Supporting Knowledge Transfer through Decomposable Reasoning Artifacts

William Pike, Richard May, and Alan Turner
Pacific Northwest National Laboratory
{bill.pike, richard.may, alan.turner}@ pnl.gov

Abstract

Technology to support knowledge transfer and cooperative inquiry must offer its users the ability to effectively interpret knowledge structures produced by collaborators. Communicating the reasoning processes that underlie a finding is one method for enhancing interpretation, and can result in more effective evaluation and application of shared knowledge. In knowledge management tools, interpretation is aided by creating knowledge artifacts that can expose their provenance to scrutiny and that can be transformed into diverse representations that suit their consumers' perspectives and preferences. We outline the information management needs of inquiring communities characterized by hypothesis generation tasks, and propose a model for communication, based in theories of hermeneutics, semiotics, and abduction, in which knowledge structures can be decomposed into the lower-level reasoning artifacts that produced them. We then present a proof-of-concept implementation for an environment to support the capture and communication of analytic products, with emphasis on the domain of intelligence analysis.

1. Introduction

Many fields in which the need for knowledge management is particularly acute are characterized by an emphasis on hypothesis generation. In these fields, such as intelligence analysis and scientific research, tools and methods to support knowledge transfer need to emphasize the primacy of the reasoning process in constructing and applying knowledge resources. Collaborating inquirers in these fields – and collaboration is often implicit in the sharing of resources – need to be able to describe, communicate, and reuse representations of their reasoning processes in order for the knowledge artifacts they produce to be useful to colleagues. "Decomposability" should be a central feature of knowledge representations produced in analytic environments; the ability to traverse levels of abstraction in the representation of knowledge means that inquirers can recover the information they need in the form they want, thereby increasing the likely utility of shared information.

Ways to measure this utility include the efficiency with which members of an organization execute tasks or the quality of their decisions [1]. In discovery-based applications, both efficiency and quality are served by preserving audit trails of reasoning. Knowledge is often constructed incrementally through cooperative belief revision; new observations are collected, organized, integrated into explanatory structures (often with the aid of supporting resources such as assumptions and models), interpreted, and the results passed on again. Auditability enhances verification and decision making during this process by supporting evaluation of knowledge structures. Audit trails also result in reproducible analyses that can be efficiently repurposed by others.

Supporting knowledge transfer in contexts founded on hypothesis generation puts the need for interpretation and evaluation at the fore. However, both institutional and technological barriers make transferring knowledge across an organization difficult. Different actor roles in an organization suggest different requirements for both the content and presentation of information – not every actor in a knowledge-transfer pathway has the same information needs, desires, or perspectives. Instead of contending with the full detail of a complex analysis, stakeholders typically only want information relevant to their goals [2]. It is therefore desirable to support expressions of belief at a range of levels of abstraction, allowing stakeholders to receive information in a manner appropriate to their role. The technical barriers to this degree of "on-the-fly" transformation derive from the need to support current tools and methods. For instance, text documents are the primary agents of information exchange in many organizations. The content of the document is up to the producer, and what is not written is not accessible to the consumer. Moreover, it can be difficult to faithfully represent an

Proceedings of the 40th Annual Hawaii International Conference on System Sciences (HICSS'07)
0-7695-2755-8/07 $20.00 © 2007 IEEE

analytic process, complete with its cycles and dead ends, in a linear document. Without imposing radically new work practices, it is desirable to enrich existing product formats and information flows such that information consumers can query them for more content than was provided by the producer, while preserving their familiar format.

Here, we address the problem of knowledge transfer in abductive environments (that is, those directed toward hypothesis generation and testing, following from Pierce [3] and discussed further in Section 2.2) comprised of stakeholders with different information needs and preferences. We identify three core functions for knowledge management (KM) technology in support of these environments:

1. Bridge the divide between collection, analysis and communication, enabling representations of the reasoning process to be captured *in situ* and associated with the knowledge products in which they result.
2. Allow consumers of shared knowledge to "unwrap" artifacts, recovering the reasoning process that produced a conclusion so that it can be evaluated or reused.
3. Apply user-specific transformations to communicate knowledge through varying levels of abstraction while preserving the capability for drill-down to more granular representations.

In Section 4, we present a proof-of-concept implementation of these ideas, with theoretical bases in hermeneutics, semiotics and abduction, that demonstrates how computational artifacts can be used as vectors for information flow in an organization.

1.1. KM needs in Intelligence Analysis

The core functions noted above, and in particular the problem of representing reasoning strategies during knowledge transfer, are materially important to the domain of intelligence analysis (IA). In weighing alternative explanations or scenarios, analysts need to remain cognizant of their assumptions and justifications as well as knowledge gaps and uncertainty. Their reports, which constitute a representation of their knowledge transferred to colleagues and decision makers, need to be interpreted in light of these factors [4]. The technique of "analysis of competing hypotheses" is sometimes used in this community to expose the implications of analysts' assumptions [4], which results in a style of inquiry similar to the Hegelian dialectic postulated for other types of inquiring organizations. Hypothesis generation in intelligence contexts aligns with theories of "learning-oriented knowledge management systems" [5] inasmuch as knowledge creation in IA hinges on representing multiple perspectives in an organizational memory.

Organizational culture is such that analysts with different domain expertise, and thus potentially different worldviews, need to be able to use each others' findings (horizontal communication), as do consumers at different positions in an organizational hierarchy (vertical communication). Communication problems are thus inherent to IA communities of interest. Context dependencies are introduced that go beyond location, time, or application and attentive state – elements of context that we might currently be able to detect [6] – to one's context within a knowledge pathway (or more likely, cycle) from collection to analysis and decision. To support cross-context and cross-role interpretation of information – the movement of information from the analytic space into an action space – the information resources that IA stakeholders use can be conceptualized as boundary objects [7]. Such objects are concepts that take on different meanings in different communities, making it necessary to clearly define the roles they play not just to make possible negotiated agreement, but communication at all.

As a domain often marked by more data than time or people to analyze it, demand-side models for knowledge sourcing are particularly germane in IA. A demand-side model looks at the design and evaluation of knowledge management tools not from the amount of content that is available or the ease with which users can contribute information (increasing the knowledge *supply*), but at how and why it is retrieved and used (increasing the knowledge *demand*) [8]. Since intelligence analysis hinges on the perspectives and beliefs of its practitioners in what Johnston [9] calls a "socio-cognitive process", satisfying knowledge demand requires mapping between sets of biases. The approach we develop in this paper is an attempt to capture and express these perspectives.

2. Foundations for cooperative knowledge construction

A theoretical basis for the three core functions identified in Section 1 can emerge from the compatibilities among hermeneutic, abductive, and semiotic models of understanding. This foundation can lead to knowledge transfer techniques that support analytic process capture, communication and audit, while acknowledging the bias and assumptions inherent to IA.

2.1. Hermeneutics as KM design criteria

Interpretation is the basis of hermeneutics, which places the individual in a historical context that colors his or her judgments [10]. The intelligence analysis process is a quintessentially hermeneutic enterprise: interpretation occurs at every stage of analysis, from the relationship between observations and reports to ground truth, and between received conclusions and the degree to which they logically follow from the process of their construction. Hermeneutic principles have been extended to knowledge-based information systems under the argument that judgment is inherent to any act of information creation and use in a digital environment [11]. Because an analyst's context is constantly evolving, hermeneutics stresses that understanding a concept comes from the ability to apply it in a current situation. Information systems to support the IA process must therefore be able to represent the recursion and discursion of understanding. In such systems the user works within a trail of concepts and interpretations that constitute the rational foundation for his or her work. During the process of inquiry, the relationships between interpretations should be preserved so that changes in one can be used to detect cascading effects in others. By examining these interpretations, the user enters a dialogue with other elements of a problem, and the other thinkers who have engaged it. The process of searching, evaluating, and reusing information over time suggests a spiral model for the development of understanding in knowledge-based systems [12]. Thus emerges the first core function introduced in Section 1: capturing the flow of reasoning so that the process of discovery can be shared and evaluated.

In a hermeneutic approach to representing the IA process, information reaches the recipient bearing a message that is not entirely apparent (and not entirely objective – one is complicit in recognizing it, so one's biases factor into the meaning that is derived from it). The consumer of a knowledge artifact interprets that message in light of his or her own experiences. In computational environments, the problem of interpreting meaning can be facilitated by an ability to encode each others' interpretations and perspectives into the digital artifacts we share. In inquiring organizations, hermeneutics has been shown to provide a foundation for this process of distributed cognition [13].

Gadamer's [10] suggestion that "we can understand a text only when we have understood the question to which it is an answer" is particularly apropos in the domain of cooperative information analysis. Aids to interpretation ought to capture the provenance and evolution of knowledge artifacts as they are manipulated by analysts; once represented, these can help subsequent discourse by exposing the analytical process of their construction. Recovery of this process is a manifestation of Gadamer's "historically effected consciousness", which is not merely an acknowledgement that a particular work has a history, but that this history is inseparable from understanding the work. Appreciating a resource's history does not suffice to put the questioner in the resource's historical context, however, nor is it even possible to transcend one's contemporary context to visit a historical resource in its native context. Rather, exposing the history of an analysis allows one to ask questions of resources in an effort to delineate what in hermeneutics are called "horizons". Both questioner and artifact have horizons, which bound the concepts important from an analyst's perspective. The hermeneutic act of fusing horizons occurs between the consumer of an information product and the expression of its producer's horizon encoded within. To Gadamer, fusion is only accomplished through questioning. Interrogating an intelligence product allows the analyst to begin eliciting the perspective of its author. Preserving these perspectives to enable interactive questioning of an intelligence product accomplishes the second goal identified in Section 1, that of allowing consumers to "unwrap" knowledge artifacts.

Gadamerian hermeneutics identifies a collection of activities involved in interpreting texts, which we can extend to apply to digital resources. Indeed, through automated versioning and provenance capture, computational environments may have a unique ability to support hermeneutics in ways previously impossible.

- **Dialogue.** Questioning an information product to elucidate its horizon involves dialogue, whether between two people attempting to achieve mutual understanding, between a reader and a text, between data user and data producer, and so on. With traditionally static information resources, the horizon of the text can only be inferred by the receiving analyst; with knowledge management tools capable of presenting the provenance of a resource and the directed aims of its creators, the implicit dialogue that leads to understanding between and among analysts and resources can be made more explicit.
- **Play.** Hermeneutic play describes the back-and-forth that occurs between questioner and text during the process of understanding. Play reinforces the free flowing nature of possibilities during the creation of meaning, but the playful

0-7695-2755-8/07 $20.00 © 2007 IEEE

exploration involved in manipulating tools and data is lost in the traditional information system. If we develop systems capable of capturing the recursive processes of collecting information and generating and evaluating hypotheses, it may become possible to both share and reuse the process of experimentation that leads to knowledge construction in analytic settings.

- **Prejudice.** In contrast to information systems founded in a Cartesian rationality in which the biases of inquirers are disregarded, hermeneutics acknowledges that pre-judgments are the basis of our ability to have experience. These judgments guide our exploration of an information resource as we attempt to elicit its horizon [14]. The horizon of an information resource includes the problems to which it is relevant, but is generally implicit in the representation of that resource. Augmenting this representation with a description of its horizon can lead to its reuse in appropriate situations.
- **The hermeneutic circle.** Following from play, the hermeneutic circle defines the process by which we experiment with possible interpretations. Making one interpretation may change the way we interpret another resource. Interpretation can thus be circular, or at least involve feedback with earlier interpretations, making it impossible to model analytic inquiry as a teleological process. If systems can track how particular resources are interpreted through use, it becomes possible for them to detect other interpretations that will be strained by new developments.
- **Breakdown.** When theories fail to explain experiences or observations, a condition of breakdown ensues. While the breaking down of a prior interpretation necessitates the construction of a new interpretation, hermeneutics does not see this as a condition of failure so much as an opportunity – breakdown is typically what creates a space for ingenuity. Tracking the use of resources helps detect conditions of breakdown, pinpointing the circumstances surrounding new insights that result from hypothesis failure.

2.2. Peircean abduction and semiotics

Many of the hermeneutic principles on which we can ground a new knowledge management approach for information analysis bear similarities to the theories of C.S. Peirce. Peirce argues for a logic of inquiry to augment Aristotelian logic of classification (deduction) and Kantian logic of judgment (induction). He argues that *discovery* is a privileged kind of cognition that supplants other modes through which humans generate beliefs about the world. Beliefs created during a process of discovery, while fallible, are self-correcting [3] – it is the nature of hypothesis construction to accommodate continuous revision. In the Peircean model, knowledge emerges from the interaction between three stages of inquiry: the first, abduction, relies on observations to stimulate possible hypotheses by what Peirce calls "an appeal to instinct." Deduction follows, in which the consequences of those hypotheses are examined. Lastly, inductive hypothesis testing selects the most likely explanations. The self-correcting nature of this inquiry model arises from the inductive step, as induction's shortcomings (it is not truth-preserving, so future observations may alter or contradict a hypothesis) turn into advantages. As a body of observations grows, induction converges toward truth. To make this long-term convergence efficient, it is important to leverage the reasoning structures of many collaborating inquirers.

Peirce's model for abductive inference aligns with subsequent developments in hermeneutics, giving us a theory of knowledge management that can support inquiring communities. Hermeneutic interpretation, particularly the circular nature of dialogue and play, is compatible with constant revision of the abduction-deduction-induction process. (Peirce uses the same concept of dialogue to describe the inquirer as engaged in dialogue with a "future self" who is coming into being [15]). Abduction also underscores the importance of pragmatic manipulation of observations (hermeneutic play). It is only through this manipulation that the spark of hypothesis generation can light. Finally, the need to devise computational methods to represent the *process* of inquiry (the first core function identified in Section 1), and reference this process in the *products* of inquiry (the second core function), follows from Peirce's view that the highest degree of explanatory power derives from those explanations that are grounded in pragmatic experience.

The manipulation of information resources in the course of hypothesis generation and testing can be further informed by Peirce's semiotics. This theory of signs can help articulate the different roles that resources play during the course of analytic inquiry as well as their meaning to different actors in a cooperative analysis. In Peircean semiotics the sign is a triadic relationship between representamen or sign vehicle, interpretant, and object. The testing of possible interpretations in both abduction and hermeneutics is reflected in the transitivity between these components. In an analytic setting the sign vehicle could be a data resource that stands for some

0-7695-2755-8/07 $20.00 © 2007 IEEE

observation on the ground (its object) and is given meaning (a unit of knowledge the analyst applies to connect representamen and object) through the interpretant. The choice of interpretant is colored by the analyst's bias; each recipient of an information resource may interpret it different and consequently use it toward different ends in the course of inquiry. Communicating knowledge among analysts in collaborative environments can thus require preservation of the *significance* each attaches to resources (in the semiotic sense, significance could indicate what digital resources are used to stand for a given concept – different analysts might use different evidence to ground the same argument, for instance). In particular, Peirce notes that semiotics is involved in the communication of analyses from one person to the next; each interpretant is a new sign in the mind of the person to whom it is communicated. To support effective interpretation as knowledge flows through an organization, knowledge management tools for cooperative inquiry should be capable of representing the signifiers and interpretants that users attach to digital resources. A semiotic model also suggests how different knowledge representations can signify the same underlying concepts, supporting the third core function for knowledge management in IA: communicating knowledge in consumer-customizable formats.

2.3. Tools in support of communicative inquiry

There is growing recognition that information systems in support of collaborative inquiry need to do more than facilitate file sharing. Knowledge management tools should reflect the situated work practices of their users [16] and accommodate the dialogical, manipulative nature of exploration [17, 18]. Abduction is becoming accepted as a useful model for computer-based hypothesis creation because it produces creative information even under conditions of hypothesis failure [19] and explicitly supports cooperative belief revision [20]. Indeed, there is a need for what might be called "creativity support systems" – tools to help collect information, analyze it, collaborate with others, and distribute resulting products [21]. Moreover, calls for a "Pragmatic Web" (that explicitly enables communities to test, refine, and implement emergent, experience-based, solutions) counterbalance the purely deductive nature of much contemporary knowledge representation (such as in ontologies for the Semantic Web) [22]. Empirical studies of collaborative inquiry in investigative settings, notably [23], have indicated user desire for (a) the ability to link information artifacts to the reasoning processes in which they figure, as an aid to keeping a community informed about the state of its knowledge, (b) the ability to generate "big picture" reports, and (c) minimizing ontological complexity. Some extant tools accomplish a few of these objectives: [24] provides a utility for describing observations, beliefs, and uncertainty associated with online documents using semi-formal semantic markup; [25] enables communication of user-created and automatically generated "knowledge maps" in document analysis applications. There is also a body of work regarding techniques for capturing analytic events in context (e.g., [26] for computational tools in general and [27] for the IA domain), although this context is not yet used as a hermeneutic aid.

In the remainder of this paper we discuss a knowledge management approach that, in meeting the three core objectives outlined earlier, uses the principles of hermeneutics and abduction to support analytic inquiry and produce flexible, "horizon-aware" representations of analytic results.

3. Production and reproduction

Our approach to modeling and communicating analytic tasks focuses on readily-interpretable communications based on a messaging model we call *rich products*. The rich product transcends traditional static messaging, where content and presentation are conflated and hermeneutic dialogue between product and consumer is passive. The rich product, by contrast, is intended as a dynamic resource that can actively engage in dialogue with the consumer, bringing previously suppressed information to the fore and granting access to representations of the analytic process so that the consumer can easily recover the author's reasoning.

Support for the hermeneutic circle is a core aspect of the rich product model. The products a community of inquirers creates are not ends in themselves but grist for continued analysis, a result of unavoidable semiotic interpretation during the communication of information. Thus the rich product is a computational reflection of the evolving nature of knowledge (Figure 1). A reporting cycle might begin with reports ingested from colleagues, transformed to accommodate the recipient's mental models, decomposed to evaluate reasoning for quality or reusability, augmented with further analysis, rolled into a new presentation format, and passed on, where the cycle continues.

The rich product makes explicit the separation between content and presentation such that the product can be modeled as a decomposable set of resources

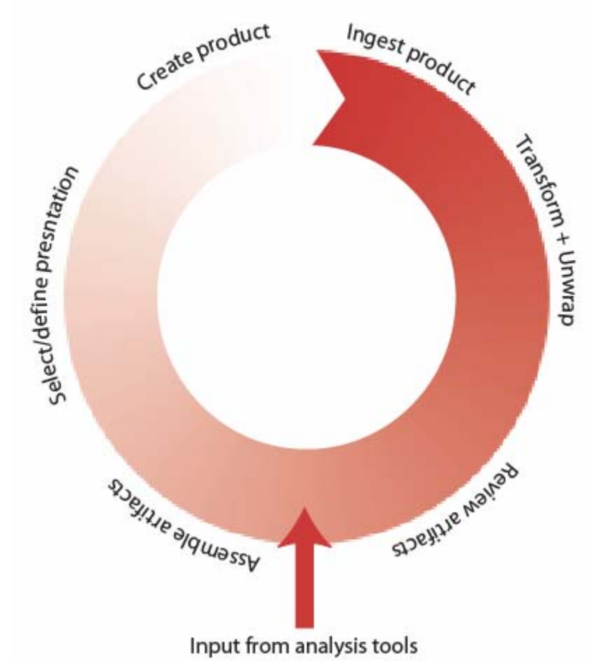

Figure 1. Rich product reporting cycle.

joined in a rhetorical structure, to which a set of presentation rules is applied. Each of the resources comprising the rich product may be an atomic unit (that is, an artifact that represents a single piece of source information) or a synthetic unit (a collection of atomic elements gathered into a rhetorical package). The product itself is a collection of references to digital artifacts that reside in a knowledge base external to the product.

Rich products allow a distinction to be made between knowledge management and simple document management; a document is but one representation for a collection of knowledge resources. A knowledge management system that follows the rich product model is not a collection of *products*, then, but a collection of artifacts that can be composed into products as required. (To enable this representation of knowledge we require an authoring environment in which the construction of knowledge is treated separately from the construction of reports based on that knowledge, and this effort is taken up in Section 4). Each of these artifacts must maintain references to its history of manipulation and application, if cooperative abductive inquiry is to be made possible. Detecting the history of play – how multiple analysts have applied a given piece of evidence, say – is difficult if knowledge is encoded in reports *before* it is stored in a knowledge management tool.

We define a rich product *transform* as a rule set for styling a set of artifacts into a presentation format on the fly (an example is given in Section 4.2). A transform defines the structure of a communication in terms of the reasoning patterns that should be contained within it. A transform service selects, from any given collection of artifacts, those components that can fill the roles required for the product. The styling of the template then determines what representational device should be used to encode the artifact. For instance, the transform can compose a textual description of the artifact and its relationships to others from the artifact's annotations and semantic markup. Alternatively, a visual thinker can apply a style guide that converts the same set of artifacts to an entity-relationship diagram. For machine-machine communication, the artifacts can be encoded into common interchange format.

Basing a knowledge transfer scheme around a rich product model enhances analytic *re*production as well. Reports composed of elemental units that can be decoupled from their presentation allow components of an analysis to be extracted and re-used. In addition, in circumstances where the veracity of a particular reasoning component comes into question, or beliefs evolve, it is possible to recreate the analysis in light of the updated information. The rich product model also reduces the problem of false corroboration among groups of analysts who are not in explicit collaboration. Such circumstances occur when, for instance, several investigators receive a report from the same source and incorporate this material into their own reports, which, when passed on to a single recipient, can appear to present corroborating support for a finding. In a rich product that enables drill-down into the analytic provenance of artifacts, it is possible to detect overlap in sources such that the likelihood of circular reasoning is reduced.

4. Implementing an environment for analytic composition

We have implemented a proof-of-concept analytic environment in which the abductive act of hypothesis construction and the hermeneutic act of play and interpretation takes place. This environment attempts to bridge the divide between information retrieval, information analysis, and information dissemination tools. By providing support for an end-to-end analytic workflow that helps situate resulting reports in the contexts of their construction, we aim to increase the interpretability of shared resources for teams of cooperating analysts (whether this collaboration is explicit or, more likely, implicit in their use of common resources and interest in common problems).

The architecture of this environment (Figure 2) broadly involves three components: a collection service, which maintains repositories of information resources and semantic markup describing their analytic relations according to individual users; a set of interfaces for user interaction; and a transform service that mediates between the repository and the user, customizing the display of analytic processes according to the consumer's role and preferences as well as the limitations of the interface (for instance, a PC client offers different representational affordances from a mobile device).

4.1. The analytic workspace

To capture the process of information manipulation that both hermeneutics and abduction hold so highly, the primary interface to our analysis environment is a visual, Web-based client for managing information collections and marshalling them into hypotheses and scenarios that can be published as rich products. These products, in turn, can be ingested into the same environment to expose their reasoning and allow the recipient to engage in the "horizon fusion" necessary for understanding.

The Web client is a workspace where information resources both tangible and abstract are given a visual representation as a "widget" that represents some facet of the resource. A widget might represent a single file, a collection of files, a more abstract entity such as an assumption, or a live information analysis tool. Analysts tag their resources according to a taxonomy of reasoning structures derived from [28]. This taxonomy accommodates artifact types ranging from elemental units such as evidence and assumptions, to reasoning aids such as arguments and models, to pattern structures such as temporal and spatial relations, and finally to knowledge structures such as hypotheses and scenarios. As information is gathered (from external applications or the user's desktop), the analyst can package lower-level artifacts together into sets representing higher-order collections, and can indicate that particular artifacts act as associations between others (for instance, a causal model might attribute an item of evidence to a spatial pattern).

Figure 3 depicts a section of a sample analysis in the domain of pandemic flu as it appears in the workspace. Each artifact in the analysis is represented as a rectangular widget that may contain other widgets nested within. Here, the "Drug resistance" artifact describes a hypothesis (indicated by the light bulb icon at lower right) about the possibility of the H5N1 influenza strain not responding to vaccines; this artifact is founded on three sub-elements (boxed icons within it) that could be opened to explore the hypothesis further. The drug resistance hypothesis is linked to a scenario about an H5N1 pandemic through an assumption about exposure to mutated virus strains (encoded in the green arrow linking the two nodes; although not depicted in this view, the assumption could be examined in detail). The pandemic scenario, in turn, defines the analyst's estimate of a temporal pattern of the onset of infection by geographic area.

Each visual representation of an artifact can be expanded to show a richer view in which a more complex analysis can be described. For instance, one expanded view of the H5N1 pandemic scenario provides a means of estimating the degree to which the components of the analyst's scenario support the likelihood of a pandemic (akin to the methodology of [29]), and the analyst's confidence in the estimate (Figure 4).

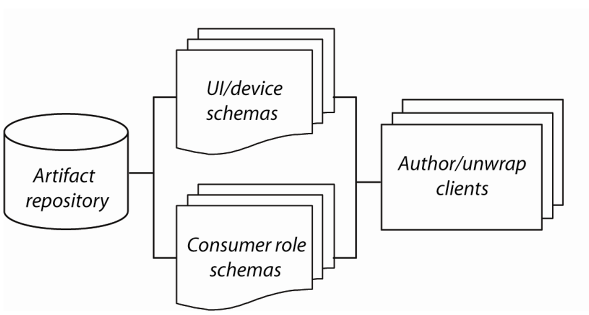

Figure 2. Simplified architecture of a system for analytic product creation and interpretation.

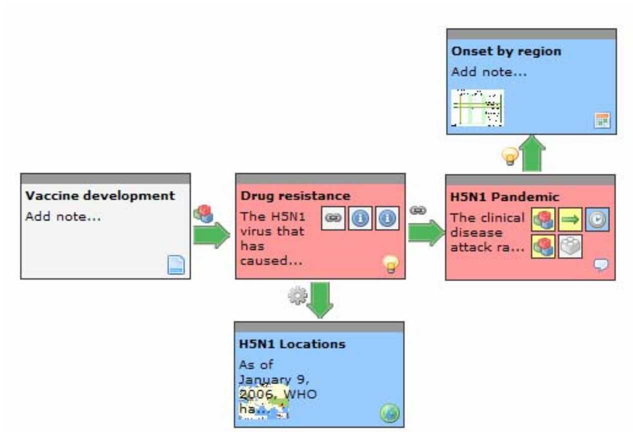

Figure 3. Sample analytic pathway for an avian flu analysis.

0-7695-2755-8/07 $20.00 © 2007 IEEE

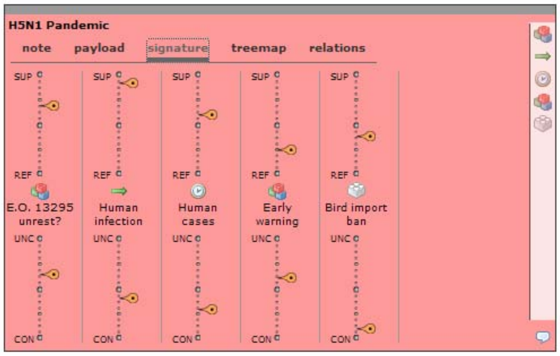

Figure 4. Describing support and confidence values for components of an analysis.

Analytic processes are captured in this environment in two ways. First, a versioning mechanism tracks the activity related to each artifact, so as items are connected, disconnected, and re-connected elsewhere, the history of this testing is maintained and auditable. Second, the widget environment is designed to reduce the computational and cognitive distance between information analysis environments and the reasoning process by allowing analytic tools to run within a widget. For example, one widget may represent a collection of thousands of documents returned in response to a Web query. Instead of simply serving as a container for these documents in the workspace and relegating their analysis to a separate environment, document clustering techniques are integrated into the workspace such that relationships between those documents can be explored. A further example of the integration of analysis tools into the environment is the "Onset by region" artifact in Figure 3, showing a thumbnail of a spreadsheet with time estimates. In the expanded view of this node the spreadsheet could be manipulated directly without leaving the analysis environment, therefore allowing the environment to capture *in situ* the history of these manipulations and reduce the loss of an analyst's situational awareness that would result from switching tools.

4.2. Building rich products

Analysts can move fluidly between constructing representations of their reasoning, as is shown in Figure 3, and creating reports summarizing their findings. Artifacts and groupings can be dragged onto reporting engines that attach those artifacts to components of a product based on roles they play. In a simple case, a text document transform contains rules for filling slots in a report according to a recipient's preferences, such as first outlining the source evidence and its provenance, then describing the analytic process, and finally present the resulting hypotheses and their likelihood. In a more complex example, a visual thinker can use a transform that structures hypotheses into a more abstract graphical depiction. Figure 5 shows one such transform, a treemap of the support and confidence encodings made in Figure 4, communicated to a user who wants the most significant contributors to the scenario to occupy more of the visual display. Each rectangular segment of this mapping can itself be depicted as a deeper treemap, revealing its lower level structure but always in the form the user requests, regardless of the form in which it was authored. The ability to communicate information from a shared knowledge base using flexible transforms responds to the semiotic nature of analysis; the form of a sign representing the same underlying evidence can differ between analysts.

4.3. Evaluating computational support for hermeneutics and abduction

Our prototype inquiry environment can be assessed by its degree of support for the principles of hermeneutics and abduction articulated in Section 2. Analysts can engage in **dialogue** with information resources by questioning them through interactive drill-down for more detail. Presenting information in consumer-customized views facilitates dialogue by giving the information at hand the ability to speak in a manner most appropriate for the recipient. An environment in which community resources are not just stored, but actively manipulated in the course of sensemaking permits **play**. **Prejudice** is addressed in both our rich products' ability to conform to the role or worldview of producer and recipient, but also in the explicit support for artifact types such as assumptions

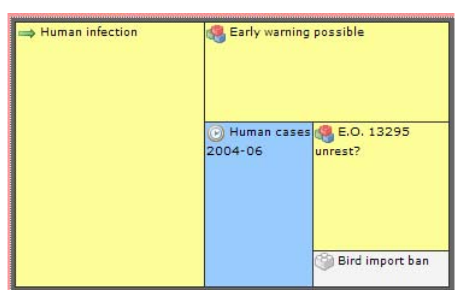

Figure 5. Transform of hypothesis structure from Figure 4 into a treemap, resulting from the application of a rich product style rule.

and arguments that help encode biases directly into a knowledge representation. A **hermeneutic circle** results from the cascade effect that interpretations made in this environment can have. Changing estimates of support or confidence can propagate to reports that had earlier been constructed; because reports are dynamic entities, new information or beliefs can be incorporated instantly, and the recipient can access the versioning history to see how the interpretations have changed. Consequently, conditions of **breakdown** – where evidence fails to support hypothesis – can be revealed through examination of the analytic pathways encoded in a product.

In Peircean abduction, truth is that to which a community of inquirers agrees in the long run. As an aid to abductive investigation, our method for capturing analyses as they occur helps communities reach common understanding by disclosing how information is used. We can detect a case where a given resource is acting as a boundary object by playing a different role in different analyses, indicating a touchstone over which negotiation can occur. Abduction also privileges pragmatics, and our representation of analyses in terms of their pedigree creates opportunities for effective information re-use and analysis re-enactment. Information is not simply made available, but made useful.

5. Conclusions

In information analysis fields where knowledge construction is largely an act of hypothesis generation, knowledge management tools can play an important role in communicating the reasoning processes necessary to evaluate those hypotheses. We present a knowledge sharing approach informed by hermeneutics and abduction, theories of inquiry that emphasize the interpretive and evolving nature of understanding. This theoretical foundation provides a set of design criteria for a prototype knowledge sharing tool that helps capture analytic processes *in situ* and that supports the construction of reports encoding the knowledge that results.

A barrier to promoting wholly new models for information flow is that entrenched tools and procedures, such as the writing of reports, still need to be accommodated. The *rich product* model offers a way for information producers and information consumers to reach shared understanding. The producer can outline the structure of an analysis in a Web based visual reasoning environment, then wrap it into a format suitable for dissemination. The consumer can apply transformations to that report to alter its presentation or level of abstraction, accessing more detail where needed without intervention from the consumer because the analysis pedigree was captured prior to report generation.

Ultimately, for such techniques to be widely adopted it is desirable to devise more implicit ways of capturing analytic context, such as closer integration with the array of analytic tools in use. There is also difficulty in encouraging the introspection necessary to expose the assumptions and interpretations inherent in an analysis. Nonetheless, we hope that the advantages gained in terms of knowledge auditability and reuse encourage analytic communities to work toward tools that better support the capture and sharing of interpretations.

6. References

[1] G. Fischer and J. Ostwald, "Knowledge management: Problems, promises, realities, and challenges," *IEEE Intelligent Systems*, vol. 16, pp. 60-72, 2001.

[2] E. Arias and G. Fischer, "Boundary objects: Their role in articulating the task at hand and making information relevant to it," Proceedings of the International ICSC Symposium on Interactive and Collaborative Computing (ICC2000), Wetaskiwin, Canada, 2000.

[3] C. Peirce, "Fixation of belief," *Popular Science Monthly*, vol. 12, pp. 1-15, 1877.

[4] R. Heuer, *Psychology of Intelligence Analysis*: Center for the Study of Intelligence, Central Intelligence Agency, 1999.

[5] D. Hall, D. Paradice, and J. Courtney, "Building a theoretical foundation for a learning-oriented knowledge management system," *Journal of Information Technology Theory and Application*, vol. 5, pp. 63-89, 2003.

[6] J. Grudin, "Desituating action: Digital representation of context," *Human-Computer Interaction*, vol. 16, pp. 269-286, 2001.

[7] S. Star, "The structure of ill-structured solutions: Heterogenous problem-solving, boundary objects, and distributed artificial intelligence," in *Distributed Artificial Intelligence*, M. Huhns and L. Gasser, Eds. Menlo Park, CA: Morgan Kaufmann, 1989, pp. 37-54.

[8] P. H. Gray and D. B. Meister, "Knowledge sourcing effectiveness," *Management Science*, vol. 50, pp. 821-834, 2004.

[9] R. Johnston, "Developing a taxonomy of intelligence analysis variables," *Studies in Intelligence*, vol. 47, 2003.

[10] H.-G. Gadamer, *Truth and Method*. New York: Continuum, 1975.

[11] F. Fonseca and J. Martin, "Toward an alternative notion of information system ontologies: Information engineering as

a hermeneutic enterprise," *Journal of the American Society for Information Science and Technology*, vol. 56, pp. 46-57, 2005.

[12] E. S. Abou-Zeid, "What can ontologists learn from knowledge management?" *Journal of Computer Information Systems*, vol. 43, pp. 109-117, 2003.

[13] R. Boland, R. Tenkasi, and D. Te'eni, "Designing information technology to support distributed cognition," *Organization Science*, vol. 5, pp. 456-475, 1994.

[14] R. Bernstein, *Beyond Objectivism and Relativism: Science, Hermeneutics, and Praxis*. Philadelphia: University of Pennsylvania Press, 1983.

[15] C. Peirce, "What pragmatism is," *The Monist*, vol. 15, pp. 161-181, 1905.

[16] U. Schultze and R. J. Boland, "Knowledge management technology and the reproduction of knowledge work practices," *Journal of Strategic Information Systems*, vol. 9, pp. 193-212, 2000.

[17] F. Nake and S. Grabowski, "Human-computer interaction viewed as pseudo-communication," *Knowledge-Based Systems*, vol. 14, pp. 441-447, 2001.

[18] S. Dustdar, "Caramba - A process-aware collaboration system supporting ad hoc and collaborative processes in virtual teams," *Distributed and Parallel Databases*, vol. 15, pp. 45-66, 2004.

[19] L. Magnani, *Abduction, Reason, and Science: Processes of Discovery and Explanation*. New York: Kluwer, 2001.

[20] P. Thagard, *Conceptual Revolutions*. Princeton, NJ: Princeton University Press, 1992.

[21] B. Shneiderman, "Codex, memex, genex: The pursuit of transformational technologies," *International Journal Of Human-Computer Interaction*, vol. 10, pp. 87-106, 1998.

[22] A. de Moor, "Patterns for the Pragmatic Web," in *Conceptual Structures: Common Semantics for Sharing Knowledge, Proceedings*, vol. 3596, *Lecture Notes In Artificial Intelligence*, 2005, pp. 1-18.

[23] R. Carvalho, J. Williams, I. Sturken, R. Keller, and T. Panontin, "InvestigationOrganizer: The development and testing of a Web-based tool to support mishap investigations," Proceedings of 2005 IEEE Aerospace Conference, 2005.

[24] Y. Gil and V. Ratnakar, "An interactive tool for capturing information analysis and decision-making," Knowledge Engineering and Knowledge Management, Proceedings of the thirteenth international conference, Siguenza, Spain, 2002.

[25] J. Novak, M. Wurst, M. Fleischmann, and W. Strauss, "Discovering, visualizing and sharing knowledge through personalized knowledge maps," Proceedings of the AAAI Spring Symposium on Agent-Mediated Knowledge Management, Lecture Notes in Artificial Intelligence 2926, 2003.

[26] K. Wideroos and S. Pekkola, "Presenting the past: A framework for facilitating the externalization and articulation of user activities in a desktop environment," Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06) Track 7, 2006.

[27] P. Cowley, L. Nowell, and J. Scholtz, "Glass Box: An instrumented infrastructure for supporting human interaction with information," Proceedings of the 38th Annual Hawaii International Conference on System Sciences, 2005.

[28] J. Thomas and K. Cook, "Illuminating the Path: The Research and Development Agenda for Visual Analytics." Los Alamitos, CA: IEEE Press, 2005, pp. 200.

[29] A. Turner, E. Hetzler, B. Cheney, L. Williams, and S. Zabriskie, "Method and system for prospective analysis of alternative futures," Proceedings of the 2005 International Conference on Intelligence Analysis, McLean, VA, 2005.