requirements. An information integration device would, by operational necessity, be programmable in the field. Further, management of the device should not be under the sole purview of experts in database administration. Therefore, the device would need sufficient intelligence, learning, and interface features to enable data users to encode the proper information transformation routines in a timely manner. Many of the recent advances in integration technology—from data model extraction to schema mapping to data visualization—would need to be built into the device in order to achieve this goal.

The appliance model of information integration is not without challenges, both technical and operational. However, as a vision for information integration it may enable a clear articulation of technical barriers, metrics, evaluation scenarios, deployment strategies, and return on investment. It may also lift the veil of abstraction that obscures the benefits of the technology, and provide a way for the data users to be directly involved in meeting their operational objectives.