Microsoft Office Dynamic Documents as R Applications

Josh van Eikeren[1,*] and Paul van Eikeren[1]

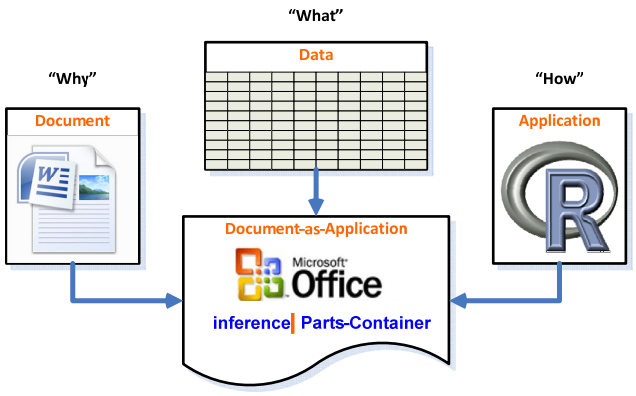
1. Blue Reference, Inc., Bend Oregon USA
* Contact author: Josh.van.Eikeren@BlueReference.com

Keywords: Dynamic Documents, reproducible research, Microsoft, Word, Excel

Business and technical luminaries argue that future competitive success will be based on the application of data, analysis, predictive modeling, and fact-based decision making. "Competing on Analytics" makes use of a class of software, like R, which enables a range of techniques from statistics to data mining to analyze historical data and make predictions about future events. However, despite these sophisticated developments, common documents remain central to business processes. Documents are portable, persistent and provide contextual views of information organized for the purpose of relating disparate pieces into actionable knowledge.

Modern business practices (e.g., Business Process Management) are driving the convergence of documents and predictive analytics. Business data is typically focused on the "what" of the business and maintained in highly structured databases. Business documents tend to focus on the "why," are unstructured and contextual, and are usually maintained in separate document repositories. And, predictive analytic applications contain the "how" of the business, are highly structured, but typically lack an appropriate storage mechanism. The reality is that business is done at the intersection of "what", "why" and "how"—where facts and context meet actionable analysis.

Traditionally, the domains of data, documents and predictive analytics application have been isolated from one another. Dynamic documents provide a means to combine the strengths of documents with the power and flexibility of predictive analytics applied to data. In fact, dynamic documents can act like situational software applications by combining the best attributes of applications and static documents. To address the needs for dynamic documents, we have extended the capabilities of Microsoft Office to enable embedding a structured entity we call a Parts-Container for holding and managing data (e.g., data frames), software objects (e.g., ASCII and binary files), code blocks (e.g., R scripts) and inline expressions (e.g., R commands). Since the elements of the Parts Container are linked to a computational engine like R, the documents become applications.



In this presentation, we will outline several case studies that illustrate the construction and application of dynamic documents in conjunction with the R statistical computation environment towards

- test-driven new method development;
- reproducible research; and
- study management in pharmaceutical development.

References

Josh van Eikeren (2009). Inference for R,
http://InferenceForR.com.

Paul van Eikeren (2009). Inference Parts-Container Explained,
http://inferenceforr.com/documentation/Documents/ProductOverviews/Inference%20Parts%20Container%20Overview.pdf