In search of missing pieces: A re-examination of trends in ICTD research

Megh Marathe[1], Priyank Chandra[1], Vaishnav Kameswaran[1], Tsuyoshi Kano[1], and Syed Ishtiaque Ahmed[2]

[1]University of Michigan, Ann Arbor
{marathem, prch, vaikam, tkano}@umich.edu

[2]Cornell University, Ithaca
sa738@cornell.edu

ABSTRACT

ICTD benefits from being a broad multidisciplinary field that unites researchers from a wide range of domains attempting to understand the role of ICTs in the context of social, economic and political development. As a new field of study, however, ICTD continuously grapples with epistemological differences and varying (and often evolving) notions of what counts as development, and there still exists a significant gap in outlining where the current boundaries of this field lie. In this paper we present "Missing Pieces," an ongoing research project with the primary aim of uncovering the constitution, development, growth and impact of ICTD. Through in-depth quantitative and qualitative analyses of the distribution of participation in and impact of ICTD research across places, people, institutions, organizations and funding agencies, this project will look beyond just trends, and instead focus on finding the missing pieces, i.e. what, or rather who is being left out.

Categories and Subject Descriptors

K 4.0 [Computers and Society]: General

Keywords

ICTD; Trend Analysis; ICTD Trends.

1. INTRODUCTION

The field of ICTD (Information Communication Technology and Development) benefits from being a broad multidisciplinary field that unites researchers from a wide range of domains attempting to understand the role of ICTs in the context of social, economic and political development. The last few decades in development has seen coalitions of researchers, bureaucrats, NGOs, and industry attempting to bridge technological divides and, through enhancing access to technology, alleviate the condition of marginalized populations. In recent times, we also see researchers eschew techno-deterministic analyses and paternalistic interventions, instead working with communities and attempting to transition ICTD to a bottom-up user-centric research domain.

As a new domain of study, a handful of academic papers have attempted to map the evolution and impact of ICTD, along with defining what truly constitutes ICTD research. ICTD is a field that continuously grapples with epistemological differences and varying (and often evolving) notions of what counts as development, and while papers have attempted to address these concerns [2], there still exists a significant gap in outlining where the current boundaries of this field lie. Hence, we have initiated a project titled "Missing Pieces" which primarily seeks to identify the constitution, growth and impact of ICTD.

One part of this project will seek to enhance the understanding of the historical development of this young discipline by looking at the literature, people, and places that it has drawn towards it. Studying the actors that constitute ICTD research along with analyzing the spaces they function in – geographical, theoretical, and methodological, allow us to map it across multiple dimensions. In doing so, we hope to identify the clusters in ICTD research and the trajectory of these clusters over the last few decades. The other part of this project will examine the influence of this discipline over theory and practices in different other domains both inside and outside academia. This will not only help investigate the inclusiveness and residuality at play in ICTD research but also provide useful pointers that can help guide the future development and growth of this discipline. More specifically, how can it be made more meaningful, effective and impactful?

Further, through in-depth quantitative and qualitative analyses of the distribution of participation in and the impact of ICTD research across places, people, institutions, organizations and funding agencies, this project will look beyond just trends, and instead focus on the missing pieces, i.e. what, or rather who, is being left out.

In this first account of our work in progress, we focus on a description of our methods and report results from our preliminary analyses. We conclude with a discussion of our findings and gesture towards their implications for our project in particular and ICTD research in general. In the long term, our hope is to create a comprehensive, scalable database of ICTD research, that can act as a repository for ICTD research from around the world.

2. RELATED WORK

This is not the first paper to study trends in past ICTD research. In some aspects, we build upon work by Patra et al. [7], who conduct an extensive literature review on ICTD research along with interviewing researchers working in the domain to examine the growth of ICTD from the 1990s to 2009.

Chepken et al. [3] conduct a review of papers by fifty ICTD researchers to identify trends in the technologies, domain areas, regions, research methods and disciplines in ICTD interventions from 1995 to 2010. While presenting a good overview, Chepken et al.'s sampling strategy is a bit restrictive, resulting in a total of only ninety-three papers. Their findings are further restricted by the coarseness of some of their categories and their focus on interventions alone.

ICTD '16, June 03-06, 2016, Ann Arbor, MI, USA.
ACM 978-1-4503-4306-0/16/06.
http://dx.doi.org/10.1145/2909609.2909644

Gomez et al. [5] adopt a mixed-methods approach and analyze 948 papers ranging from 2000 to 2010 to identify trends in ICTD research. Of the many points of analysis, emphasis was placed on technological objects of study and ICTD domains. We extend this extensive study by adopting a critical lens to look beyond trends in data. By doing so, we seek to map the domain of the ICTD, specifically its nature and areas of impact, and subsequently identify missing pieces.

Dodson et al. [4] present an analysis of ITID articles with the objective of evaluating how development objectives are implemented in the field of ICTD. However, ITID in itself is not a complete representation of work in ICTD. We adopt a broader perspective by looking at work in other conferences and thus build a more holistic picture of ICTD and what it is missing.

We, however, also seek to identify existing biases in the research and broadly investigate the following:

a) Who are the people contributing to ICTD scholarship?

b) Whose problems are being addressed and who are left out?

c) To what extent is ICTD scholarship participatory and situated?

The project also seeks to assess the influence of this discipline over theory and practices in other domains both inside and outside academia. In this note, we describe the the "Missing Pieces" project and detail how we intend to proceed in the future.

3. THE MISSING PIECES PROJECT

This project is the result of a collaboration between five graduate students from two universities in the United States (institution names removed to preserve anonymity during the review process). Our team collectively represents a diverse mix of research interests, research methodologies, nationalities, educational backgrounds and work experience in academia, industry and development agencies. As researchers deeply invested in the field of ICTD, we initiate this project as a form of sensemaking – understanding the historical development and impact of ICTD and the literature, people, and places that constitute it, but also as an opportunity to look back and reflect on what is missing, in terms of populations, locations, kinds of problems, methodologies and specific technologies. A straightforward way to achieve this could be to compare the collective past and status quo of ICTD research with global development aims and recommendations such as the United Nation's Sustainable Development Goals [9]. This will not only help investigate the inclusiveness and residuality at play in ICTD research but also provide useful pointers that can help guide the future development and growth of this discipline.

While the following sections detail our methods and preliminary findings, this is where the project currently stands: we have created a corpus of ICTD research articles including automatically-extracted metadata and for around one-third of the articles, qualitatively coded and triangulated categories. As part of the creation of this corpus, we have developed a set of scripts that can be easily extended to include papers from other ICTD conferences and journals in future. In future, the coded subset of the corpus will serve as gold-standard training data for natural language processing algorithms that will automatically infer categories via topic modeling or classification. Automating this process will allow for the creation of a comprehensive, open repository of ICTD research articles, which summarizes past trends and updates itself automatically following the publication of new issues of journals or conference proceedings.

4. METHODS

4.1 Data Collection

We created a repository of 532 full papers and notes from three leading ICTD venues – the journal Information Technologies and International Development (ITID), the International Conference on Information and Communication Technologies in Development (ICTD), and the Symposium on Computing for Development (DEV), from 2003 to 2015. ITID is a peer-reviewed quarterly journal, which has published 44 issues containing 189 full papers or notes. We excluded ITID articles that were position papers or opinion pieces, in line with our aim of studying the trends in ICTD research. The ICTD conference has met seven times since its inception (2006, 2007, 2009, 2010, 2012, 2013, 2015) and published a total of 237 full papers, notes and demos. DEV is currently in its sixth year, with a total of 106 full papers or notes until 2014. These venues were chosen as a starting point with the goal of extending this study to ICTD research in other venues, such as EJIS DC, ITD, CSCW and CHI once we had a proof of concept of the repository and analysis processes in place.

The Association for Computing Machinery (ACM) kindly provided us with proceedings of the ICTD and DEV conferences in XML format. We wrote automated scripts to parse the proceedings and populate our repository with metadata for ICTD and DEV articles. Since ITID issues were not available in a similar format, we implemented automated scripts to download papers and scrape metadata from the ITID website. Our scripts assign a unique identifier to each paper and extract its title, author names and affiliations, abstract, and keywords (if present).

4.2 Qualitative Coding

To present a proof-of-concept of our work in progress, we selected a random sample 150 of the total corpus of 532 papers for qualitative coding and analysis. Our sample contains 67 articles from ICTD, 46 articles from ITID, and 37 articles from DEV. Each member of the research team independently coded 30 papers in two phases. For the first phase, in keeping with prior work on the analysis of trends in ICTD research (e.g. [6, 7]), we chose the following categories:

a) domain of work (e.g. agriculture, governance, health)

b) location of study (e.g. Nairobi, Kenya; Xekong, Laos)

c) problem being addressed (e.g. breast milk pasteurization, use of Nepali in software, ludic design in ICTD)

d) kind of work (theoretical, intervention, or analysis)

e) research methodology (qualitative, quantitative, or mixed)

f) specific research methods (e.g. interview, survey, discourse analysis, social network analysis)

g) type of ICT, if applicable (e.g. mobile phone, tele-center)

h) target population (e.g. farming communities in rural central India, students in Laos, hunters in the rainforests of Congo)

This first round of coding gave us a bottom-up understanding of the kinds of themes that are common to each of these categories.

4.3 Triangulation

Based on a first round of coding, we came up with a shared rubric so as to establish some uniformity between coders in subsequent coding, paving the way for quantitative analyses, but also creating gold standard training data for automated analyses in future. Once the coding process was complete, four of the team members convened to triangulate the coded data by checking it for

inaccuracies and re-investigating papers that did not fit into the shared rubric.

5. PRELIMINARY FINDINGS

5.1 Research Methodologies

The multidisciplinary nature of ICTD research has brought in a wide variety of methodological approaches and our analysis attempts to map ICTD in these methodological spaces. In our preliminary analysis, we classified the nature of the study in each of the sampled paper. The paper was manually coded as either a technological design intervention, an empirical analysis of an existing technology, or a theoretical analysis. We further looked at the research methods used - dividing it up into quantitative studies, qualitative studies, or mixed methods studies with an additional category for purely technical solutions (*Engineering*) and for papers that described demonstrations (*Demo*).

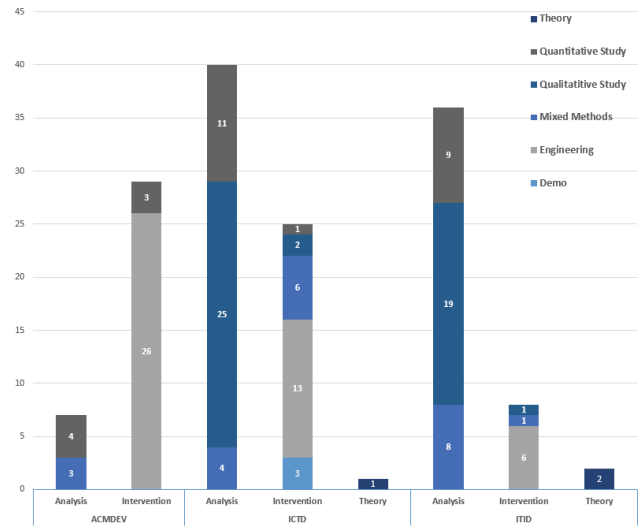

Figure 1. A break-up of ICTD research by venue, type of study and methodology.

As we see from our preliminary coding (Figure 1), papers from the ACMDEV conference were primarily technological interventions (81%) with the rest analyzing the impact of existing interventions. This is a direct consequence of it being the primary venue for presenting technologies and architectures related to development. In contrast, both the ITID (78%) journal and ICTD (60%) conference had a relatively larger proportion of papers empirically analyzing interventions, and the general adoption and use of technologies. The difference in proportions between the ITID journal and ICTD conference is driven by the fact that the notes published in the ICTD conference proceedings predominantly present interventions. Future analysis of this data will seek to include notes as a distinct category to refine these results.

With respect to the primary methodology used by researchers, we find that the majority of analysis papers in both the ITID journal (53%) and ICTD conference (63%) follow qualitative research methods. This is driven by semi-structured interviews and relatively more in-depth ethnographies being an important method in understanding the perceptions of a community and individuals with respect to technology use. Quantitative analyses are restricted by what kind of data is available to researchers. Most of such analyses has thus either focused on macro-level analysis of country-level data or survey data. Some research has also been able to analyze user behavior through analyzed quantified data that is generated by information systems.

5.2 Domains

The field of socio-economic development can potentially encompass a wide range of domains. Our analysis looks specifically at where ICTD has focused on so we can better understand how it intersects with the development goals of policy makers and what domains are potentially under-represented.

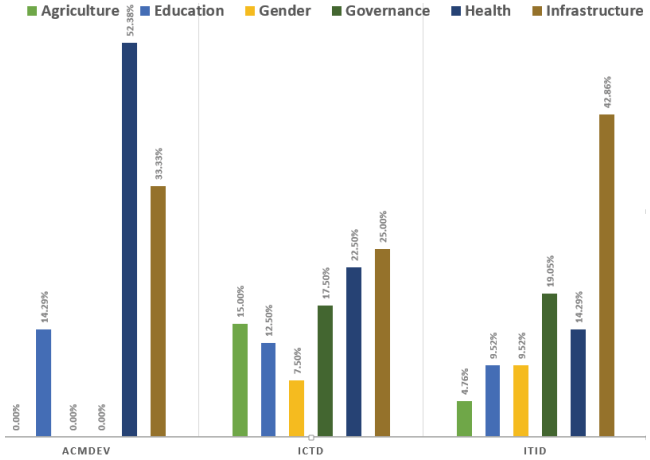

Figure 2. A break-up of ICTD research by venue and domain.

As seen in the above graph (Figure 2), there are six primary domains that ICTD research has focused on – agriculture, education, gender, governance, health and infrastructure. Studies on ICT infrastructures have a significant presence in all the venues (ACMDEV – 33%, ICTD – 25%, ITID – 43%). This is a broad domain that includes the systems and structures that deliver ICT applications and services to developing communities. While ACMDEV has primarily focused on technological interventions in the domain of health and education, the ICTD conference and ITID journal are relatively more distributed with respect to these 5 domains. Future analyses would be focus on comparing these domains with development goals and identifying the reasons that these domains receive greater attention than others.

5.3 Affiliations

A critique of development has often been the distance between researchers and sites of underdevelopment, both geographically and culturally [8]. As a preliminary analysis, we looked at the affiliations of researchers who contribute to ICTD research. For this, we parsed the ACM dataset consisting of data from the ACMDEV and ICTD conferences, and extracted the authors and affiliations for each publication. We subsequently took a count of the institutions that are present for each publication.

Table 1. The ten most frequent affiliations of authors

Top 10 Affiliations	Count
University of Washington, Seattle	38
University of California, Berkeley	30
Microsoft Research India	22
University of Michigan	14
University of Cape Town	12
Carnegie Mellon University	12
Georgia Institute of Technology	12
New York University, New York	10
Unaffiliated	9
Michigan State University	8

As we see in Table 1, out of the top 10 affiliations, 8 are universities in the United States with only University of Cape Town and Microsoft Research India are located in the Global South. An extension of this study will be to identify the targeted populations of the research studies

With the objective of our analysis to understand the people and places that are drawn to this field, studying the spatial distribution of research will be an important means of mapping ICTD clusters and understanding who are part of the transnational networks of ICTD research.

6. DISCUSSION

6.1 Classification is not trivial

A major challenge we faced during qualitative coding was standardizing the coding categories across venues, domains, and methodologies. This problem is exacerbated by ICTD research inhabiting multiple intersecting research disciplines. This boundary object-like [1] nature of the field makes it harder to evaluate and classify ICTD research under categories and keywords borrowed from other disciplines such as engineering or development studies. We might instead need to think about creating new taxonomies that can capture ICT innovations that often seek to fracture older existing categorization of research.

Like any other knowledge institution, journals and conferences are about more than just venues for publishing research papers; they play a very important role in building a shared understanding of a research field. The classification and organization schemes that these institutions must be carefully designed so that they don't replace unique local labels familiar to ICTD researchers with more universal taxonomies. One starting point could be a more nuanced scheme of ICTD-specific categories and subject descriptors.

6.2 The way forward

The "Missing Pieces" project's long-term objective is to create a comprehensive scalable annotated database that can act as a repository for ICTD research from around the world. While our preliminary analysis focused on three important venues of ICTD, this will be scaled to include other conferences and journal in the near future. We also hope to create a web interface that can allow anyone to propose a paper to be added to this ICTD database.

There are two significant directions that we seek to go forward – the first is to comprehensively code existing ICTD research papers while the second seeks to situate the ICTD research with respect to other existing research domains through a bibliometric analysis.

The coding of the papers will itself be divided into two primary categories – an automated coding that extracts meta-data from the paper and a manual coding that is a more thorough annotation done by expert ICTD researchers. The automated coding will extract fields such as the title, abstract, author names and their affiliations, country of analysis, research methodologies used, and other easily extractible fields. The codes are currently being written and as we do so, we are increasingly finding the need to rethink how content is published in online journals and conference proceedings. Portable Document Format (PDF) files are harder to parse than text files, and online repositories need to keep this in mind if they intend to create more accessible research.

The subsequent manual coding will validate the fields extracted in the automated process and involved ICTD scholars carefully reading the papers and providing more comprehensive annotations. For this, we have ICTD research students from varying methodological and theoretical backgrounds who have volunteered to curate papers in the variety of domains – from human rights to media studies that will allow us to comprehensively annotate existing ICTD research.

The second part of the project will examine the influence of this nascent discipline over theory and practices in other domains both inside and outside academia. Through a bibliometric analysis of the ICTD research, we will explore its transition with respect to time and its methodological and theoretical intersections with other domains.

A theme that shall run through the entire project is to continuously critically analyze ICTD research – both its historicity and current status and seek to identify spaces where it needs greater reflection and attention.

ACKNOWLEDGEMENTS

We are indebted to the ACM for providing us with past conference proceedings. We thank our anonymous reviewers for their comments and feedback.

7. REFERENCES

[1] Bowker, G.C. and Star, S.L. 1999. *Sorting Things Out: Classification and its consequences*.

[2] Burrell, J. and Toyama, K. 2009. What Constitutes Good ICTD Research? *Information Technologies and International Development*. 5, 3 (2009), 82–94.

[3] Chepken, C. et al. 2012. ICTD interventions: Trends Over the Last Decade. *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development - ICTD '12*. (2012), 241.

[4] Dodson, L. et al. 2013. Considering Failure: Eight Years of ITID Research. *ITID*. 9, 2 (Jun. 2013).

[5] Gomez, R. et al. 2012. The changing field of ICTD. *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development - ICTD '12*. December 2015 (2012), 65–74.

[6] Gomez, R. and Pather, S. 2011. ICT Evaluation: Are We Asking the Right Questions? *EJISDC*. 50, (Oct. 2011).

[7] Patra, R. et al. 2009. ICTD state of the union: Where have we reached and where are we headed. *2009 International Conference on Information and Communication Technologies and Development, ICTD 2009 - Proceedings* (2009), 357–366.

[8] Taylor, A.S. 2011. Out There. *CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.* (2011), 685–694.

[9] United Nations General Assembly 2015. *Transforming our world: the 2030 Agenda for Sustainable Development*.