14. Ethnographic Techniques for Information System Design

15 October 2008

Bob Glushko

Plan for Today's Lecture

Document Anthropology and Archeology

Finding, Organizing, and Understanding Documents

"Staple Yourself to an Order"

SITREP "Document Ethnography" (Elisa Oreglia)

Reminder: The "Ethnography Continuum"

(Today's lecture is the contrast/complement to lecture #9 on 29 September)

Design contexts range from "experience-intensive" to "information-intensive"

On this continuum "documents" and other other information sources go from being incidental or occasional to being ubiquitous and intrinsic to the goals and activities of the stakeholders and actors

Put another way, on the "experience-intensive" send of this continuum the most important things to study are (human) participants, and on the "information-intensive" end we need to pay most attention to the documents

Put another way, we can learn mostly from "informants" or we can learn mostly from "information"

Document Anthropology and Archeology

Locating and understanding documents always requires a mixture of "anthropology" ("observing their use") and "archeology" ("digging into their history")

Document designs are often the most enduring aspects of business processes, lasting for years or decades, far longer than the tenures of the specific people who produce and use documents

Document implementation or management technology often changes, but the logical model of a document is often preserved (even if it should have changed to improve the efficiency or effectiveness of a process)

How Information-Intensive is the Business?

Some business models are almost entirely about information: insurance, finance, accounting, education...

Others have a "tangible" or "product" component but have information as a primary focus: transportation, logistics, healthcare...

Even business models that are highly "tangible" or "experiential" can have a significant emphasis on information because they need documents for planning and accounting as records of inputs and outputs

Example: Restaurants, which we've used as a canonical type of "experiential" service, have many documents types that are the tangible evidence of the experience or essential to operations

Contexts of Document Use

How we conduct "document anthropology and archeology" differs substantially among three different contexts:

- Within an organization or enterprise
- Between organizations or enterprises
- Within an information ecosystem

What a "Document" Is [1]

Every major advance in transportation, communications, manufacturing, financial technology or "governance" has required new types of documents

But the basic idea of a document has been surprisingly stable for a couple of millennia

- A document is a *self-contained package of related information*
- Documents organize business interactions around the information needed to carry out transactions
- Documents are the inputs and outputs of business processes

What a "Document" Is [2]

A critical step in understanding an information-intensive design context is creating a document inventory and classifying the "documents" you locate

You need to take a very broad view about what's a document because much of what's important to analyze isn't a traditional document

Much of what we analyze comes from people or systems or machines, and the lines between "requirements analysis," "document analysis," and "customer-centered design" aren't always sharp

You can think of what you learn from people as instances of "interview" or "observation" document types

Recognizing Documents [1]

Documents are packages used for exchanging information.

Packages may be:

- Paper form (printed/written, formal/informal)
- Digital form (computer files, structured/unstructured, databases)

Exchanges may be:

- Messages (emails, EDI)
- Online or Web
- Postal, Fax

Recognizing Documents [2]

Sets of data in databases, spreadsheets, accounting systems

Completed Printed forms

Job aids, "cheat sheets," sticky notes and other informal or unofficial documents

Lots of undocumented information in people's heads that you write down after talking to them

Document Types (vs Instances)

Blank Printed forms

Web forms

Database schemas

Documents that describe APIs or maybe even the code that implements them

Style sheets or templates in office applications

Finding the Essential Documents

Not all document types are equally important; is a document intrinsic to a business process or a derivative/aggregate of it?

If there are many instances of a particular type, we might have to be concerned about representiveness and selection biases

Don't assume that the names given to documents fit the people, tasks, and organizations in which we locate them

Regardless of its title, make sure a document is being used before you conclude it is important

Document Types that Help Locate Documents

Directories

Organizational Charts

Organizational Charters, Project Plans

Process / Compliance Manuals, Procedure Guides

Organizational Issues in Document Ethnography

Organizational charts can suggest business processes (and their associated documents), people who can tell us about them, and the context boundaries we can enforce

Be careful in assuming that job titles and formal organizational structure reflect what people actually do

The level at which you interact with an organization - the kinds of people you interact with - strongly shapes what you learn about it

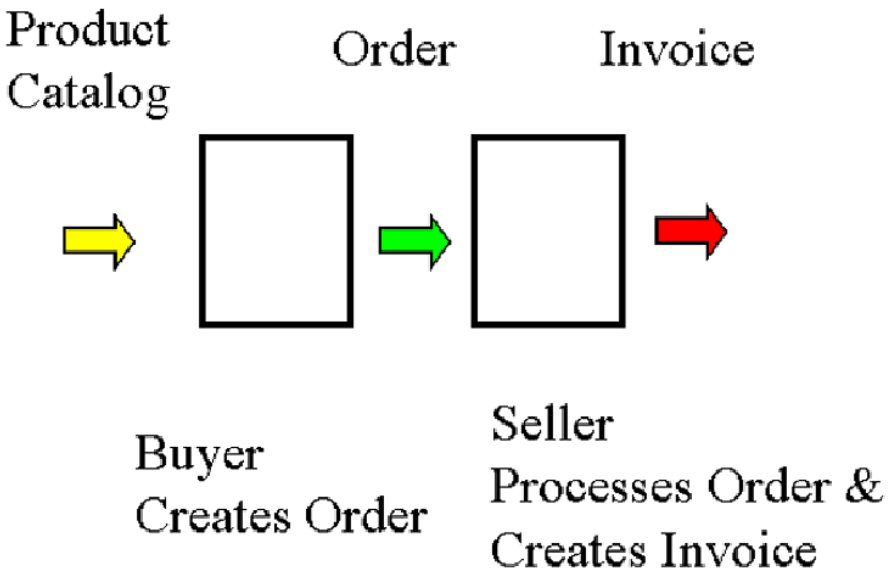
Analyzing Documents {and,vs,or} Analyzing Processes

To understand a business activity you need to study documents and the processes that produce and consume them

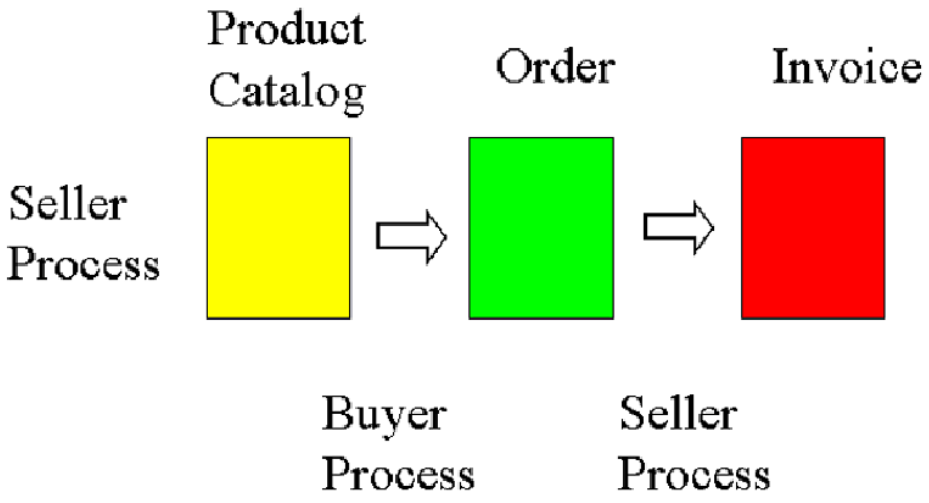
By understanding the information in the documents, we learn what kinds of processes (or services) are possible

By understanding the processes (or services), we learn what kinds of information are needed

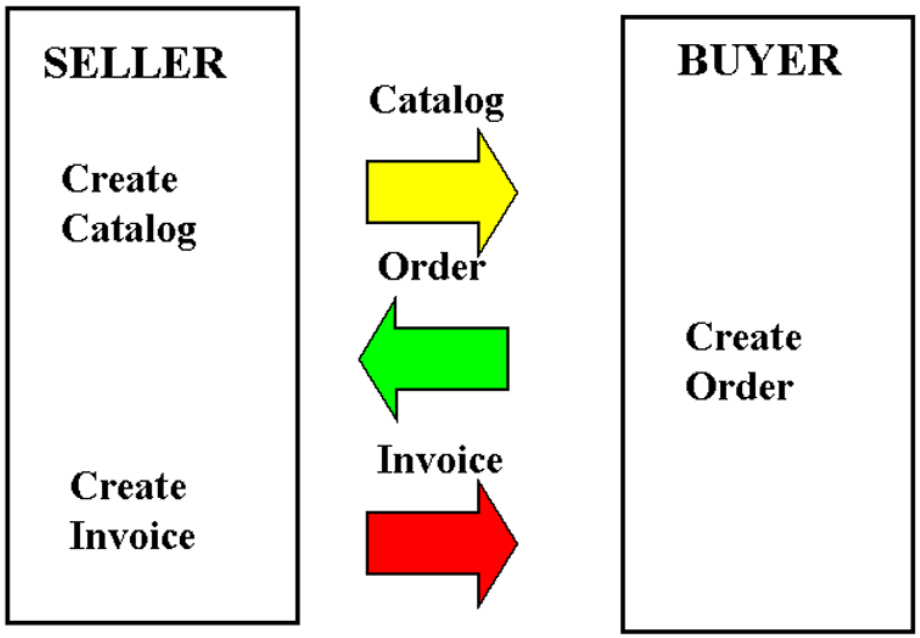
A Process-Centric Depiction



A Document-Centric Depiction



The Fundamental Difference is Visibility



Implications of a Document-Centric Analysis Approach

Documents are more tangible than processes, easier to analyze and communicate

The concreteness of document ethnography makes it more "bottom up" than business process analysis

This means that document analysts typically work at lower levels in organizations, observing and interviewing "ordinary workers" rather than high-level managers or executives

Iteration in Document Inventory

Identifying all the potentially relevant documents or information sources is inherently an iterative task

- Documents may refer or link to other documents
- Documents may refer to people, who can refer to other documents or people

Developing a causal model of the domain can help identify the intrinsic documents

- What high level business goals need documents as part of the activities to satisfy them?
- Where are the "headwaters" for the information -- what events or processes cause it to be created?
- A causal analysis can suggest other correlated information "streams" that merge with the primary source you've identified

Standard Questions About Documents You Receive or Use

What is its name? Any aliases?

From which person or process do you receive it?

Why do you receive or use it? What triggers the sender's action?

What are you expected to do with it?

Does the document contain sufficient information for this expected use? Do you need additional documents or information?

Does the document contain unnecessary information for this expected use?

What do you do with the document after your process is finished?

Standard Questions About Documents You Create or Send

What is its name? Any aliases?

To which person or process do you send it?

Why do you create or send it? What triggers your action?

What do you expect the receiving person or process to do with it?

Does the document need to conform to any standards or rules for content, structure, or presentation?

What does the receiving person or process do with the document after their process is finished?

Names for Document Types and Instances

Sometimes there are rules for names of document types

Sometimes there are rules for names of document instances

Sometimes the names of document types or instances aren't informative

Names are just one kind of metadata attached to document instances; there is lots more

Using Business Model and Process Patterns to Find Documents

Businesses / governments have long dealt with each other / their people through documents

We use concepts like "supply chains" and "distribution channels" as metaphors for the coordinated or choreographed flow of information and materials/products between businesses

These are complex patterns composed from simple binary document exchange patterns

Document Exchange: The Mother of All Patterns

Document exchange is the "mother of all patterns" for business models, business processes, and business information

- *Business model or organizational* patterns: marketplace, auction, supply chain, build to order, drop shipment, vendor managed inventory, etc.
- *Business process* patterns: procurement, payment, shipment, reconciliation, etc.
- *Business information* patterns: catalog, purchase order, invoice, etc. and the components they contain for party, time, location, measurement, etc.

Using Process Patterns to Find Documents: The Document Checklist

	Procurement Process Pattern				
	Collaborations	Request Service	Contract Formation	Payment	Distribution and Fulfilment
Function	Key Information Components				
Request details of item(s)	Product Identifier, Customer Identifier	X			
Provide ordering details	Product Description	X			
Request item(s)	Product Identifier, Customer Identifier, Quantity		X		
Acknowledge commitment to fulfill request	Order Identification, Status		X		
Request delivery of item(s)	Product Identifier, Customer Identifier, Destination, Quantity				X
Confirm sending of item(s)	Product Identifier, Customer Identifier, Destination, Quantity, Shipment Identifier, Date				X
Confirm delivery of item(s)	Product Identifier, Customer Identifier, Destination, Quantity, Shipment Identifier, Date				X
Enquire about status of item(s)	Shipment Identifier, Status	X			X
Advise on status of item(s)	Shipment Identifier, Status	X			X
Notify of required payment	Product Identifier, Customer Identifier, Quantity, Amount			X	
Request funds transfer	Invoice Identifier, Customer Identifier, Amount			X	
Acknowledge funds transfer	Invoice Identifier, Customer Identifier, Amount, Status			X	

"Staple Yourself to an Order"

Customer Participation	Steps in the Order Management Cycle
plans to buy	1. Order planning
gets sales pitch	2. Order generation
negotiates	3. Cost estimation and pricing
orders	4. Order receipt and entry
waits	5. Order selection and prioritization
waits	6. Scheduling
accepts delivery	7. Fulfillment
pays	8. Billing
negotiates	9. Returns and claims
complains	10. Postsales service

10 steps in the Order Management Cycle define a company's way of doing business

Because the OMC is an end-to-end system, every employee who affects an Order is the equivalent of a frontline worker

But most companies don't view the OMC as a system, and most executives have a simplistic or incorrect model

"Staple Yourself to an Order" -- Organizational Responsibilities

● leading role ⊙ supporting role — no role

Customer Participation	Steps in the Order Management Cycle	Sales	Marketing	Customer Service	Engineering	Purchasing	Finance	Operations	Logistics	Top Management Participation
plans to buy	1. Order planning	⊙	●	⊙	⊙	⊙	⊙	●	⊙	coordinates
gets sales pitch	2. Order generation	●	⊙	⊙	—	—	—	—	—	some
negotiates	3. Cost estimation and pricing	⊙	●	⊙	⊙	⊙	⊙	⊙	⊙	some
orders	4. Order receipt and entry	⊙	⊙	●	⊙	—	—	⊙	⊙	none
waits	5. Order selection and prioritization	⊙	●	⊙	⊙	—	⊙	⊙	⊙	some
waits	6. Scheduling	⊙	⊙	⊙	⊙	⊙	—	●	⊙	none
accepts delivery	7. Fulfillment	⊙	⊙	⊙	⊙	⊙	—	●	⊙	none
pays	8. Billing	⊙	⊙	⊙	—	—	●	—	⊙	none
negotiates	9. Returns and claims	●	—	⊙	⊙	—	⊙	⊙	⊙	some
complains	10. Postsales service	⊙	—	●	⊙	—	—	⊙	⊙	none

"Staple Yourself to an Order" -- Cracks in the OMC

What are the (horizontal) cracks described in the article and why do orders fall through them?

What are the (vertical) knowledge gaps and why do they exist?

What does it mean that "not all orders are created equal" and why does it matter?

Why is it important to chart the OMC?

Why is it important to look at OMC priorities from customer's perspective?

Analyzing the Document Inventory

You need to arrange the results of your inventory so you can think about it as a whole and in parts

What aspects of documents vary systematically across the categories in the inventory?

What other aspects of documents vary, but not systematically across the categories?

We need some concepts and vocabulary for answering these questions

Categories of Document Types

There are a few hundred common types of documents used in business transactions

But transactions are just one category of document types

Other categories with many distinct types include:

- Software and system documentation
- Procedures, policies, laws, and regulations
- Reference books, encyclopedias, dictionaries
- Catalogs

Organizations often use or produce multiple document types within the same category

Document "Collections" or "Chains" or "Clusters" or "Complements"

Some sets of document types in an inventory are related to each other

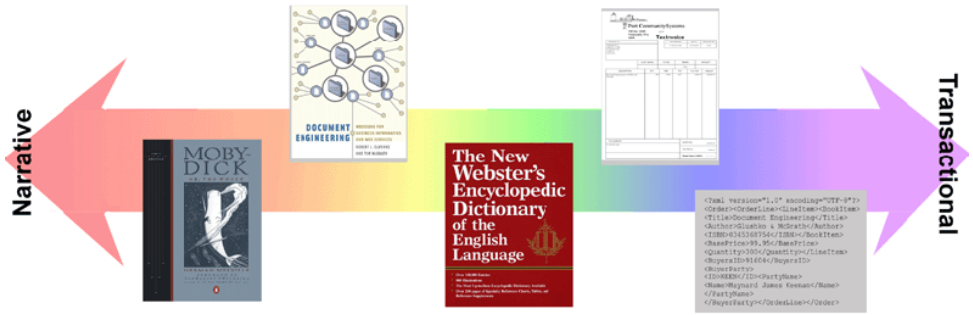
Some document types are themselves sets of documents of another type

Other document types fit together in a kind of sequential or process relationship where information flows from one to another in the normal way in which they are used or created

Transactional documents often come in pairs that must be correlated

Documents can have complementary (they are useful together) or uncomplementary (they are not useful together) relationships, and the relationships aren't necessarily symmetric

The Document Type Spectrum



Systematic Variation in Document Types Across the Spectrum

Instances more heterogeneous on narrative end

Types are "broader" and more descriptive, less prescriptive on narrative end

The set of content types within a document type is much greater on the transactional end because the leaves aren't "just text"

More need for "metadata" augmentation of documents on narrative end, because on transactional end what would be metadata is more likely to be explicitly contained in the content already

Presentational information more likely to be correlated with content and structure on narrative end

Readings for 20 October

Robert J. Glushko & Tim McGrath, Document Engineering, Chapter 9, "Analyzing business processes," 2005.