Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning Across Multiple Devices

Frederik Brudy  
University College London  
London, United Kingdom  
f.brudy@cs.ucl.ac.uk

Christian Holz  
Microsoft Research  
Redmond, WA, USA  
cholz@microsoft.com

Roman Rädle  
Aarhus University  
Aarhus, Denmark  
roman.raedle@cs.au.dk

Chi-Jui Wu  
Lancaster University  
Lancaster, United Kingdom  
c.wu10@lancaster.ac.uk

Steven Houben  
Lancaster University  
Lancaster, United Kingdom  
s.houben@lancaster.ac.uk

Clemens Nylandsted Klokmose  
Aarhus University  
Aarhus, Denmark  
clemens@cavi.au.dk

Nicolai Marquardt  
University College London  
London, United Kingdom  
n.marquardt@ucl.ac.uk

ABSTRACT
Designing interfaces or applications that move beyond the bounds of a single device screen enables new ways to engage with digital content. Research addressing the opportunities and challenges of interactions with multiple devices in concert is of continued focus in HCI research. To inform the future research agenda of this field, we contribute an analysis and taxonomy of a corpus of 510 papers in the cross-device computing domain. For both new and experienced researchers in the field we provide: an overview, historic trends and unified terminology of cross-device research; discussion of major and under-explored application areas; mapping of enabling technologies; synthesis of key interaction techniques spanning across multiple devices; and review of common evaluation strategies. We close with a discussion of open issues. Our taxonomy aims to create a unified terminology and common understanding for researchers in order to facilitate and stimulate future cross-device research.

CCS CONCEPTS
• Human-centered computing → HCI theory, concepts and models; Ubiquitous and mobile computing systems and tools; Interaction paradigms; Ubiquitous and mobile computing design and evaluation methods.

KEYWORDS
Cross-device interaction; cross-device computing; survey; multi-device; taxonomy; cross-surface; distributed user interfaces

ACM Reference Format:

1 INTRODUCTION
The way we interact with computers has fundamentally changed in the last 30 years. Never in the history of computing did we have access to so many powerful computing devices with different forms factors, affordances, functionalities, and technical capabilities [250]. Since Weiser’s seminal paper [355] on Ubiquitous Computing, we have seen an explosion of research into these device form factors to envision new interaction paradigms that transcend the individual device and user. From early smart-space to multi-display, distributed surface, and multi-device; to cross-display, trans-surface, and cross-device interaction: research involving an
understanding of people’s interactions with various device configurations and ecologies is at the forefront of modern Human-Computer Interaction (HCI) research [17, 26, 83, 107, 168, 273, 285]. However, this explosion of research topics has led to disconnected terminology, techniques, and systems creating a fragmented research landscape. In this paper, we unify this fragmented research under the umbrella of cross-device computing. To inform future research, we analysed 510 papers in the cross-device interaction domain to synthesise the state of the field. Based on our analysis of terminology, applications, systems, techniques, and evaluations, we present a unified taxonomy and overview of the state of the art. We discuss open issues and challenges in the cross-device domain, and outline an agenda for future research in the space of cross-device interaction.

The goal of this taxonomy is to contribute a general working- and discussion-ground for the wider area of HCI. We provide newcomers to cross-device research with a starting point, but also create a common ground for discussions with current researchers in the field and to allow for a time of reflection within the cross-device community. This work directly incorporates and extends earlier taxonomies of multi-device research. Our taxonomy spans across the research areas of distributed user interfaces [76, 86], second-screen and multi-device research [229], and mobile multi-device ecosystems [111]. We build upon other related taxonomies (all with their own specialised focus): for example, covering the scale of multi-display systems [330], display switching for multi-displays [275], interaction techniques for spontaneous device association [62], or characteristics of devices’ ownership, access, and distance [292]. Our goal is to unify the terminology and strategies in the cross-device space, extending the scope of earlier taxonomies. In particular, we focus on themes around interactions within cross-device settings. We cover relevant technical aspects (e.g. tracking systems and evaluation strategies for cross-device interactions), but other engineering aspects are outside of the scope of this taxonomy, for example cross-device architectures or development frameworks (an overview is available in [149]).

In summary, our paper contributes: (1) a unified taxonomy and terminology for cross-device research; (2) the mapping and reflective discussion of the current design space (across application domains, tracking systems, interaction techniques, and evaluation strategies); and (3) the identification of opportunities and challenges informing a future research agenda of cross-device computing. Our data and the complete tagging tool are released as open-source 1, and we invite the research community to contribute to the collection and to extend the taxonomy.

1https://github.com/frederikbrudy/cross-device-taxonomy

2 METHODOLOGY

Creating Corpus of Relevant Publications: In order to create the collection of papers for this analysis, we conducted a systematic search in the ACM Digital Library (May 2018). Our search terms included all possible combinations of cross- and multi- with each of device, surface, monitor and display; and distributed user interfaces as well as acronyms. By looking at references within our corpus as well as using our own expertise of the cross-device research domain, we identified an additional 48 articles that were missing from the search results (which is a common strategy for survey and taxonomy papers, e.g. [111, 184]). After merging duplicates, our selection comprised a set of 738 papers.

Filtering and Inclusion Criteria: Papers had to be concerned with interaction tasks or techniques, tracking technology for people and/or devices, or technologies involving multiple devices. We excluded papers without a contribution (e.g. proceedings front matter, keynotes, workshop calls for participation) or short contributions with a full paper follow-up. Our resulting corpus included 510 tagged papers in total, which are incorporated in our taxonomy.

Tagging: We iteratively developed the tagging categories for our taxonomy (through both top-down specification of categories and bottom-up trial-tagging). Final tagging categories included: contributions, application areas, interaction techniques, deployment details, evaluation techniques, definitions, technological aspects, and several further fields. We tagged all 510 papers with a custom-built web-based paper-tagging system. The front-end includes forms for tagging and annotation, pre-set tags, free text fields, and auto completion. During tagging we frequently discussed emerging patterns, revisited tagging schemes, and iterated on previously reviewed papers. The final tagging dataset was cleaned up (e.g. filling missing tags, correcting misspellings, and merging identical tags) and exported as JSON for analysis.

Analysis: We primarily analysed three data collections: (1) the tagging dataset with all data fields, (2) the complete set of all document PDFs, and (3) extracted text from all PDFs excluding references. The analysis was done iteratively by all authors individually, in pairs, and in group discussions. We used open coding for the analysis of our tagging dataset as well as axial coding for the identification of higher level themes. In addition, we used Tableau, R, and Excel to visualise the development of key topics over time and to support uncovering trends and patterns. We used these visualisations as pointers to where we should dive deeper into our dataset, as well as into the papers themselves. Frequent full-text searches complemented these explorations. Emerging patterns, findings, and key categories for the papers were discussed with all co-authors and refined over time.
Three Areas of Cross-Device Research

We begin by unpacking a brief history of the earliest work in the area, and by highlighting the main trends over time. We distill this work into three key areas of cross-device work: (1) early work on multi-monitor workstations, (2) emergence of multi-display/surface environments, and (3) the increasingly mobile, ad hoc cross-device use.

**Area 1: Multi-Monitor Workstations.** One of the earliest visions for a personal workstation, the Memex [35], featured a setup consisting of multiple monitors, two for display and one for pen input. Early work included in our survey covers multi-screen systems [201, 202] from as early as 1981, where the effect of having four monitors at a workstation was studied. Wellner [356] took this notion further with DigitalDesk. Similar – albeit later – work includes Grudin’s [112] work on peripheral awareness in multi-monitor use, and the advantages of spreading information across connected, but distinct, output screens. Prior work has extensively studied multi-screen setups [82, 153, 225, 276, 290].

**Area 2: Multi-Device Environments and Spaces.** Weiser’s “The Computer for the 21st Century” [355] inspired research on computing that went beyond the single user at a single computer. Notably, Rekimoto’s work from the late 90s [278, 281] explored interaction techniques that crossed device boundaries. After Fitzmaurice introduced the notion of spatially aware palmtop devices a few years earlier [90]. Around the same time, research on large interactive spaces started to appear, with seminal work like i-LAND [324]. Different from fixed spaces and ecologies, the third area of cross-device work focuses on mobile and flexible ad hoc cross-device setups. Enabled by ubiquitous availability of smartphones and tablets, this research strives towards individual or collaborative applications spanning across portable devices, providing a digital information space to support the task at hand [11, 31, 52–54, 109, 111, 128, 129, 204, 206, 270, 272, 280, 362]. Ad hoc portable setups lead to new challenges for technology, tracking, and field studies, as we explain in more depth shortly.

Towards Unified Terminology

When sampling the research space, it became clear that there is a need to unify the terminology used to describe ongoing research. Not entirely unexpectedly for a large research field, in some cases the different introduced terms describe

---

2https://github.com/frederikbrudy/cross-device-taxonomy
identical research areas. At other times, it is the other way round, when identical terms in fact refer to different areas of research. We untangle the diverse terminology of cross-device work and map terms out into a single ontology in Figure 1. The main goal of this synthesis is to provide better guidance about scope and specialisations of research within our field. The unified ontology is formed of two parts of key terms used to refer to cross-device literature:

- The top part of Figure 1 categorises key terms of cross-device sub disciplines. The nested categories begin at the bottom with dual-display and multi-monitor work (1a), extend to work with multiple mobile devices, tablets or slates (1b), further to the category of cross/multi-display (1c) and cross/multi/trans-surfaces (1d), and finally to cross/multi-device and distributed covering the broadest scope (1e). The nested structure in the figure indicates focus areas associated with each term as well as the often included device form factors. It is important to note that a large subset of the terms are used interchangeably.

- The bottom part of Figure 1 includes a list of terms indicating focus areas of research projects (1f): interactions and collaboration; interfaces; applications or systems; platform or middleware; environments or ecologies; and computing. We annotated these terms with the most common research focus of papers using each term.

### How to use the ontology

For anyone entering the field it clarifies terminology used, while at the same time highlights that some terms are used interchangeably. For current researchers framing their work, it can help to identify the best terms to describe the work. And within the cross-device community, the ontology can provoke reflection on the use and appropriateness of terminology.

### 4 TAXONOMY OF KEY CHARACTERISTICS

In this section, we delve deeper into our derived taxonomy of key characteristics of the cross-device design space. In particular, our taxonomy in Figure 2 is a fine-grained synthesis of the disparate but interwoven threads of the research field. We explain the six key dimensions and discuss how to use this taxonomy as a lens to look at existing - or inform new - research.

**Dimension 1: Temporal.** Cross-device work can be classified as either synchronous (when interactions happen at the same time) or asynchronous (with a sequential flow of interactions across devices). The large majority of the work in our survey falls into the former category.

Figure 2: Taxonomy of cross-device design space dimensions: temporal, configuration, relationship, scale, dynamics and space.
Dimension 2: Configuration. This dimension classifies the actual setup of the cross-device system as well as its use of input and output modalities. The main categories within synchronous use are mirrored and distributed user interfaces. Most active research is done within the distributed UI category, investigating the spatial and logical distribution of interfaces. The asynchronous work is divided across two categories: interfaces that allow migration across devices, and cross-platform research to make applications run consistently across diverse operating systems. Related taxonomies align with this dimension, in particular Elmqvist’s taxonomy of distributed UIs [86] and Rashid’s focus on multi-device attention switching [276].

Dimension 3: Relationship. Research addresses different people-to-device relationships. While one person interacting with a single device (1..1) is usually not part of cross-device work, one person interacting with two or more devices (1..m) covers work on cross-device workstations. Collaborative settings fall mostly in two categories: group activities where each person primarily uses a single device (1..1) and collaborative settings with n-people and m-devices (n..m). Examples of the last two categories relate closely to research and studies done in the CSCW community.

Dimension 4: Scale. Interactions can vary across the dimension of scale: from near, to personal, social, and public rooms or buildings. Edward Hall’s proxemics is a commonly used anthropological lens for the scale of interactions [115], which was later operationalised for cross-device work as proxemic interactions [12, 107, 205]. Scale dimension relates closely to Terrenghi’s taxonomy of display ecosystems across scale [330] and the progression of Weiser’s Tab/Pad/Yard computing [355].

Dimension 5: Dynamics. Dynamics can vary between setups, and the categories of ad hoc/mobile, semi-fixed, and fixed spaces closely relate to the phases of cross-device research we introduced earlier. Fixed spaces often include larger-scale wall displays and tablets, while semi-fixed spaces allow a certain degree of reconfigurability, and ad hoc/mobile spaces focus on portable devices, allowing dynamic changes and re-configuration.

Dimension 6: Space. The last dimension differentiates between co-located and remote interactions (and corresponds to Johnson’s CSCW matrix [162]). The large majority of cross-device work covers co-located scenarios, but few examples address the challenges of providing cross-device interactions across remotely distributed locations [13, 174].

Note: research projects do not necessarily have to fall into one single category. Instead, it is common that projects address different areas across this design space. Furthermore, research might follow lateral transitions, where the scope of the research shifts over time across any of these dimensions.

How to use this taxonomy: This taxonomy has different functions helping researchers: (1) it compresses the large research field and synthesises seminal work to ease entry into an unfamiliar research domain, (2) the specific dimensions (and subcategories) can support framing and scoping of new and ongoing research, and (3) it allows discussing research in the context of the major related categories within the six dimensions. This taxonomy works in conjunction with the analysis across applications (starting in the next section), technology, interaction techniques, and evaluation strategies.

### Table 1: Cross-device application domains: Nine application categories (and sub categories) with examples use cases.

<table>
<thead>
<tr>
<th>Knowledge work (62)</th>
<th>Games and Installations (31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>→ Project management: task management [265], productivity [266], project management [17], ad hoc sharing of resources [252], job scheduling [263]</td>
<td>→ Museum experiences: museum activities and simulations [154], museum guide and artwork [93], spatial music experience [203]</td>
</tr>
<tr>
<td>→ Presentations: presentation software [279], projection-screen presenter [239], multi-display environment [201]</td>
<td>→ Playing and dancing: interactive dance class [80], interactive experience design [205], arcade games [147], Nike’s Walls games [189], lower-division game [147], multi-device AR game [232], VR skydiving experience [125]</td>
</tr>
<tr>
<td>→ Searching: search [250], finding information [96]</td>
<td>→ Guides: city guide [201], visitor guide [119], conference guide [204], exploring neighborhood [179]</td>
</tr>
<tr>
<td>→ Browsing: cross-screen browsing [164]</td>
<td>→ Businesses: retail [208], restaurant menu [98]</td>
</tr>
<tr>
<td>→ Application examples: note-taking[44], scientific work [140], activity-aware computing [140], police analysts station [4]</td>
<td>→ Collaboration (21)</td>
</tr>
<tr>
<td>→ Mobile Computing (61)</td>
<td>→ Mediating discussions: brainstorming and discussions [200], meeting room [105], interactive brainstorming [203], voting system [164]</td>
</tr>
<tr>
<td>→ Creativity: drawing [5], photo editing [219], comic strips and collage [24], scrapbooks [236]</td>
<td>→ Content creation: creative collaborative work [354], collaborative synthetic闵式-design [149], collaborative writing [174]</td>
</tr>
<tr>
<td>→ Media consumption: mixed-reality TV [8], TV access [269], active-reading [52]</td>
<td>→ Planning: emergency response planning [58], university campus planning [54]</td>
</tr>
<tr>
<td>→ Planning: travel-planning [113], calendar [204], online shopping [205]</td>
<td>→ Other examples: collaborative browsing [144], collaborative cognitive walkthrough [261], scientific collaboration [148], collaborative data visualization [7]</td>
</tr>
<tr>
<td>→ Smart environments: child monitoring [131], smart home [165], IoT, coffee machines [166]</td>
<td>→ Data Exploration (39)</td>
</tr>
<tr>
<td>→ Other examples: card game [36], polling [254], advertising [11]</td>
<td>→ Medical data: 3D medical data [9], brain scans [112], synthetic bio-design rules [105]</td>
</tr>
<tr>
<td>→ Scientific visualization: molecular visualization [26], astrophysics [35, 7]</td>
<td>→ Learning and educational games: games for training math [27], simulated classroom-distributed applications [140], learning task for children [99], games for training memory [21], biology simulation [106]</td>
</tr>
<tr>
<td>→ Spatial data: terrain analysis [245], geospatial disease data [215], oil and gas [206], city maps [49], interior rendering and design [319]</td>
<td>→ Presenting and critiquing: class presentation [117], critique student portfolio [165], electronic blackboards [57]</td>
</tr>
<tr>
<td>→ Financial data: financial computing [204]</td>
<td>→ Managing classroom activities: classroom capture [244], classroom polling [205], collaborative financial activity in groups [179]</td>
</tr>
<tr>
<td>→ Other data explorations: scatter plot [125], learning about dataset [206]</td>
<td>→ Mobile Computing (33)</td>
</tr>
<tr>
<td>→ Data from sensors: data from sensors and sensors connected to other devices [206], data from sensors connected to other devices [206]</td>
<td>→ Public displays and advertisements: communicating with public displays [57], interaction with displays at a distance [29], tracking engagement with public displays [161], bus-stop display [269], airport public display [205]</td>
</tr>
<tr>
<td>→ Sharing information: business card sharing [57], scanning receives [71], data transfer to the fly [177], ad-hoc connection [250]</td>
<td>→ UI design: UI design [219], UI builder [219], multi-user interface design [219], debugging and cross-device [105], Java applications [219]</td>
</tr>
<tr>
<td>→ Input redirection: UI redirection [154], content distribution across keyboard and display [170], resource consumption [205], secure entry on public display [271]</td>
<td>→ Other examples: extended screen space [267], authentication [130, 133, 134]</td>
</tr>
<tr>
<td>→ Other examples: extending screen space [267], authentication [130, 133, 134]</td>
<td>→ Mobile Computing (33)</td>
</tr>
</tbody>
</table>

5 APPLICATION DOMAINS

Our survey revealed a range of different application domains for cross-device computing. Although only 63/510 papers were tagged as making an ‘application’ contribution, 361/510 papers included some form of application use cases to motivate or frame the main contribution, such as new system designs, interaction techniques, tracking technologies, or interaction concepts. We identified nine high-level application-type clusters: (i) knowledge work, (ii) home
The largest category of applications focuses on knowledge work with 62 papers. Typical knowledge work is information management across various displays and devices [90, 201, 283], sharing information and resources across devices [164, 252], multi-device activity and task management [13, 140, 265, 353], or productivity and creativity tasks [148, 205, 246, 356, 379]. Other domains include police analyst workstations [4], industrial facilities [180], aviation cockpits [88], and collaboration between aerospace scientists and engineers [148].

Starting in the year 2000, we see a growing interest in applying cross-device interaction to data exploration. Information visualisation leveraging mobile devices including tablets and smartphones has called for new interaction vocabulary [18, 138, 171, 322]. In particular, Horak et al. [138] described interaction concepts for a smartwatch-display configuration in a crime analysis scenario using a real city criminal dataset. Other kinds of data sets that has been explored including geography [210, 249], physics [18, 315, 357], life science [66, 92, 102, 108, 296], city planning [46, 340], energy [43, 306], and finance [204].

Education [37, 179, 365] and health [3, 87, 217, 363] have also been popular application domains, where there has been an emphasis on collaborative and distributed work. Education applications have primarily focused on supporting classroom capture [344], classroom presentation [37, 117], educational games [29, 71, 195], and simulations of classroom activities [163, 232, 262, 345, 365]. Kreitmayer et al. [179] present one of the few in-the-wild studies, where they observed collaborative activities in the classroom to inform the design of a group finance management activity with a shared tablet and large display. In the health domain there have been a few studies in the wild, including work on exploring pre-surgery scenarios [3], supporting patient registration [363], and designing distributed user interface systems in surgery practice [87, 217]. Other health applications are designed for personal use, namely memory games [72], cognitive training [73], and physical activity monitoring [172].
light-weight but lower fidelity inside-out tracking technology. This classification shows the breadth of state-of-the-art research of cross-device tracking technologies, including those off-the-shelf. It serves as a reference table to help inform which tracking technologies might be best suited for a particular usage scenario.

### 7 INTERFACE TECHNIQUES

The fundamental method by which people use cross-device computing is through interaction techniques. In our sample, 130/510 explicitly mention interaction techniques as the main contribution of the paper. Further in-depth analysis reveals that another 221 papers introduce interaction techniques. The interaction techniques can be grouped into five distinct categories: (i) on-screen interaction, (ii) around the device interaction and gestures, (iii) device motion, (iv) changing the shape of the devices, and (v) using body gestures.

**Phase 1: Configuration of Devices.** The first phase focuses on setting up cross-device configurations of devices including pairing, combining, connecting, or coupling multiple devices. The purpose of this configuration phase is to establish a meaningful semantic relation between devices that enables cross-device activity. Examples of cross-device interaction techniques can be found in all six modalities. For example, using on-screen interactions, techniques such as stitching [129] or pinching [192, 238, 243] multiple display together, or performing synchronous tapping touches [280] have been used to pair devices into one configuration. Pairing techniques have also leveraged on-device pointing [259, 349] or different finger posture and identification [144] to implicitly create cross-device configurations. For around-the-device interactions, examples include knocking to pair [103], or even taking a picture to recognize other devices [299].

### Table 2: Tracking characteristics and modalities of the cross-device papers with tracking as a main contribution. Our tracking classification directly relates to surveys of tracking technologies in ubicomp [126, 137].

<table>
<thead>
<tr>
<th>Modality</th>
<th>Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic</td>
<td>standalone</td>
<td>Distance without cross-device synchronization [358], motion-resistant distance calculation [370], relative positioning using custom ultrasonic sensors [122, 377].</td>
</tr>
<tr>
<td>user-generated</td>
<td></td>
<td>Acoustic 3D positioning on unmodified mobile phones [156], conceptual relative 3D positioning on static mobile phones [265].</td>
</tr>
<tr>
<td>capacitive</td>
<td></td>
<td>Doppler-based gesture recognition between 2 phones [30], Doppler-based multi-device selection [4], keeping between devices on surface [150].</td>
</tr>
<tr>
<td>magnetic</td>
<td></td>
<td>Detect upper/lower device in device stacks [78], detect magnetic field change to define neighbours [147].</td>
</tr>
<tr>
<td>optical</td>
<td>fiducial markers</td>
<td>Detect configuration of devices [187].</td>
</tr>
<tr>
<td></td>
<td>HGA</td>
<td>Detection of neighbouring subtitles [216].</td>
</tr>
<tr>
<td>RF-based</td>
<td>RGB camera</td>
<td>Detects display position by searching for display content in the camera image [8], relative positioning from inside-out tracking [87], compare camera images across devices [74].</td>
</tr>
<tr>
<td>any input</td>
<td></td>
<td>Through synchronously providing (tap) input to two or more devices [80].</td>
</tr>
<tr>
<td>capacitive</td>
<td></td>
<td>Tracks proximity of person with wearable device through BLE [134], tracks wearable device through body-coupled communication on touch [134], tracks device's capacitive footprints on touchscreens [371], tracks tangible and device footprints on capacitive surface [346], capacitive markers on phone base for touch on large screen [374], read biometric data through body-coupled communication on touch [154].</td>
</tr>
</tbody>
</table>
| optical | | Tracks screens in image [271], overhead Kinect camera [200], surrounding Kinect camera [304], depth camera to track mobile devices in 3d space [320], compare device IMU with depth camera optical flow [350], computing device orientation using a 6-DOF IMU [350], displays code on screens [370], identifies tracked devices by flashing a marker on the screen [206], detects display position using matrix of NFC tags [195],detecting 2d position using matrix of NFC tags [195], RF/dual camera [196]
| RF-based | | Detects presence/absence devices (RFM) [317], custom RF trackers integrated into phones [196]. |
| markerless IR | | Measure distance between devices using FSRs [272]. |

Legend: Proximity/distance O, Relative locations (top, right, bottom, left) L, Location 2D, Location 3D, Orientation O, Identity id, and people P, tracking T

Phases of Cross-Device Interaction Techniques

We identified three phases of cross-device interaction techniques (see a complete overview in Table 3): (Phase 1) the configuration phase, (Phase 2) the content engagement

---

3The table is not an exhaustive or complete list of all interaction techniques, but a representative sample from our dataset that is exemplary for the three stages and six modalities identified in our analysis.
Gestural pairing techniques include Rekimoto’s seminal tapping [280], and techniques such as roll-and-pointing [61] or hold-and-flipping [54] to combine devices.

Most techniques in the configuration phase are designed with 2D or 3D device motion as the main input modality. As seen in Table 3, examples of pairing techniques using motion include bumping [127], stacking [61], or shaking devices [131, 211]. The few shape-changing techniques examined here are concerned with modifying the physical shape of devices through bending, sandwiching, or stretching [61] can be used to relate devices to each other.

Using eye, gaze, or head orientation, techniques such as perspective-aware interfaces [226], perspective-aware cursor [227], or display change visualisations [81] are used to select the right device or screen. Finally, techniques such as gradual engagement [203] and proxemic interactions [107] leverage the location, position, and orientation of the entire body to create semantic relations between devices. As combining various devices into a cross-device configuration or ecology is a central precondition for any application or technique to work across devices, it is unsurprising that so many techniques explicitly focus on this pairing or configuration phase.

**Phase 2: Content Engagement.** The second phase occurs after devices have been configured for cross-device usage, and includes direct or indirect interaction with content, data, visualisations, applications or interfaces that are spread across multiple devices. Content engagement encapsulates the actual consumption of content across various changing device configurations. Many of the classic cross-device interaction techniques – inspired by the seminal Hyperdrag [281] and Pick-and-drop [222, 278] – use direct touch or mouse interaction with the displays to move content from one device to another. Examples include using drag-and-drop across the bezel of multiple screens [312], swiping in the direction of another device [159, 348], as well as panning [270], tapping [310], and flicking gestures [277]. Direct interaction with content across devices has been supported through drag-and-drop proxy icon portals [64, 123, 206], pressure-based press-and-flick techniques [223], pinch-to-zoom across multiple displays [206], or drag-and-pop and drag-and-click techniques for multi-screen environments [15]. Around-the-device interactions are predominately based on interactive surfaces (like PhoneTouch [294] or ActivitySpace [143]), or

---

**Table 3: Overview of interaction techniques for cross-device computing.**

<table>
<thead>
<tr>
<th>Input Modalities</th>
<th>Touch, Devices, or Device Manipulation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>On-Screen</strong></td>
<td><strong>Touch Around Device</strong></td>
</tr>
<tr>
<td>Pair, select or engage with device by:</td>
<td>Pair, select or engage with device by:</td>
</tr>
</tbody>
</table>
| - virtual-based handshake [350] | - taking picture to recognize device [296] | - snapping a picture [60] | - detect head position [110] | - -
| - finger postures [141] | - holding-and-flicking [54] | - placing down device [143, 270, 303] | - - | - -
| - painting-on surface [378] | - pointing-to-open [54] | - recognizing motion correlation [350] | - - | - -

| **Phase 1** | **Configuration** | Setting up cross-device configurations of devices including pairing, combining, connecting devices. |
| **Phase 2** | **Content Engagement** | Techniques aimed at interacting with, transferring or exploring content, data, visualisations, or interfaces that spread across multiple devices. |
| **Phase 3** | **Disengagement** | Techniques to disconnect cross-device setups or configurations, interactions, and applications. |
projection systems that extend the interaction space to visual proxies next to a device [311], extended projected displays [57], touch-enabled surfaces around the device [362], or even augmented mouse, touch, and keyboard input [23].

Mid-air gestures have also been considered for cross-device interaction. Some of these are variations of waving such as waving between devices [50] or waving above devices [141]. Other mid-air examples are performed after a touch or 3D motion interaction, such as point-and-grab [198], lift-and-drop [8], or grasp and micro-mobility [375].

The majority of interaction techniques for content engagement use the 2D or 3D device motion modality. A first category of techniques focuses on 2D movement on a flat surface. Techniques include rotating devices to interact with content [143, 270, 288], moving devices to explore spatially aware maps [90, 270], or to explore information visualisations [367]. The second category focuses on advanced 3D motions with devices for content interaction. Examples include pouring [159] or throwing content onto a display [69], and tilting actions to pan a map [69]. Further techniques include rotating [188], throwing and chucking [122], shoot-and-copy [27], and tilting [191] techniques to interact across devices (see the full list in Table 3).

There are only a few full-body gestures, such as content transfer propagated through F-formations [206], or gaze and head gestures to select devices or screens [183, 303], and pick and drop content [333]. Many of these techniques leverage the physicality and affordances of the devices to enable expressive 3D device motion to receive, use, or send data to other devices that are very easy and intuitive to perform.

**Phase 3: Disengagement.** The last phase covers interaction techniques for a person to stop cross-device content engagement on a device, infrastructure, or application level. While the first configuration phase has received much attention in earlier work, the disengagement phase remains less well explored. Few examples for on and around-the-device include cross-device session management [116] or covering a smartwatch to reset the cross-device configuration [32]. Using 3D motion, there are techniques to break connection by moving the device [127, 298], picking up the device from the tracked area [143, 190, 270], tilting devices vertically towards oneself to stop sharing [190], shaking to break connection [243], or implicit disconnection of the device by breaking physical contact [371]. Finally, proxemic interaction supports implicit disengagement by leaving the operation-space [206] or by turning away from the display [369].

It is important to note that interaction techniques can occur in – or combine – multiple functions from different phases at the same time. For example, techniques discussed in PhoneTouch [294], WatchConnect [141], Gradual Engagement [206], or Gluey [303] combine Phase 1 configuration and Phase 2 content engagement functions in one interaction technique. However, applying this taxonomy can be a helpful analytical lens to understand the breadth and focus of most cross-device interaction techniques.

<table>
<thead>
<tr>
<th>1</th>
<th>Managing the applications, information and spaces</th>
</tr>
</thead>
<tbody>
<tr>
<td>→ Space Window Manager [24] and Layout Manager [357]</td>
<td></td>
</tr>
<tr>
<td>→ RadarMap [2, 97], Map [24], MiniMap [19]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2</th>
<th>Content visualisation and distribution strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>→ Multiple coordinated views [48, 91, 92, 120, 263, 322]</td>
<td></td>
</tr>
<tr>
<td>→ Brush and linking, across devices [7, 63]</td>
<td></td>
</tr>
<tr>
<td>→ Overview and detail [16, 136, 287, 322]</td>
<td></td>
</tr>
<tr>
<td>→ Central overview device [31]</td>
<td></td>
</tr>
<tr>
<td>→ Focus and Context [98, 322]</td>
<td></td>
</tr>
<tr>
<td>→ Adaptive views, applications, adaptive UIs [21, 150, 372]</td>
<td></td>
</tr>
<tr>
<td>→ Overlaying information [357]</td>
<td></td>
</tr>
<tr>
<td>→ MagicLens views [16, 28]</td>
<td></td>
</tr>
<tr>
<td>→ Dynamically corrected perspective views [206]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3</th>
<th>Feedback and feedforward approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td>→ Object shadow [23]</td>
<td></td>
</tr>
<tr>
<td>→ Gesture shadows/ghosts [22, 188, 301]</td>
<td></td>
</tr>
<tr>
<td>→ Temporary preview shadows of content [206]</td>
<td></td>
</tr>
<tr>
<td>→ Perspective cursor [227, 349]</td>
<td></td>
</tr>
<tr>
<td>→ Multi-display pointers [163, 298]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4</th>
<th>Explicit linking between devices</th>
</tr>
</thead>
<tbody>
<tr>
<td>→ Coloured physical borders, frames or cases around device [116]</td>
<td></td>
</tr>
<tr>
<td>→ Coloured digital borders around screen/display [116, 206, 271, 351, 358]</td>
<td></td>
</tr>
<tr>
<td>→ Coloured lines between digital content across devices [69]</td>
<td></td>
</tr>
<tr>
<td>→ Visual proxy icons or portals [24, 64, 89, 97, 123, 203, 357, 358]</td>
<td></td>
</tr>
<tr>
<td>→ Proxy icons around devices [143]</td>
<td></td>
</tr>
<tr>
<td>→ Cross-device portals, borders [206]</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Cross-device visualisation and management.

**Visualisation and Cross-Device Management**

Related to the interaction techniques, we identified four major categories of visualisation and feedback that have been used to help users understand how a particular cross-device interaction technique or application works (Table 4).

First, a number of techniques provide users with an overview of the cross-device interactions across devices and space. Examples include the use of a layout or window manager [24, 357] and the use of an overview or mini-map [2, 16, 24, 97, 358]. Second, to increase the overview and understanding of where information and applications are located across devices, a number of content visualisations and distribution approaches have been developed. This category includes real-time coordinated views across devices [91, 92, 120, 263, 322], brushing-and-linking between devices [7, 63], overview and detail on demand or on other devices [16, 31, 136, 287, 322], magic lens views [16, 28], or dynamically corrected perspective views [226]. Third, to increase the intelligibility of cross-device systems, feedback and feedforward mechanisms have been developed. Examples such as object preview and gesture shadows [22, 188, 206, 301], perspective cursor [227, 349] or multi-display pointers [163, 286] help users understand how information is travelling between devices. A final strategy is in explicitly visualising the links between devices. Both
Informative studies (e.g. observational studies) [112, 148, 153, 252, 290] gesture elicitation studies [236, 304] qualitative user study based on heuristic criteria [236, 351, 357, 358] Observations: Observational studies show what a cross-device system or interaction technique supports and how it is utilised by users for certain tasks [184].


Table 5: Evaluation methods used in our corpus (in round brackets are the number of papers employing each strategy).

<table>
<thead>
<tr>
<th>Evaluation Method Used</th>
<th>Usage (219)</th>
<th>Informative (167)</th>
</tr>
</thead>
<tbody>
<tr>
<td>qualitative lab study</td>
<td>[112, 148, 153, 252, 290]</td>
<td>gesture elicitation studies [236, 304]</td>
</tr>
<tr>
<td>mixed method lab study</td>
<td>[7, 54, 151, 191, 206, 225, 254, 255, 258, 280]</td>
<td>technical evaluation (66)</td>
</tr>
<tr>
<td>deployment in</td>
<td>social setting, conference [204, 217, 325]</td>
<td>performance, compared to other systems [6, 7, 47, 54, 55, 103, 110, 123, 145, 189, 227, 230, 258, 267, 289, 292, 303, 305, 379]</td>
</tr>
<tr>
<td>lab study with experts</td>
<td>[14, 58, 99, 197, 202, 205, 236, 261, 273]</td>
<td>system performance (e.g. time, frequency, FPS, round trip time, memory, etc.) [6, 47, 103, 265, 298, 270, 282, 293, 304, 358, 374]</td>
</tr>
</tbody>
</table>

Table 5: Evaluation methods used in our corpus (in round brackets are the number of papers employing each strategy).

8 EVALUATION STRATEGIES

In this section we report on our analysis of evaluation strategies for cross-device work. While technical research might have a performance evaluation or a preliminary expert evaluation with developers, other work is evaluated through lab studies or real-world deployments with users.

Evaluation Methods Used

In summary, 317/510 of the papers in our corpus reported on a study, which we clustered into five main evaluation strategies for cross-device work. As much of the cross-device research can be considered a "constructive problem" (constructive—conceptual or constructive—empirical [251]), we extended Ledo et al.’s evaluation strategies of technical toolkits for HCI [184] for the purpose of our classification with "informative", resulting in the five main categories: informative (observational and elicitation), demonstration, usage (qualitative and quantitative), technical, and heuristic evaluation (Table 5). We would like to point to Ledo et al.’s in-depth discussions about evaluation strategies of toolkits for the latter four [184], which similarly surfaced in our analysis. In the following we will give a brief summary of each strategy, and explain our newly added fourth strategy in more detail.
how cognitive load can be reduced through usage of multiple displays [153] and multiple devices [75, 252, 290]. Similarly, other observational studies have been used to investigate current multi-device utilisation [41, 42], barriers for true multi-device usage [263], or the effects of display sizes in collaborative work [148, 377].

Gesture Elicitation Studies: Cross-device interaction techniques frequently involve on-, around-, and mid-air gestures to connect devices or manipulate content. Such gestures designed by the researcher building the system are not always reflective of users’ preferred choice. Wobbrock et al. [364] proposed gesture elicitation studies as a tool when designing novel interaction techniques: participants are presented with the effects of an action (referent), and they are asked to perform the signs which could cause those actions. Through the weighing and calculation of agreement scores, the proposed gestures can be mapped to user-friendliness and acceptance [343]. Through sensible crafting of elicitation tasks, such studies can generate a gesture vocabulary which allows for re-use of gestures for similar tasks.

Technical Evaluation: Technical evaluations are used to show how well a system works [184]. 66 papers in our corpus report on a technical evaluation, which is exclusively used to evaluate a tracking technology (28) or other cross-device toolkits (38). Again, we point to Ledo et al. for more details [184].

How to use the evaluation classification: The classification of evaluation strategies in the cross-device domain can help researchers find an appropriate evaluation method. Even though usability evaluation is (sometimes) considered harmful [106], we hope to show that a multitude of methods are available to evaluate cross-device interactions, systems, and interfaces, and we encourage a discussion within the HCI community about these methods.

Studying cross-device interactions: Studying cross-device interactions often consist of video recording lab studies or in-the-wild deployments, especially for qualitative studies. However, only a few systems support researchers in analysing this video material: VICPAM [222], EXCITE [207], and EagleView [34] are systems enabling researchers to visualise and/or query spatial interaction with multiple devices. Ultimately, more work is needed about how to best evaluate cross-device interactions and systems.

9 KEY CHALLENGES AND RESEARCH AGENDA

Research on cross-device computing and interaction has shown that there are tremendous benefits to be gained by breaking the confinements of solitary computers, devices, and users. Through our survey of cross-device literature, we identified open challenges and issues during tagging of our corpus of papers and our own reflections on these findings. We synthesised all tagged entries (e.g. explicitly mentioned key challenges from 67 papers) into the following nine themes. We combine them with challenges identified in prior work (e.g. [111, 142]).

Bridging the Gap between Studies and Systems: To support human activities in a cross-device ecology or to meaningfully compare cross-device interaction techniques and approaches, we need to develop testable design patterns [203] by making applications and scenarios the central focus. Although some work has attempted to compare techniques [223, 271], the underlying fundamental problem is that there is no frame of reference to compare and evaluate cross-device techniques and systems. While many of the technical contributions re-envision and push the boundaries of interaction possibilities, they are often disconnected from findings using empirical studies. More work is needed to unify empirical and technical cross-device work into one stream of research.

Conveying Cross-Device Interaction Capabilities: While cross-device capabilities have in recent years started to appear in commercial products, several studies have contributed sometimes controversial findings about device utilisation. A study of Apple’s Continuity, for example, showed that users have challenges in understanding its features, being aware of its presence and effects, and mitigating significant privacy issues when devices are shared [273]. And while Räddle et al. argue for spatially aware cross-device interactions, they remark that such interactions have to be designed with great care to reduce users’ cognitive load and mental demand [271]. The underlying challenge is in communicating the action possibilities and benefits of cross-device interaction in systems and applications. More specifically, we need new concepts, feedback and feed-forward mechanisms, and user interface patterns that are designed specifically for cross-device computing [31].

Mitigating the Effects of Legacy Bias: The phenomenon of legacy bias, where users resort to well-known interaction styles even when more effective and novel techniques are available, has been documented in studies of cross-device sensemaking [263], note-taking [157], and curation tasks [32]. Although workplace and user experience studies consistently report that many people are already struggling with multi-device fragmentation [42, 75, 290], it remains an open issue to what extent users will adopt new multi-device systems. More research into the mental models of individual devices and larger ecologies is needed to provide an empirical ground for new technical cross-device research.

Addressing Social Challenges: Designing new cross-device systems involves tackling the challenges of social relations and norms [77], privacy [33, 337], authentication, as well as providing support for the configuration work [143].
needed in the engagement and disengagement stages of the cross-device interaction. Our analysis finds the systemic lack of interaction techniques for disengaging from cross-device interaction (Table 3). Users need the ability to configure (or reconfigure) cross-device functionality [273], and Greenberg et al. argue for explicitly supporting opt-in and -out of interactive systems [104]. Although some initial work has been conducted in this space, fundamental issues around the entire cross-device interaction model remain.

**Enabling Proactive Cross-Device Interaction:** Although Weiser’s vision [355] is foundational for cross-device computing research, it has also elicited critical reflection. For example, Rogers proposes a “shift from proactive computing to proactive people” [284] in which purposefully built experiences engage people while leaving them in control of their interactions with the world [284]. Similarly, Oulasvirta summarises that it is users who are “doing the ubicomp” [250] and Dourish argues that users, not designers, appropriate technology and thus create meaning for their interactions [83]. Therefore, instead of blending devices together and trying to hide the boundaries between them, designers should embrace and leverage the heterogeneity and flexibility of devices and their “seams” [45, 250] – ultimately creating an ecology of devices that build the conceptual foundation of cross-device computing. The current move to mobile, ad hoc, and re-configurable device configurations is reflective of this shift, but the context of use, the user’s action, and specific applications and scenarios need to be considered in a much stronger way.

**Building and Deploying Cross-Device Systems In-the-Wild:** Many cross-device systems and interactions were built and tested in controlled lab setups (Table 5), often involving only a small set of simultaneous users. It is unclear how well the systems and interactions transition and scale to environments that are more representative of everyday interaction [135], what users’ actual cross-device interactions are in their everyday lives [368], or how they may change their use of multiple devices outside the lab [151]. Researchers have therefore argued for in-the-wild deployments in users’ typical environments [30, 285] to better understand and support their actual tasks in their settings [179]. However, this opens up new challenges about technical capabilities and the infrastructure problem.

**Improving Tracking Technology and Infrastructure:** Much of the enabling technology for cross-device interactions (Table 3) is prototypical, difficult to set up, expensive, or requires a lot of space. There is a need for such systems to be more reliable [181], to improve speed and accuracy during regular use and motion [47, 145, 258, 267, 299], and to bring cross-device capabilities to unmodified devices outside the research space [243]. On a technical level, cross-device research needs more practical testing [160] and refinement for situations outside the lab [187], to support wider-scale use and in-the-wild deployments. Outside-in-tracking often requires markers attached to the tracked object or is easily confounded with uncertain conditions (light, noise). And while inside-out-tracking has in the past always been used for in-the-wild deployments – due to its robustness, mobility, and support of ad hoc situations – it lacks the fidelity and details (e.g. tracking of users) of outside-in. On the other hand, devices themselves can be aware of their context of use. Researchers in academia and industry have begun to point out and tackle this infrastructure problem [143, 250], but only a few efforts have focused on minimising setup and configuration time on the user’s part to enable interactions out-of-the-box (e.g. [159, 378]).

**Bespoke Solutions vs. Platforms:** Commercial attempts at cross-device computing are limited to a single user managing their personal device ecology within a particular manufacturer’s ecosystem, with little support for real collaborative activities. However, the technological innovations that succeed in breaking the barriers of commercial applications and ecosystems are most often built on open standards, notably the World Wide Web, e-mail, and open file formats. Few standards do exist that support cross-device computing, and the integration of access to technologies like Bluetooth in modern web-browsers points in a direction that opens up for exploiting cross-device interactions outside the commercial or research silos. However, design of meaningful standards must be informed by rigorous studies of use, and not only confined to the laboratory [142, 250]. This presents a chicken-and-egg problem as cross-device interaction techniques and applications are notoriously difficult and costly to build, deploy, and test.

**Towards a Symbiosis between Cross-Device Capabilities and Human Activities:** Cross-device research is diverging further with new interaction techniques for mobile, wearable, and tangible devices, with various input and output modalities. While cross-device computing essentially focuses on ‘the device’, it is in itself also ‘device-agnostic’. With new device form factors, materials, mixed-reality technology, IoT devices, and shape-changing interfaces there is a renewed challenge and opportunity to rethink the boundaries, purposes, and scope, of devices within a complex ecosystem.

10 **CONCLUSION**

Over the past three decades, interaction with computers has progressed from single-screen mainframe computing, to dual-screen desktop PCs, to advanced multi-display devices with gesture interactions, to the proliferation of today’s mobile and wearable devices. Multi- and cross-device computing has become a fundamental part of human-computer interaction research. Despite the great variety in research
agendas and focus points, the common ground in our community is to understand, create, and deliver experiences that transcend the individual device. Surveying over 30 years of cross-device research in a single paper does not do justice to the many researchers who have actively created, developed, contributed to, and shaped this field. We acknowledge that there are many different ad to be drawn from this survey and we hope that our interpretation is one further step for a wider discussion within the HCI community. By reflecting on the terminology we are using, and by identifying and addressing the fundamental challenges, the community can shape the future of cross-device computing together. We are looking forward to continuing discussions about where we are heading, and invite cross-device researchers and practitioners – new and established alike – to contribute to our open dataset: https://github.com/frederikbrudy/cross-device-taxonomy.

11 ACKNOWLEDGEMENTS

We thank our anonymous reviewers for their suggestions on improving this paper. Frederik Brudy’s work has been supported by Microsoft Research through its PhD Scholarship Programme. Roman Rädle and Clemens Klokmose’s work has been supported by the Aarhus University Research Foundation.

REFERENCES


