Developing a user-centred system for long-term elephant monitoring

Vicki Fishlock[1,2], Fran Michelmore Root[3], Norah Njiraini[1], Katito Sayialel[1], Cynthia Moss[1], Phyllis Lee[1,4]*

[1]Amboseli Elephant Research Project, Amboseli Trust for Elephants, PO Box 15135, Nairobi, Kenya
[2]Centre for Ecology and Conservation, University of Exeter, Exeter EX4 4PY, UK
[3]Lewa, PO Box 1, Private Bag, Isiolo 60300, Kenya
[4]University of Stirling, Department of Psychology, Stirling, FK9 4LA, Scotland

*corresponding author: vfishlock@elephanttrust.org

Introduction

Originally envisaged as a three-year project, the Amboseli Elephant Research Project (AERP) has proved a labour of love for a small, dedicated team who have followed the life trajectories of more than 3,900 individual elephants over five decades. AERP's unique knowledge base is derived from tracking individually recognized animals in a small well-protected population studied continuously since 1972, providing an important baseline for a free-ranging elephant population with an intact age structure. This dataset forms the basis of our demographic analyses (Moss 2001; Lee et al. 2013), and through training, research, books and films has contributed to global understanding of elephant reproductive biology and behaviour; *musth*: (Poole and Moss 1981; Poole 1987: Poole 1989; Hollister-Smith et al. 2007); *oestrus*: (Moss 1983); *mothering and grandmothering*: (Lee et al. 2016; Lee et al. 2022), *elephant cognition*: (McComb et al. 2001; Bates et al. 2008; McComb et al. 2014) and communication: (Poole et al. 1988; McComb et al. 2000; McComb et al. 2003).

How can we now, after a half century of effort, make our data accessible and of continued use to the global elephant community? To try to answer this demand, AERP has moved beyond a simple relational database into a data management system that allows users to add, map, inspect, edit, and extract data. Here we share some of the key concepts that drove this process and outline our hopes for making elements of the system available to other projects that may benefit from similar capacities.

Problem statement

In such a long-term study we have faced many challenges in maintaining records on individual elephants, due to changing technologies (from hand-tallied summaries to IBM punch cards to cloud storage, each with a finite lifespan) and the sheer volume of data collected. Inevitably, data protocols evolved to encompass more elements of elephant biology, although the basic type of records made, and our definitions of behaviour have remained consistent. This consistency is vital for long-term monitoring (LTM) and is in part thanks to long-serving team members making many of those records; it is also due to careful training of research collaborators and their inclusion in LTM record-keeping alongside their specific research projects.

Technology has been the largest transformation challenge over the project's lifetime. Landsat for satellite imaging the earth's habitats was launched in the same year as the project began, and its resolution has been continually upgraded. Computing has evolved from room-sized mainframes to nanochips in smartphones. And, like any project spanning decades, we have had to balance new technology with investment in financial and staff resources that major system changes entail. Growing apace with technology was the elephant population itself; thanks to community endeavour and research presence, the number of elephants has increased slowly over the decades (Moss et al. 2011). This conservation success has presented new challenges; elephants have shifted their ranging patterns to take advantage of the larger safe landscape, while high survivorship has meant that the overall number of individuals to be tracked has increased significantly from ~700 early on to over 1,900 in 2022. Individual elephant identification lies at the very heart of the project and is constantly updated through a photographic dataset of ears, tusks, tails and body markings. Maintaining identifications, and transmitting this knowledge to others, is key to the project's enduring success.

In 1997 AERP constructed the first Access database, capturing the elephant sightings data (Table 1), enabling tracking and analysis of elephant occupancy of the ecosystem in time and space, and the varied social opportunities these groups represent for elephants over the ecological year. This huge step nonetheless left out key LTM components, namely demographic data (births, deaths, musth, oestrus and mating), within-family dynamics, and key ecosystem variables (rainfall and vegetation), because computers of the time simply could not cope with the size and complexity of the full dataset. By 2015 computing power had advanced and the size and ranging patterns of our study population made it clear we needed a data management system that integrated all LTM elements and followed our actual data structure, with individuals at the heart (Fig. 1).

The build process

We rebuilt the entire data capture system (Fig. 1) using Microsoft Access and Excel interfaced with QGIS open-source GIS software (https://www.qgis.org). Given unpredictable and unstable internet access in the field, we retained an offline system that has slightly more complex file sharing and backup procedures, but which allows AERP users to work on data regardless of network connectivity. Key to our needs is the flexibility to build multiple databases that interact with each other, allowing the system to grow and change; and separating the user interface from the raw data (an early step in the process; Fig. 2) allows for constant design improvements without interrupting data flow for users in the field.

We followed a collaborative process between the designer (FR) and project manager (VF), coordinating each stage to include end-user feedback (Fig. 2). We used feedback at the design stage to determine what users felt was missing from the previous database, and in the build and launch stages to build and refine the queries (data inspection tools) that users needed. Some of these tools were only possible for users to

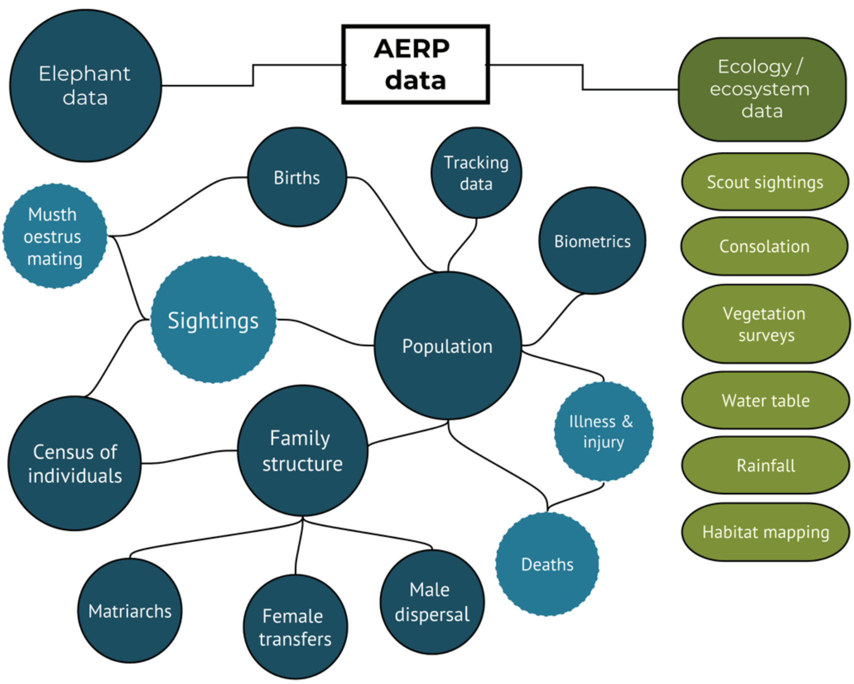

Figure 1: The data types included in the AERP long-term dataset. Dark blue circles show ID-dependent data areas, light blue circles with dotted borders indicate elephant data not dependent on known IDs. See also Table 1 for some data definitions and sample sizes.

Table 1. Details of selected data sections from Fig. 1, and current dataset sizes

Data Area	Details	Current Dataset Size
Encounter Data	**Sightings:** • Date, time, location, observer • Group type - male-only, female-only, mixed • Group activity - travelling, feeding, socialising, etc. • Group size and quality of count (exact, good estimate, approximate) • Quality of recognition, for males and females separately • Independent males present • Musth and oestrus individuals present (yes/no, and identity) • Families present, and portion present (all, >half, < half)	N = 54,244
	Musth, oestrus, mating • dyadic male-female interactions (guard, mount, successful mate)	N = 2,215 males for 1,884 oestrus females
	Census (of families) • Presence/absence data for all family individuals • Demographic event reporting e.g. birth, missing individuals (inferred mortality)	N = 10,989 census groups for 3,012 individuals since 1999
Life History Tracking	**Population** • ID code • Birth family • Sex • Birth month, year and accuracy (within 1mo, 6mo, 2yrs, 5yrs, 10yrs) • Mother ID, accuracy (unknown, estimated, good estimate, known) • Grandmother ID • Death month, year, and accuracy (unknown, within 2 yrs, 1yr, 3mo, 1wk) • Death cause, and accuracy (unknown, estimate, good estimate, known)	Total N= 3,906, then per population item; • 3,906 • 3,719 • 1,949 females, 1,910 males, 44 unknown • 2712, 577, 270, 243, 204 • 466, 77, 77, 3286 • 2,478 • 1,891 (216, 171, 351, 753, 398) • 1,891 (496, 293, 551, 549)
	Mortality Data • inferred i.e. missing from family, or disappearance • carcass (identified/not)	• 1,176 • 782

visualize once they had seen earlier versions of the system, so building and integrating the data entry and outputs was a stepwise process, which we continue to refine. Two key parts of our success to date were that the designer had a full understanding of the data flow from field to computer, and that the data capture forms we designed for the system always mirrored datasheets used in the field, to make it easier for users to become familiar and confident with the new system.

Key system features

Data entry is as simple as possible for users with data categories separated into different areas with “Add”, “View” or “Edit” options available once users navigate into the chosen area (Fig. 3). A series of controls reduce user errors and streamline workflows, e.g. when entering elephant ID code, the elephant’s name is always displayed so users can immediately recognize and correct typos. Where connections exist between different data areas, background code creates automatic lists of pending entries, so users can complete one data type at a time, e.g. when a family census is recorded during a sighting, the system allows the users to complete all the sightings data first, then go to the census area and select from a list of groups where census data is pending. The system also pulls real-time information on births and deaths, allowing for reliable and fast entry of individuals present in a census (Fig. 4).

Change log

Data can change as further observations are made on individuals, or input errors are corrected. Although all users can see the full population list, free editing is not permitted by all users, instead changes to key data are requested and then approved by an administrative user. We therefore built a Change Log, where changes to key fields are logged with the user’s identity, date, and the old and new values. For example, if the sex of a calf was incorrectly logged or initially unknown, users can

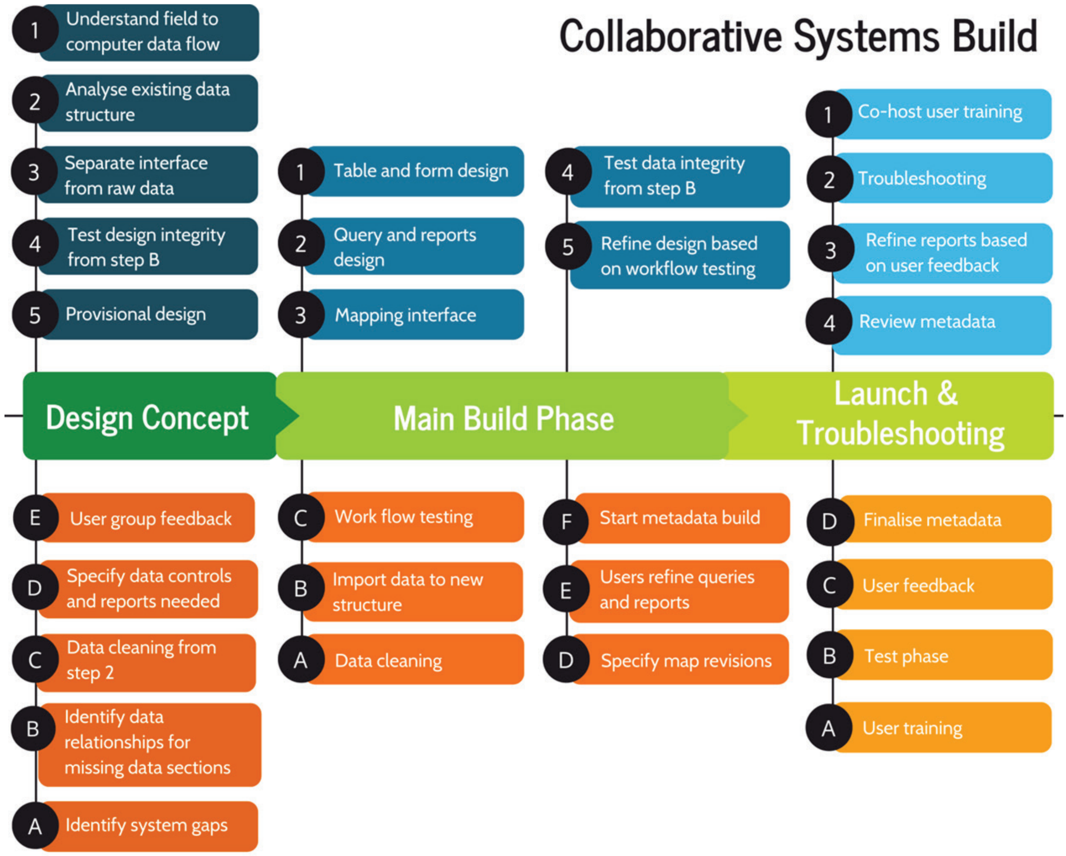

Figure 2. The interaction between the systems designer (top section, numbered 1, 2, 3) and the data users (bottom section, marked A, B, C) ensured user needs remained at the heart of the process. 'Metadata' is a full description of how the system uses and stores data, and creates a reference manual for the system, including a definition of each data type and the relationships between data sections.

report the change and the old values are preserved. This kind of logging makes it possible to examine conflicting observations and reconstruct data sequences rather than have them overwritten.

Reporting functions

To enable users to interact with the data we built an extensive set of pre-defined reports, allowing users to interrogate the dataset without having any programming knowledge. These reports include, for example, population size over time, sightings of families and individuals and ecological data (Fig. 5). For each report or query, users can select the desired time range, individual or other parameters of interest, and the results can be exported to Excel or as a PDF. Users can thus produce regular reports easily and examine individual life history data whenever they wish, without having to ask for technical assistance. Administration users also have a special set of data management queries. We expect to further refine all the reports and queries as part of the ongoing evolution of our system.

Further work

Our data management system is still evolving with aims to integrate tracking data (from live GSM collars and historical datasets), remote sensing (NDVI) and photo datasets. Functionality for photos of carcasses, wounds or treatments is already built in and will be tested over the following months. We have not yet embarked on integrating a photo identification library with our database, although we are aware that others are tackling these questions (Poole et al. 2022; this volume pp 72–90). The database developed by Poole et al. shares much of the AERP approach on capturing

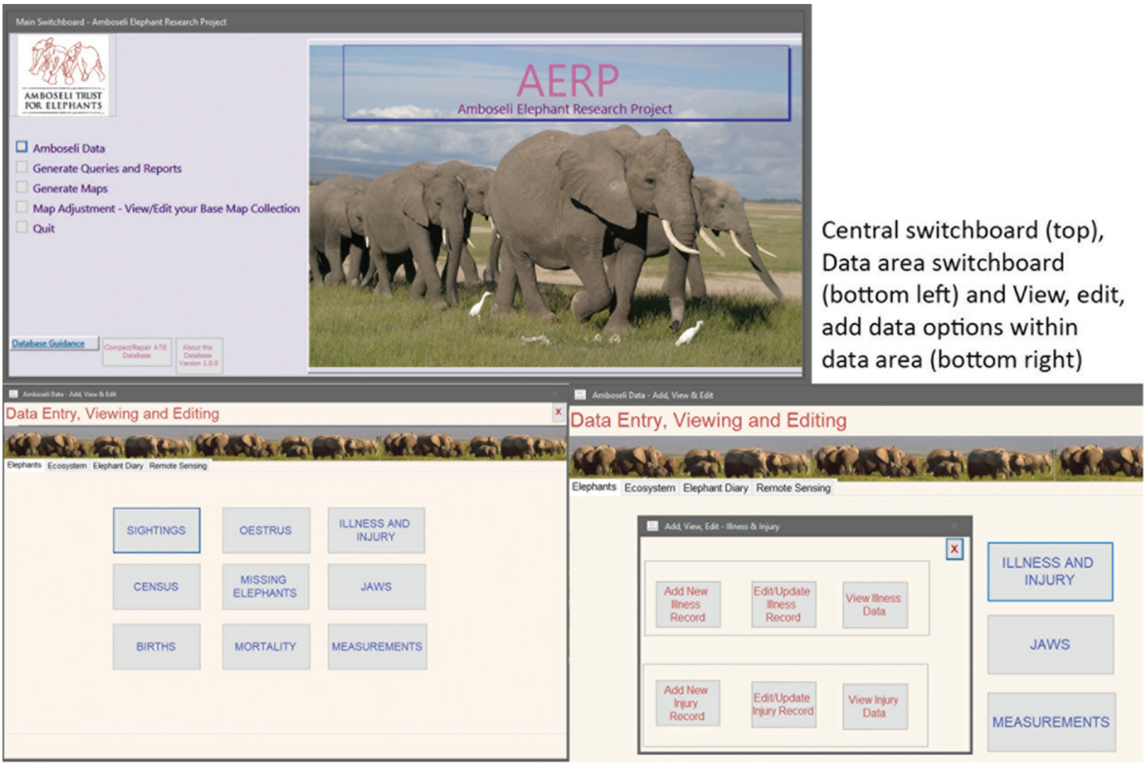

Figure 3. Screenshots of the main switchboard and system organization by data category. Above. Central switchboard; (Below left) Data area switchboard and; (Below right) view, edit add data options within data area.

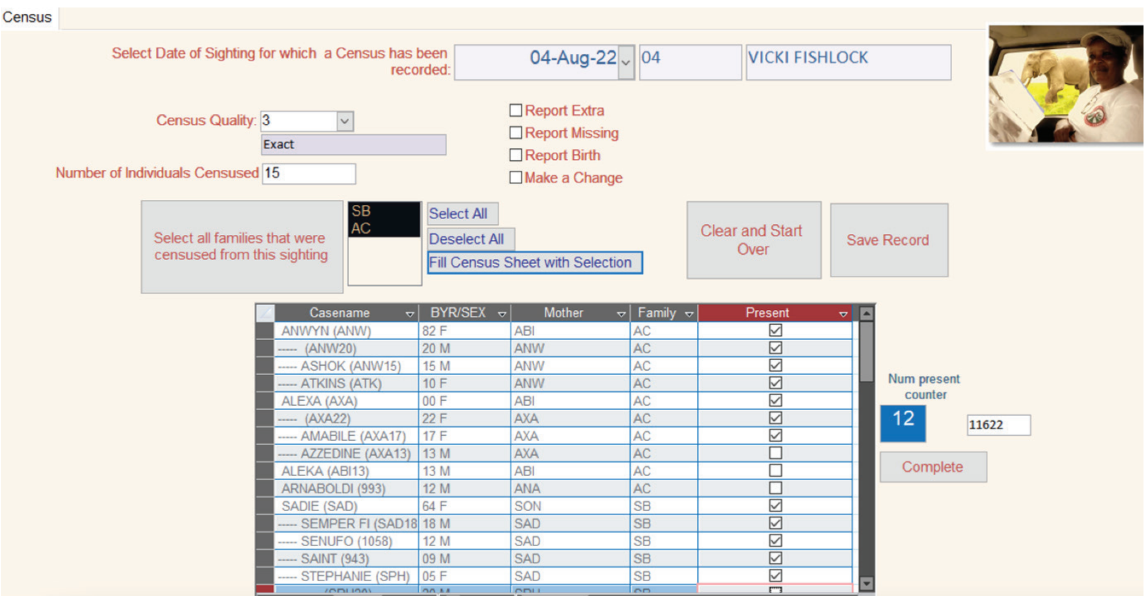

Figure 4. Census data entry form, where a group is from an auto populated drop-down list (top), users select those families censused, and then the “fill census” command draws up-to-date information from the population Table. A counter (blue box) helps users quickly stay on track with the number of individuals identified.

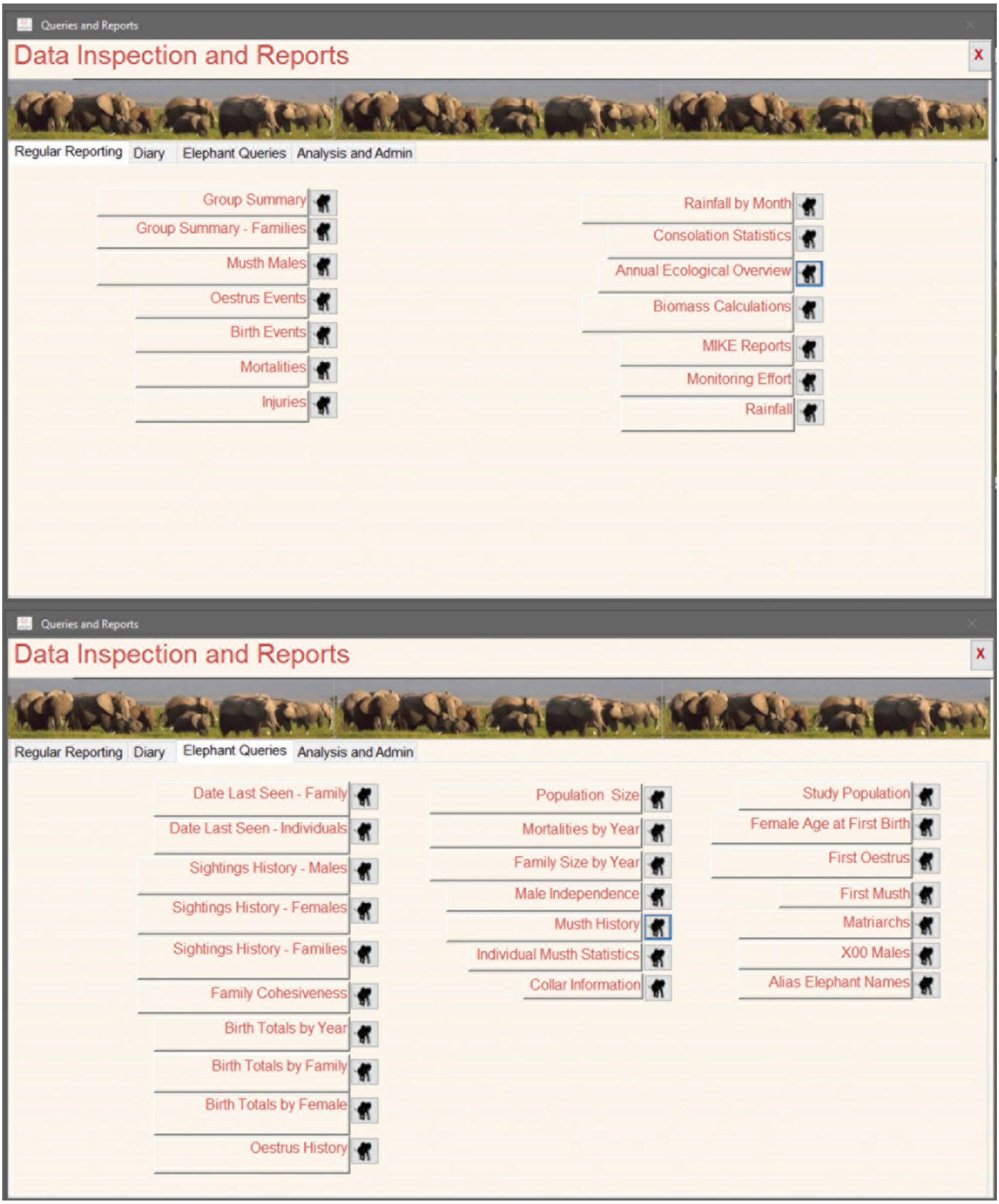

Figure 5. Examples of the query and reporting areas of the system; clicking on the desired report brings further pop-up boxes asking users to specify the time range or individual(s) of interest. Reports can be exported to Excel or PDF.

encounter data, thanks to our shared history, and integrates ID management (photographs and ID characteristics) with basic life history data, whereas the AERP design captures detailed demographic data at the individual and population levels and allows users to interrogate through pre-defined reports that don't require programming expertise. Some of the system's design complexity is generated by the way data protocols have evolved for our project, e.g. sightings data are taken at the level of family, and then detailed information on which females and offspring are present is added when possible (a census). However, we also recognize that many of the challenges that we have faced will be shared by colleagues with shorter-duration projects, where multi-level datasets on elephant sightings and life history data are maintained for and by multiple user groups. We would therefore like to appeal to any others who would be interested in collaborating to build a simple, standardized and freely available version of our system, using the framework of solutions that we have developed here.

References

Bates LA, Sayialel KN, Njiraini NW, Poole JH, Moss CJ, Byrne RW. 2008. African elephants have expectations about the locations of out-of-sight family members. *Biology Letters* 4: 34–36. doi.org/10.1098/rsbl.2007.0529

Hollister-Smith JA, Poole JH, Archie EA, Vance EA, Georgiadis NJ, Moss CJ, Alberts SC. 2007. Age, musth and paternity success in wild male African elephants, *Loxodonta africana*. *Animal Behaviour* 74: 287–296. https://doi.org/10.1016/j.anbehav.2006.12.008

Lee PC, Bussière LF, Webber CE, Poole JH, Moss CJ. 2013. Enduring consequences of early experiences: 40 year effects on survival and success among African elephants (*Loxodonta africana*). *Biology Letters* 9: 20130011. https://doi.org/10.1098/rsbl.2013.0011

Lee PC, Fishlock V, Webber CE, Moss CJ. 2016. The reproductive advantages of a long life: longevity and senescence in wild female African elephants. *Behavioral Ecology and Sociobiology* 70: 337–345. doi.org/10.1007/s00265-015-2051-5

Lee PC, Moss CJ, Njiraini N, Poole JH, Sayialel K, Fishlock VL. 2022. Cohort consequences of drought and family disruption for male and female African elephants. *Behavioral Ecology* 33: 408–418. https://doi.org/10.1093/beheco/arab148

McComb K, Moss C, Sayialel S, Baker L. 2000. Unusually extensive networks of vocal recognition in African elephants. *Animal Behaviour* 59: 1103–1109. doi.org/10.1006/anbe.2000.1406

McComb K, Moss C, Durant SM, Baker L, Sayialel S. 2001. Matriarchs as repositories of social knowledge in African elephants. *Science* 292: 491–494. doi.org/10.1126/science.1057895

McComb K, Reby D, Baker L, Moss C, Sayialel S. 2003. Long-distance communication of acoustic cues to social identity in African elephants. *Animal Behaviour* 65: 317–329. doi.org/10.1126/science.1057895

McComb K, Shannon G, Sayialel KN, Moss C. 2014. Elephants can determine ethnicity, gender, and age from acoustic cues in human voices. *Proceedings of the National Academy of Sciences* 111: 5433–5438. doi.org/10.1073/pnas.1321543111

Moss CJ. 1983 Oestrous behaviour and female choice in the African elephant. *Behaviour* 86: 167–195. https://doi.org/10.1163/156853983X00354

Moss CJ, Croze H, Lee, PC (Eds). 2011. *The Amboseli elephants: a long-term perspective on a long-lived mammal*. University of Chicago Press.

Moss CJ. 2001. The demography of an African elephant (*Loxodonta africana*) population in Amboseli, Kenya. *Journal of Zoology* 255: 145–156. https://doi.org/10.1017/S0952836901001212

Granli P and Poole J. 2022. Who's Who & Whereabouts: an integrated system for reidentifying and monitoring African elephants. *Pachyderm* 63: (72–90).

Poole JH and Moss CJ. 1981. Musth in the African elephant, Loxodonta Africana. Nature 292: 830–831. doi.org/10.1038/292830a0

Poole JH. 1987. Rutting behavior in African elephants: the phenomenon of musth. *Behaviour* 102: 283–316. doi.org/10.1163/156853986X00171

Poole JH. 1989. Mate guarding, reproductive success and female choice in African elephants. *Animal Behaviour* 37: 842–849. https://doi.org/10.1016/0003-3472(89)90068-7

Poole JH, Payne K, Langbauer WR, Moss CJ. 1988. The social contexts of some very low frequency calls of African elephants. *Behavioral Ecology and Sociobiology* 22: 385–392. doi.org/10.1007/BF00294975