What We Know About Traditional MERL Tech:

Insights from a Scoping Review

Zach Tilton, Michael Harnar, Linda Raftree, Paul Perrin, Gretchen Bruening, Soham Banerji, John Gordley, Manon McGuigan, Hanna Foster, and Michele Behr

2

June 2020

Contents

What We Know About Traditional MERL Tech: Insights from a Scoping Review

Authors: Zach Tilton, Michael Harnar, Linda Raftree, Paul Perrin, Gretchen Bruening, Soham Banerji, John Gordley, Manon McGuigan, Hanna Foster, and Michele Behr

This research was partially funded by the Interdisciplinary PhD in Evaluation program at Western Michigan University and the Pulte Institute for Global Development at the University of Notre Dame.

Abstract

This paper explores the peer-reviewed evidence base of "traditional" technology-enabled monitoring, evaluation, research, and learning (MERL Tech) in international development assistance from 2015 to 2019. The authors conducted a scoping review that searched seven databases, screened 3,054 reference titles and abstracts, coded 886 abstracts, and extracted and analyzed conclusions and recommendations from the full texts of 256 studies. The findings reveal the most frequently reported technologies, MERL activities, and the sub-sectors, and the geographies where those tech-enabled activities occur. Gaps in the evidence for specific technologies, MERL activities, and sectors are mapped. The data reveals which technologies are trusted more than others and reported barriers to effective MERL Tech implementation and areas that researchers suggest for further investigation. The results suggest that the evidence from peer-reviewed studies is not proportional to estimated MERL Tech activity, significant publication bias exists, and further knowledge synthesis of unindexed grey literature is needed to provide a more comprehensive and possibly accurate description of MERL Tech practice.

Introduction

International development, with its disparate geographic, social, and political contexts, is a challenging field. Its complexity has tested the utility of traditional monitoring and evaluation (M&E) approaches. More than ever, there is increasing emphasis on real-time feedback, rigorous data collection, and quantifiable results.[1]

Advancements in data and technology over the past decade have fundamentally changed how practitioners conduct systematic inquiry and increased the demand for and utilization of innovative techniques for technology-enabled monitoring, evaluation, research, and learning (MERL Tech). These developments have in turn increased the number of pioneering new platforms and data collection tools. MERL Tech is a relatively new sub-field that has garnered increased interest over the past five years. Despite the reported benefits of integrating technology into MERL processes, there is ongoing debate about the impact of technology on MERL.[2] The tension is exacerbated by the fact that, although MERL Tech practitioners are charged with using technology to develop evidence bases for various sectors of sustainable development, there has been little effort to systematically develop an evidence base for MERL Tech.

The field of MERL Tech has grown in the five years since Linda Raftree and Michael Bamberger's 2014 report, Emerging Opportunities: Monitoring and Evaluation in a Tech-Enabled World[3], in large part due to technological advances. MERL Tech innovators were experimenting then and continue to do so now to harness the power of technology to help overcome common MERL challenges and real-world evaluation constraints. At the time, tech-enabled MERL was new territory, and Raftree and Bamberger reflected on how early adopters were navigating this space while advancing traditional evaluation methods.

Since 2014, practitioners, researchers, technologists, and evaluators have been working together to improve their practice-based fields by integrating information and communication technologies (ICT) into their respective transdisciplinary evidence-based work. The assumption among these actors is that technology enhances monitoring, evaluation, research, learning, and related activities. We hear of more accurate data collection and lower costs. Practitioners use data analytics, visualization, dashboards, and mapping to help make sense of collected data and to consume it more efficiently and encourage data-driven decision-making. MERL Tech has opened a universe of richer data that helps reveal previously obscured patterns and develop deeper and better understanding of communities. Whether through mobile data collection via smart phones or tablets, short message service (SMS)-based surveys, global positioning system (GPS)-enabled tracking, intuitive dashboards, or voice-enabled beneficiary interactions, MERL Tech has been playing an integral role in advancing the application of technology for the pursuit of the sustainable development goals. Indeed, a recent global digital ecosystem study from the Digital Impact Alliance found that the majority of information communication, technology for international development (ICT4D) solutions are typically internally focused, supporting organizational technology and enterprise needs, rather than externally focused end-user solutions.[4] In other words, the lion's share of digital development or ICT4D solutions are MERL Tech solutions.

[1] United States Agency for International Development (USAID) Evaluation Policy (2016); USAID Digital Strategy 2020–2024 (2020).

[2] Linda Raftree (2020). MERL Tech State of the Field: How is Innovation Happening in MERL Tech?"

[3] Linda Raftree and Michael Bamberger (2014). Emerging Opportunities: Monitoring and Evaluation in a Tech-Enabled World. https://www.rockefellerfoundation.org/report/emerging-opportunities-monitoring/

[4] Digital Impact Alliance (2018). DIAL Baseline Ecosystem Study, 11. https://digitalimpactalliance.org/research/digital-impact-alliance-2018-baseline-ecosystem-study/

With this growth, there is increasing consensus among practitioners that MERL Tech can serve as a solution to many of the challenges related to MERL — but it can also become a source of new problems. For instance, a Mobenzi workshop report from the MERL Tech conference in Johannesburg in 2018 found that many MERL practitioners and users are unaware of MERL use cases and unsure of how to implement solutions. The report also found that one barrier to implementing MERL Tech solutions was the availability and quality of training material and support from knowledgeable consultants.[5] This comes as no surprise, as the community of professionals who recognize and identify with MERL Tech as an emerging interdisciplinary professional field remains small, and the evidence base available to the community is limited. While MERL Tech practitioners convene, share insights, and reflect through blogging, annual conferences, and the occasional white paper, there has been no attempt to date to systematically synthesize and distill what might be considered an emerging evidence base of MERL Tech activities.

A 2019 systematic evidence mapping study of ICT4D interventions using impact evaluations by Annette Brown and Hannah Skelly of FHI 360 comes closest to a systematic formation of a MERL Tech evidence base.[6] Brown and Skelly identified and mapped 253 studies that use ICT4D interventions across multiple sectors. The review looked at impact evaluations across nine development sectors and 11 intervention types that use or promote mobile and internet technologies such as mobile health and digital inclusion. However, the study did not utilize a framework of specific technologies such as mobile devices or global positioning systems and the respective MERL activities that those technologies enabled. While this study has been instrumental in describing the nature of the causal evidence in ICT4D broadly, due to the focus on ICT4D intervention types as opposed to tech and MERL combinations and the exclusive inclusion criterion for randomized controlled trails, few (if any) inferences about the large share of internally focused ICT4D or MERL Tech evidence can be drawn from the study.

Given these opportunities and developments, we conducted a scoping review with the goal of identifying gaps in the knowledge base, clarifying key concepts, reporting on the types of evidence that inform practice, and beginning to synthesize conclusions and recommendations related to MERL Tech. In service to those aims, we formulated the following research question:

Research Question

"What types of evidence, activities, and conclusions related to MERL Tech for development assistance in lower to middle-income countries were reported in academic literature between January 2015 and May 2019?"

[5] Mobenzi, MERL Tech Jozi (2018). 2018 Workshop Report: Potential gaps in the MERL Tech sector, https://www.mobenzi.com/merl-tech-report-2018.html (requires registration).

[6] Annette N. Brown and Hannah J. Skelly (2019). "How Much Evidence Is There Really? Mapping the Evidence Base for ICTD Interventions," *Information Technologies and International Development* 15.

Summary of Methods

A team of 10 researchers conducted the scoping review between mid-2019 and early 2020. The two components of traditional MERL Tech are 1) the specific information and communication technologies that do not constitute emerging or big data technology[7] and 2) the myriad activities found within the monitoring and evaluation (M&E) lifecycle that the technologies enable. Analyzing five years of MERL Tech conference data, we identified the 25 most frequently occurring technology categories (see Table 4, page 9). The 10 categories of MERL activities modified from Raftree and Bamberger's M&E Lifecycle[8] are diagnosis; design; planning; data collection; implementation and monitoring; data analysis; evaluation; reporting, sharing, and learning; prediction and forecasting; and decision facilitation.

> **The two components of traditional MERL Tech are 1) the specific information and communication technologies that do not constitute emerging or big data technology[7] and 2) the myriad activities found within the monitoring and evaluation (M&E) lifecycle that the technologies enable.**

For inclusion in our review, references needed to be *related to* or *focused on* traditional MERL Tech. We deemed a reference *related to* traditional MERL Tech if the authors reported technologies enabling MERL activities as contributing to answering research questions or reports on interventions within relevant sectors, although not in and of themselves focusing on MERL Tech. We considered a reference *focused on* traditional MERL Tech if the main object of the study or report was research questions or reported activities specifically about MERL Tech. In effect, in focused studies MERL Tech is under the magnifying glass; in related studies MERL Tech *is* the magnifying glass. This distinction is an important analytical framework used throughout this paper. The research team validated the distinction through multiple rounds of piloting screening and extraction forms.

Inclusion Criteria

Table 1: Traditional MERL Tech Inclusion Criteria

Population	Individuals, households, organizations, communities, and governments in low- to middle-income countries
Concept	Any combination of 25 "traditional" information communication technology categories that enable any combination of 10 MERL activity categories
Context	10 subsectors of international development; any study design published in English-language peer-reviewed journals from January 2015 through May 2019

The Methods section (page 23) describes the methods and inclusion criteria used for the scoping review.

[7] See the companion papers in this series for an exploration of big data technologies and emerging approaches and technologies.
[8] Raftree and Bamberger, op. cit.

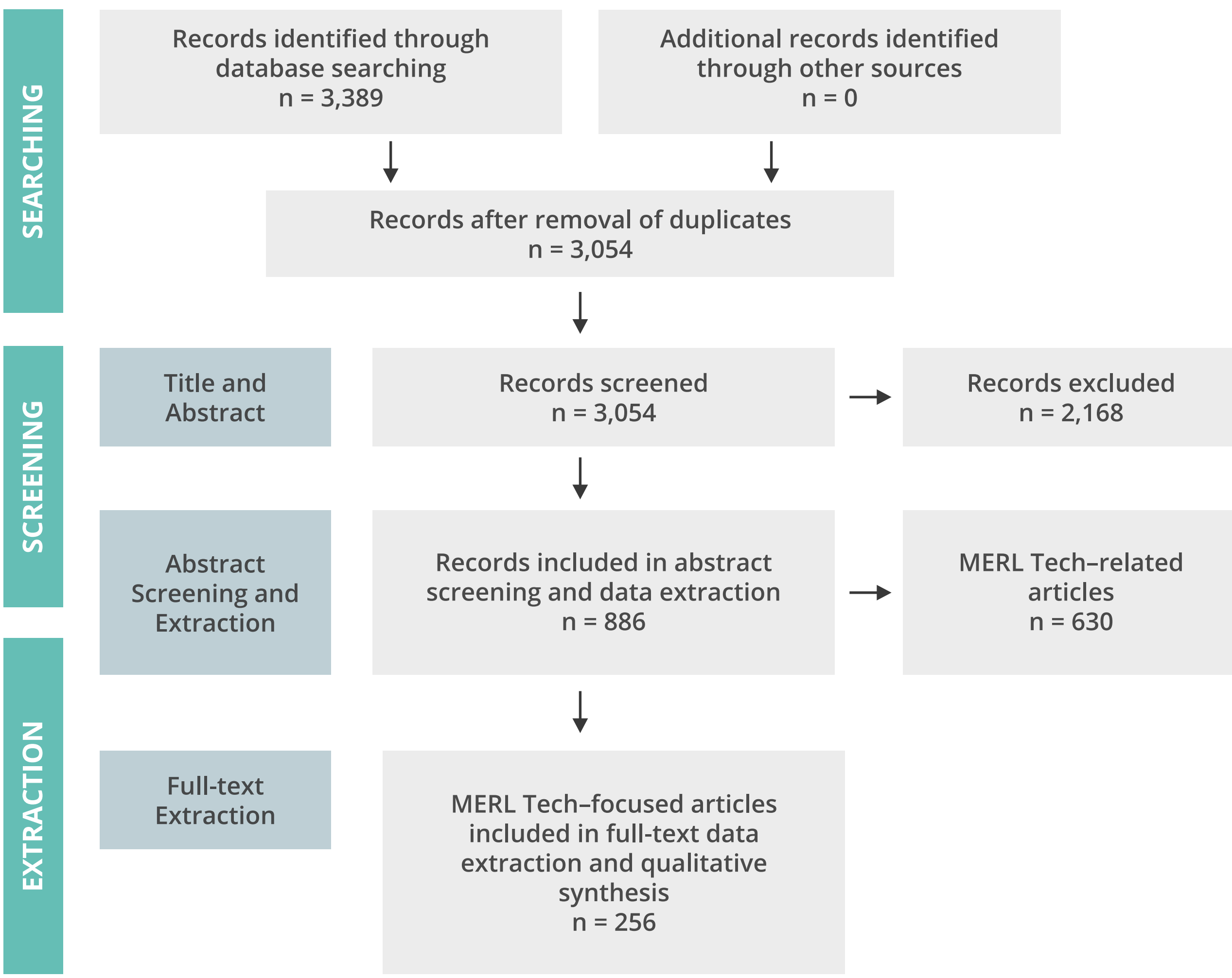
Findings

Characteristics of the MERL Tech Evidence Base

Initial Results

The team screened 3,054 studies to identify 886 references that reported use of MERL Tech. Of those, 630 (75 percent) were related to MERL Tech or reported the use of some type of MERL Tech to investigate other phenomena related to development assistance, while 256 studies made the various applications of tech-enabled monitoring, evaluation, research, and learning the object of investigation. Figure 1 depicts the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) flow chart for the search and screening results of the scoping review.

Figure 1: PRISMA Flow Chart: MERL Tech State of the Field Scoping Review



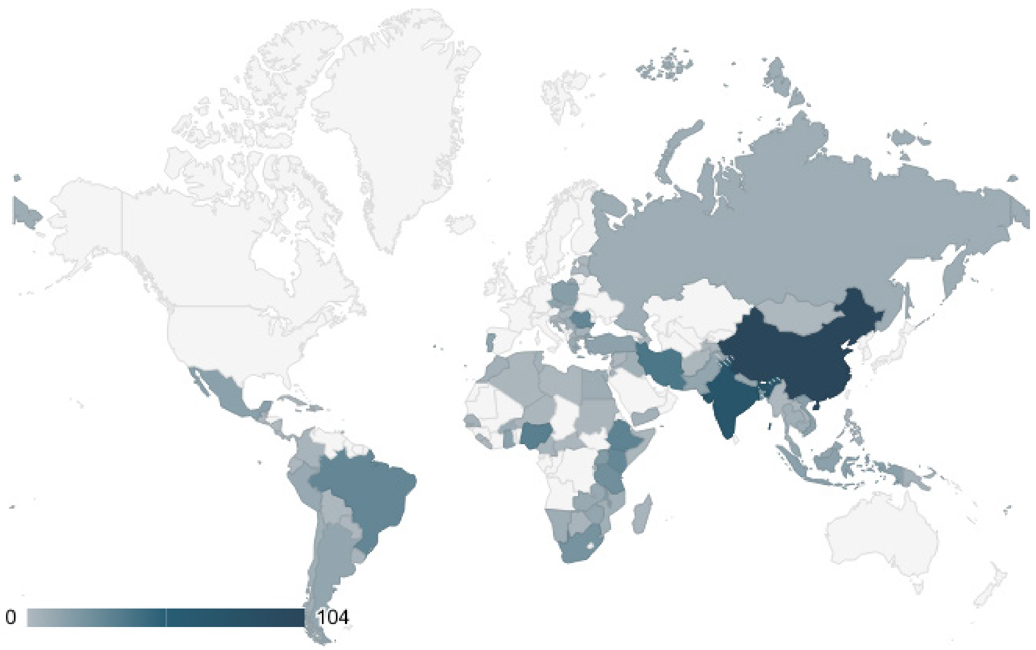
Region

Most reports of MERL Tech practice describe work in the sub-Saharan Africa or Asia and the South Pacific regions. Of the 886 references included in the study, 38 reported on research and programmatic activities in more than one setting for a total of 973 countries, regions, or contexts reported. Table 2 shows the largest number of mentions of activity in sub-Saharan Africa, ranging down to the fewest in South Asia. This distribution would be nearly identical to the regional composition of a recent ICT4D evidence mapping study,[9] had the authors included studies in low- to middle-income countries in Europe and Central Asia. A third category reflects non-specific low- to middle-income countries or resource-deprived settings (n = 80, 9 percent). The heat map (Figure 2) shows the prevalence of refences by country.

Table 2: Number of Studies by Region

Region	Number	Percentage
Sub-Saharan Africa	238	26%
East Asia and Pacific	181	20%
Europe and Central Asia	136	15%
South Asia	111	12%
Latin America and the Caribbean	84	9%
Middle East and North Africa	82	9%
Non-specific	80	9%

Figure 2: Heat Map of Number of References *Related to* or *Focused on* MERL Tech of the Field Scoping Review



[9] Brown and Skelly, op. cit.

Study Design

Less than 10 percent of identified studies are experimental. As listed in Table 3, 791 of the 886 studies included were coded as non-experimental. Of those, 451 were coded as case studies, 220 as correlational, 25 as systematic reviews, and 95 with other categories (grounded theory, historical, ethnographic, and others). Fifteen studies were coded as quasi-experimental and 60 as experimental.

Following the practice of most scoping reviews, the team did not interrogate the quality of these studies and methods, as in systematic reviews or meta-analyses. However, some readers may draw inferences about the state of the evidence generally from the distribution of reported study designs.

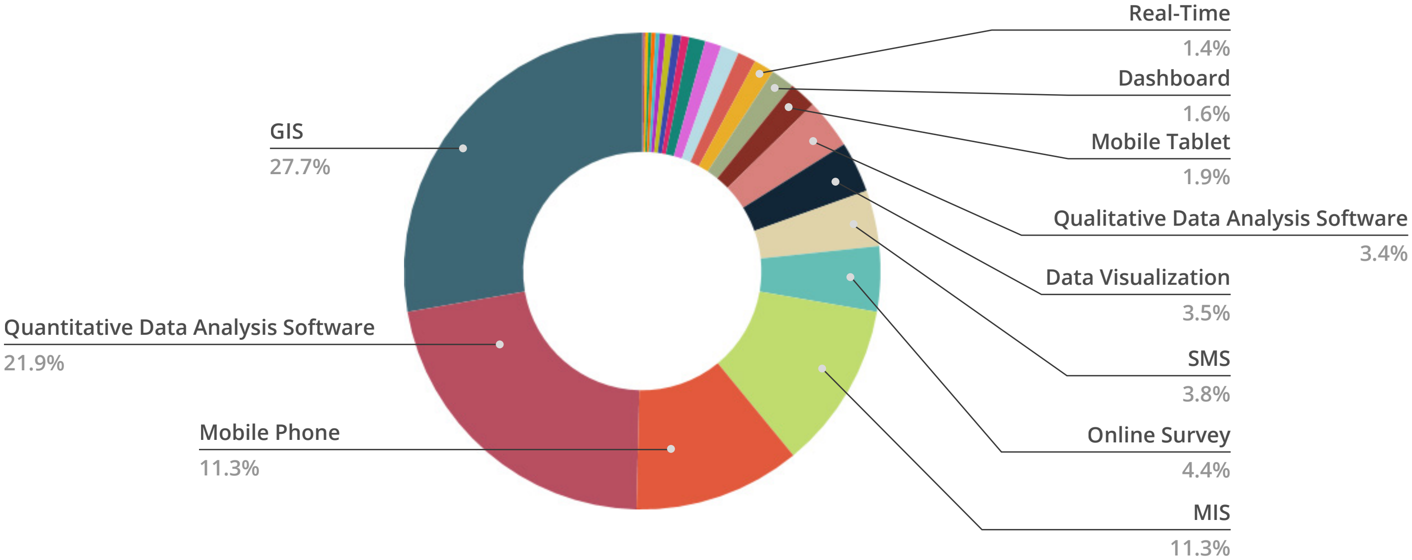
Table 3: Study Design Composition

Study Design	Number	Percentage
Experimental	60	7%
Quasi-experimental	15	2%
Non-experimental	791	89%
Not specified	20	2%
Total	886	100%

Technology

Geographic information system (GIS) and quantitative data analysis software account for half of all reported technologies. Figure 3 reveals the composition of reported technologies. GIS (n = 402), computer-assisted quantitative data analysis software (CAQDAS, n = 317), mobile phones (n = 164), and management information systems (MIS, n = 164) were the most frequently reported technologies among included references (n = 886). Many references reported the use or investigation of multiple technologies per reference, with 1,449 technologies reported for 886 references (an average of 1.6 technologies per reference). The most commonly reported paired technologies were GIS and quantitative data analysis software (n = 162), mobile phones and SMS (n = 39), and GIS and data visualization (n = 38). The least frequently reported technologies were Bluetooth (n = 1), computer-assisted interviewing (n = 3), and voice messaging (n = 3).

Figure 3: Percentage of Reported Technologies in MERL Tech–*Related* and –*Focused* References



These findings show that many technologies that were popular in 2014 remain popular today. MIS played a critical role in enabling organization-wide M&E capacity up to 2014, as underscored by Raftree and Bamberger and confirmed by our study. MIS is the third-most frequently reported technology (n = 164, 11 percent), the second-largest sector-specific reported technology (health), and the largest cluster of studies reported to enable a particular MERL activity for "monitoring and implementation" (n = 88, 34 percent). MIS, along with GIS, is utilized for every MERL activity throughout the MERL lifecycle. These findings suggest that young and emerging MERL Tech practitioners might do well to become familiar with and proficient in these more frequently reported technologies.

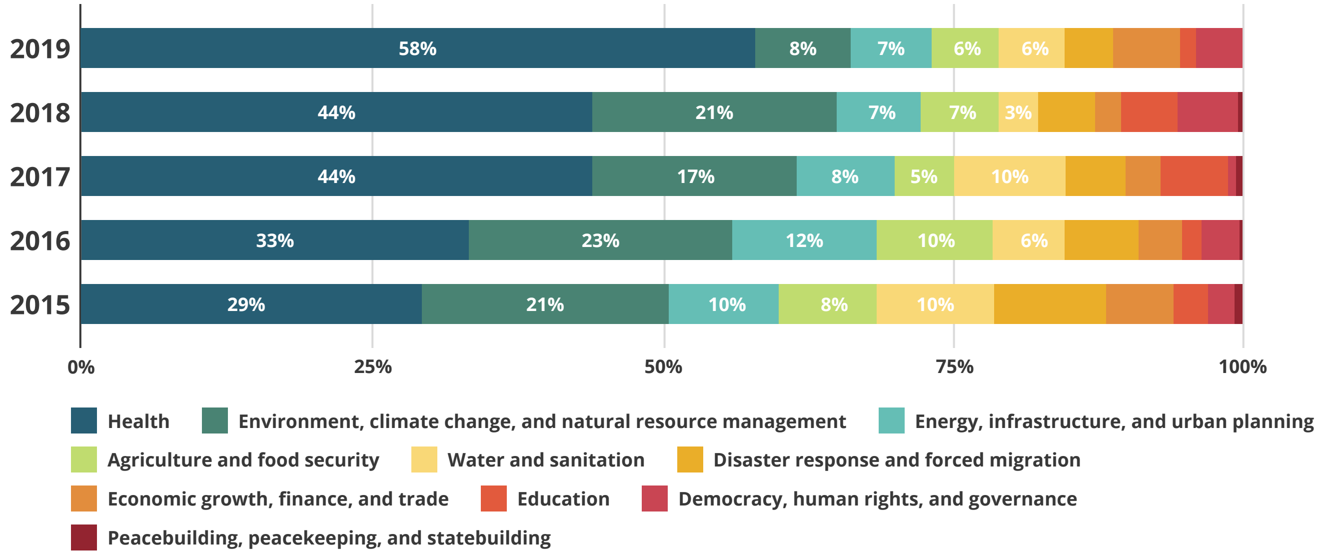
Real-time designation was mostly absent from evidence, despite 2014 hype. Other technology classifications that seemed to hold much promise, or received considerable hype, were markedly absent from the included studies. These classifications include real-time technology, especially after significant use of real-time technology for the 2010 Haiti earthquake disaster response and in the midst of the Ebola outbreak in 2014. Consider the view from 2014 that "ICTs are allowing for the collection of real-time data on participant experiences, behaviors and attitudes, meaning that analysis can be conducted early in the process and course corrections can be made to improve interventions and outcomes."[10] In stark contrast, the term real-time featured in only 20 abstracts, or 2 percent of total *related* and *focused* studies. This suggests that 1) either the promise of real-time in 2014 was overhyped; 2) the real-time qualifier quickly became less useful as this tech category became more institutionalized; 3) real-time tech is widely used but not being reported in academic literature; or 4) real-time technology is widely used and reported, but the scoping review search strategy was not sensitive enough to identify that discourse. While some combination of these may be accurate, we believe the real-time label has lost some purchase as more MERL Tech applications and use cases have approached what could be considered real-time.

Sector

The health sector increased its evidence share over the years observed and alone accounts for 39 percent of all identified evidence. Figure 4 shows the sector share of reported aid sub-sectors among references. Health dominates the sector share in all five years, ranging from nearly one in three sector codes in 2015 to over half in 2019, with a total of 412 references (39 percent) of all sector category codes. The environment, climate change, and natural resource management sector category was the second-most represented sector, although it was just over half the size of health with a total of 208 references (20 percent of sector category codes). The third-most frequently coded sector was energy, infrastructure, and urban planning, with less than half of the former at 99 references (9 percent of total sector category codes). The peacebuilding, peacekeeping, and state-building category had only six references, or less than 1 percent, through all five years. The sub-sector environment, climate change, and natural resource management category appeared in the three most common pairings of two or more sector categories, with energy, infrastructure, and urban planning (n = 28); water and sanitation (n = 25); and agriculture and food security (n = 22).

[10] Raftree and Bamberger, op. cit.

Figure 4: Sector Share: Proportion of Total Sector Code Categories by Year



Although the share of studies within the health sector in this review (39 percent) appears high, Brown and Skelly's evidence mapping study of ICT4D shows nearly 170 of 260 impact evaluations (65 percent) were from the health sector.[11] MERL Tech practitioners might consider what factors within this sector contribute to such a high proportion of total scholarly evidence and impact evaluations in both reviews, especially since official development assistance (ODA) allocated just 11 percent of total funding to the health sector for 2018.[12]

GIS in the environment sector and MIS in the health sector reflect the largest clusters of evidence, while the peacebuilding sector reflects the smallest. The two-variable evidence gap map (Table 4) shows the frequency of references coded by a specific technology and sector among all 886 *related* and *focused* references included in this study.

Table 4: Tech and Sector Evidence Map: Number of References that Report Each Technology, by Sector

	Agriculture	Democracy	Disaster	Economic	Education	Energy	Environment	Health	Peace	Water
audio recording	0	0	1	0	0	0	0	6	0	0
bluetooth	0	0	0	0	0	0	0	1	0	0
CAPInterview	0	0	0	0	0	0	0	3	0	0
CAQaulDAS	4	3	3	2	1	2	4	33	0	0
CAQaunDAS	26	5	26	21	13	34	93	121	2	33
CATele-Interview	0	0	0	0	0	0	0	4	0	0
crowdsourcing	1	4	4	0	0	2	1	7	0	0
dashboard	1	2	0	0	1	0	2	18	0	1
data viz	4	1	4	3	1	8	13	20	0	4
digital photography	1	0	2	0	0	0	0	12	0	1
GIS	53	10	52	21	5	76	183	65	1	55

[11] Brown and Skelly, op. cit.

[12] "Sectoral allocation of ODA: Aid by Major Purposes in 2018," Organisation for Economic Co-operation and Development. Accessed May 13, 2020. http://www.oecd.org/dac/financing-sustainable-development/development-finance-data/TAB19e.xls

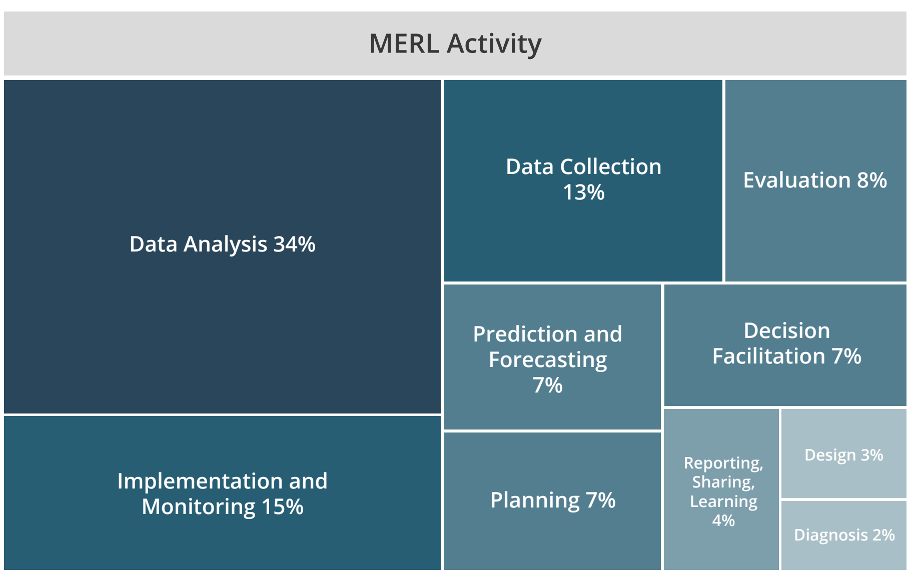
GPS	1	0	1	0	0	0	0	5	0	0
int. voice resp.	0	1	0	0	0	0	0	7	0	0
MIS	6	4	3	1	8	3	6	137	1	3
mobile phone	9	5	3	3	8	3	5	133	3	9
mobile tablet	4	1	1	2	0	2	1	20	2	1
online design prog.	1	0	0	0	1	0	0	0	0	0
online survey	2	6	1	3	12	6	8	33	2	3
open source	2	0	1	0	0	1	1	3	0	2
real-time	3	0	0	0	0	0	1	17	0	1
SMS	1	0	1	0	1	0	0	52	0	0
social media	1	1	3	1	1	0	0	11	0	0
telephone	0	0	0	0	0	0	0	18	0	0
voice messages	0	0	0	0	0	0	0	3	0	0
voice recording	1	0	0	0	2	0	0	1	0	0

The map includes 1,690 codes from 886 references. Although the counts in each cell represent unique references, counts across rows and down columns may likely refer to multiple sectors and technologies. For example, focusing on the most populated cell, 183 unique references pertain to both the GIS technology row and the environment, climate change, and natural resource management sector column. In contrast, although the peace sub-sector column reflects only six studies, 10 technology types were reported as common features.

MERL Activity

The most frequently reported MERL activity is data analysis. Figure 5 depicts the proportion of MERL activities reported in all 886 studies. Among 1,644 codes in this category across 886 studies, the average reference reported 1.86 types of MERL activities. Data analysis (n = 552), implementation and monitoring (n = 250), and data collection (n = 216) were the most frequently reported MERL activities.

Figure 5: Reported MERL Activities in MERL Tech–*Related* and –*Focused* References



Isolating these activities further helps identify the most frequent co-occurring technology categories. The most frequently co-occurring technologies for data analysis were GIS (n = 343), quantitative data analysis software (n = 302), and MIS (n = 55); data collection showed mobile phone (n = 71), online survey (n = 56), and quantitative data analysis software (n = 54); and implementation and monitoring showed MIS (n = 130), mobile phone (n = 95), and quantitative data analysis software (n = 48). Since this tree map and the narrative findings about co-occurring technology correspond to studies both related to and focused on MERL Tech, we are less likely to assume that all the technologies reported for this larger sample actually enable or facilitate the reported MERL activities. To make that inference, we would need to filter only the 256 focused references that made MERL Tech the object of investigation, as shown in Table 5.

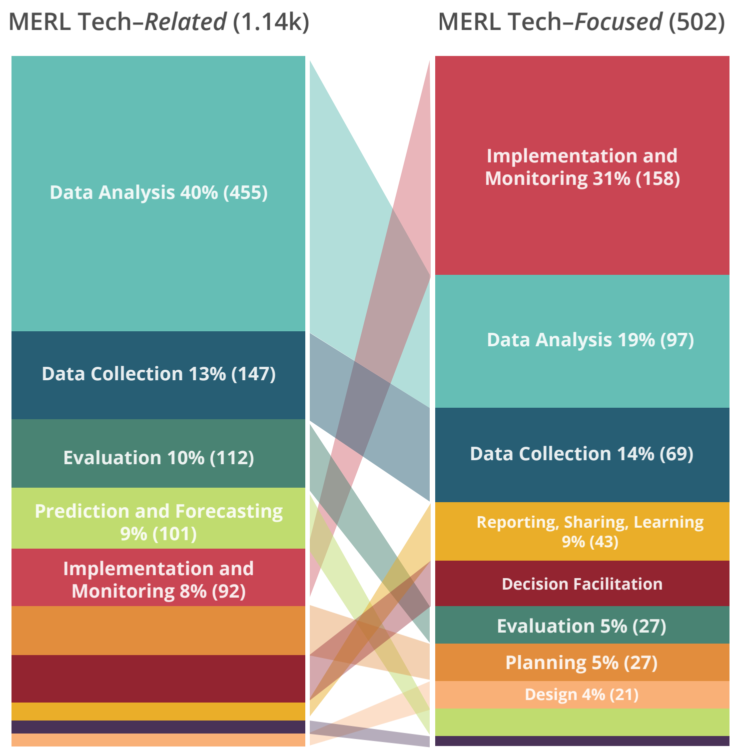
Table 5: MERL Tech Evidence Map: Number of *Focused* References that Report Tech–Enabled MERL Activities

Technology Category	Diagnose	Design	Plan	Collect	Monitor	Analyze	Decide	Forecast	Evaluate	Report
audio recording	0	0	0	2	2	1	0	0	0	0
bluetooth	0	0	0	0	0	0	0	0	0	0
computer-assisted personal interviewing	0	0	0	0	0	0	0	0	0	0
computer-assisted telephone interviewing	0	0	0	1	0	0	0	0	1	0
crowdsourcing	0	1	2	9	4	4	3	0	1	4
dashboard	0	3	3	4	12	6	4	2	4	4
data visualization	0	5	5	4	8	9	7	3	1	4
digital photography	4	0	0	5	3	4	1	0	1	0
geographic information system	0	5	13	9	14	36	12	11	4	7
global positioning system	0	0	1	1	2	2	1	0	0	2
interactive voice response	0	0	1	3	1	1	1	0	3	0
management information system	1	9	11	20	88	32	15	3	9	22
mobile phone	6	5	6	41	62	23	9	3	10	15
mobile tablet	0	1	2	9	13	5	8	0	3	5
online design program	0	0	0	0	2	1	0	0	0	0
online survey	0	0	0	7	1	3	0	0	3	4
open-source	1	0	0	1	4	4	2	1	0	1
qualitative data analysis software	0	3	1	3	11	12	2	1	4	3
quantitative data analysis software	1	5	6	9	25	43	6	8	7	5
real-time	0	1	2	10	11	5	3	2	3	5
short messaging service (SMS)	2	3	2	13	25	9	2	3	4	4
social media	0	0	0	3	2	3	1	1	0	2
telephone	0	0	0	3	5	2	0	0	1	0
video recording	0	0	0	0	0	0	0	0	0	0
voice messages	0	0	0	0	1	0	0	0	0	0

MIS and mobile phones enabling monitoring and implementation are the most frequently reported MERL Tech among *focused* studies. The *focused* two-variable evidence map in Table 5 shows that the majority of technology-enabled MERL activities occur in the middle of the MERL lifecycle during the collection, monitoring, and analysis phases. Use of quantitative analysis software and GIS feature less predominantly in this sample of studies; both appear to facilitate the study of other phenomena in MERL Tech-*related* studies rather than serving as the object of study in MERL Tech-*focused* studies. Diagnostic and forecasting stages show the least amount of evidence. Instead, we see MIS and mobile phones facilitating much of the data collection and monitoring and implementation activities that were reported.

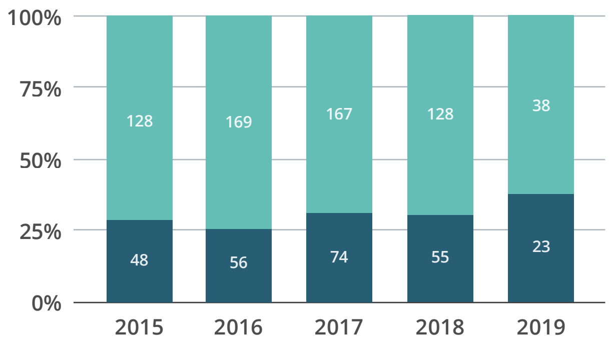
Differences in technology categories between *related* and *focused* studies is further highlighted in Figure 6, which shows the shifting category proportions. The data analysis and the implementation and monitoring activities reflected the greatest proportional differentials between *related* and *focused* studies. Data analysis, which received the largest number of codes (40 percent) among MERL Tech–*related* studies, shifted to second place among *focused* studies (19 percent). Implementation and monitoring jumped from the fifth-highest number of codes in *related* studies (8 percent) to first among *focused* studies (31 percent). While the decrease in data analysis codes among *focused* studies could be reasonably attributed to systematic coding of any studies that describe some statistical test or procedure in *related* studies, it is not entirely clear why there were so many more code occurrences for implementation and monitoring among *focused* studies.

Figure 6: Comparing Reported Activities Between MERL Tech–*Related* and –*Focused* References



Some kinds of MERL Tech are trusted more and hyped less than others. Figure 7 shows that *related* and *focused* references have remained relatively close to 70 percent and 30 percent, respectively, within the past five years. These findings indicate that ICT4D researchers and practitioners are using ICT to enable systematic inquiry about specific phenomena *related* to international aid more than they are conducting systematic inquiry about or *focused* on ICTs enabling systematic inquiry.

Figure 7: Percentages and Absolute Numbers of MERL Tech–*Related* and –*Focused* References Over Time



Given our definitions of MERL Tech–*related* and –*focused* studies in the context inclusion criterion (see page 23), we can interpret these codes and code ratios by technology types as a trust-in-tech proxy indicator among researchers and practitioners. For example, Table 6 lists GIS as the most frequently reported technology among references (n = 402). Of those reporting the use of GIS, 87 percent (n = 350) were coded as MERL Tech–*related*, and 13 percent (n = 52) as MERL Tech–*focused*. The observed proportion of use of GIS to inquire about other phenomena (87 percent) versus making GIS-enabling-MERL the object of inquiry (13 percent), raises the uncontroversial suggestion that GIS is among the most trusted technologies within the ICT4D sector for delivering MERL Tech results.

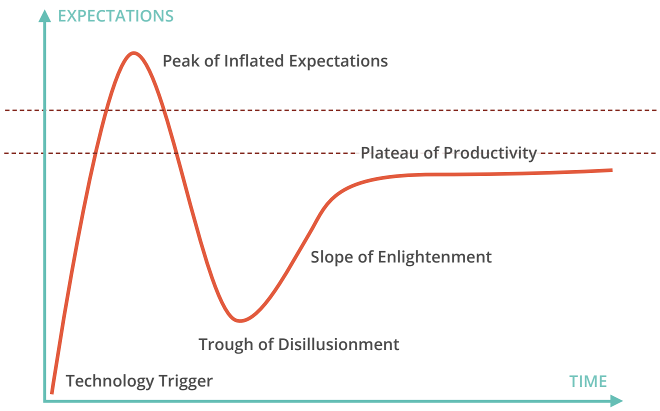
Table 6: MERL Tech–*Related* or –*Focused* Proportions as Proxies for Trust in Tech for MERL Results[13]

Technology	MERL *Tech–Related*	Count and Percentage of MERL *Tech–Focused*	Level of Trust to Enable MERL Activities
Online survey	57 (89%)	7 (11%)	High
GIS	350 (87%)	52 (13%)	High
Quantitative data analysis software	270 (85%)	47 (15%)	High
Data visualization	36 (72%)	14 (28%)	Moderate-High
SMS	23 (42%)	32 (58%)	Moderate
Mobile Phone	68 (41%)	96 (59%)	Moderate
MIS	62 (38%)	102 (62%)	Moderate-low
Dashboard	8 (35%)	15 (65%)	Moderate-low
Mobile tablet	8 (30%)	19 (70%)	Moderate-low
Real-time	6 (30%)	14 (70%)	Moderate-low

[13] Total percentages correspond to each individual technology type and are additive horizontally in the table. Cut-off scores for trust levels are as follows: low (0-20% related studies), moderate-low (21-40% related studies), moderate (41-60% related studies), moderate-high (61-80% related studies), and high (81-100% related studies).

Online surveys, GIS, and quantitative analysis software were used more to study than were studied. In contrast, dashboards, mobile tablets, and real-time technologies were studied more than they were used to study other phenomena, suggesting less trust in them. Another interpretation of these observations is the evidence of various technologies as they move along the Gartner Hype Cycle,[14] a conceptual framework for representing the maturity and adoption of technologies and applications (Figure 8). The Hype Cycle visualizes a common pattern in the maturity and adoption of technologies and applications.

Figure 8: Gartner's Hype Cycle (1995)



Technologies garnering high levels of trust, as shown in Table 6 are likely situated on Gartner's Plateau of Productivity and those with low levels of trust somewhere around the Peak of Inflated Expectations. Conversely, some technology that was less frequently explicitly reported in either category could be further along the Plateau of Productivity. For example, just 11 of 886 studies mentioned either voice or audio recording. We can safely assume many MERL practitioners use handheld or computer-based digital recording devices to collect and transcribe interview data. Reports of quantitative data analysis software would also technically fall into this category of "underreported though likely trusted," as the team systematically coded all reports of statistical testing with the qualitative data analysis software code whether or not a specific quantitative data analysis software package was reported.

Evidence of MERL Tech effectiveness is suggestive — but inconclusive. The logic of the trust-in-tech proxy is reminiscent of a key point that Raftree and Bamberger made in 2014: "ICTs are being used throughout the planning, monitoring and evaluation cycle, but there is little hard evidence of their effectiveness."[15] Where do we stand five years later? A few observations from this study might begin to test or contribute to this claim.

The fact that 70 percent of studies used ICT to enable systematic inquiry into some non-MERL Tech phenomena suggests that MERL Tech practitioners trust technologies to be effective in enhancing inquiry. Recalling that related studies use MERL Tech as the "magnifying glass" and focused studies place MERL Tech under the "magnifying glass" of inquiry, the observed 3:1 ratio of MERL Tech–related to MERL Tech–focused studies can be interpreted as a "soft" effectiveness metric. However, this conclusion is tentative given the number of assumptions needed to hold it to be true. Conversely, the fact that only 256 of the 886 studies investigate MERL Tech explicitly, with under 10 percent of them coded as quasi-experimental or experimental designs, and only 106 causal conclusions coded across 87 of those 256 documents, suggests that the claim of little evidence of effectiveness holds up today.

[14] Gartner Hype Cycle. Accessed May 14, 2020. https://www.gartner.com/en/research/methodologies/gartner-hype-cycle
[15] Raftree and Bamberger, op. cit.

Finally, we can view the total number of identified studies (n = 886) from 2015 to 2019 as a proxy for research activity on MERL Tech and compare it with a proxy for MERL Tech activity — the total number of ODA flows[16] ($691 billion) from Organisation for Economic Co-operation and Development member countries within an overlapping five-year span (2014 to 2018). The comparison shows that, for every $787 million of ODA funded, this review identified one peer-reviewed study published in English reporting use of MERL Tech.

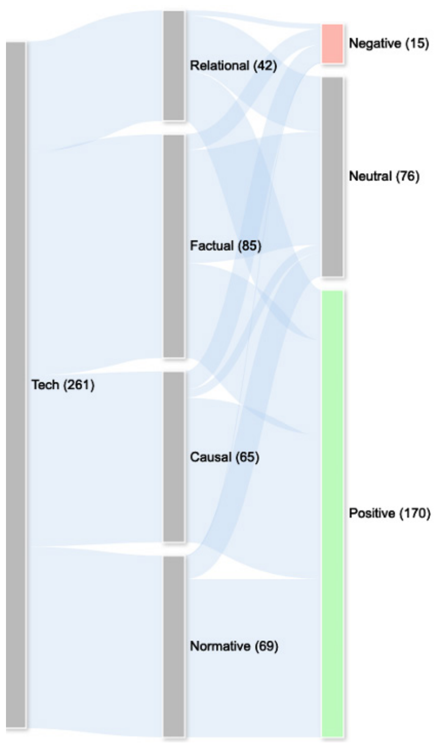
Taken together, these findings appear to support the 2014 claim about the paucity of evidence, although our study did not include evidence before 2015 or grey literature, so we cannot make direct and equivalent comparisons. However, it is safe to conclude that there is more evidence about the practice and effectiveness of ICT now than there was in 2014. Still, there is a yet-to-be-known (although hypothetically larger) amount of informal, non-academic, unindexed grey literature that should receive systematic synthesis. Given this recognition, it might be most appropriate to suggest that the identified body of evidence from peer-reviewed studies reporting the use of MERL Tech is not proportional to estimates of actual MERL Tech activities.

> **there is a yet-to-be-known (although hypothetically larger) amount of informal, non-academic, unindexed grey literature that should receive systematic synthesis.**

MERL Tech–*Focused* Studies

Synthesized conclusions and recommendations about MERL Tech have a high degree of publication bias, with only 10 percent negative in sentiment. As discussed in the Methods section, the 256 MERL Tech–*focused* studies received additional qualitative analysis. From these studies, the research team extracted 407 conclusions and 282 recommendations. These received additional first-cycle attribute and sentiment coding and second-cycle pattern and hypothesis coding. Among all conclusions and recommendations, 63 percent pertained to technology, 21 percent to MERL activities, and 15 percent to contextual factors (1 percent were not coded). The 407 conclusions comprised five sub-categories: 130 factual (32 percent), 67 relational (16 percent), 99 causal (24 percent), 97 normative (24 percent), and 14 (3 percent) non-coded data segments for conclusion type. The 282 recommendations also comprised five sub-categories: 59 tactical (21 percent), 70 strategic (25 percent), 42 policy-oriented (15 percent), 79 further-research (28 percent), and 32 (11 percent) non-coded data segments for recommendation type. At a high level, the conclusions and recommendations offered in these studies were significantly biased toward positive and neutral sentiments, with only 10 percent registering as negative (see Figure 9).

Figure 9: Sentiment of Conclusions About Technology



[16] Using the more conservative pre-grant equivalent "cash basis" or "flow basis" calculation from "Net ODA," Organisation for Economic Co-operation and Development. Accessed March, 18, 2020. https://data.oecd.org/oda/net-oda.htm

Qualitative Synthesis

Reporting data that describes the nature of evidence at a high level, as in an evidence map, is the typical terminus of scoping reviews. However, the our team determined early in the review to attempt a full-text data extraction and qualitative synthesis for MERL Tech–focused studies. We made this decision in part to explore the extent to which it is possible to answer specific questions from the synthesis of disparate conclusions and to increase the relevance and utility of study findings for practitioners, not just researchers. Descriptive pattern coding enabled us to group and identify thematic data sets in order to answer select, high- demand questions identified by practitioners at the 2018 MERL Tech conference in Washington, D.C., during a pairwise comparison voting exercise. In that event, practitioners wrote their most important curiosities or questions on cards that were allotted a certain number of votes in head-to-head match-ups among participants.

Barriers to effective MERL Tech implementation. A highly prioritized question from the pairwise comparison voting exercise was, "What are the barriers to effective MERL Tech implementation?" Filtering the conclusions and recommendations coded as negative sentiments, we identified a possible data corpus to begin to answer this question.

> ### Research Question
> **"What are the barriers to effective MERL Tech implementation?"**

A total of 45 unique barrier codes were generated and applied to 70 negative sentiment data segments. The five most frequently occurring code categories (applied to four or more data segments) were:

1. Capacity gap (11 segments)
2. Poor data quality (7 segments)
3. Lack of systems integration (5 segments)
4. Lack of collaboration (4 segments)
5. Lack of incentivization for MERL Tech use (4 segments)

Three of these barriers are associated directly or indirectly with the five new challenges Raftree and Bamberger identified in 2014. First they flagged low institutional capacity in organizations that lack the budget to train and integrate ICTs into their operations. Many capacity gap codes that we identified corresponded to diagnoses or calls for training and capacity development. Second, overreliance on digital tools was suggested in the 2014 report as leading to loss of data quality control measures and overcollection of data, which we coded as "poor data quality" in this sample. Third, common challenges related to incentivization for MERL Tech tool usage may follow from what Raftree and Bamberger call tool or technology-driven M&E processes that are mandated or result from the adaptation of MERL activities and plans into ICT tools, rather than tools being selected to meet MERL needs. Each of the two remaining challenges, selectivity bias and loss of privacy, was coded only once. Other real-world and methodological challenges flagged in the 2014 report and coded infrequently from this dataset included data collection costs, data availability, cost of quality control of data, broader contextual factors affecting program outcomes, and the complexity of programs necessitating specialized evaluation applications. These findings underscore that most challenges foreseen in 2014 were still relevant in the five years that followed. Other barriers coded from this sample but not highlighted in the 2014 report were multiple conflicting MERL systems, conflicting stakeholder expectations, and issues in scaling MERL Tech from pilot programs.

MIS and decision facilitation. Raftree and Bamberger reported, "Data are also being aggregated more quickly and shared at various levels to improve participation in the planning process and to make better decisions."[17] The finding that MIS was the third-most frequently reported technology, along with mobile phones, suggests that this claim remains true five years later. Although the data collected from the composition of reported studies could not speak to the quality of data aggregation and management, we did investigate the priority question identified at the MERL Tech 2018 conference in Washington, D.C., "What elements of digital information systems facilitate regular use of data in decision-making?" We combined the extracted texts coded as MIS and decision facilitation to create a data corpus of 50 text segments of conclusions and recommendations across 16 studies.

> ### Research Question
> **"What elements of digital information systems facilitate regular use of data in decision-making?"**

Overall, the corpus was mostly neutral in sentiment and pertained to technology as opposed to MERL or contextual factors. Conclusions were mostly descriptive in nature, and recommendations were predominantly strategic. From this corpus, we determined that improvements in scalability, personnel training, and sustainability were overarching elements of information systems that facilitated use of data in decision-making. These elements coded across the studies in this corpus were also reported from a study included in this review, a systematic review of 14 studies for decision-making for health programs in low-income settings.[18] Those studies identified three categories of challenges that affected decision-making: 1) the availability and quality of health facility data, 2) human dynamics, and 3) financial constraints.

Real-time data and data sources. Real-time technology has been mentioned above. However, the combined codes real-time and dashboard generated 50 conclusions and 36 recommendations across 28 references and provided a sufficient corpus to attempt to answer the priority question "How is MERL using real-time data and what are [...] data sources?"

> ### Research Question
> **"How is MERL using real-time data and what are [...] data sources?"**

From the qualitative synthesis, we observed that use of real-time and dashboard technologies in included studies were most frequently applied in the health and agriculture sectors across sub-Saharan Africa, often in tracking vaccination and medical teams and monitoring the health of crop production. The sources for this data was predominantly GIS records from farmers and community health workers, aggregated across provinces or regions. The combined data set clarified that not all dashboards and real-time data are created equally. Factors such as general data quality, the relative frequency of real-time collection, and technical capacity in using dashboards all contributed to reported successes and failures with these MERL technologies.

[17] Raftree and Bamberger, op. cit.

[18] Deepthi Wickremasinghe, et al. (2016). "District decision-making for health in low-income settings: a systematic literature review" Health Policy and Planning, 31, ii12–ii24.

Cost-efficiency of MERL Tech. Of the 256 studies related to MERL Tech, 43 (16 percent) mentioned improved cost-efficiency as a research outcome. This finding relates most to two other, related high-priority questions, "To what extent is MERL-enabled technology making data collection more cost-efficient/cost-effective? Can we generalize how much money it is saving compared to paper?"

Research Question

"To what extent is MERL-enabled technology making data collection more cost-efficient/cost-effective? Can we generalize how much money it is saving compared to paper?"

The comparison of digital and paper-based MERL activities was a common theme among the coded ranked questions from the 2018 conference in Washington, D.C., suggesting that, despite anecdotal reports of cost savings from the 2014 report and the widely held view that technology saves time and money, many practitioners still demand empirical evidence to support these notions and to advocate for MERL Tech among decision-makers. The included studies reported some instances of cost savings, although there were also reports of unforeseen costs (similar to hidden costs Raftree and Bamberger allude to, such as those for training, technical support, or software licensing fees). Ultimately, the studies that this review identified reveal insufficient data to satisfy the demand for cost-effectiveness and efficiency data about MERL Tech.

Calls for further research from the evidence base. Among the 256 references that received a full-text data extraction, we analyzed 70 references and 79 text segments containing the words further research for any trends significant for the MERL Tech community. Some calls for further research were overly specific, some were too general, and some lacked adequate context to be meaningful. After filtering this noise, we identified two categories of codes: 1) calls for improved technical quality in future studies and 2) investigations into the impact of technology on development outcomes. Technical quality focused mainly on different or more representative sampling methods, as well as improvements for data quality and validation studies. Calls to explore the impact of technology ranged from broad calls for evaluations to exploring the influence, effectiveness, efficiency, systemic changes, and exploration of unintended consequences.

Limitations

While many decisions about the scope of this study may have contributed to its limitations, one of the most significant limitations is reporting bias. There are many types of reporting biases,[19] and a number are relevant in this study. First, publication bias is certainly a factor both generally and in this sample. Close to 90 percent of conclusions and recommendations in this review were either positive or neutral. It is likely that many studies or findings were not published due to their nature and direction. Also, there is likely an unknown time lag bias. While there are some assurances of study quality in indexed, peer-reviewed articles, there are often significant levels of latency between data collection and publication. This is less of a factor for a review in traditional MERL Tech, where the focus is not on emerging applications, such as this one. Beyond these biases, there is a significant limitation in omitting a systematic grey literature or unindexed search strategy. This is especially relevant where much of what practitioners consult as evidence is likely shared in sources that are not peer-reviewed, thus vectoring the studies toward a research audience rather a practitioner audience and perhaps limiting the relevance of conclusions for practitioners. Future reviews of MERL Tech should build systematic grey literature search strategies into their sampling methodologies.

We also note a limitation pertaining to data extraction and analysis. Due to the large number of relevant studies, and the researchers' limited time, it was not feasible to conduct full-text reviews of 886 studies. Accordingly, we coded all attributional data items where studies were the unit of analysis — as opposed to conclusions and recommendations— from reference abstracts, conserving time in retrieval and full-text review. Consequently, for studies with multiple codes of technology and MERL activities, it is not always clear which technology is enabling which MERL activity. In practice, multiple technologies often enable multiple MERL activities within the same project and lifecycle; this is difficult to parse in abstracts. This limitation informed our decision to keep the two variables in the *related* and *focused* evidence map focused on technology and sector, and to use the *focused* evidence map only to compare technology and MERL activity. Full-text data extraction might have provided marginal gains for coding these attributes but would have been inefficient given the time required to retrieve reference copies and extract data.

Finally, the breadth of studies, the heterogeneity of conclusions and results, and the often decontextualized extracted data segments from the full-text review may have limited our ability to draw meaningful conclusions or answer additional, specific questions from aggregated conclusions and recommendations. Scoping reviews do not typically attempt to synthesize findings, unlike qualitative reviews, and it was challenging to build data sets that would lend themselves to substantive qualitative analysis. We suggest caution if future reviewers attempt qualitative syntheses of findings, conclusions, or recommendations across a broad range of subject matter with limited resources.

[19] "Reporting Bias," Cochrane Methods, accessed May, 14, 2020. https://methods.cochrane.org/bias/reporting-biases

Conclusions and Next Steps

Over five years after Raftree and Bamberger reported that "ICTs are being used throughout the planning, monitoring and evaluation cycle, but there is little hard evidence of their effectiveness,"[20] little seems to have changed — at least within the scholarly evidence base. Of the 886 peer-reviewed studies reporting use of MERL Tech in international development that were published in the interim, fewer than one in three focused on investigating MERL technologies themselves. Those 256 studies made just 106 conclusions that could be considered causal claims, and less than 10 percent of those studies were experimental (n = 24) or quasi-experimental (n = 5).

Ironically, given the charge of MERL Tech practitioners to help their practice-based fields use technology to develop their own evidence bases for what does and doesn't work, the state of scholarly evidence pertaining to the efficacy of MERL Tech is underdeveloped. The dissonance of a paucity of formal evidence about the effectiveness of MERL Tech — a field focused on developing credible evidence for decision-making — is heightened when we consider that MERL Tech applications are seen as constituting the majority of digital development.[21] Further, if ODA flows used as a crude proxy for digital development and MERL Tech activity, and the number of total identified studies for roughly the same time used as a proxy for and MERL Tech research are both warranted, then it is jarring that every $787 million in ODA funding equates to just one English peer-reviewed study focused or related to MERL Tech. This suggests that the formal, scholarly, indexed evidence base of peer-reviewed studies significantly underrepresents the universe of learning and evidence about MERL Tech.

> **...given the charge of MERL Tech practitioners to help their practice-based fields use technology to develop their own evidence bases for what does and doesn't work, the state of scholarly evidence pertaining to the efficacy of MERL Tech is underdeveloped. The dissonance of a paucity of formal evidence about the effectiveness of MERL Tech — a field focused on developing credible evidence for decision-making — is heightened when we consider that MERL Tech applications are seen as constituting the majority of digital development.**

[20] Raftree and Bamberger, op. cit.

[21] Digital Impact Alliance, DIAL Baseline Ecosystem Study (2018). Accessed May 20, 2020. https://digitalimpactalliance.org/research/digital-impact-alliance-2018-baseline-ecosystem-study/

Despite the evidence gaps, we can derive a description of what existing MERL Tech practice actually looks like from this scoping review. MERL Tech practitioners use a range of technologies to enable a multitude of MERL activities across sectoral and geographic contexts, and the findings of this review confirm what many MERL Tech practitioners see in the field. We found the three most frequently reported technologies, MERL activities, sectors, and regions to be MIS, quantitative analysis software, and mobile phones and GIS (tied for third). These enabled implementation and monitoring, data analysis, and data collection across the health, environment, and energy sectors within the sub-Saharan Africa, East Asia and Pacific, Europe and Central Asia regions. Among these, GIS, data analysis, the health sector, and sub-Saharan Africa were the most frequently reported technology, MERL activity, sector, and region — although not necessarily in that combination. While the health sector evidence share could be viewed as disproportional to other sectors in this study (39 percent) and relative to proportions of health sectoral ODA flows for an equivalent time frame (11 percent), this health bias is less than the 65 percent of total experimental studies found in the Brown and Skelly digital development evidence mapping study.[22]

Despite the dearth of formal evidence about MERL Tech effectiveness, comparing the reported technology frequencies between MERL Tech–*related* and MERL Tech–*focused* studies can serve as a proxy for MERL Tech effectiveness. Technologies such as online surveys, GIS, and quantitative data analysis software are trusted most among MERL Tech practitioners and are often used to investigate other MERL technologies that practitioners view more skeptically. These include dashboards, mobile tablets, and real-time technology. This is not to say these latter technologies are not being used, or do not deliver results — only that, relative to the other technologies in the high to moderate trust categories, they were shown as objects of investigation in MERL Tech–*focused* studies more than they facilitated investigations into other phenomena as in MERL Tech–*related* studies. However, consideration of these proxies should be tempered with recognition of publication biases discussed above.

Traditional MERL Tech — tech-enabled activities for systematic inquiry — has weathered the peaks and troughs of the Gartner Hype Cycle and are close to or already on the Plateau of Productivity. They are integral for many practitioners in international development and the social sector broadly, especially those tasked with coordinating evaluation, research, and learning. These combinations of technology and MERL activities will likely continue to provide value for practitioners and communities looking to work more effectively. Yet the field needs to apply the reflective imperative it espouses to development practitioners inwardly, possibly using trusted MERL Tech applications to expand the nascent evidence base for what works and what doesn't within the MERL Tech space.

> **Traditional MERL Tech — tech-enabled activities for systematic inquiry — has weathered the peaks and troughs of the Gartner Hype Cycle and are close to or already on the Plateau of Productivity.**

[22] Brown and Skelly, op. cit.

What is next?

Our findings and conclusions support seven recommendations for researchers, practitioners, and policymakers.

1. Future knowledge syntheses should include, if not focus solely on, grey literature.

We highly recommend a comparable knowledge synthesis of non-indexed, secondary source grey literature for MERL Tech researchers. We anticipate that much of the knowledge translation within the MERL Tech space occurs outside the paywalls of the academy and in the reports and white papers of frontline organizations (and practitioners' blogs). The description of MERL Tech practice is likely to be more comprehensive, if not more accurate, and the aggregated conclusions and recommendations likely to be of greater relevance to practitioners.

2. Focus the scope of future knowledge syntheses.

The conceptually broad transdisciplinary professional space of MERL Tech invites researchers to use caution in conducting any form of mixed-method review or hybrid qualitative synthesis stages for broad scoping reviews. This study faced challenges in aggregating heterogeneous conclusions and recommendations from studies focused on MERL Tech.

3. Take small steps to reduce publication bias.

For any MERL Tech researchers looking to publish in peer-reviewed journals on aspects of MERL Tech practice and effectiveness, we recommend pre-publishing study protocols to mitigate publication bias in the nascent scholarly evidence base.

4. Pursue practice-based evidence.

For practitioners, we recommend integrating and publishing more practitioner-based research[23] on MERL Tech, especially for submission to peer-reviewed journals. We recognize that the distinction between research and practice for a profession based on technology-enabled systematic inquiry may be less helpful. Nonetheless, there appears to be an opportunity for M&E professionals who use tech to complement existing systematic inquiry within their organizations or their practices to take small measures to investigate questions focused on issues and aspects of MERL Tech and to share those with the broader research community. While the incentives for practitioners to invest time and energy in the peer-review process are low, conducting this type of research would buttress the formal evidence base.

5. Ask questions of cost efficiency and effectiveness.

Practitioners should look for opportunities to weave cost analysis into the various uses of MERL Tech in their work. Questions of cost efficiency and effectiveness were high priorities for the MERL Tech community in 2018 and likely remain salient — and not just for practitioners. Many decision-makers who fund and manage further research or digital development projects want data about how investing in MERL Tech will translate to better results. However, cost analysis of MERL Tech remains scarce.

[23] Christina Christie (2012). "Advancing Empirical Scholarship To Further Develop Evaluation Theory and Practice," *The Canadian Journal of Program Evaluation*, Vol. 26, No. 1, 1–18.

6. Invest in building the evidence base for MERL Tech.

Especially in forward-operating digital development organizations, practitioners and researchers demand more high-quality studies on the effectiveness of MERL Tech. Policymakers should consider how marshaling resources at an organization to ask these questions would help both the organization and the field at large to understand what does and doesn't work in MERL Tech.

7. Foster capacity to improve MERL Tech data quality.

Reports from this review suggest capacity building limitations and poor data quality are significant barriers to realizing the full potential of MERL Tech. Policymakers should consider how current or future investments in MERL Tech may be underleveraged by inadequate training and up-skilling in MERL Tech use. They should use various evaluation standards and data quality frameworks to match MERL Tech investments with end-user data quality needs.

Methods[24]

A team of 10 researchers conducted the scoping review from mid-2019 to early 2020. The research protocol was informed by participatory data collection exercises from the 2018 MERL Tech conference in Washington, D.C. The team screened 3,054 study titles and abstracts, coded 886 abstracts, and extracted 256 full texts to describe the state of reported MERL Tech practice in development assistance in English-language peer-reviewed studies from 2015 to 2019. The team analyzed the extracted data quantitatively and qualitatively and reported the results in multiple tables, figures, evidence maps, and narratives.

Inclusion Criteria

As listed in Table 1 (page 4) and discussed below, we used the traditional MERL Tech inclusion criteria to operationalize the research question and guide the searching, screening, and extraction phases of this scoping review.

Population

The review considered studies and references that included any and all populations within lower to middle-income countries. We used World Bank designations from a recent evidence mapping study of ICT4D[25] to generate the list of low- to middle-income countries.

[24] For readability, the authors abbreviated this methods discussion from a methodologically detailed forthcoming scoping review paper that follows the Preferred Reporting Items for Systematic Reviews and Meta-Analysis extension for Scoping Reviews. The scoping review protocol and data will be published on Open Science Framework. Until then, interested parties can request a copy by emailing the corresponding author, Zach Tilton, at zachary.d.tilton@wmich.edu.

[25] Brown and Skelly, op. cit.

Concept

MERL Tech is a shorthand for technology-enabled monitoring, evaluation, research, and learning. Specifically, MERL Tech refers to the combination of digital information and communication technologies and specific activities that facilitate systematic inquiry to support and improve interventions for social amelioration. This scoping review focuses on what we term traditional MERL Tech to distinguish from emerging approaches and big data for MERL Tech, addressed in companion papers.

The two components of traditional MERL Tech are 1) the specific information and communication technologies that do not constitute emerging or big data technology and 2) the myriad activities within the M&E lifecycle that the technologies enable. Analyzing five years of MERL Tech conference data, we identified the 25 most frequently occurring technology categories (see Table 4, page 9). The 10 categories of MERL activities modified from Raftree and Bamberger's M&E Lifecycle[26] are diagnosis; design; planning; data collection; implementation and monitoring; data analysis; evaluation; reporting, sharing, and learning; prediction and forecasting; and decision facilitation.

Context

We used Skelly and Brown's evidence mapping study on ICT4D[27] to derive a modified list of 10 sub-sectors of international development (see Figure 4, page 9). We considered all study designs and publications types eligible for inclusion so long as they were published in an English-language peer-reviewed journal from January 2015 through May 2019. Finally, to be included in the review, references needed to be *related to* or *focused on* traditional MERL Tech. We deemed a reference *related to* traditional MERL Tech if technologies enabling MERL activities were reported as being in service to answering research questions or reports on interventions within relevant sectors, although not in and of themselves focusing on MERL Tech. We considered a reference *focused on* traditional MERL Tech if the research questions or reported activities specifically about MERL Tech were the main object of the study or report. This distinction is an important analytical framework used throughout this paper. The research team validated the distinction through multiple rounds of piloting screening and extraction forms.

Procedures

Information sources and search strategy

The research team searched seven bibliographic databases with the assistance of a research librarian using a thematic Boolean syntax of keywords corresponding to the elements of our review question: technology, evaluation, sector, and geographic context.[28]

[26] Raftree and Bamberger, op. cit.
[27] Brown and Skelly, op. cit.
[28] Search strategy and sources are available upon request from the corresponding author at zachary.d.tilton@wmich.edu.

Selection of sources of evidence

We imported sources from each database directly via Research Information System files into the web-based systematic review software DistillerSR[29] for reference management, deduplication, screening, and data extraction. Table 7 is an overview of the screening and extraction stages.

Table 7: MERL Tech State of the Field Scoping Review Levels

Level 1: Title and Abstract Screening	Retrieved reference titles and abstracts were dual-screened for relevance
Level 2: Abstract Data Extraction and Filtering	Data items were extracted from the abstracts of references included from Level 1 screening; a conditional workflow was used to advance MERL Tech–focused studies to Level 3
Level 3: Full-text Data Extraction	Full texts of *focused* studies were retrieved and reviewed, and conclusions and recommendations were extracted

Data items and charting process

Screening and extraction forms were developed for all three levels of the review. Levels 2 and 3 had extraction forms for first-cycle attribute and sentiment coding. Level 3 had a codebook and additional form developed for second-cycle pattern and hypothesis coding. The data items extracted pertained to aspects of the research protocol, study type, geography, sector, technology, MERL activity, conclusion and recommendation content, type, and sentiment.

Synthesis of findings

Synthesis of findings. Data extracted from abstracts and full texts underwent an initial round of inductive and deductive attribute coding in the qualitative data analysis software, MAXQDA.[30] Coded data items were charted on evidence maps and in tables for high-level quantitative synthesis. Data extracted from the conclusions and recommendations of full texts received an additional round of pattern and hypothesis coding for qualitative synthesis.[31]

[29] Evidence Partners Incorporated, Ottawa, Canada.
[30] VERBI Software, MAXQDA 2020, software (2019). www.maxqda.com
[31] Johnny Saldaña (2016). The Coding Manual for Qualitative Researchers, Sage.