Local-Global: Reconciling Mismatched Ontologies in Development Information Systems

Jessica Seddon Wallack[1]
Ramesh Srinivasan[2]

Abstract

This paper extends pre-existing digital divide conceptualizations to further investigate the important issue of mismatches between the ontologies of state-created information systems and local communities' representation of their contexts. Comparability of data across time and place, as well as compatibility of data with state administrative needs come at a cost of information loss about the setting and individuals that policymakers are trying to impact. We argue that the reconciliation of community and state logics and framings is critical for effective engagement with communities as well as formulation and implementation of development policies. We suggest several paths toward overcoming mismatched ontologies: education and communications strategies to enable communities and states to translate across ontologies and fill in significant gaps; re-assignment of policy responsibilities to minimize information loss; and several mechanisms that would enable communities to be directly and productively engaged in developing shared ontologies.

1. Introduction

"Waterlogging" is a perennial complaint in cities in Karnataka, India. A few hours of rain can turn a dry street to a rushing torrent, while a burst pipe or a blocked drain can turn a pedestrian crossing into a treacherous lake. Local newspapers are full of photos, bus stops and public places full of discussion. Yet data on public grievances in 57 cities over nearly two years (2005-2007) contains only 203 records of "waterlogging" – instead there are recorded incidents of storm drains in need of desilting, storm drains in need of repair, leaking pipes, and choked underground drains. These are the various categories that citizens can choose from to report the puddles – which may very well look the same regardless of origin - via cities' Public Grievance and Redressal Systems. [73] No individual type of complaint accounts for more than 9% of complaints (most types are much less frequent), but the total of all complaints related to puddles – the actual thing that citizens can see and complain about – comes to 26%.

Bus stop conversations about waterlogging and reported complaints about pipes and drains refer to the same occurrence. Yet the community's understanding of the situation and the Karnataka government's record of the event label, catalogue, and interpret the event in divergent manners. The first characterization represents local knowledge, the picture that communities hold of the event, its significance, and its possible solutions. The person affected by waterlogging contextualizes it directly into existing activities and categories of their community. The second 'sees like a state' [52]: the problem is absorbed and recorded according to the logic of administrative efficiency and/or its policy ramifications. The designated categories list efficiently routes complaints to the concerned departments, but it may also fragment and obscure citizens' concerns.

This example, we believe, is emblematic of a key attribute of the digital divide: mismatched ontologies that impede communities' ability to impart and communicate information and states' ability to fully understand the territories they govern. Communities and states (we use the term generically to refer to subnational, national, and international governing institutions) represent the realities around them through distinct ontologies, or systems of categories and their interrelations by which groups order and manage information about the people, places, things, and events around them.

On the one hand, ontologies provide an essential shared infrastructure for individuals to function as a group. Communities may in fact form and self-define around shared ontologies, constructed and re-constructed fluidly [57] through shared social and cultural activities and the ever-changing lived experiences of their members. Ontologies work to create and enact worldviews within the social group and situate knowledge [63,64] within the organizational or community setting.

This is as true for states' ontologies as it is for community ontologies. State data systems are the infrastructure of administration: the structured, comparable information about the area under jurisdiction enables resources to be identified and taxed, needs to be assessed, and returns on public expenditure or investment to be monitored.

[1] *Director, Ctr for Development Finance, Institute for Financial Management and Research, Chennai, India & Research Scholar, Global Information Industry Center, UC San Diego, California.*
[2] *Assistant Professor, Department of Information Studies, UC Los Angeles.*

Policymakers seeking to direct resources to "the poor" and monitor the impact, for example, need to have figures that enable comparisons of poverty across their territory and over time even if the local perceptions of poverty vary. Current standards for "good policy," including evidence-based decisionmaking and monitorability of policies also implicitly rely on meta ontologies' ability to produce the kind of quantitative indicators and large-scale datasets required for "objective" evaluation of performance and research that meets current standards for evidence of effectiveness. See www.paris21.org for an example of an influential group of development agency's call for this type of information as a step toward improved policymaking.

Nevertheless, ontologies also act as objects that create and negotiate boundaries between groups [9,61]. The state 'meta ontology' sheds much of the local context in order to ensure tractable management for policy purposes including taxation, defense, provision of infrastructure and service, and economic management[76,1,52]. This tractability comes at the cost of information loss due to mismatched ontologies. While any group's ontology is unlikely to match that of every individual within the group, the extent of mismatch tends to increase with the scale of the group and the differences between the purpose of individual and group ontologies. Note that our use of ontology does not imply a reified nor exoticized model of 'pastness' or 'locality' that ignores flows of interaction that shape communities over time [2], but merely implies a distinction between groups' mental maps of their surroundings.

We argue that states cannot ensure the outcomes that they seek if their policies are targeting a skewed picture of communities' reality. Even the most benign development-seeking state may go awry if its actions are guided by its own ontology and indicators, while levels of and progress toward development are individual or at most community-specific concepts.

This paper characterizes an important and overlooked element of the digital divide: information loss due to mismatch between community ontologies and 'meta ontologies.' We argue that the digital divide may be usefully framed around a model of communication, of reconciliation of ontological frameworks that enables information flow between communities and states. The divide is not strictly digital since the issues with mismatched ontologies we elaborate on below would occur between any small and large scale ontology, but it has been exacerbated as ICTs designed around state meta-ontologies are increasingly deployed to impact development in diverse communities. This view

extends the digital divide literature focused on barriers between infrastructure and technology access [13, 16, 17, 18, 12, 28].

This discussion also extends research on literacy and contextual information uses [65, 56, 38, 7, 31] to offer an expanded argument for the importance of systems and institutional mechanisms that enable effective communication between differing ontologies. Mismatched ontologies contribute to: (a) ineffective delivery of information services to communities; (b) insufficient participation and interaction with local communities; and importantly, and (c) 'information loss' that affects states' abilities to effectively deliver goods, services, and development-supporting interventions. Much cited research has pointed to the issues within (a) and (b). This paper focuses on (c), to make the argument that the issue of mismatched ontologies is a lose-lose proposition for communities and states.

The gap between community and meta ontologies is not necessarily purposeful, but rather is often a symptom of the fundamental difficulty of incorporating local, contextualized knowledge into large scale, comparable-across-time-and-place datasets. We recognize that states must be able to compare and aggregate across communities so that resources can be allocated, scalable policies can be developed, and effectiveness of any interventions evaluated. Research on development demands data that are comparable across communities and time and specifically collected to clarify hypothesize categories and relationships. Yet at the same time this aggregation leads to information loss, not just in terms of overlooked entities but more importantly in overlooked or misjudged semantic relationships between these entities.

Finally, we also propose some ways to mitigate the information loss by identifying ways in which systems and technologies can collect and represent information meaningfully for communities and states alike, reconciling mismatched ontologies. These types of solutions necessitate equitable forms of access, however, and cannot resolve all aspects of ontology mismatch. Institutional and policy design may also need to be reconsidered to decentralize more decision making in light of the costs of information loss associated with larger-scale meta ontologies.

We devote the next section of this paper to underscore the relevance of ontologies, characterize information loss with further specificity, and outline some consequences of information loss in the creation of state meta ontologies. The third section proposes some mechanisms of reconciliation and alleviation of the mismatch. The final section

concludes with a discussion of a broader research agenda on these issues.

2. The power of ontologies, the problem of information loss

Ontologies represent reality, but this representation of information may in turn become the basis for actions that in turn shape reality. The actions may be shaped in response to ontologies and aimed at problems defined by ontologies, but their impacts depend on the actual reality rather than the representation. Any actor's effectiveness in achieving their goals thus depends on the quality and completeness of their ontology. Incomplete or inadequate state meta ontologies are especially problematic because of the power that states' actions have to affect communities.

The problem is especially pervasive for economic development policy, in which states' goals are (at least normatively) defined in terms of individuals' utility, or sense of wellbeing. Some of the most prominent formulations of "development" measure progress in terms of achievements that only make sense with reference to individuals' or communities' ontologies. Nobel Laureate Amartya Sen's concept of "development as freedom," [54] for example, argues that development consists of ensuring that people have the capability to do and be what they value. The notion of "human development" considers health and access to knowledge, both of which are best assessed by the individual, to be part of development along with income. (The concept has been implemented in practice [68] with simplistic measures of observable levels of health, educational attainment, and gross domestic product per capita, but that is another matter.) Economics has traditionally resolved the tension between the need to measure wellbeing in order to design and assess policies and the acknowledgement that wellbeing is a subjective concept dependant on an individual's worldview by assuming a representative 'rational' individual with a well-defined and known ontology, but the increasing presence of the "economics of happiness" [22] and "behavioral economics" [11] in mainstream journals and top academic venues suggests that the compromise is fraying. States' attempts to promote "development" are thus limited by the information loss between the community ontologies that define development and meta ontologies that guide their actions.

More concrete consequences of information loss can be seen in a variety of policy areas. Urban planning, for example, seeks to devise guidelines to ensure that communities have desirable places to live, work, and play [27]; but often ends up creating empty monstrosities such as Brasilia's Plaza of the Three Powers [48] or dysfunctional living spaces the destroy rather than support community. The official guidelines are often ignored by communities operating on their own principles of order, leading to inevitable conflicts between states and citizens who essentially seek the same outcomes [44, 72]. Many failed efforts to provide infrastructure and services can be traced to similar misunderstandings. Affordable drinking water projects have been rejected by communities because their pricing models fail to recognize that while communities are willing to pay for services such as water purification or water delivery, they see the water itself as a right [44]. Roads have been built and left unused because they failed to connect places of local import. Free schooling has failed to educate half of some areas' population because classroom design did not include separate toilets for girls [44]. Technologies that could ostensibly improve living conditions and economic opportunities have been rejected because they were inconvenient for some other community practice: many rural Indian households continue to cook over smoky indoor fires while free clean-burning stoves distributed by government programs go unused because they did not accommodate locally used cooking pots or could not be used to produce local specialties. Public IT kiosks sit idle even as the benefits of IT skills and access become clearer. [70]

The fate of the clean-burning cook stoves and the IT kiosks is hardly surprising to students of "ICT for Development." Significant research has revealed that an understanding of community ontologies, particularly in the context of information systems and technologies, is fundamental to the ability of new innovations to insert themselves sustainably within the lives of local communities [58,59,60].

There is thus a far-reaching and fundamental policymaking problem due to the inadequate meta ontologies that states often rely on. We now turn to be more specific about nature of the information loss between community and meta ontologies. First, what, specifically, is the "information loss" that leads to inadequate meta ontologies? Second, how does it arise and why is it perpetuated? Understanding a problem is the first step toward devising a solution.

We characterize information loss as a consequence of mismatches between components of community and meta ontologies. It thus has various interlinked dimensions: entities that are included in one ontology but excluded from another; categories that create different groupings of these entities for communities and for states; discrepancies between

the attributes that are recorded and attached to entities or categories; divergence in the recognized potential changes in categories or attributes; differences among the relations used to link entities or categories in each ontology; and variations (and perhaps contradictions) in the restrictions and rules that define allowable inputs and inferences about the world represented by the ontology. (We are formalizing this notion of information loss and its implications for the design of technology to cross the digital divide in concurrent work as it is beyond the scope of this paper.)

Most of these dimensions of information loss could, in principle, be resolved at some cost. Any object, attribute, category, or relation included within a local ontology could be included in a meta ontology. Governments often collect data on food consumption to measure poverty; they could (and sometimes do) just as well collect data on clothing, appliance ownership, extent of private living space, and other ways in which communities understand poverty. Censuses often group individuals as employed or unemployed, there is no reason that they could not also include categories for happily employed and unhappily employed as well. Governments often base the relations in their ontologies on those derived by the scientific method; there is no reason that they could not also incorporate folkloric relations that guide community perceptions. The only fundamental obstacles to eliminating information loss between ontologies stem from potential conflicts between rules and restrictions in different ontologies and disagreement over the nature of events that determine transitions in attributes or objects.

Why would states purposely create, rely on, and perpetuate ontologies that are inadequate for characterizing the terrain they administer and the goals they seek to achieve? The political economy literature offers various explanations for the nature of states' logical frameworks and resulting policies. [21] Some authors focus on ideology or scientific/academic reasoning (in development, for example, the reigning schools of thought about how growth happens) as the main influence ([14], for example). Others characterize states' logic as an emergent property of the political interactions within the institutional setting, as a function of whether countries are democracies or autocracies, presidential or parliamentary democracies, federal or unitary states, for example [50, 19, 64]. The third main school of thought sees states and their meta ontologies as reflecting some aggregate summary of citizens' views and preferences as expressed at the ballot box or through lobbying and "insider dealing."

(See, for example [20] and the vast literature on Public Choice that this work has inspired on voting, or [34] on lobbyist networks.) We note that all but the last involve forces that have no necessary linkage to communities' mental maps of their surroundings, and the link between mechanisms for aggregation of citizens' preferences and community ontologies is weak at best. Voting, rioting, and lobbying provide limited opportunities for conveying information. The limitations on how much information can feasibly be aggregated through group decision making to determine social choices have been formally and extensively explored in social choice theory [3, 5].

We argue that states' meta ontologies are also at least in part a function of the logistics of their efforts to collect data. Every explicit effort to document a territory, such as a census, is based on particular claims of how a community is to be measured, how the boundaries of a community are to be determined, what counts as an activity, and how these collected data points are to be connected and compared [69, 62]. These claims may be motivated by politics [49, 32] or determined by administrative and technological feasibility of data collection, storage, and retrieval. Scott also shows that administrative exigencies related to the division of the responsibilities among and within bureaucratic entities have been influential in determining data collection norms [52]. (He also shows that administrative norms have also driven policies to forcibly alter reality to make it more 'countable': assigning surnames, for example, or removing bio-diversity enhancing and forest-sustaining underbrush in order to better count trees useful for lumber and shipbuilding.) Communities' documentation efforts may very well be driven by the same factors, but on a local rather than cross-community scale so there is little reason to expect intersection or reconciliation between the statistical representations.

Once in place, state meta-ontologies tend to be self-reinforcing or at least unlikely to converge with community ontologies. States are one of the largest sources of credible data on a variety of economic, political, and social processes and outcomes. Their collection and presentation of information determines the ways in which empirical relationships can be discovered, verified, proved, used for public debate.

Researchers and citizens are less able to challenge the meta ontology when they cannot model and demonstrate the validity of local restrictions, practices, events, and entities according to community ontologies. States' dominant position in the supply of data will likely change over time as the costs of collection, compilation, storage, and dissemination of community-produced data continue

to decline. But even then, states' authority may privilege conclusions drawn from "official" versus non-state produced data.

Second, communities may be less willing to interact with the state to provide data if they feel that the way that this data is aggregated, presented, and used has no connection to the questions that they would like to ask. As one professional demographer reports in personal communication with the author "In surveys that I have handled … the resistance to provide data was huge. Most respondents asked first, 'What do I gain by all these?' They did not believe research could be productive for them" [55]. Non-response is a widespread problem for censuses and other surveys [25].

Third, communities are less able to engage with states to influence ontologies or information contained in them when they do not find the starting points intelligible. Notable research in Science and Technology Studies [15, 72] has demonstrated that without an approach that directly asks stakeholders to collaboratively work to reflectively create data models, often information systems end up disorienting the groups at hand. Summary statistics from the public grievance and redressal system mentioned in the introduction, for example, shows that citizens in some cities in Karnataka are very bothered by "biomedical waste." A closer look at the (harder to aggregate and summarize) text of the grievances shows that many people had selected the "biomedical waste" category to actually refer to dead pigs, dogs, and other animals [73].

Unintelligible or incomplete ontologies also weaken the basis for collective action that may be required to draw attention to citizens' voices. This is particularly the case for problems that tend to be observed by individuals rather than groups. Citizens are unlikely to band together and lobby for the state to recognize and address sexual harassment, for example, if each person affected has no idea that his or her own experience is not unique because there are no data on prevalence of cases.

This discussion, therefore, leaves us with some important questions concerning the extent to which sociotechnical systems and institutions optimize between local sustainability and cross-community scaleability? Or, is there a way in which community activities can be viewed and monitored from the birds-eye by the states while still preserving the local nuances?

3. Bridging local and global

We close this paper by reflecting on several ways in which information loss can be reduced. This section offers three possibilities, each the basis of ongoing research:

1) Developing collaborative and inclusive ontologies. Systems that engage communities to dynamically model their relationship to the information they are provided, around local categories, and fluid relationships between these, have been used sustainably and innovatively in cross-cultural local community contexts [e.g; 37, 47, 38]

2) Educational and communications strategies to empower and encourage communities to interpret state-provided data and provide feedback including alternative representations of their contextual knowledge into the ontology underlying official "development data."

3) Rethinking policymaking and institutional design along two dimensions to a) include more alternative channels for information flow to fill in gaps due to mismatched ontologies, and b) decentralize more decision making to subnational and local governments that may operate on 'less-meta' ontologies that are better matched to community ontologies.

3.1. Collaborative and inclusive ontologies

Socio-technical systems can enable communities to occupy a variety of roles, falling on a spectrum from passive consumer to co-designer and co-creator, to collaborate in creating ontologies. Moreover, in relevant cases, communities can be involved in actually developing, in decentralized manners, their own fluid ontologies, by being exposed to the raw information behind the system and perhaps through semi-automated models, such as rating similarities between entities and evaluating semi-automated models, such as hierarchical clustering and multidimensional scaling [56]. Fluid ontologies, in their most localized form, involve content creators and multiple stakeholders in the direct crafting of categories and data representations so as to ensure that the information they interact so as to ensure that information is presented, retrieved, preserved, and shared around relevant categorical and relational attributes that are sensible to the community in particular [57]. These stakeholders are presented with adaptive and dynamic possibilites to continuously re-craft such ontologies as their reflection of community practices shifts over time. For example, a local

village community can access an information kiosk of state services, rate the relative importance of each of these within their own community, re-categorize (via tags) services provided to follow local vernaculars, draw semantic connections between different services by attributing a connection with a particular term, and blog by commenting or submitting a video in response to something they experience via the system. We believe that these creative and local uses of tagging, rating, and other types of Web 2.0 technologies present powerful opportunities to adapt and edit a meta ontology and reconcile it with local practices. Validating this hypothesis is the subject of current research. We believe that this approach can be inherently collaborative, as it asks for reaction and re-crafting of an ontology based around a type of conversation that occurs between the policymaker/system and local community members. These social web 2.0 type uses can ask community members to submit their own categories, comment on information they are presented, and actively work with policymakers, even if not in real time, to develop and identify adequate ontologies that reconcile these different perspectives.

Designing more inclusive ontologies that can record more objects and attributes, tag them as belonging to any category any community wished to include, link them to other objects, attributes, and categories in any relationship any community cared to define, and store all of this information for ready retrieval, is an important challenge. We recognize that for this to occur, new experiments and prototypes need to be created that engage important work on collaborative and inclusive ontologies conducted by such scholars of information science, as Jens-Erik Mai, Birger Hjorland, and others. Mai's work on pragmatic approaches toward determining 'likeness' in knowledge organization argues that classification is fundamentally interpretive [40]. Instead, presenting different 'documents' into the local social and cultural context could enable them to be reflectively classified based around the community's own pragmatics [39]. Likewise, Hjorland, across an extensive body of work, has argued that "different paradigms or conceptualizations in discourses should be made visible in information repositories, and that semantic tools should not just support the navigation between topics but also views and interests. This approach denounces an a priori 'universalist' approach toward defining how knowledge shoul be organized, and instead decentralizes the project to allow different discourses to be shared based around multiple ontological viewpoints. Mai and Hjorland's arguments are well-taken, and remind us that the question of localization in such systems is a social

and linguistic problem, rather than solely technical [4], one that must necessarily involve local communities to articulate their own vocabularies and contextualize information in meaningful manners.

We are developing a field study to investigate possibilities for building a larger scale information sharing network with a foundation of collaborative and inclusive ontologies centered on community economic development initiatives.

3.2. Education and Communications Strategies

Meta ontologies that are more transparent and communicable through relevant forms of communication and outreach to communities may also reduce the dynamics of divergence described in Section 2. NGOs and in some cases states themselves have filled the role of translating meta ontologies into language more familiar to communities through services such as literally translating from one language to another or helping citizens to represent their circumstances on forms for loans, benefits, taxes, etc. While these are important "bridging" services, they are not a full solution as the power to build a bridge lies only with those who are conversant in the meta ontology.

Technology and education could combine to create a more inclusive solution in which communities can reorder state-collected information according to their local ontologies. Database design and interface determine in part how communities can interact with and reorganize raw data to retrieve the objects and attributes they felt were important, choose the categories to sort and compare, and select the relationships they wished to use for ordering the observations. "Numeracy" empowers citizens to take advantage of these flexible modes of dissemination. Citizens who are able understand statistics, manipulate them, interpret and re-represent them are better able to adjust data for their own needs as well as communicate local reality to the state authorities and point out gaps that the state data do not fill.

This approach to resolving information loss has the potential to create its own form of digital divide, however, as digital representation of information and interactivity with the underlying data warehouse are necessary. Users cannot manipulate data presented on a printed page, for example, in the same way that they can work with an excel spreadsheet or, better yet, a database with opportunities to construct queries to retrieve data. Users cannot readily visualize information for quick comparisons and checks without having GIS or some web-based applications. It is far easier to assess the quality of city services if

one can look at the location of bus shelters on a city map of bus routes than if one has to compare a list of written addresses with a list of bus routes defined by their origin and destination for example.

3.3. Institutional design

Rethinking institutional design to weigh the costs of information loss relative to the benefits of larger-scale decisionmaking systems is also essential. Technology access is not universal and relying on technology can create its own second-order digital divide between those who have the technology means to cross the communication divide we have identified here, and those who don't. Technology fixes can be costly and difficult to implement over legacy systems of information management and organizational resistance. Finally, as discussed above, technology fixes can only solve some dimensions of the information loss.

The inclusive ontologies discussed in section 3.1 would only be possible with changes in public sector information management to improve its ability to qualitative information, multiple characterizations of the same event, and evolving categorizations, relationships, and other aspects of ontologies. The business literature and basic textbooks have long emphasized the importance of organizations' ability to manage information, particularly information generated by actual and potential customers [53, 66], but this has been less emphasized in public sector.

The political economy of such reorganizations is likely to be challenging, never mind the sheer technical feat of incorporating legacy systems into new means of information sharing and the training required to update civil servants' capacity to use the new systems. There are no quick fixes or easy suggestions for how this can happen.

"Fire-alarm" [43] mechanisms, provisions for citizen feedback to sound the alarm about poor service, corruption, problems with policy implementation, etc, may be useful in the meantime. Structures such as public grievance systems or ombudsmen and practices such as comment periods for regulation, politicians' "constituency service," or "town hall" fora allow citizen feedback. It can be costly for states to absorb this less structured dialogue, but "fire alarms" have the advantage of being self-targeting. The information fills the most problematic gaps because these are the ones that the community deems most urgent and bothers to take extra effort to report.

Fire alarms do not entirely avoid political economy problems, however, since they do change individual policymakers' and bureaucrats' work contexts. The state may be interested in knowing where citizens are displeased; the individual who displeases the citizen may want this information to remain hidden.

Policy decentralization can also mitigate information loss by empowering decision makers with 'less meta' ontologies to respond to community needs. Local governments are of course embedded in the overall fiscal federal framework and therefore their ontologies are likely to be influenced by revenue transfers and mandates from the central government as well as inputs from communities, but they are still likely to be more in touch with community perspectives due to sheer proximity [46]. This proximity and increased possibility of information transfer between communities and states also has a dark side, however, in that it can allow locally powerful groups to unduly influence policies in their favor rather than the community interest. [24]

Decentralization, like improving public sector information management, is easier said than done. Policymakers and bureaucrats at one level of government are often protective of their turf. Looking across countries, there has been significant nominal devolution of responsibilities to local government, but often without the funds and human resources to support the new tasks. Local autonomy over expenditure decisions – the heart of actual policymaking - is often severely constrained under the premise that these lower levels of government need time to build up capacity before they can be trusted to operate independently. Centralization of control over revenues is a notable overall trend in federations around the world [78].

4. Conclusions

This paper has identified a new form of divide that has been exacerbated by the increasing reliance on and use of information technology for growing markets and states. We have argued that information loss due to mismatch between community ontologies and the meta ontologies that states act upon has serious consequences for the efficacy of state policies, especially those aimed at accelerating development.

We are not the first to point out the defects of centralized planning and the hubris of states. Friedrick Hayek won the Nobel Prize in Economics in 1974 for pointing out the value of local, practical knowledge, some of the the pathologies of states' approach to 'scientific' planning, and the need for any functional system to maintain a space for people to use their local, practical knowledge [29,30] .

The paper does offer a new perspective on this long-recognized problem, however, by re-conceptualizing information loss as a kind of communication failure that can be increasingly mitigated through technology as well as addressed through institutional redesign to reallocate decisionmaking authority and establish alternative channels for communication. This conceptualization of information loss from ontology mismatch offers a new criteria for evaluating institutional design.

This paper represents a first step in a larger conceptual and empirical research agenda that we are working toward. First, this paper uses logic and anecdotes to identify the information loss due to mismatched ontologies as a potentially significant challenge for development policymaking and a priority for further research. We have not yet rigorously established the extent of information loss nor the financial or human costs of mismatched ontologies. Our first step will be a meta-analysis of existing studies: measurable discrepancies between community understandings and the pictures portrayed in policy-relevant official statistics have been studied for various data series and settings [47, 6], but the collective implications have not been studied.

We are also looking at possibilities for field research in India that would document discrepancies between official portrayals of a development context and citizens' understanding of the context. In one, we are working with an ongoing initiative to disseminate census data about Indian village to add a feature requesting and collecting peoples' updates and comments on how those data portray their village.

Second, rigorously establishing the extent of information loss and assessing the efficacy of proposed solutions requires a measurement concept. We are currently formalizing the concept of information loss that we have introduced in this note. We conceptualize information loss as occurring in dimensions corresponding to mismatch between various components of local and meta ontologies, but we have yet to develop a sensible way to weight each of these dimensions in any summary of information loss. The effort to weigh the relative importance of the various dimensions requires additional empirical work.

The study of whether and how effectively collaborative and inclusive ontologies can be developed is also the subject of ongoing research. Hjorland and Mai's arguments have been mentioned briefly in this paper, and inspire a few important projects that we are currently implementing, focused on studying how multiple communities can develop tags and vocabularies that reconcile differing perspectives and worldviews, and implementing

these via mixed 'fluid' and semi-automated approaches [56, 57, 58]. Finally, practical progress in reducing information loss ultimately relies on our ability to demonstrate the impact of mismatched ontologies on success or failure of specific policy initiatives. As discussed above, states do not always have strong incentives to adopt the kind of solutions that we suggest, especially when the "problem" is murky and current notions of "good policy" rely on the comparability and monitorability than meta ontologies can deliver. We have discussed cases of information loss and are improving our ability to quantify it, but the next step of linking information loss to economic or human loss will be more challenging. It will be difficult to conclusively disentangle mistaken assumptions or understandings of development challenges from other sources of policy failure, but we are optimistic that some carefully selected cases developed as theoretically-grounded analytic narratives [8] will help us to tease out the effects of poor information from the effects of institutions, ideology, and other features of the decision making environment. We will be collecting focus group information on public sector decision making practices in various levels of government in India in order to design these studies.

5. References

[1] Alonso, W. and P. Starr, eds. The Politics of Numbers. Russell Sage Foundation, New York, 1987.

[2] Appadurai, Arjun. "Disjuncture and Difference in the Global Cultural Economy." Modernity at Large: Cultural Dimensions of Globalization. Minneapolis: University of Minnesota Press, 1996. 27-47.

[3] Arrow, K. Social Choice and Individual Values. John Wiley & Sons, New York. 1951

[4] Atzeni, P. "Schema and Data Translation," m/s, 2005.

[5] Austen-Smith, D., and J. Banks. Positive Political Theory I. Michigan University Press, Ann Arbor, MI, 1998.

[6] Banerjee, A., Duflo, E., Chattopadhyay, R., and J. Shapiro. "Targetting Efficiency: How Well Can We Identify the Poor?" Institute for Financial Management and Research Centre for Micro Finance Working Paper Series, no. 21, December 2007.

[7] Barzilay-Nahon, K. 2006. Gaps and Bits: Conceptualizing Measurements for the Digital Divide. The Information Society 22: 269-277.

[8] Bates, R.H.,Greif, A., and M. Levi. Analytic Narratives. Princeton University Press, Princeton, 1998

[9] Bowker, Geoffrey C., and Susan Leigh Star. "How Things (Actor-Net)Work: Classification, Magic and the Ubiquity of Standards." Philosophia 25.3-4 (1996): 195.

[10] ---. Sorting Things Out: Classification and its Consequences. Cambridge, MA: MIT Press, 1999.

[11] Camerer, C., Loewenstein, G., and M. Rabin. Advances in Behavioral Economics. Princeton University Press: Princeton, NJ, 2004.

[12] Chambers, Robert. 1994. Participatory Rural Appraisal (PRA): Challenges, Potentials and Paradigms. World Development 22(10): 1437-1454.

[13] Chen W., and B. Wellman. 2003. Charting and Bridgind Digital Divides: Comparing Socio-economic, Gender, Life Stage and Rural-Urban Internet Access and Use in Eight Countries. AMD Global Consumer Advisory Board (GSAB)

[14] Chibber, V. Locked in Place: State Building and Late Industrialization in India. Princeton University Press, Princeton, NJ, 2003.

[15] Christie, Michael, and Helen Verran. "Using/Designing Digital Technologies of Representation in Aboriginal Australian Knowledge Practices." Human Technology 3.2 (2007): 214-27.

[16] Compaine B.M. 2001. The digital divide : facing a crisis or creating a myth? Cambridge, Mass.: MIT Press

[17] DiMaggio P., E. Hargittai. 2001. From the 'Digital Divide' to Digital Inequality: Studying Internet Use as Penetration Increases. Center for Arts and Cultural Policy Studies, Princeton University Working Paper #15

[18] DiMaggio P., Hargittai, Eszter, Celeste, Coral and Shafer, Steven. 2004. Digital Inequality: From Unequal Access to Differentiated Use. In K Neckerman (Ed.), Social Inequality: 355-400. Russell Sage Foundation: New York

[19] Dixit, A. The Making of Economic Policy: A Transactions Cost Perspective. MIT Press, Cambridge, MA, 1996.

[20] Downs, A. An Economic Theory of Democracy. Harper Brothers: New York, 1957.

[21] Evans, P., Rueschmeyer, D., and T. Skocpol, eds. Bringing the State Back In. Cambridge University Press, New York, 1985.

[22] Graham, C. "The Economics of Happiness: Insights on Globalization from a Novel Approach," World Economics, v.6, No.3, July-September 2005.

[23] __ "Insights on Development from the Economics of Happiness," World Bank Research Observer, 20(2), 2005:201-231.

[24] Grindle, M. Going Local: Decentralization, Democratization, and the Promise of Good Governance. Princeton University Press, Princeton, 2007.

[25] Groves, R., Dillman, D., Eltidge, J., and R. Little. Survey Nonresponse. John Wiley and Sons, New York, 2001.

[26] Hargittai E. 1999. Weaving the Western Web: Explaining Differences in Internet Connectivity Among OECD Countries. Telecommunications Policy 23(10/11): 701-718

[27] Hall, Peter. Cities of Tomorrow. Blackwell Publishing, Boston, 2002.

[28] Hargittai E. 2002. Second-Level Digital Divide: Differences in People's Online Skills.First Monday 7(4). http://firstmonday.org/issues/issue7_4/hargittai/

[29] Hayek, F. "The Use of Knowledge in Society," American Economic Review, XXXV, No. 4; September, 1945, pp. 519-30.

[30] Hayek, F. "The Pretence of Knowledge," in Lindbeck, A., ed Nobel Lectures, Economics 1969-1980, World Scientific Publishing Co., Singapore, 1992

[31] Heeks, R. 1999. Information and communication technologies, poverty, and development. Development Informatics Working Paper Series, Institute for Development Policy and Management. University of Manchester.

[32] Herrera, Y. Imagined Economies: The Sources of Russian Regionalism. Cambridge University Press, New York, 2005.

[33] Herz, B., & G. Sperling. Evidence and policies from the developing world: What works in girls' education. Executive summary. Council on Foreign Relations, New York, 2004.

[34] Hjørland, Birger (2007) Semantics and Knowledge Organization. Annual Review of Information Science and Technology 41:367 - 405.

[35] Hjørland, Birger (2004) Theory of knowledge organization and the feasibility of universal solutions. Presented at the Eighth International ISKO Conference, London, July 13-16, 2004.

[36] Jenkins, R. Democratic Politics and Economic Reform in India. Cambridge University Press, New York, 1999.

[37] Kanungo, S. 2004. On the emancipatory role of rural information systems. Information Technology and People. 17(4): 407-422.

[38] Loegelin, M. 1992. Interformation et development: etude synthetique des lignes de force du discours universitaire. Institut Universitaire de Technologie B, Department Carrieres de l'information et del al communication, Universite de Bourdeaux.

[39] Mai, J-E. "Likeness: A Pragmatic Approach" In Dynamism and Stability in Knowledge Organization. Proceedings of the Sixth International ISKO Conference. Advances in Knowledge Organization, 7: 23-27, 2000.

[40] Mai, J-E. "The Concept of Subject: On Problems in Indexing" In Knowledge Organization for Information Retrieval. Proceedings of the 6th

International Study Conference on Classification Research. 6: 60-67, 1997.

[41] Mas-Colell, A., Whinston,M., and J.Green. Microeconomic Theory. Oxford University Press, New York, 1995.

[42] Madon, S. and Sahay, S. 2002. An Information-based Model of NGO-mediation for the Empowerment of Slum Dwellers in Bangalore. The Information Society:18(1): 13-20.

[43] McCubbins, M., and T. Schwartz. "Congressional Oversight Overlooked: Police Patrols vs. Fire Alarms," American Journal of Political Science, February 1984.

[44] McKenzie, D., and D. Mookherjee. "Distributive Impact of Privatization in Latin America: An Overview of Evidence from Four Countries," Economia, vol. 3, no. 2, Spring 2003, pp. 161-218.

[45] Menou, M.J. 1985. An overview of social measures of information. Journal of the American Society of Information Science and Technology. 36(3): 169-177.

[46] Oates, W. "An Essay on Fiscal Federalism," Journal of Economic Literature, 1999, 37, (3), 1120-1149

[47] Olken, B. "Revealed Community Equivalence Scales," Journal of Public Economics 89 (2005): 545-566.

[48] Neuwirth, R., Shadow Cities: A Billion Squatters, Routledge, London, 2006.

[49] Perlman, M., "Political Purpose and National Accounts," in Alonso, W. and P. Starr, eds. The Politics of Numbers, Russel Sage Foundation, New York, 1987.

[50] Persson, T. "Do Institutions Shape Economic Policy?" Econometrica, v.70, no.3, 2002.

[51] Puri, S.K. and Sahay, S. 2003. Participation through communicative action: A case study of GIS for addressing land/water management in India. Information Technology for Development. vol 10: 179-199.

[52] Scott, J. Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed. Yale University Press, New Haven, 1998.

[53] Seely-Brown, John. and S. Duguid. The Social Life of Information. Harvard Business Press, Cambridge, 2002

[54] Sen, A. Development as Freedom. Oxford University Press, New York, 1999.

[55] Somasundaram, L. personal correspondence, Chennai, India, December 20, 2007.

[56] Srinivasan, R., Pepe, A., and Rodriguez, M. "Eliciting Cultural Ontologies: A comparison between Hierarchical Clustering Methods and Participatory Design Processes.", under review.

[57] Srinivasan, Ramesh, and Jeffrey Huang. "Fluid Ontologies for Digital Museums." International Journal of Digital Libraries.5 (2005): 193-204. .

[58] Srinivasan, Ramesh. "Ethnomethodological Architectures: Information Systems Driven by Cultural and Community Visions." Journal of the

American Society for Information Science and Technology 5.5 (2007): 723-33.

[59] ---. "Indigenous, Ethnic and Cultural Articulations of New Media." International Journal of Cultural Studies 9.4 (2006): 497-518.

[60] ---. "Where Community Voice and Information Society Intersect." The Information Society 22 (2006): 355-65.

[61] Star, Susan Leigh, and James R. Griesemer. "Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39." Social Studies of Science 19 (1989): 387-420.

[62] Starr, P, "The Sociology of Official Statistics,"

[63] Suchman, Lucy A. Human-Machine Reconfigurations: Plans and Situated Actions. 2nd ed. ed. Cambridge: Cambridge University Press, 2006.

[64] ---. "Located Accountabilities in Technology Production." Scandanavian Journal of Information Systems 14.2 (2002): 91-105.

[65] Stepan, A. State and Society: Peru in Comparative Perspective. Princeton University Press, Princeton, NJ, 1978.

[66] Turban, E., Aronson, J., Lian, T., Sharda, R. Decision Support and Business Intelligence. Prentice-Hall, New Jersey, 2006 (8th Edition)

[67] Ulman, J. Future Tendencies in Computer Science, Control and Applied Mathematics. Springer-Verlag, Heidelberg, 1992.

[68] United Nations Human Development Report. http://hdr.undp.org/en/.

[69] Vernon, R., "The Politics of Comparative Economic Statistics: Three Cultures and Three Cases," in Alonso, W. and P. Starr, eds. The Politics of Numbers, New York, 1987.

[70] Warschauer, M. "Reconceptualizing the Digital Divide," First Monday, v.7, no.7, July 2002.

[71] ___. Demystifying the digital divide. Scientific American 289(2), August 2003, 42-47.

[72] Watson-Verran, Helen, and Leon White. "Issues of Knowledge in the Policy of Self-Determination for Aboriginal Australian Communities."

[73] Wallack, J.S., and N.K. Singh. "Of Demolition Drives and Creative Destruction," Indian Express, October 15, 2006.

[74] Wallack, J.S. "Fire Alarms in Action: Making Sense of User Feedback on City Services in Karnataka, India," under review, 2008.

[75] Wallack, J.S, and S. Nadhamuni, "User Innovation and eGovernance Design," forthcoming in Hidden Successes: Urban Reforms in India, MIT Press, Cambridge, MA,

[76] Ziegler, E., and J.S. Wallack, "State Capacity," m/s, UCSD, San Diego, 2008.

[77] Wallack, J.S., and T.N. Srinivasan. Federalism and Economic Reform. Cambridge University Press. Cambridge, 2006.