

DOI:10.1145/3231588

The system transforms raw telemetric data into engaging and informative blog texts readily understood by all.

BY ADVAITH SIDDHARTHAN, KAPILA PONNAMPERUMA, CHRIS MELLISH, CHEN ZENG, DANIEL HEPTINSTALL, ANNIE ROBINSON, STUART BENN, AND RENÉ VAN DER WAL

Blogging Birds: Telling Informative Stories About the Lives of Birds from Telemetric Data

BLOGGING BIRDS is a novel artificial intelligence program that generates creative texts to communicate telemetric data derived from satellite tags fitted to red kites — a medium-size bird of prey — as part of a species reintroduction program in the U.K. We address the challenge of communicating telemetric sensor data in real time by enriching it with meteorological and cartographic data, codifying ecological knowledge to allow creative interpretation of the behavior of individual birds in respect to such

enriched data, and dynamically generating informative and engaging data-driven blogs aimed at the general public.

Geospatial data is ubiquitous in today's world, with vast quantities of telemetric data collected by GPS receivers on, for example, smartphones and automotive black boxes. Adoption of telemetry has been particularly striking in the ecological realm, where the widespread use of satellite tags has greatly advanced our understanding of the natural world.^{14,23} Despite its increasing popularity, GPS telemetry involves the important shortcoming that both the handling and the interpretation of often large amounts of location data is time consuming and thus done mostly long after the data has been gathered.^{10,24} This hampers fruitful use of the data in nature conservation where immediate data analysis and interpretation are needed to take action or communicate to a wider audience.^{25,26}

The widespread availability of GPS data, along with associated difficulties interpreting and communicating it in real time, mirrors the scenario seen with other forms of numeric or structured data. It should be noted that the use of computational methods for data analysis per se is hardly new; much of science depends on statistical analysis and associated visualization tools. However, it is generally understood that such tools are mediated by human operators who take responsibility for identifying patterns in

» key insights

- **Summarizing environmentally enriched satellite-tag data in the form of informative, engaging, and fluent blogs is a challenge, even for trained ecologists, and computer-generated blogs were preferred by readers.**
- **Natural language generation, specifically data-to-text technology, is sufficiently advanced to achieve more than just factual summarization of data for professional use.**
- **This development opens new avenues for addressing societal challenges related to communicating data effectively and engaging the public with scientific research.**



IMAGE BY ANDRZJ BORYS ASSOCIATES/SHUTTERSTOCK

data, as well as communicating them accurately. An important but relatively recent addition to the growing field of data science is a technology called natural language generation¹⁵ that automates the entire data pipeline to produce textual reports from data, whether numeric or structured. Originally developed to offer decision support in the workplace, natural language generation has generated textual summaries of technical data for professionals, including engineers, nurses, and oil-rig workers,^{5,9,13,21} and is increasingly mainstream. Gartner, Inc. forecast in 2017 that 90% of business intelligence systems will incorporate natural language generation by 2019.¹¹ Companies like Arria (<https://www.arria.com/>), Narrative Science (<https://narrativescience.com/>), and Automated Insights (<http://automatedinsights.com/>) have developed

software that summarizes data as textual reports; indeed, print-media organizations are increasingly turning to robo-journalism, and many routine data-driven news stories that are time consuming and mundane for professional journalists to write are being written entirely by computer programs. Such data-to-text applications require accuracy and clarity first and foremost, and it has been noted that for workplace applications consistency in language use is the main reason why computer-generated output is preferred to text produced by humans.¹⁶

At the other end of the spectrum of computer-generated language is the discipline of computational creativity, whereby computer programs attempt to construct jokes,¹ short stories,⁷ and poetry.⁸ Here, we use the term “creativity” in the context of “creative writing,”

defined by the *Oxford Dictionary* as “writing, typically fiction, or poetry, which displays imagination or invention (often contrasted with academic or journalistic writing).”^a It is frequently said that creativity, especially in relation to design, requires the work to not just be imaginative or inventive but also “appropriate,” as in Sternberg.¹⁹ In his account of writing as design, Sharples¹⁸ related the idea of appropriateness to “constraints,” which provide the framework and context for creative expression and can be imposed either by the literary genre or by the conceptual space in which the writer is working.

Computer programs for computational creativity use static knowledge sources, typically manually construct-

a https://en.oxforddictionaries.com/definition/creative_writing

ed, to source joke templates, narrative plots, story grammars, and characters. In the storytelling domain, creativity manifests itself through emergent narratives dynamically created through the interactions of characters modeled as intelligent agents,²⁰ construction of different narratives from the same underlying plot representation,¹⁷ or the tailoring of linguistic components to generate human-like narrative prose.³ Deep neural networks have recently been applied to the generation of poetry by predicting likely word sequences fitting a mood or theme while also modeling tonal and structural constraints imposed by specific genres like Chinese quatrains.^{28,29}

Earlier work exists on communicating spatiotemporal data in the form of stories²² to help children with complex communication needs describe their school day to their parents. Here, microphone and radio frequency identification (RFID) readers were mounted on wheelchairs to make audio recordings by teachers or interactions with RFID-tagged locations, people, and objects. In this work, the computer-generated text was restricted to a factual summary of interactions recorded by RFID, while creativity was incorporated either through voice recordings provided by teachers or through functionality that

allowed the children to personalize their stories by editing system output.

The body of work summarized earlier either generates factual reports from real-world data, with creativity introduced through direct human input²² or generates creative texts from formal representation without recourse to real-world data. We are unaware of any previous computer program that generates creative texts from real-world data without human input. Addressing this gap, we describe Blogging Birds, which we designed to generate creative texts from data generated by satellite tags fitted to animals. The focal species for Blogging Birds is the red kite (*Milvus milvus*). This bird of prey was once widespread in the U.K., but prolonged and intense persecution led to its near extinction by the 1940s.

In 1989, the Royal Society for the Protection of Birds (RSPB) started a scheme to reintroduce the species in various locations across the U.K.⁴ In one of these locations, the Black Isle near Inverness in the north of Scotland, several birds were equipped with solar-powered satellite tags. Limited human resources meant the tags were used mainly to locate birds that had died to foster detection and prosecution of possible wildlife crimes. However, it was felt there was scope for us-

ing data from these tags for public-engagement activities surrounding the reintroduction initiative, to communicate ecological insight that enhances people’s understanding of the species, and to create a positive image of the species to harness public support for the reintroduction.²⁴ RSPB staff were themselves also keen to gain a better understanding of the lives of reintroduced birds, particularly how they recolonized a landscape that held precious few red kites for well over a century. They appreciated the inherent limitations in the data generated by the tags and were open to imaginative interpretations of the data, so long as the behaviors being narrated were ecologically plausible.

These requirements allowed us an opportunity to investigate data-driven generation of creative texts by computers, something we believe Blogging Birds is so far unique in its ability to achieve. The generated texts are creative in that they display imagination and inventiveness in how they interpret and report data under constraints imposed by kite ecology and the data itself. We sought to answer two research questions through experiments: Would the computer-generated blogs be well perceived by readers in comparison to blogs written by humans based on the same data?; and How important would the creative narration of ecological insight be to readers’ perceptions of computer-generated blogs?

The Blogging Birds System

The starting aim of Blogging Birds was to bring satellite-tagged individuals of a species (such as the red kite) “to life” by constructing ecologically sound narratives describing their movements. Conservationists fitted satellite tags—PTT-100 22-gram Solar Argos/GPS

Figure 1. System architecture.

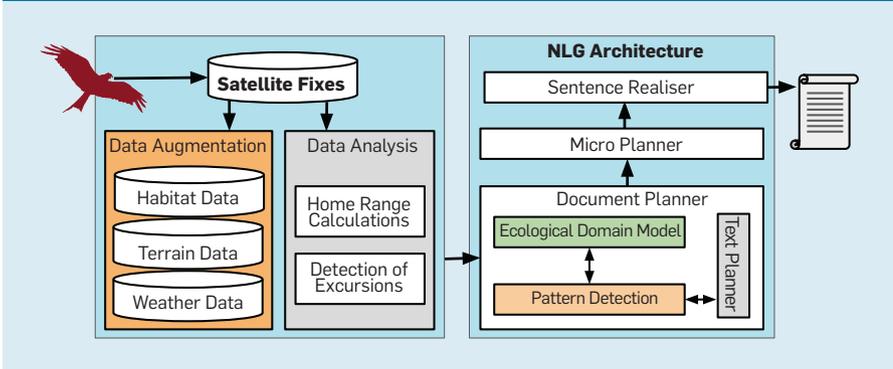


Table 1. Example augmented data used for pattern mining for one day of one week for a particular bird.

Day of Week	Hour	Habitat	Weather	Temp (C°)	Visibility (meters)	Wind Speed (miles per hour)	Location	Features	Distance Flown (miles)	Other Kites
Friday	08:00	Coniferous woodland	Overcast	13.0	24,000	3	East Croachy	Loch Ruthven	0	
Friday	10:00	Rough grassland	Heavy rain	13.9	5,000	2	Torness	Loch Ruthven	4	
Friday	12:00	Rough grassland	Heavy rain	16.0	3,600	2	Torness	Loch Ruthven	2	
Friday	14:00	Rough grassland	Heavy rain	16.0	3,600	2	Torness	Loch Ruthven	2	Merida
Friday	16:00	Improved grassland	Overcast	18.4	45,000	2	Torness		3	

PTT—to red kite chicks immediately prior to fledging, using a backpack harness designed for minimal hindrance. The tags were solar-powered and programmed to record up to six location fixes per day. Although this maximum could indeed be achieved during the summer months, a lack of sunlight in Scotland meant fewer fixes (a maximum of four per day) were obtained in spring and autumn and only the occasional fix during winter. To further preserve battery power, data was transmitted from the tag to the satellite only once per week. We thus configured Blogging Birds to produce a blog every week, or each time data was received from a bird.

Figure 1 outlines the overall architecture of the Blogging Birds system. We next describe the main components; see also Ponnampetuma et al.¹²

Data augmentation. The system processes an email messages with GPS fixes from the tags fixed to the red kites and enriches that data from readily available online sources about the local weather (<https://www.metoffice.gov.uk/datapoint>), habitat (such as different types of grassland and forests, <https://eip.ceh.ac.uk/lcm>), and geographic features (such as rivers, lochs, roads, and location names, <https://www.ordnancesurvey.co.uk/>). Table 1 presents a sample of the enriched data used by Blogging Birds.

Data analysis. The system then applies data-analysis procedures for identifying home ranges and patterns of movement with respect to these temporary settlement areas. Home ranges are identified as polygons using the Adehabitat package for R² by clustering the previous locations of an individual using 90% kernels. As described by van der Wal et al.,²⁴ we modeled local movement patterns as angular and radial velocity vectors to identify excursions, characterized by travel in relatively straight lines at higher speeds. This data analysis allows the document planner (described next) to detect the three prototypical patterns of movement in Figure 2, whereby the kite remains within a home range, explores an area outside its home ranges, or moves from one home range to another. Figure 3 shows the calculated home ranges for a bird (gray polygons), as well as the fixes classified as excursions

Figure 2. Prototypical red kite movement patterns: C1 is small and constricted movements within an area of intense usage (home range); C2 is exploratory movement from a home range (round trip); and C3 is direct movements between separate home ranges.

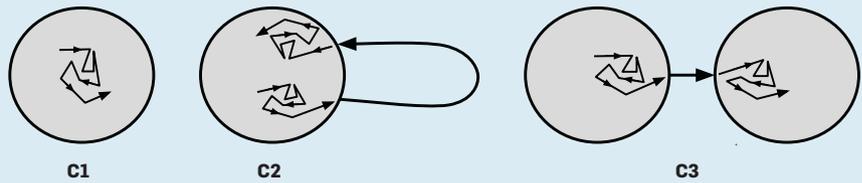
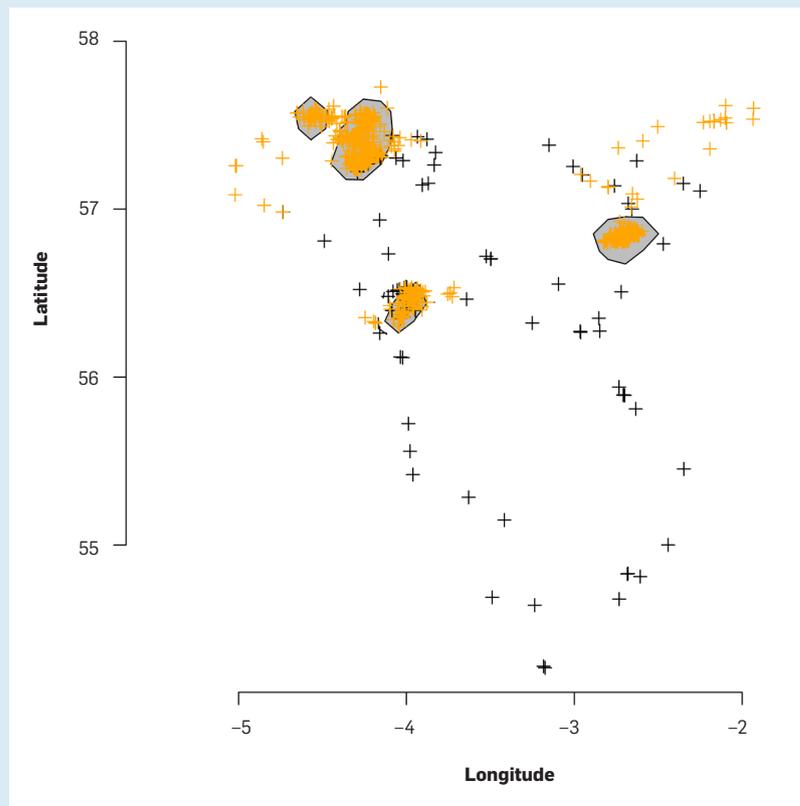


Figure 3. Calculated home ranges (gray polygons) and classification of fixes as excursions (black crosses) or non-excursions (amber crosses) for a particular bird.



(black crosses) and non-excursions (amber crosses).

Document planner. The document planner in Figure 3 identifies patterns in the data that signal different red kite behaviors and creates “messages” (implemented as Java classes) that encode these behaviors for use by the “micro planner” and “sentence realiser,” which then generate sentences in English.

The data analysis allows us to detect the three prototypical patterns of movement outlined in Figure 2, whereby the kite remains within a home range, explores an area outside its

home ranges, or moves from one home range to another. An ecological-domain model further defines different travel, foraging, and social behaviors as rules that can apply under specific environmental and geographic conditions; for instance, following heavy rain, a kite observed on any of the grassland habitats might feed on earthworms or a kite observed near a woodland habitat late in the afternoon is likely to be preparing to roost. These rules are implemented as JBoss Drools (<http://www.jboss.org/drools>), a business-logic-integration platform that allows us to instantiate messages when

Figure 4. Screenshot of the Blogging Birds Web interface.



particular patterns are detected in the data. In total, the system implements Drools for 26 movement behaviors (such as flying along a coast or over a landmark like a castle or loch and the home-range-related movement patterns in Figure 2); 33 foraging behaviors, mostly detailing the food available for a kite in different habitats at different times of the year but also sometimes related to specific features (such as when a red kite near a road might be looking for roadkill); and six social behaviors (such as roosting and nesting); see the online appendix “Example Rules” (dl.acm.org/citation.cfm?doi=3231588&picked=formats).

The pattern-detection module then exhaustively applies the rules to the satellite fixes to produce a list of all observed movement behaviors and all possible foraging and social behaviors consistent with known environmental and geographic conditions. The latter is the first step in the creative process, whereby the program explores the conceptual space to “imagine” how the kite might have been behaving.

Blogging Birds uses a rule-based text planner for dynamic text generation. The planning rules decide how information is ordered, but what information to include and how to organize it into sentences is determined at runtime in a data-driven manner.

The blogs are always planned as three paragraphs, the first describing the overall trends, the second providing more detail on a day-to-day basis, and the third posing a question about what the kite might do next, as well as occasionally offering a conclusion.

The content is selected through a process of summarization and aggregation of information. This is the second creative aspect of the blog generation (the first involved imagining a wide range of possible behaviors), as it plans what story to tell from the imagined behaviors. Blogging Birds aims to provide an overview of the main behaviors and highlight aspects that might be interesting to the human reader. Movement behaviors are considered more interesting than foraging behaviors, and rarer foraging behaviors are prioritized over more frequent ones. Each blog attempts to inform the reader about different aspects of red kite ecology by selecting different behaviors from different days. The main steps are as follows:

Paragraph 1

Movement pattern. Generate a message based on the detected movement pattern—C1, C2, or C3 in Figure 2; if the age of the bird can be used to interpret this pattern, add such an interpretation message;

Habitats visited. Generate a message summarizing the habitats visited; and

Other kites. Generate a message about other kites recorded nearby, if any.

Paragraph 2

Days of the week. Iterate over each day of the week (Monday to Friday):

► If the bird remained relatively static—C1 in Figure 2—then generate a message about nearby places or generate a message about any movement behavior detected; and

► Generate a message about a new (not previously used) possible foraging behavior, if any deduced; unusual (historically infrequent) behaviors are selected over common ones.

Remove redundancy. Aggregate the messages generated for the week through these two steps to remove redundancy (such as by grouping together days with similar behaviors).

Paragraph 3

Movement pattern. Generate a message for a question or comment based on the movement pattern—C1, C2 or C3—with the aim of intriguing the reader.

Micro planner and sentence realiser. The micro planner takes the messages generated by the document planner, implements aggregation through a variety of linguistic devices (such as ranges, coordination, and subordination), and limits linguistic repetition by varying the vocabulary. It provides sentence specifications to the “sentence realiser,” which then generates sentences using the SimpleNLG library.⁶

Figure 4 is a screenshot of the Blogging Birds interface in which an automatically generated weekly blog for a kite is overlaid on a Google map of the bird’s whereabouts marked with blue polygons. In this example, Wyvis, one of five red kites being blogged about, has traveled between two home ranges (movement pattern C3), and an explanation for the observed movement pattern is provided based on the age

Table 2. Baseline computer-generated blog without reference to ecological concepts for the week outlined in Figure 4.

Wyvis had enough of the area around Teavarran and decided to move to Crieff approximately 73 miles away. No doubt Wyvis had a social week, as kites Moray and Millie were often seen in the vicinity.

Monday, Wyvis spent most of her time around Torness, Errogie, and Teavarran. On Tuesday evening, she reached moorland near Crieff flying approximately 65 miles amid cloudy conditions and averaging a remarkable 11 miles per hour. The next five days Wyvis spent most of her time around Edinample, Tullybannoch, and St Fillans. During this time, she was seen mainly on acid grassland while making occasional journeys to farmland.

Will Wyvis settle down here?

of the bird. The system emphasizes the social side with reference to roosting and encounters with other tagged kites. The second paragraph is narrated chronologically, with care taken again to emphasize any unusual behaviors (such as the long distance flown on Tuesday) and to reference weather conditions (“cloudy”) to make the text more engaging. Information is also provided about the foraging potential of the different habitat types visited. Aggregation is used to avoid repetition, using linguistic devices (such as range “Wednesday to Sunday,” coordination “St Filans, Tullybannoch, and Edinample,” and subordination “mainly on acid grassland, while making odd journeys to arable land”). The question posed in the final paragraph is selected based on the movement pattern detected.

Here, we focus on situations where the timeframe covered by each blog is set at one week, as this is the frequency at which the tags are programmed to transmit data. However, the system architecture is sufficiently generic to be able to handle other timeframes, and the interface also allows the user to select a day of the week and read a blog composed for that day. Blogs could in theory also be provided for longer timeframes, but as the goal of the project was to allow readers to monitor or follow the birds on a continuous basis, this option was not implemented.

Evaluation

We investigated both how computer-generated blogs are appraised by readers in comparison to human-written blogs based on the same data and the contribution of the generated ecological insights to such appraisals. To this end, we designed studies to evaluate the quality of the computer-generated blogs for different patterns of movement, first through comparison with blogs written manually, then through comparison with baseline computer-generated blogs that report the data factually without ecological insights.

Method. We focused on the three prototypical movement patterns outlined in Figure 2 as conditions C1, C2, and C3. For each condition, we identified 12 weeks of data such that the focal red kite’s movements broadly matched this condition (for example, the week



The generated texts are creative in that they display imagination and inventiveness in how they interpret and report data under constraints imposed by kite ecology and the data itself.



in Figure 4 would correspond to C3), giving us 36 weeks of data in total.

Comparison with human-written blogs. We recruited 12 post-graduate master’s-level ecology students from the University of Aberdeen in Scotland (representative of those who might be hired by a conservation charity) to take part in a two-hour session on “digital media in nature conservation” outside teaching hours. We told them they would be writing three short blogs on the basis of environmental data we would provide, saying it would take them approximately 1.5 hours, that partaking would benefit our research while giving them unique insights into new technologies, and we would compensate them £15 cash to express our gratitude for helping us while learning.

We provided each writer with access to a one-page information sheet about red kites that summarized the typical movement patterns and foraging and social behaviors that were encoded in the Blogging Birds system. They were also free to consult any online sources they preferred. We also provided them with the enriched data available to the system for the week, presented in both tabular form (as in Table 1) and overlaid on a map showing home ranges and fixes (as in Figure 4, but without the blog). The information we provided to the 12 student writers was sufficient to allow them to make the same inferences as the system. However, in order to grant full creative freedom to the writers and avoid priming them to write similar blogs to the system, we avoided giving them direct access to the inferences made or used by the system. They were further informed about the intended purpose of the blogs and the target audience, and each was asked to write three 200-word blogs; that is, for data from three different weeks, one in each condition (C1–C3 in Figure 2) such that for each of the 36 weeks selected for the study we had one manually written blog. The order in which writers encountered each condition was randomized and writers not made explicitly aware of the existence of these conditions in the study, though the patterns were clearly visible on

the respective maps and described on the information sheet. These 36 manually written blogs were compared to computer-generated blogs for the same weeks in the evaluation.

As our goal was to investigate Blogging Birds not just as a tool for those with an interest in nature conservation but as a resource to engage those interested in new technologies. We ran evaluations with two distinct groups of participants: 93 undergraduate biology students enrolled in a second year “community ecology” course and 49 first- and second-year undergraduates from across disciplines enrolled in a course entitled “digital society,” both at the University of Aberdeen. In each trial, a

participant sitting at an individual workstation was shown an interface with a map with home ranges and fixes of a kite for one of the weeks, as well as two blogs, one written manually and one computer-generated, without any information about their provenance. Participants said what blog they preferred (or expressed no preference) and also rated each blog on how informative, fluent, and engaging they found it on a seven-point Likert scale. Each participant evaluated three pairs of blogs. We designed the study to test three specific hypotheses:

- H1. Computer-generated blogs are preferred to human-written blogs;
- H2. Computer-generated blogs are

more informative, fluent, and engaging than human-written blogs; and

H3. The differences in ratings for computer-generated and human-written blogs are conditional on the movement pattern of an individual bird C1, C2, or C3, as in Figure 2.

Comparison with baseline. To directly evaluate whether communicating ecological insights through the blogs is important to readers, we compared Blogging Birds to a computer-generated baseline that blogs about the movement patterns without reference to ecological concepts; see Table 2 for an example. These baseline blogs were entirely factual and reported behaviors only directly observed in the data but that otherwise followed the same format as the full-system blogs. An additional 27 undergraduate students enrolled in the digital society course, but who had not participated in the earlier experiment, evaluated the full vs. the baseline system using the same methodology and interface as before. We designed this study to test two specific hypotheses:

- H4. Computer-generated blogs with ecological insights are preferred to baseline computer-generated blogs without ecological insights; and
- H5. Computer-generated blogs with ecological insights are more informative and engaging than baseline computer-generated blogs without ecological insights, while their fluency is comparable.

Figure 5. Preferences for human-written and for computer-generated blogs by movement condition, as in Figure 2: C1 is movement within a home range; C2 is a round trip; and C3 is movement between home ranges.

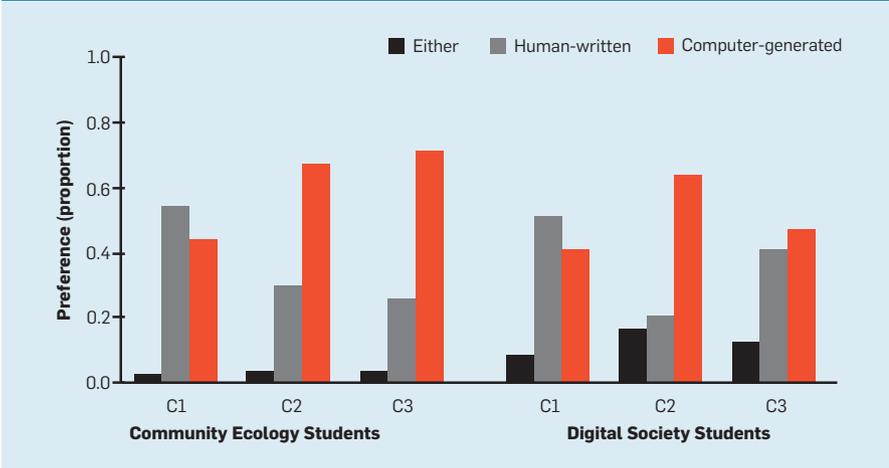
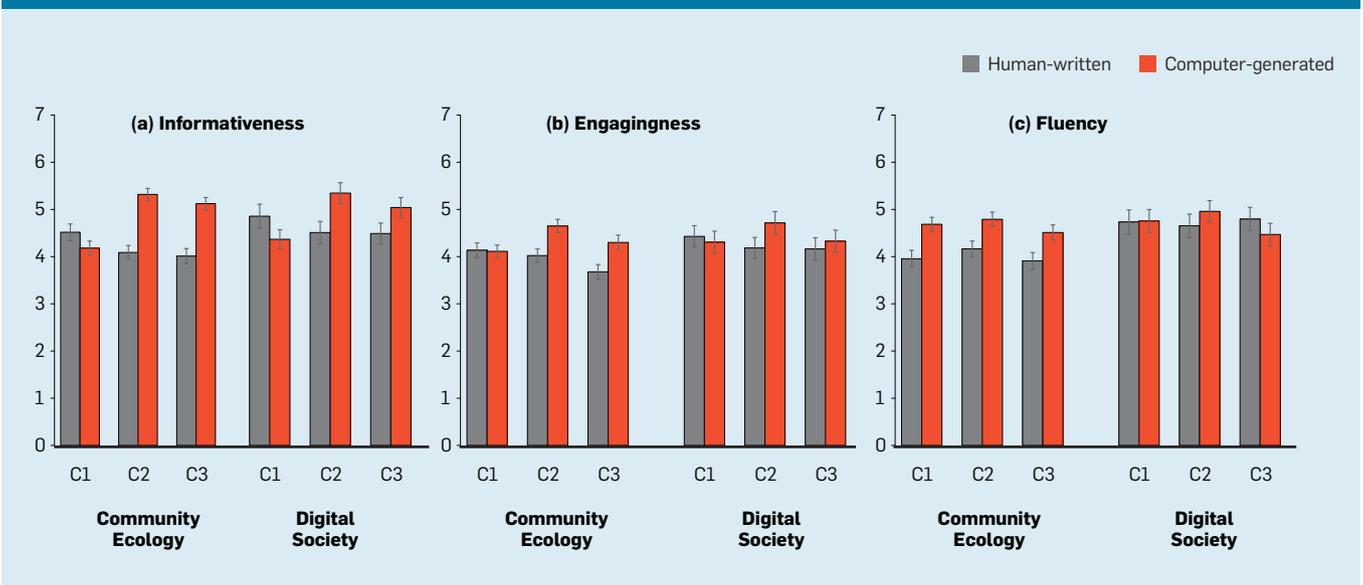


Figure 6. Average ratings for human-written and for computer-generated blogs by movement condition, as in Figure 2: C1 is movement within a home range; C2 is a round trip; and C3 is movement between home ranges.



Results

Evaluation against human-written blogs. Both sets of students showed an overall significant preference for the computer-generated blogs (238 trials vs. 153 trials in which human-written blogs were preferred; $\chi^2 = 18.5$; $p < 0.001$), confirming hypothesis H1. However, a more complex pattern emerged (see Figure 5), with this preference being dependent on the type of kite movement covered in the blog—C1, C2, or C3—and the orientation of the course—ecology or technology.

Across the community ecology students, there was a strong preference for computer-generated blogs when they captured more extensive movement by the kites—round trips (C2) and movement between home ranges (C3)—while there was little difference in preference between the two blog types when kite movement was limited; that is, small movements within home ranges (C1). Digital society students showed an overall clear preference for the computer-generated blogs only when they described round trips (C2). Combined, our findings indicate Blogging Birds is particularly skilled at handling cases where the focal bird shows substantial movement. Average ratings for how fluent, engaging, and informative the blogs were (see Figure 6) showed the main perceived advantage of the computer-generated blogs pertains to their “informativeness,” with smaller improvements visible for how engaging and fluent they were.

We ran a MANOVA, with informativeness, engagingness, and fluency as the dependent variables and blog type (computer or human), kite movement-pattern (C1, C2, or C3), student group (community ecology or digital society), and their interactions as fixed effects, and writer ID and evaluator ID as random effects. We found the following main effects and interactions at $p < 0.01$: computer-written blogs were rated significantly higher ($p < 0.0001$) than human-written blogs (confirming hypothesis H2); students in the digital society course gave higher ratings overall than students in community ecology ($p < 0.01$); and there was interaction between blog type and movement pattern ($p < 0.0001$), confirming hypothesis H3. Post-hoc analysis using the Tukey HSD test on the individual ANOVAs with Bonferroni-correction re-

vealed this interaction came about because the computer-generated blogs capturing conditions with more movement by kites (C2 and C3) were more informative than the human-written blogs for the same conditions and more informative than computer-generated blogs capturing constricted movement (C1) ($p < 0.0001$ for each comparison).

To better understand these described effects, we compared the distribution of ratings obtained by each human writer (H1–H12) and the computer (Comp) in Figure 7. Only two of the blog writers (H3 and H10) were deemed to write more informative blogs than the computer, and both of them were considered less engaging and fluent than the computer-generated blogs. Likewise, H4, who wrote more fluent and engaging blogs than the computer, was rated rather low

Figure 7. Computer-generated blogs (Comp) vs. human-written blogs (H1–H12).

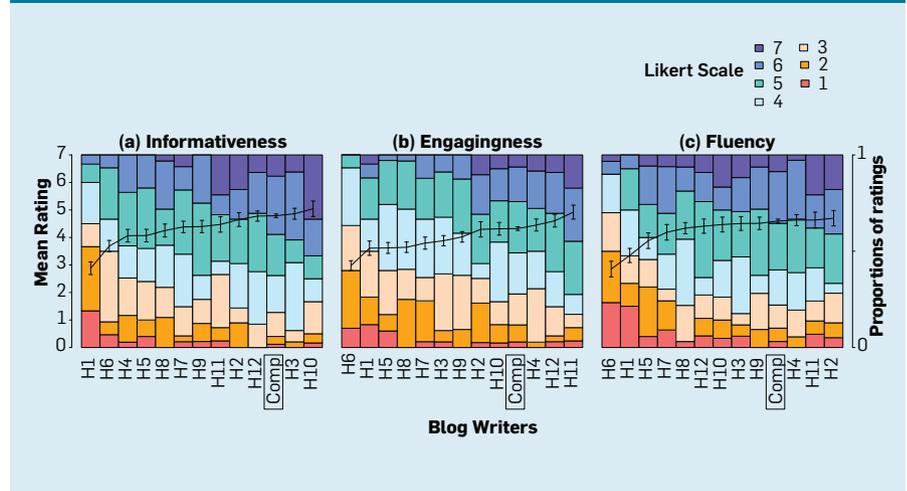
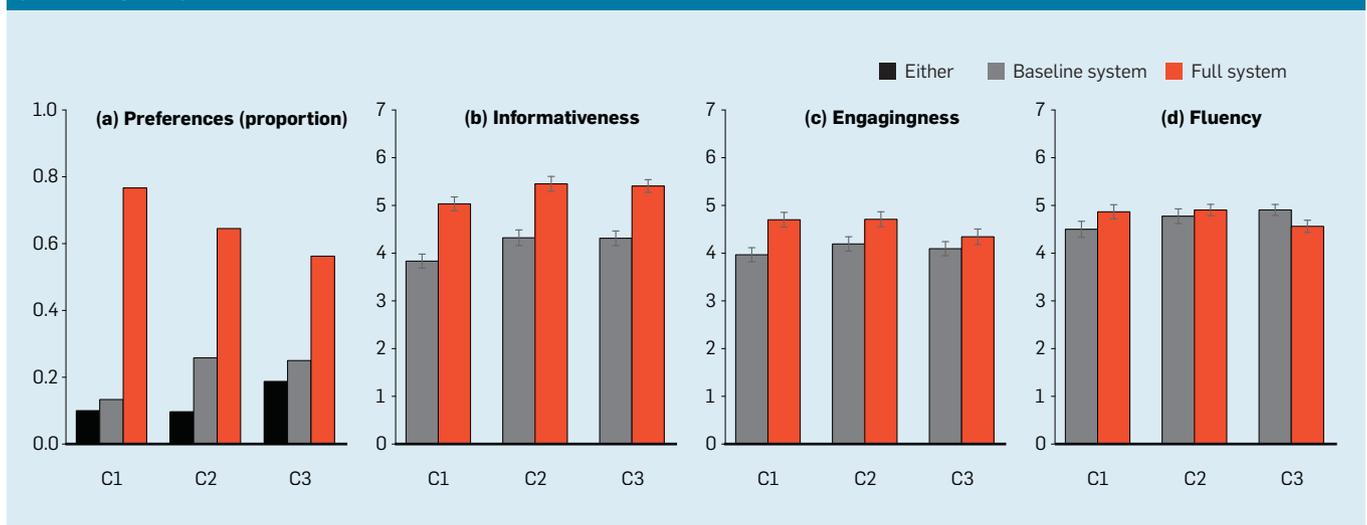


Figure 8. Computer-generated blogs with ecological insights (full system) vs. computer-generated blogs describing movement patterns only (baseline system).



for “informativeness,” thus illustrating the difficulty of being informative, engaging, and fluent at the same time, even for a human writer. Indeed, all the writers were committed and used the full 1.5 hours for composing the blogs, yet most were outperformed by the computer on each of the three metrics. For examples of human-written and computer-generated blogs, as well as details of how they were appraised by evaluators, see the online appendix.

A questionnaire filled out by the blog writers provided many interesting insights. In general, they found it difficult to comprehend and summarize the sheer amount of data in fewer than 200 words but also felt the process became easier the more they did. There was, however, concern from many that the blogs were becoming repetitive, especially if there was little variation in what the red kites were actually doing, stemming largely from a lack of knowledge of kite ecology and behavior. Summarizing the range of data in different formats was certainly challenging, and some enjoyed the process more than others. There was considerable variability in how the blog writers used the materials provided them to create the blogs. Some concentrated mostly on the visible patterns on Google maps, others looked in more detail at the map data by clicking on individual map points to find out more, and yet others found inspecting the data in a tabular format was most useful. Asked whether they would like to write the red kite blogs as a job, the consensus was that, although initially enjoyable, it would quickly get tedious and increasingly more difficult to write non-repetitive material.

Evaluation against baseline. Participants demonstrated a conclusive preference for the full system with ecological insights, preferring it in 61 trials compared to only 20 trials in which the baseline was preferred ($\chi^2 = 21.5$; $p < 0.001$), confirming hypothesis H4. Interestingly, this effect was strongest when blogs described situations with little movement by the birds during those weeks (C1); here, the full-system blogs were preferred in 23 trials compared to just four baseline blogs ($\chi^2 = 13.4$; $p=0.0002$). For C2 and C3, the corresponding values were preferences for the full system in 20 and 18 trials, compared to preferences for the baseline in eight trials each ($\chi^2 =$



Telemetric data is ubiquitous, captured by smartphones and other mobile devices, as well as through GPS sensors embedded in vehicles used by the transportation industry and others.



5.1, 3.8; $p = 0.0233, 0.0499$). The absence of ecological interpretation by the baseline system was thus judged adversely for all movement patterns, particularly so when the birds were relatively static. We also found the full blogs were rated as more informative ($p < 0.0001$) and more engaging ($p=0.0215$) but not more fluent ($p=0.825$) (see Figure 8), confirming hypothesis H5.

The two studies we have presented here demonstrate that computer-generated blogs are appraised more positively than human-written blogs, and that computer-generated blogs with creatively generated ecological insights are preferred overwhelmingly to blogs generated from the same data but without inclusion of these insights.

Conclusion

The Blogging Birds system shows that raw satellite tag data can be transformed into fluent, engaging, and informative texts directed at members of the public and in support of nature conservation.

We demonstrated that computers can compete with human experts in generating creative stories from numerical data. Unlike natural language generation systems that generate texts for news reporting or for decision making in the workplace, Blogging Birds’s narratives are not entirely factual. Though the system is constrained by the observed data and its ecological domain model, the red kites’ reported foraging and social behaviors are only *imagined* to have taken place. However, including these behaviors in the narratives allows us to communicate red kite ecology to the reader, and the blogs are better appraised as a consequence. Our work thus simultaneously addresses the societal challenges of communicating data effectively and engaging the general public with scientific research.

Blogging Birds composes blogs by combining texts produced through three different types of analysis: The first is a generic factual summarization of telemetric data enriched with location-specific information about weather conditions, habitat type, and geographic features, and can be readily adapted for use in other domains. The second is the processing and ecological interpretation of movement

data in the context of home range use, and the third is the exploitation of domain knowledge encoded as a collection of rules that help the system imagine possible foraging and social behaviors from environmental and geographic parameters. Much of what is creative and interesting about the blogs derives from the latter domain-specific types of data analyses. Although the developed principles apply more broadly, new applications would require construction of knowledgebases pertinent to the domain of use. While this is a clear limitation of our approach, note our ecological interpretation of movement data in particular would be applicable to several other species. For example, we have already developed a version of Blogging Birds for golden eagles (*Aquila chrysaetos*) for use by RSPB conservation officers, successfully reusing the second, as well as the first, type of analysis.

During the course of the project, we also discovered ecologists had limited knowledge of the foraging behavior of red kites in Scotland, as they had not been studied extensively following their relatively recent reintroduction. We could thus encode only a limited number of rules per habitat type. The absence of any large-scale corpus of texts in this domain also meant we could not apply the deep learning methods that are rapidly gaining popularity for generating linguistic variation in computer-generated texts.²⁷ In future work, we plan to invite Blogging Birds' users to contribute behavioral observations from across the U.K., enabling us to simultaneously curate a larger set of rules and further public engagement.

Finally, our ideas demonstrated here are applicable more generally. Telemetric data is ubiquitous, captured by smartphones and other mobile devices, as well as through GPS sensors embedded in vehicles used by the transportation industry and others. Even albums of time-stamped and geo-tagged photos provide data similar to what we used here. The nature of the blogs, along with the information sources used for data enrichment, would depend on the application, to blog about a holiday or reveal the provenance and journey of a food item in a

supermarket. In effect, we have demonstrated it is possible to blog about such data through a process of data enrichment and natural language generation, opening up new avenues for using AI to engage people through data.

Acknowledgments

This research was supported by an award from the RCUK Digital Economy Programme to the dot.rural Digital Economy Hub, award reference EP/G066051/1. C

References

1. Binsted, K. and Ritchie, G. Computational rules for generating punning riddles. *International Journal of Humor Research* 10, 1 (July 1997), 25–76.
2. Calenge, C. The package 'adehabitat' for the R software: A tool for the analysis of space and habitat use by animals. *Ecological Modelling* 197, 3 (Apr. 2006), 516–519.
3. Callaway, C.B. and Lester, J.C. Narrative prose generation. *Artificial Intelligence* 139, 2 (Aug. 2002), 213–252.
4. Carter, I. *The Red Kite*. Arlequin Press, Chelmsford, Essex, U.K., 2007.
5. Gatt, A., Portet, F., Reiter, E., Hunter, J., Mahamood, S., Moncur, W., and Sripada, S. From data to text in the neonatal intensive care unit: Using NLG technology for decision support and information management. *AI Communications* 22, 3 (third quarter 2009), 153–186.
6. Gatt, A. and Reiter, E. SimpleNLG: A realisation engine for practical applications. In *Proceedings of the 12th European Workshop on Natural Language Generation* (Athens, Greece, Mar. 30–31). Association for Computational Linguistics, Stroudsburg, PA, 2009, 90–93.
7. Gervás, P. Computational approaches to storytelling and creativity. *AI Magazine* 30, 3 (Fall 2009), 49–62.
8. Ghazvininejad, M., Shi, X., Choi, Y., and Knight, K. Generating topical poetry. In *Proceedings of Empirical Methods in Natural Language Processing* (Austin, TX, Nov. 1–5). Association for Computational Linguistics, Stroudsburg, PA, 2016, 1183–1191.
9. Goldberg, E., Driedger, N., and Kittredge, R.I. Using natural language processing to produce weather forecasts. *IEEE Expert* 9, 2 (Apr. 1994), 45–53.
10. Hebblewhite, M. and Haydon, D.T. Distinguishing technology from biology: A critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365, 1550 (July 2010), 2303–2312.
11. Panetta, K. *Neural Networks and Modern BI Platforms Will Evolve Data and Analytics*. Gartner, Inc., Stamford, CT, Jan. 16, 2017; <http://www.gartner.com/smarterwithgartner/nueral-networks-and-modern-bi-platforms-will-evolve-data-and-analytics/>
12. Ponnampuruma, K., Siddharthan, A., Zeng, C., Mellish, C., and Wal, R. Tag2Blog: Narrative generation from satellite tag data. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (Sofia, Bulgaria, Aug. 4–9). Association for Computational Linguistics, Stroudsburg, PA, 2013, 169–174.
13. Portet, F., Reiter, E., Gatt, A., Hunter, J., Sripada, S., Freer, Y., and Sykes, C. Automatic generation of textual summaries from neonatal intensive care data. *Artificial Intelligence* 173, 7–8 (May 2009), 789–816.
14. Pschera, A. *Animal Internet: Nature and the Digital Revolution*. New Vessel Press, New York, 2016.
15. Reiter, E. and Dale, R. *Building Natural Language Generation Systems*. Cambridge University Press, Cambridge, U.K., 2000.
16. Reiter, E., Sripada, S., Hunter, J., Yu, J., and Davy, I. Choosing words in computer-generated weather forecasts. *Artificial Intelligence* 167, 1–2 (Sept. 2005), 137–169.
17. Rishes, E., Lukin, S.M., Elson, D.K., and Walker, M.A. Generating different story tellings from semantic representations of narrative. In *Proceedings of the International Conference on Interactive Digital Storytelling* (Istanbul, Turkey, Nov. 6–9) Springer, New York, 2013, 192–204.

18. Sharples, M. An account of writing as creative design. In *The Science of Writing*. Lawrence Erlbaum, Hillsdale, NJ, 1996.
19. Sternberg, R.J. *Handbook of Creativity*. Cambridge University Press, Cambridge, U.K., 1999.
20. Theune, M., Faas, S., Heylen, D.K.J., and Nijholt, A. The virtual storyteller: Story creation by intelligent agents. In *Proceedings of the Conference on Technologies for Interactive Digital Storytelling and Entertainment*, S. Göbel et al., Eds. (Darmstadt, Germany, Mar. 24–26). Fraunhofer IRB Verlag, Stuttgart, Germany, 2003, 204–215.
21. Theune, M., Klappers, E., de Pijper, J.-R., Kraemer, E., and Odijk, J. From data to speech: A general approach. *Natural Language Engineering* 7, 1 (Mar. 2001), 47–86.
22. Tintarev, N., Reiter, E., Black, R., Waller, A., and Reddington, J. Personal storytelling: Using natural language generation for children with complex communication needs, in the wild. *International Journal of Human-Computer Studies* 92 (Aug. 2016), 1–16.
23. Tomkiewicz, S.M., Fuller, M.R., Kie, J.G., and Bates, K.K. Global positioning system and associated technologies in animal behaviour and ecological research. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365, 1550 (July 2010), 2163–2176.
24. van der Wal, R., Zeng, C., Heptinstall, D., Ponnampuruma, K., Mellish, C., Ben, S., and Siddharthan, A. Automated data analysis to rapidly derive and communicate ecological insights from satellite-tag data: A case study of reintroduced red kites. *Ambio* 44, 4 (Oct. 2015), 612–623.
25. Verma, A., van der Wal, R., and Fischer, A. Microscope and spectacle: On the complexities of using new visual technologies to communicate about wildlife conservation. *Ambio* 44, 4 (Oct. 2015), 648–660.
26. Wall, J., Wittmyer, G., Klinkenberg, B., and Douglas-Hamilton, I. Novel opportunities for wildlife conservation and research with real-time monitoring. *Ecological Applications* 24, 4 (June 2014), 593–601.
27. Wen, T.-H., Gašić, M., Mrkšić, N., Su, P.-H., Vandyke, D., and Young, S. Semantically conditioned LSTM-based natural language generation for spoken dialogue systems. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (Lisbon, Portugal, Sept. 17–21). Association for Computational Linguistics, Stroudsburg, PA, 2015.
28. Yan, R. I. Poet: Automatic poetry composition through recurrent neural networks with iterative polishing schema. In *Proceedings of the International Joint Conference on Artificial Intelligence*. New York, July 9–15). AAAI Press, Palo Alto, CA, 2016, 2238–2244.
29. Zhang, X. and Lapata, M. Chinese poetry generation with recurrent neural networks. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (Doha, Qatar, Oct. 25–29). Association for Computational Linguistics, Stroudsburg, PA, 2014, 670–680.

Advait Siddharthan (advait.siddharthan@open.ac.uk) is a Reader in the Knowledge Media Institute at The Open University, Milton Keynes, U.K.

Kapila Ponnampuruma (kapila.ponnampuruma@arria.com) is the lead natural language engineer at Arria NLG plc, Aberdeen, Scotland, U.K.

Chris Mellish (c.mellish@abdn.ac.uk), now retired, was a professor of computer science at the University of Aberdeen, Scotland, U.K., at the time this research was conducted.

Cheng Zeng (zengc@hotmail.co.uk) was a research assistant on the Blogging Birds Project at the time this research was conducted.

Daniel Heptinstall (dheptinstall@gmail.com) is a senior international biodiversity adviser on the U.K. government's Joint Nature Conservation Committee.

Annie Robinson (annierobinson@abdn.ac.uk) was a Research Fellow on the Blogging Birds Project at the time this research was conducted.

René van der Wal (r.vanderwal@abdn.ac.uk) is a professor of ecology at the University of Aberdeen, Scotland, U.K.

Copyright held by authors.