Open Research Online

Blogging Birds: Telling informative stories about the lives of birds from telemetric data

Journal Item

For guidance on citations see FAQs.

Version: Accepted Manuscript

oro.open.ac.uk

Blogging Birds: Telling informative stories about the lives of birds from telemetric data

Advaith Siddharthan
Knowledge Media Institute
The Open University
Milton Keynes, UK
advaith.siddharthan@open.ac.uk

Kapila Ponnamperuma, Chris Mellish, Chen Zeng
Computing Science
University of Aberdeen, UK
k.ponnamperuma,c.zeng,c.mellish@abdn.ac.uk

Danny Heptinstall, Annie Robinson
Biological Sciences
University of Aberdeen, UK
d.heptinstall,annierobinson@abdn.ac.uk

Stuart Benn
Royal Society for the Protection of Birds (RSPB), Inverness, UK
stuart.benn@rspb.org.uk

René van der Wal
Biological Sciences
University of Aberdeen, UK
r.vanderwal@abdn.ac.uk

ABSTRACT

We introduce and evaluate Blogging Birds, a novel artificial intelligence program that generates creative texts to communicate telemetric data derived from satellite tags fitted to red kites as part of a species reintroduction programme. We address the substantial challenge of effectively communicating telemetric sensor data in real time by (a) enriching this data with meteorological and cartographic data, (b) codifying ecological knowledge to allow creative interpretation of the behaviour of individual animals with respect to such enriched data, and (c) dynamically generating informative and engaging data-driven blogs aimed at the general public. Our evaluation found that summarising environmentally-enriched satellite-tag data as informative, engaging and fluent blogs was highly challenging even for trained ecologists, and that the computer-generated blogs were preferred by readers, notably where movements of focal animals were wide-ranging. We conclude that Natural Language Generation, and specifically data-to-text technology, has come of age sufficiently to achieve more than just the factual summarisation of data for professional use, thus opening up new avenues for addressing societal challenges related to communicating data effectively and engaging the public with scientific research.

1 INTRODUCTION

Geospatial data is ubiquitous in today's world, with vast quantities of telemetric data collected by GPS receivers on, for example, smart phones or automotive black boxes. The 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 brings with it the important short-coming that both the handling and interpretation of often large amounts of location data is time consuming, and therefore mostly done long after the data have 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 to communicate to wider audiences [25, 26].

The widespread availability of GPS data, and the associated difficulties with interpreting and communicating them in real time, mirrors the scenario unfolding with respect to 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 visualisation tools. However, it is generally understood that such tools are mediated by human operators, who take responsibility for identifying patterns in 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 Generation (NLG) [15], which automates the entire pipeline to produce textual reports from data, numeric or structured. Originally developed to offer decision support in the workplace, NLG has been successful in generating textual summaries of technical data for professionals, such as engineers, nurses and oil rig workers [5, 9, 13, 21], and is now becoming more mainstream. Gartner, Inc. forecasts that 90% of business intelligence systems will incorporate NLG by 2019 [11]. Companies such as ARRIA NLG (http://www.arria.com), Narrative Science (https://www.narrativescience.com) and Automated Insights (http://automatedinsights.com/) already summarise numerical data, such as financial and accounting data or sports statistics, in the form of brief texts. Indeed, print media organisations are increasingly adopting robo-journalism, and many routine data-driven news stories that are time consuming and mundane for human journalists to write are now written entirely by computer programs. Such data-to-text applications require accuracy and clarity first and foremost, and it has been noted that for work place applications consistency in language use is the main reason why computer-generated output is preferred to that of humans [16].

At the other end of the spectrum of computer-generated language is the discipline of computational creativity, where computer programs attempt to construct, for example, jokes [1], short stories [7] or poetry [8]. In this article we use the term 'creativity' in the context of 'creative writing', defined in the Oxford Dictionary as "Writing, typically fiction or poetry, which displays imagination or invention (often contrasted with academic or journalistic writing)"[1]. It has frequently been noted that creativity, especially in relation to design, requires the work to not just be imaginative or inventive, but also 'appropriate' [e.g. 19]. In his account of writing as design, Sharples [18] relates 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 the writer works within.

Computer programs for computational creativity use static knowledge sources, typically manually constructed, to source joke templates, narrative plots, story grammars, or characters. In the storytelling domain, creativity manifests itself through emergent narratives dynamically created through the interactions of characters modelled as intelligent agents [20], the construction of different story tellings from the same underlying plot representation [17], or the tailoring of linguistic components to generate human-like narrative prose [3]. Recently, deep neural networks have been applied to poetry generation to generate likely word sequences fitting a mood or theme, while also modelling tonal and structural constraints imposed by specific genres such as Chinese Quatrains [28, 29].

There also exists previous work on communicating spatio-temporal data in the form of stories [22], to help children with complex communication needs describe their school day to their parents. A microphone and Radio Frequency Identification (RFID) readers were mounted on wheelchairs to record events such as 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. Creativity was incorporated either through voice recordings provided by teachers or through functionality that allowed the children to personalise their stories by editing system output.

The body of work summarised above either generates factual reports from real world data, with creativity introduced through direct human input in [22], 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 generates 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 UK, 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 this species in various locations across the UK [4]. In one of these locations, the Black Isle near Inverness in the north of Scotland, several individuals were equipped with solar-powered satellite tags. Human resource constraints meant that the tags were mainly used to locate birds that had died to foster the detection and prosecution of possible wildlife crimes. However, it was felt that there was scope

for using data from these tags for public engagement activities surrounding this reintroduction initiative, to communicate ecological insights that enhance 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 re-introduced birds and particularly how they recolonised 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 such data, so long as the behaviours narrated were ecologically plausible.

These requirements allowed us an opportunity to investigate the data-driven generation of creative texts by computers, and we believe Blogging Birds is so far unique in its ability to achieve this. 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: (a) would the computer-generated blogs be well perceived by readers in comparison to blogs written by humans from the same data, and (b) how important was the creative narration of ecological insights to readers' perceptions of computer-generated blogs.

2 THE BLOGGING BIRDS SYSTEM

The aim of Blogging Birds was to bring satellite-tagged individuals of a species (e.g. the red kite) "to life" by constructing ecologically sound narratives describing their movements. Conservationists fitted satellite tags (PTT-100 22 gram Solar Argos/GPS 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 a day. Although this maximum could indeed be acquired 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 the battery, data was transmitted from the tag to the satellite only once a week. Our system was therefore configured to produce a blog every week; i.e. each time data was received from a bird.

The overall architecture of the Blogging Birds system is shown in Fig. 1. We next describe the main components (see also [12]).

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

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 R [2] by clustering the previous locations of an individual using 90% kernels. As described previously [24], we model local movement patterns as angular and radial velocity vectors to identify excursions (characterised by travel in relatively straight lines at higher speeds).
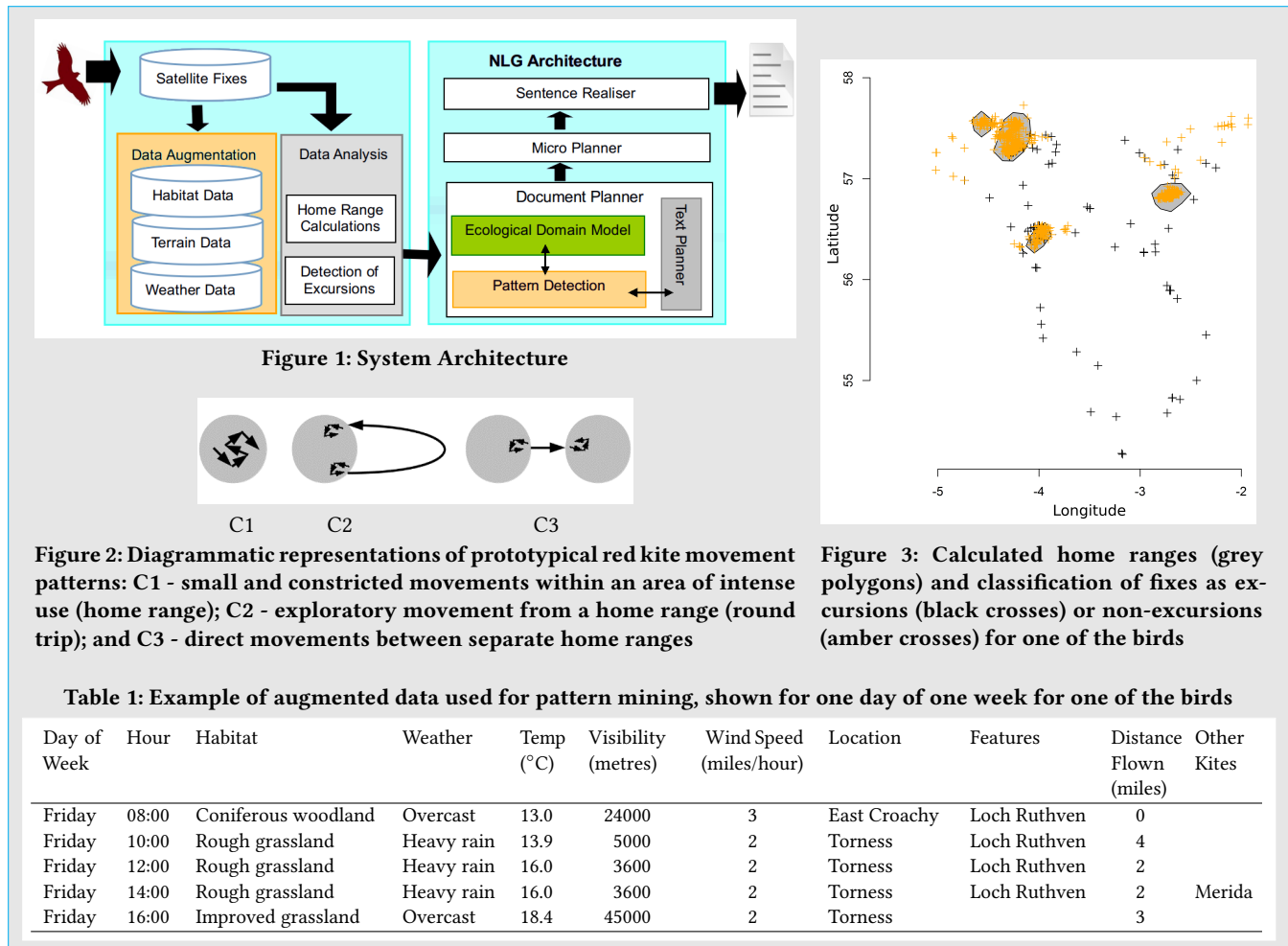
---

[1]https://en.oxforddictionaries.com/definition/creative_writing (Accessed 2/5/2018)

Figure 1: System Architecture

Figure 2: Diagrammatic representations of prototypical red kite movement patterns: C1 - small and constricted movements within an area of intense use (home range); C2 - exploratory movement from a home range (round trip); and C3 - direct movements between separate home ranges

Figure 3: Calculated home ranges (grey polygons) and classification of fixes as excursions (black crosses) or non-excursions (amber crosses) for one of the birds

Table 1: Example of augmented data used for pattern mining, shown for one day of one week for one of the birds

Day of Week	Hour	Habitat	Weather	Temp (°C)	Visibility (metres)	Wind Speed (miles/hour)	Location	Features	Distance Flown (miles)	Other Kites
Friday	08:00	Coniferous woodland	Overcast	13.0	24000	3	East Croachy	Loch Ruthven	0	
Friday	10:00	Rough grassland	Heavy rain	13.9	5000	2	Torness	Loch Ruthven	4	
Friday	12:00	Rough grassland	Heavy rain	16.0	3600	2	Torness	Loch Ruthven	2	
Friday	14:00	Rough grassland	Heavy rain	16.0	3600	2	Torness	Loch Ruthven	2	Merida
Friday	16:00	Improved grassland	Overcast	18.4	45000	2	Torness		3	

Figure 3 shows the calculated home ranges for a bird (grey polygons) as well as the fixes classified as excursions (black crosses) and non-excursions (amber crosses).

Document Planner. The Document Planner identifies patterns in the data that signal different behaviours, and creates 'messages' (implemented as Java classes) that encode these behaviours for use by the Micro Planner and Sentence Realiser.

The Data Analysis allows us to detect the three prototypical patterns of movement depicted in Fig. 2, where the kite either 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 foraging and roosting behaviours 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 any 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 particular patterns are detected in the data. In total the system implements Drools for (a) 26 movement behaviours, e.g. flying along a coast or over a landmark such as a castle or loch, and the home range related

movement patterns as depicted in Fig. 2; (b) 33 foraging behaviours, mostly detailing the food available for a kite in different habitats at different times of the year, but also sometimes related to specific features, e.g. a red kite remaining near a road might be looking for roadkill; and (c) 6 social behaviours, such as roosting and nesting. Example rules are provided in Appendix A.

The Pattern Detection module then exhaustively applies the rules to the satellite fixes to produce a list of all observed movement behaviours and all possible foraging and social behaviours consistent with known environmental and geographic conditions. The latter is the first step in the creative process, where 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 organise this 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, or occasionally offering a conclusion.

Figure 4: Screenshot of the Blogging Birds web interface

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

Paragraph 1

(1) Generate a message based on the detected movement pattern (C1, C2 or C3 in Fig. 2). If the age of the bird can be used to interpret this pattern, add such an interpretation message.
(2) Generate a message summarising the habitats visited.
(3) Generate a message about other kites recorded nearby, if any.

Paragraph 2

(1) Iterate over each day of the week (Mon–Fri):
    (a) If the bird remained relatively static (C1 in Fig. 2), then generate a message about nearby places, else generate a message about any movement behaviour detected.
    (b) Generate a message about a new (not previously used) possible foraging behaviour, if any deduced. Unusual (historically infrequent) behaviours are selected over common ones.
(2) Aggregate the messages generated for the week through Steps 1a and 1b to remove redundancy, e.g. by grouping days with similar behaviours together.

Paragraph 3

(1) 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 (see paragraph below for examples), and limits linguistic repetition by varying the vocabulary. It provides sentence specifications to the Sentence Realiser, which generates sentences using the SimpleNLG library [6].

Figure 4 shows a screenshot of the Blogging Birds interface, where a weekly blog for a kite is overlaid on a Google map of the bird's whereabouts with its historical home ranges marked as blue polygons. In this example, Wyvis, one of five red kites being blogged about, has travelled between two home ranges (movement pattern C3) and an explanation for the observed movement pattern is provided based on the age of the bird. The system emphasises the social side with references to roosting and encounters with other tagged kites. The second paragraph is narrated chronologically, with care taken again to emphasise any unusual behaviours (e.g. 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 ranges ("Wednesday to Sunday"), coordination ("St Filans, Tullybannocher 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.

In this paper we focus on the situation 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 handle other time frames, 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.

3 EVALUATION OF BLOGGING BIRDS

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, and then through comparison with baseline computer-generated blogs that report the data factually without the ecological insights.

3.1 Method

We focussed on the three prototypical movement patterns depicted in Fig. 2 as conditions C1, C2 and C3. For each condition, twelve weeks of data were identified such that the focal red kite's movements broadly matched this condition (e.g. the week in Fig. 4 would correspond to C3), giving us 36 weeks of data in total.

Comparison with human-written blogs. Twelve post-graduate MSc Ecology students – representative of those who might be hired by a conservation charity – were recruited to take part in a creative 2-hour-long session on 'digital media in nature conservation' outside teaching hours. We informed them that they would be writing three short blogs on the basis of environmental data (warning them that this would take about 1.5 hours), that partaking would benefit our research on this topic whilst giving them unique insights into new technologies, and that we would compensate participants £15 cash to express our gratitude for helping us whilst learning. We provided each writer with access to a one-page information sheet about red kites that summarised the typical movement patterns and foraging and social behaviours that were encoded in the Blogging Birds system. They were also free to consult any online sources as they saw fit. We also provided them with the enriched data available to the system for the week, presented in both tabular form (cf. Table 1) and overlaid on a map showing home ranges and fixes (cf. Fig. 4, but without the blog). The information provided to writers was sufficient to allow them to make the same inferences as the system. However, in order to allow full creative freedom to the writers and to 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. Writers were informed about the intended purpose of the blogs and the target audience and each was asked to write three 200 word blogs, i.e. for data from three different weeks, one in each condition (C1–3 in Fig. 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 randomised, and writers were 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 in the information sheet). These 36 manually-written blogs were compared to computer-generated blogs for the same weeks in the evaluation.

As the goal was to investigate Blogging Birds not just as a tool for those with an interest in nature conservation, but also as a resource to engage those interested in new technologies, we ran evaluations with two distinct groups of participants: 93 undergraduate Biology students enrolled on a second year "Community Ecology" course and 49 first and second year undergraduates from across disciplines enrolled on a course entitled "Digital Society", both at the University of Aberdeen. In each trial, a participant sitting at an individual work station 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 manually-written and one computer-generated – without being given any information about their provenance. Participants directly stated which blog they preferred (or expressed no preference), and also rated each blog on how informative, fluent and engaging they found it on a 7-point Likert scale. Each participant evaluated 3 pairs of blogs. The specific hypotheses tested in this study were that:

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

H2 Computer-generated blogs are rated as more informative, fluent and engaging than human-written blogs.

H3 The differences in ratings for computer-generated and human-written blogs are conditional on the movement pattern of the bird (C1, C2 or C3 in Fig. 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 (cf. Table 2 for an example). These baseline blogs were entirely factual and only reported behaviours directly observed in the data, but otherwise followed the same format as the full system blogs. An additional 27 undergraduate students enrolled on the Digital Society course, but who had not participated in the earlier experiment, evaluated the full versus the baseline system using the same methodology and interface as before. The specific hypotheses being tested in this study were that:

H4 Computer-generated blogs with ecological insights are preferred to baseline computer-generated blogs without ecological insights.

H5 Computer-generated blogs with ecological insights are rated more informative and engaging than baseline computer-generated blogs without ecological insights, whilst their fluency is comparable.

Table 2: Baseline computer-generated blog without reference to ecological concepts for the week shown in Fig. 4.

> Wyvis had enough of the area around Teavarran and decided to move to Crieff about 73 miles away. No doubt Wyvis had a social week as kites Moray and Millie were seen in the vicinity often.
>
> Monday Wyvis spent most of her time around Torness, Errogie and Teavarran. On Tuesday evening she reached moorland near Crieff flying about 65 miles amid cloudy conditions and averaging a remarkable 11 miles per hour. The next 5 days Wyvis spent most of her time around Edinample, Tullybannocher and St Fillans. During this time she was seen mainly on acid grassland while making occasional journeys to farmland.
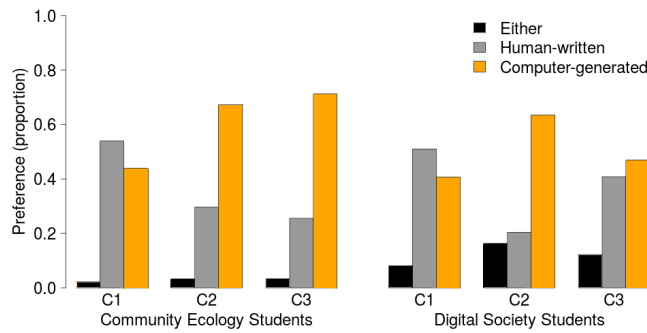>
> Will Wyvis settle down here?

Figure 5: Preferences for human-written and computer-generated blogs by movement condition (see Fig. 2; in short: C1 - movement within a home range; C2 - round trip; C3 - movement between home ranges)

3.2 Results

Evaluation against human-written blogs. Both sets of students showed an overall significant preference for the computer-generated blogs (238 trials vs. 153 in which human-written blogs were preferred; $\chi^2 = 18.5$; $p < 0.001$), confirming hypothesis H1. However, a more complex pattern emerged (Fig. 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 these captured more extensive movement by the kites (round trips (C2) and movement between home ranges (C3)) whilst there was little difference in preference between the two blog types when kite movement was limited (small movements within home-ranges). Digital Society students showed an overall clear preference for the computer-generated blogs only when these described round trips (C2). Combined, our findings indicate the system is particularly skilled at handling cases where there is substantial movement by the bird.

Average ratings for how fluent, engaging and informative the blogs were (Fig. 6) showed that 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: the computer-written blogs were rated significantly higher (p<0.0001) than human-written blogs (confirming hypothesis H2); students in Digital Society gave higher ratings overall than students in Community Ecology (p<0.01); and there was an interaction between blog type and movement pattern (p<0.0001), confirming hypothesis H3. Post-hoc analysis using the Tukey HSD on the individual ANOVAs with Bonferroni-correction revealed that this interaction came about because the computer-generated blogs capturing conditions with more movement by kites (C2 and C3) were (a) more informative than the human-written blogs for the same conditions,

and (b) 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 (H1–12) and the computer (Comp) in Fig. 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. Similarly, H4, who wrote more fluent and engaging blogs than the computer, was rated rather low for informativeness, thus illustrating the difficulty of being informative, engaging and fluent at the same time, even for human writers. Indeed, all the writers were committed and utilised the full 1.5 hours for composing the blogs, and yet most were outperformed by the computer on each of the three metrics. Examples of human-written and computer-generated blogs as well as details of how they were appraised by evaluators can be found in Appendix B.

A questionnaire filled out by the blog writers revealed many interesting insights. In general, they found it hard to comprehend and summarise the sheer amount of data in under 200 words, but felt that 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 doing, stemming from a lack of knowledge of kite ecology and behaviour. Summation of the wide range of data in different formats was certainly challenging and some enjoyed the process more than others. There was considerable variability in how people used the different materials provided to 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 as to whether they would like to write the red kite blogs as a job, the general consensus was that although initially enjoyable it would quickly get tedious and it would get increasingly harder 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 ($\chi^2 = 21.5$; $p < 0.001$), thus 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 4 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 8 trials each ($\chi^2 = 5.1, 3.8$; $p = 0.0233, 0.0499$). Thus, the absence of ecological interpretation by our system was adversely judged for all movement patterns, but particularly so when the birds were relatively static. We also found that the full blogs were rated as being more informative (p<0.0001) and more engaging (p=0.0215) but not more fluent (p=0.825) (Fig. 8), confirming hypothesis H5.

To summarise our results, the two studies 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 overwhelmingly preferred to blogs generated from the same data but without the use of these insights.
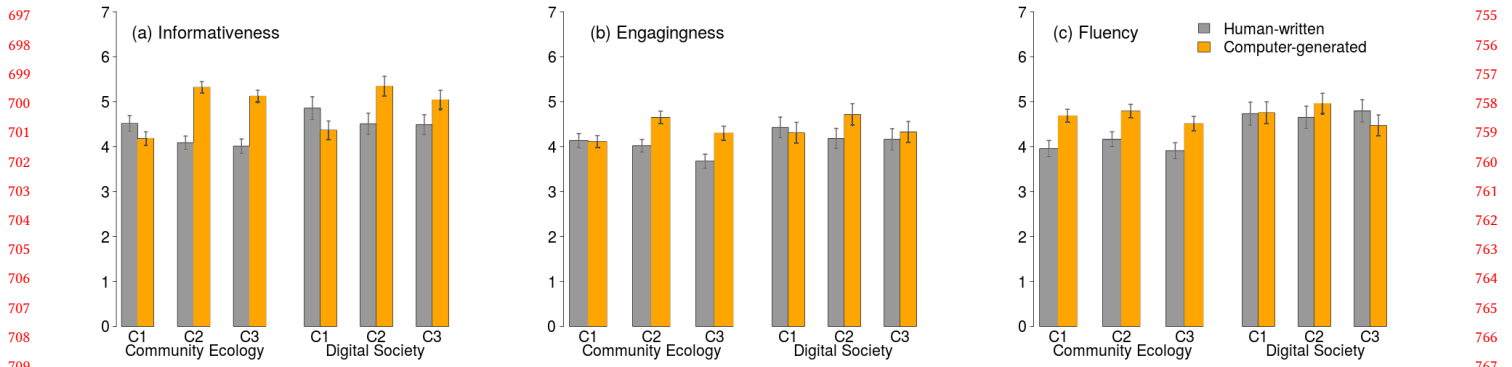
Figure 6: Average ratings for human-written and computer-generated blogs by movement condition (see Fig. 2; in short: C1 - movement within a home range; C2 - round trip; C3 - movement between home ranges)
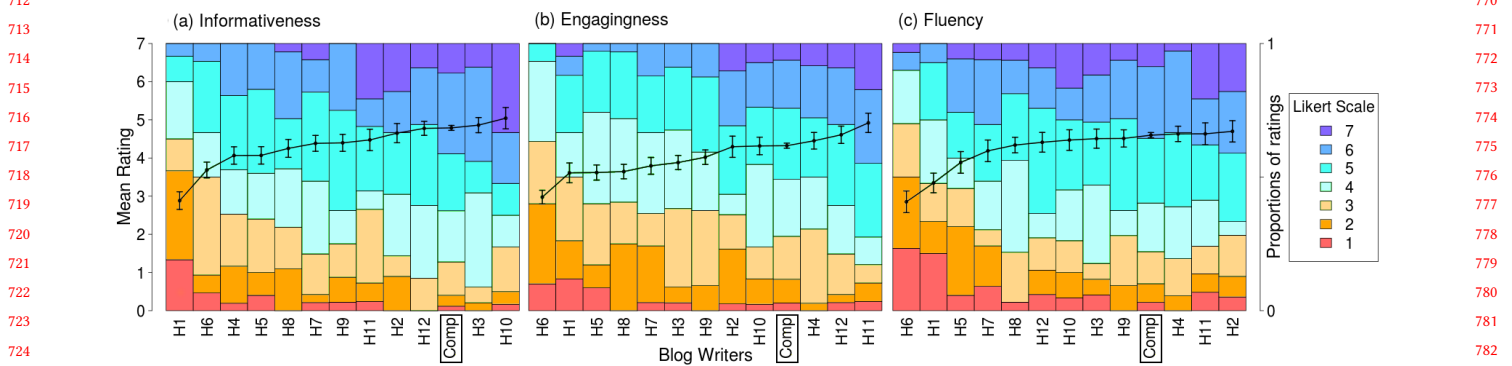

Figure 7: Comparison of computer-generated blogs (Comp) with human-written blogs (H1-12)
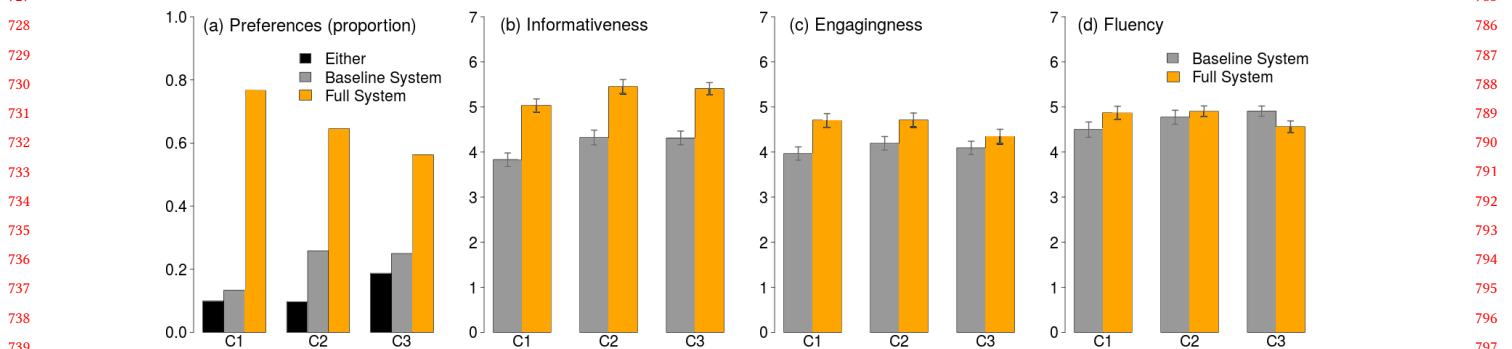

Figure 8: Comparison of computer-generated blogs with ecological insights (Full System) to computer-generated blogs describing movement patterns only (Baseline System)

4 CONCLUSIONS

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 demonstrate 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' narratives are not entirely factual. Though the system is constrained

by the observed data and its ecological domain model, the reported foraging and social behaviours are only imagined to have taken place. However, including these within 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 summarisation of telemetric data enriched with location-specific information of weather conditions, habitat type and geographic features, and this can be readily adapted for use in other domains. The second type of analysis 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 behaviours 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 the construction of knowledge-bases pertinent to the domain of use. While this is a clear limitation of our approach, note that our ecological interpretation of movement data in particular is applicable to several other species. A version of Blogging Birds for golden eagles (*Aquila chrysaetos*) has already been developed for use by RSPB conservation officers, which successfully reuses the second as well as the first type of analysis.

During the course of the project, we also discovered that ecologists had limited knowledge about the foraging behaviour of red kites in Scotland, as they had not been extensively studied following their relatively recent reintroduction. Therefore, we could only encode a limited number of rules per habitat type. The absence of any large scale corpus of texts in this domain also meant that 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 will invite Blogging Birds users to contribute behavioural observations from across the UK. This would allow us to simultaneously curate a larger set of rules and further public engagement.

Finally, we believe the ideas demonstrated in this article are applicable more generally. Telemetric data is ubiquitous, captured by smart phones and other mobile devices, as well as GPS sensors embedded in vehicles used by the transportation industry and several others. Even albums of timestamped and geo-tagged photos provide similar data to that used by us. The nature of the blogs, and the information sources used for data enrichment would of course depend on the application, be it to blog about a holiday, or the provenance and journey of a food item in a supermarket. In effect, we have demonstrated that it is possible to blog about such data through a process of data enrichment and NLG, opening up new avenues for using AI to engage people through data.

5 ACKNOWLEDGEMENTS

REFERENCES

[1] Kim Binsted and Graeme Ritchie. 1997. Computational rules for generating punning riddles. *HUMOR-International Journal of Humor Research* 10: 1, 25–76.
[2] Clément Calenge. 2006. The package "adehabitat" for the R software: a tool for the analysis of space and habitat use by animals. *Ecological modelling* 197: 3, 516–519.
[3] Charles B Callaway and James C Lester. 2002. Narrative prose generation. *Artificial Intelligence* 139: 2, 213–252.
[4] Ian Carter. 2007. *The red kite*. Arlequin Press.
[5] Albert Gatt, Francois Portet, Ehud Reiter, Jim Hunter, Saad Mahamood, Wendy Moncur, and Somayajulu Sripada. 2009. From data to text in the neonatal intensive care unit: Using NLG technology for decision support and information management. *AI Communications* 22: 3, 153–186.
[6] Albert Gatt and Ehud Reiter. 2009. SimpleNLG: A realisation engine for practical applications. In *Proceedings of the 12th European Workshop on Natural Language Generation*. 90–93.
[7] Pablo Gervás. 2009. Computational approaches to storytelling and creativity. *AI Magazine* 30: 3, 49–62.
[8] Marjan Ghazvininejad, Xing Shi, Yejin Choi, and Kevin Knight. 2016. Generating Topical Poetry.. In *Proceedings of Empirical Methods in Natural Language Processing (EMNLP 2016)*. 1183–1191.
[9] Eli Goldberg, Norbert Driedger, and Richard I Kittredge. 1994. Using natural-language processing to produce weather forecasts. *IEEE Expert* 9: 2, 45–53.
[10] Mark Hebblewhite and Daniel T Haydon. 2010. 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, 2303–2312.
[11] K. Panetta. 2017. Neural Networks and Modern BI Platforms will evolve data and analytics. (2017). http://www.gartner.com/smarterwithgartner/nueral-networks-and-modern-bi-platforms-will-evolve-data-and-analytics/ [Online; accessed 8-November-2017].
[12] Kapila Ponnamperuma, Advaith Siddharthan, Cheng Zeng, Chris Mellish, and René Wal. 2013. Tag2Blog: Narrative generation from satellite tag data. In *Proceedings of the 51st annual meeting of the association for computational linguistics: System demonstrations*. 169–174.
[13] François Portet, Ehud Reiter, Albert Gatt, Jim Hunter, Somayajulu Sripada, Yvonne Freer, and Cindy Sykes. 2009. Automatic generation of textual summaries from neonatal intensive care data. *Artificial Intelligence* 173: 7-8, 789–816.
[14] Alexander Pschera. 2016. *Animal Internet: Nature and the Digital Revolution.* New Vessel Press.
[15] Ehud Reiter and Robert Dale. 2000. *Building natural language generation systems.* Cambridge University Press.
[16] Ehud Reiter, Somayajulu Sripada, Jim Hunter, Jin Yu, and Ian Davy. 2005. Choosing words in computer-generated weather forecasts. *Artificial Intelligence* 167, 137–169.
[17] Elena Rishes, Stephanie M Lukin, David K Elson, and Marilyn A Walker. 2013. Generating different story tellings from semantic representations of narrative. In *Proceedings of the International Conference on Interactive Digital Storytelling*. Springer, 192–204.
[18] Mike Sharples. 1996. An account of writing as creative design. In *The Science of writing. Hillsdale, NJ: Lawrence Erlbaum.* Citeseer.
[19] Robert J Sternberg. 1999. *Handbook of creativity.* Cambridge University Press.
[20] Mariët Theune, Sander Faas, DKJ Heylen, and Anton Nijholt. 2003. The virtual storyteller: Story creation by intelligent agents. In *Proceedings of the Technologies for Interactive Digital Storytelling and Entertainment*. 204–215.
[21] Mariët Theune, Esther Klabbers, Jan-Roelof de Pijper, Emiel Krahmer, and Jan Odijk. 2001. From data to speech: a general approach. *Natural Language Engineering* 7: 1, 47–86.
[22] Nava Tintarev, Ehud Reiter, Rolf Black, Annalu Waller, and Joe Reddington. 2016. Personal storytelling: Using Natural Language Generation for children with complex communication needs, in the wild. *International Journal of Human-Computer Studies* 92, 1–16.
[23] Stanley M Tomkiewicz, Mark R Fuller, John G Kie, and Kirk K Bates. 2010. 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, 2163–2176.
[24] Rene van der Wal, Cheng Zeng, Danny Heptinstall, Kapila Ponnamperuma, Chris Mellish, Stuart Ben, and Advaith Siddharthan. 2015. Automated data analysis to rapidly derive and communicate ecological insights from satellite-tag data: A case study of reintroduced red kites. *Ambio* 44: 4, 612–623.
[25] Audrey Verma, René van der Wal, and Anke Fischer. 2015. Microscope and spectacle: On the complexities of using new visual technologies to communicate about wildlife conservation. *Ambio* 44: 4, 648–660.
[26] Jake Wall, George Wittemyer, Brian Klinkenberg, and Iain Douglas-Hamilton. 2014. Novel opportunities for wildlife conservation and research with real-time monitoring. *Ecological Applications* 24: 4, 593–601.
[27] Tsung-Hsien Wen, Milica Gašić, Nikola Mrkšić, Pei-Hao Su, David Vandyke, and Steve Young. 2015. Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems. In *Proceedings of Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics.
[28] Rui Yan. 2016. i, Poet: Automatic Poetry Composition through Recurrent Neural Networks with Iterative Polishing Schema. In *Proceedings of the International Joint Conference on Artificial Intelligence*. 2238–2244.
[29] Xingxing Zhang and Mirella Lapata. 2014. Chinese Poetry Generation with Recurrent Neural Networks. In *Proceedings of Empirical Methods in Natural Language Processing (EMNLP 2014)*. 670–680.

A   EXAMPLE RULES

The conditions for a Drool to apply are expressed as logical combinations of the values in the columns in Table 1. The earthworm example in the main text ("following heavy rain a red kite observed on any of the grassland habitats might feed on earthworms") is implemented as the rule below, which defines a variable "$a", referencing an object of the "FeedingMessage" class, and sets the value of the prey within the object to "earthworms" when the habitat is any type of "grassland" and either the current or immediately previous weather is "heavy rain". Note that the notation has been simplified somewhat for ease of reading: The system uses ordnance survey habitat codes 1–23, which we have replaced in the examples with more descriptive English labels such as "grassland"', and variables shown as "weather" and "previousWeather" are implemented within the system as calls to methods within objects of type geoLocation.

> **following heavy rain a red kite observed on any of the grassland habitats might feed on earthworms**
>
> ```
> (1) rule ``FeedingOnEarthworms''
>
>     when
>       $a: FeedingMessage(
>               (habitat.contains(``grassland'')
>                   )
>               && (weather == %``heavy rain''
>                     || previousWeather == %``heavy rain''
>                   )
>           )
>     then
>         $a.setPrey(``earthworms'');
>     end
> ```

We give some more examples of feeding rules:

> **between April and September and between 10am and 6pm, a redkite observed in bracken or heath might be feeding on insects and worms**
>
> ```
> (2) rule ``FeedingOnInvertebrates''
>
>     when
>         $a : FeedingMessage (
>               (month >=4 && month<=9)
>                 && (hour >=10 && hour <=18)
>                 && (habitat ==``rough grassland'' || habitat == ``acid grassland'')
>             )
>     then
>         $a.setPrey( ``insects and worms'' );
>     end
> ```

> **between January and April and between 10am and 4pm, a redkite observed in grazed turf or tilled land might be feeding on sheep carcasses**
>
> ```
> (3) rule ``FeedingOnDeadSheep''
>
>     when
>         $a : FeedingMessage (
>               (month >=1 && month<=4)
>                 && (hour >10 && hour <16)
>                 && (habitat == ``farmland'' || habitat == ``rough grassland'')
>             )
>     then
>         $a.setPrey( ``sheep carcasses'' );
>     end
> ```

> a redkite observed in the vicinity of a road might be looking to scavenge on roadkill

```
(4) rule ``FeedingCloseToRoad''

    when
        $a : FeedingMessage(
            features.contains(``road'')
             )
    then
        $a.setPrey( ``roadkill'' );
        end
```

We next give some examples of rules about movement, nesting and roosting:

> between 10am and 4pm, a redkite observed over coastal water is exploring coastal waters

```
(1) rule ``ExploringCoastal''

    when
        $a : MovementMessage(
                (hour >= 8 && hour <= 16)
                && ( habitat==``coastal water'' )
             )
    then
            $a.setBehaviour( ``exploring coastal waters'' );
        end
```

> if the first fix in the week is within a homerange and the final fix in the week is within a homerange and the two homeranges are not the same then the week has seen a change of home range

```
(2) rule ``ChangeOfHomeRange''

    when
        $a : FlyingPatternMessage(
                ( geoLocation[first].isInHomeRange() && geoLocation[last].isInHomeRange() )
                && ( geoLocation[first].getHomerangeId() != geoLocation[first].getHomerangeId() )
             )
    then
        $a.setFlyingPattern( ``changehr'' );
        end
```

> if the first fix in the week is within a homerange and the final fix in the week is within a homerange and the two homeranges are the same, and the kite has left the home range in the middle and the total distance travelled in the week is over 40km, then the kite has been exploring around the home range

```
(3) rule ``ExploringOutsideHomeRange''

    when
        $a : FlyingPatternMessage(
                ( geoLocation[first].isInHomeRange() && geoLocation[last].isInHomeRange() )
                && ( geoLocation[first].getHomerangeId() == geoLocation[first].getHomerangeId() )
                && hasLeftHomeRange()
                && ( totalDistanceTravelled > 40km )
             )
    then
        $a.setFlyingPattern( ``exploration'' );
        end
```

between March and April, a red kite over two years of age seen in woodlands between 9am and 4pm might be seeking a territory for nesting

```
(4) rule ``LookingForNestingTerritory''

    when
        $a : BehaviourMessage(
                age >= 2 && (month >=3 && month <= 4 )
                && (hour >= 9 && hour <= 16)
                && (habitat == ``woodlands'' )
            )
    then
        $a.setBehaviour( ``nestingteritory'' );
    end
```

a redkite in woodlands at or after 6pm is roosting

```
(5) rule ``Roosting''

    when
        $a : BehaviourMessage(
                hour >= 18  && habitat == ``woodland''
            )
    then
        $a.setBehaviourtype( ``roosting'' );
    end
```

B EXAMPLE BLOGS FROM EVALUATION

Tables 3–5 show three experimental trials involving blogs by the better performing human writers (H10, H12 and H4, cf. Figure 7), one in each of the conditions C1, C2 and C3 (cf. Figure 2). Each table shows the human-written and computer-generated blogs, as well as their appraisal by the evaluators. The columns I, E and F report scores for Informativeness, Engagingness and Fluency. Each table also shows the overall preference and the comments provided by evaluators to justify their scores and preferences. We have used italics to annotate the comments, which refer to version one and two, to indicate which references are to computer-generated blogs and which are to human-written blogs.

Tables 6–8 show the same material for three experimental trials involving blogs by the worst performing human writers (H6, H1 and H5).

Table 3: Comparison of Blogs by human writer H10 and Comp about a week in condition C1

Author	Blog
H10	Millie is spending her week at the banks of Loch Ness. Her nest must be is just on the other side of the lake, opposite to the ruins of Urquhart castle. She makes a 12 mile loop around countryside looking over bogs and moorlands. She tends to come back to her starting point at the end of the day. Regardless of the continuous good weather she doesn't cover more than 10-15 miles per day. It looks like Millie is making these regular short trips towards Glen Mor and Glen Albyn this week. She might have found some of good sources of prey at the moorland and grassland slopes of Cairngorms. She tends to use the part of the day with the highest visibility to hunt for food. All-together Millie covered just above 70 miles this week, which is about her average (compared to her max weekly distance of 250 miles).
COMP	This week Millie was active. She predominantly flew between Inverfarigaig and Easter Aberchalder. Millie has been feeding in a variety of habitats this week while she roosted in woodlands around Inverfarigaig. Millie was seen together with kite Moray.
	The entire week Millie spent most of her time around Inverfarigaig and Easter Aberchalder. During this time she was seen mainly on acid grassland while making occasional journeys to farmland. Acid grassland is generally a species poor habitat. However there must have been enough worms and insects for Millie to feed on.
	Will Millie continue exploring the same area next week as well?

Evaluation Scores

Comp I E F	H10 I E F	Preferred	Comments
4 4 4	6 3 4	Human	Version 1 *[Comp]* is more engaging, using emotive language and questions, but version 2 *[Human]* is far more informative - the fact that even in good weather Millie does not cover more than 15 miles a day, or 70 miles in the week, despite the good weather. Also the assumption that Millie must have a nest on the banks of Loch Ness.
3 1 2	7 6 7	Human	the person *[Human]* seemed to have a better understand of how to write the blog. they also seemed to understand the kite much better and know more about it. The way the writer *[Human]* gave us distances of her traveling and all round more information about the kites travels over the week. also how they *[Human]* talked about when they believed she hunted most.
4 5 7	7 4 1	Human	The atrocious grammar of the blog *[Human]*, however, makes it really hard to read and enjoy. It *[Human]* does seem to describe the kites week in more details, and put in into context with her habits. By just looking at this week, I got an overview of her usual lifestyle and feeding habits, rather than just a snapshot.
5 5 7	6 5 4	Comp	More precise detail could be given about Millies exact movements on the map. Detailed exposition of Millies behaviors and the analysis for the potential reasons *[Human]*. Really paints a picture of what is going on *[Human]*
3 2 2	4 5 5	Human	Version 1 *[Comp]* is not entirely accurate and is written to maybe suit younger audiences as it is very simple. Version two *[Human]* elaborates on the data and provides some insight on why it is the way it is. Its more detailed and has a better writing style.
5 6 5	3 2 3	Comp	The second blog *[Human]* was not as good in my opinion as it never flowed well and was all clumped together. It just seemed to be bits of information listed one after the other with little structure. In my opinion the first blog *[Comp]* was more interesting than the second *[Human]* as i found it more engaging than the second blog as it actually seems to engage the reader at the end by asking a question. The information on the first blog *[Comp]* also seems a bit easier to understand in relation the the map itself.

Table 4: Comparison of Blogs by human writer H12 and Comp about a week in condition C2

Author	Blog
H12	Wyvis fledged from near Inverness but has been more recently in the Stirlingshire area in the hills West of Stirling, South of the Argaty red kite center. With Wyvis being a young bird she will not start breeding for a few years but while she matures she will spend time traveling though the countryside looking for good territories to raise young and other populations of kites that might have potential mates. Red kites can fly quite a distance which Wyvis proves by flying 69 miles down to Lochmaben near lockerbie in just 4 hours.She appears to rest here for the rest of the day. she then moves to the area round the Leadhills at Sanquhar where she stays for the rest of the week. she moves around this area a lot with regular short flights. we know this is not a usual home range for her but she could be scouting out the area and mapping whats available to her. she visits a broad range of habitats that she is likely to find several potential food sources. she also makes visits most days to a woodland in the area and might be using this as a resting area or roost.
COMP	Wyvis must have been bored of the area this week as she started off from Kilsyth but ended up at Sanquhar about 54 miles away. Wyvis is only one year old. An incredible journey for such a young kite. Wyvis has been observed feeding mainly on acid grassland and farmland during this week while she roosted in woodlands around Lochmaben.
	On Monday by mid day she reached woodland near Lochmaben outside her usual home ranges flying about 69 miles amid overcast conditions and averaging an incredible 17 miles per hour. On Tuesday morning she reached acid grassland near Durisdeer flying about 14 miles amid overcast conditions. Wednesday to Sunday Wyvis spent most of her time around Durisdeermill. During this time she was seen mainly on farmland while making occasional journeys to heather. She must have been feasting on worms and insects which are in abundance in farmland. But what could she be looking for in heather?
	Will Wyvis settle down in her new found territory or will she be stretching her wings next week as well?

Evaluation Scores

Comp I E F	H12 I E F	Preferred	Comments
4 6 6	3 5 5	Comp	It *[Comp]* was more engaging and fluent and presented the information more interestingly.
5 4 5	4 5 5	Comp	The first blog *[Human]* is more general in its description of wyvis movements. Ver2 *[Comp]* describes the environmental conditions especially the different habitats in detail. It also ends with two questions which gives the reader something to think about over the next week before the next blog post.
4 6 5	5 6 6	Human	version 2 *[Comp]* was a bit boring when trying to think of reasons for the kites behavior. Ver1 *[Human]* explained the data on the map well and kept it interesting
5 4 4	3 1 1	Comp	Version 1 *[Human]* is clearly written by someone less experienced than who wrote version 2. in Ver2 *[Comp]*, the observer is asking and answering questions about Wyvis behavior.
4 6 3	5 4 5	Comp	I enjoyed both versions of the Wyvis blog although version 1 *[Human]* lacked personality and spirit. Version 1 *[Human]* was well written although very formal. I preferred version 2 *[Comp]* of the blog as it was more engaging with the reader and had more valuable information, for example, averaging an incredible 17 miles per hour and Acid grassland is generally a species poor habitat
4 6 5	6 5 5	Human	The writing style in version 2 *[Comp]* is very engaging and interesting, however not as much detail is given on what Wyvis could be up to as the blog focuses mainly on her movements. In ver1 *[Human]*, more information is given on why Wyvis has moved such a distance in one day and on what she is doing there."

Table 5: Comparison of Blogs by human writer H4 and Comp about a week in condition C3

Author	Blog
H4	Wyvis, who turned one year old in 2013 still has another 1-2 years to mature before she will breed for the first time. She began the week of 20-26 May at Torness before traveling much farther south to Crieff and Callander. In just one day she flew 67 miles! She spent most of her time in woodland, heather and grassland, but occasionally ventured into farmland habitats. She is an opportunistic feeder with a varied diet which primarily includes small mammals and sometimes amphibians and birds. Wyvis has also been known to scavenges dead animals or roadkill. This scavenging can get her into trouble though if she eats an animal that has been poisoned.
COMP	Wyvis had enough of the area around Teavarran and decided to move to his other home range around Crieff about 73 miles away. Wyvis is only one year old, but is already exploring the landscape to identify a good territory. Wyvis has been foraging in many different habitats this week and she roosted largely in woodlands around Errogie and Crieff. No doubt Wyvis had a social week as kites Moray and Millie were seen in the vicinity often.
	Monday Wyvis spent most of her time around Torness, Teavarran and Errogie. On Tuesday evening she reached moorland near Crieff flying about 65 miles amid cloudy conditions. Wednesday to Sunday Wyvis spent most of her time around St Fillans, Tullybannocher and Edinample. During this time she was seen mainly on acid grassland while making odd journeys to arable land. Acid grassland is generally a species poor habitat. However there must have been enough worms and insects for Wyvis to feed on.
	Will Wyvis settle down here or will she return to her other homerange next week?

Evaluation Scores

Comp I E F	H4 I E F	Preferred	Comments
3 3 3	6 5 5	Human	Although version 2 *[Human]* is a much smaller blog, its by far the better written one due to its informative and concise nature. I didnt know that scavenging for dead animals could kill a bird through poisoning.
5 5 4	4 3 6	Comp	There is more detail *[Comp]* in the fact about where the redkite was and the writing style is much more engaging than the other *[Human]*.
5 6 3	6 6 6	Human	*[Human]*Tells you when the kite set off on the journey and what kind of habitats she stayed near also how far she went in just one day which she traveled from Farraline to just outside crieff in one day. Version one *[Human]* because it is informative and doesnt drag out the information and tells you exactly what you need to know. Version two *[Comp]* tells it like a story.
4 2 2	3 5 6	Comp	Version 2 *[Comp]* really needs to improve its writing style because it has very useful information but it is a bit confusing at some points. While version 1 *[Human]* has a very engaging style of writing but not enough information. Version 1 *[Human]* has a better writing style and text quality but it does not goes into detail about the places Wyvis was during her travels, which version two *[Human]* does cover.
3 3 4	4 3 4	Human	Wyvis has 2 home ranges and is a varied feeder *[Comp]*. *[Human]* More informative about habitats traveled to and varied diet.
6 5 7	5 5 5	Comp	It *[Comp]* told you more information about other kites in the area as-well. it also let you know distances that she had been flying.

Table 6: Comparison of Blogs by human writer H6 and Comp about a week in condition C1

Author	Blog
H6	Millie is a year old, female kite born in 2012. We are going to go through Millie's journey around it's home ground from March 4th to March 10th. Millie's Monday started from Errogie where it then travelled to Farraline and down southeast to Torness and back up to Errogie. For the next 2 days, Millie was located flying at Errogie and flew to Strone, Farraline and back to Errogie in the evening. On March th, it just lingered around Errogie arable land area. March 8th activities were just the same as what happened on March 5th and 6th.
COMP	This week Millie was active. She predominantly flew between Errogie and Farraline. Millie has been feeding in a variety of habitats this week and she roosted largely in woodlands around Errogie and Torness.
	The entire week Millie spent most of her time around Farraline and Errogie. During this time she was seen mainly on heather while making occasional journeys to built up area. Heather is generally considered to be a species poor habitat. However there could have been some small mammals like voles for Millie to feed on.
	Will Millie continue exploring the same area next week as well?

Evaluation Scores

Comp I E F	H6 I E F	Preferred	Comments
4 5 5	3 2 1	Comp	*[Comp]* Let down by penultimate sentence, almost colloquial and by question at end, apart from that a good piece of text. Finally a concise and well-written blog! Enjoyable to read.
6 6 5	3 2 2	Comp	Version 1 *[Human]* is hard to read and does not engage at all. Version 2 *[Comp]* using a question to engage reader at end to want to follow her journey.
5 5 6	6 4 2	Comp	Version 2 *[Human]* had quite a lot of mistakes and bad grammar. e.g. "On March th, it just lingered around Errogie arable land area." March what? The red kite is a girl called Millie, not "it." Lingered around is not descriptive or informative on what Millie was doing. Version 1 *[Comp]* flowed better but didnt have much detail about Millies activities or weather etc. Nice how it ended with a question to keep the reader interested for the next blog entry.
4 4 5	1 1 1	Comp	Version 2 *[Human]* is often confusing, with lots of grammatical errors and poorly written phrases. There are lots of days and dates used which get confusing. Version 1 *[Comp]* is more grammatically correct, which makes it easier and more interesting to read. The question at the end makes us want to find out more about Millie. The key points are included in the introduction, and these are expanded on further in the second paragraph. It is more informative as we also find out about the habitat of the bird.
2 3 2	3 5 6	Human	Blog 1 *[Comp]* repeats itself over and over and doesnt cover the information very well. However, the extra information about foraging and the question at the end keeps it interesting. Blog 2 *[Human]* was easier to understand and covered more information than blog 1 *[Comp]*. However, it *[Human]* is not overly engaging.
6 7 7	1 1 1	Comp	Version 1 *[Human]* referred to Millie as "it", even though she had been given a female name so the gender was obvious. There was a couple of typing errors in Version 1 *[Human]* also. Version 2 *[Comp]* was interesting because it was short and to the point. The use of a rhetorical question at the end of the blog also left you thinking about the blog even though you had finished reading. Version 2 *[Comp]* was also put into short paragraphs which made it easy to read.

Table 7: Comparison of Blogs by human writer H1 and Comp about a week in condition C2

Author	Blog
H1	3.7.2013: After taking it slow on Monday, I was ready to get up earlier today flying to Farraline heather catching breakfast and getting all the energy I need for the 18 miles trip to Inverness. It is so nice to fly over the world thinking about what everyone down there might be doing. At lunchtime im arriving in Inverness and meeting Lewis. After having a long chat about all the "Red Kite Gossip" and live in particular we decided to meet again tomorrow morning before I'm heading back.
	4.7.2013: After picking some worms on the arable field together with Lewis, I was still a bit hungry and as I did not need to fly back home directly I spontaneously thought about flying west to Dingwall farmland, which supposed to be a delightful hunting ground.
	5.7.2013: How lovely to wake up close by the sea for change. The sound of the ocean and the maritime wind in my feathers, but now Im looking forward to go back home and seeing Beauly...
COMP	This week, Moray flew around 82 miles and made a journey to Dingwall outside her home range and back to Farraline. An incredible journey for such a young kite; Moray is only one year old. Moray has been foraging in many different habitats this week and she roosted largely in woodlands around Farraline and Inverness.
	Monday Moray spent most of her time around Farraline. On Tuesday by mid day she reached farmland near Inverness outside her usual home ranges flying about 18 miles amid cloudy conditions. In the afternoon she was spotted in rough grassland close to Taindore Wood outside her home range perhaps searching for small rats. Wednesday Moray spent most of her time around Inverness and Dingwall. On Thursday morning she reached surface water near Torness flying about 12 miles. In the evening she was seen in acid grassland close to Woodside while enjoying the sunny day maybe feeding on insects and worms. Friday to Sunday Moray spent most of her time around Farraline. During this time she was seen mainly on farmland while making occasional journeys to heather. She must have been feasting on worms and insects which are in abundance in farmland. But what could she be looking for in heather?
	Clearly Moray is fond of her home as she was swift to return back to Farraline.

Evaluation Scores

Comp I E F	H1 I E F	Preferred	Comments
7 5 6	5 5 6	Comp	Perhaps its just the way my mind has been trained, but the more I read blogs written from the point of view of the Red Kite *[Human]* the more I dislike them. I still feel that it *[Human]* would be appropriate for a younger audience. *[Comp]* Detailed description of the Red Kites activities, for example her visit to the farmland.
7 7 6	3 3 4	Comp	*[Comp]* It was good that factors such as weather and distances flown had been included. It *[Comp]* reads very well and was very detailed.
6 4 6	4 5 4	Human	birds point of view is more fun *[Human]*
6 5 3	3 2 3	Comp	A lot of short sentences were used *[Human]*, making the reader read it fast but not taking it in. Would be good to have longer sentences. This would help to make it more coherent as well, as now its day after day description not much linkage going on. *[Comp]* The easy to read explanation of the data on the map, along with the "82 miles" total. Very descriptive, and a lot of information but still engaging.
7 6 5	4 3 2	Comp	The first one *[Comp]* is more descriptive, it tells you where she has been and what she has been doing. Its also much better grammatically. The second one *[Human]* sounds like it was written for a childs story book.
7 6 6	2 4 5	Comp	Version 1 *[Human]* was not in anyway scientific and although is an interesting approach to the blog, does not give accurate information of the kite or its journey that week. Version 2 *[Comp]* gives detail in distance flown and where and why she was flying away from her home range.

Table 8: Comparison of Blogs by human writer H5 and Comp about a week in condition C3

Author	Blog
H5	Moray was born in 2011 close to North Kessock.During the one week monitoring she was mostly moving around Inverness and Torness.In the first 4 days she was spotted in Inverness, Farraline, Dingwall and she flew to Torness every day. During the first 3 days she spent almost every night in Inverness but after this time she was mostly spotted in Torness. Which means, that she found a new habitat for herself.Very likely, that her habitat changed in Inverness, maybe some disturbances occured at those times.She probably found out that the circumstances are better in Torness. In the last 3 days she was only spotted there, where she spent the evening in a woodland area, called Loch Ruthven. The reason that she moved here that this area can be a better breeding place in the future. Sheltered and several different habitats can be found around it.
COMP	This week Moray was active. She predominantly flew between Farraline and Torness. Moray's foraging patterns this week have been varied and she roosted largely in woodlands around Farraline and Torness.
	On Monday morning Moray was observed in farmland close to Black Wood outside her home range maybe feeding on earthworms. In the evening she was observed flying passing Beauly Firth before reaching Torness. On Tuesday night she was observed flying down to Dingwall over Beauly Firth. On Wednesday morning she was observed in farmland close to Bog Fada Wood amid heavy rain and poor visibility maybe feeding on earthworms. In the afternoon she reached heather near Torness flying about 14 miles amid cloudy conditions. Next she was observed in acid grassland close to Woodside perhaps looking for insects and worms. Next she reached surface water near Inverness outside her usual home ranges flying about 12 miles. On Thursday morning she was seen in heather close to Taindore Wood outside her home range while enjoying the sunny day maybe feeding on voles. Then she reached woodland near Dingwall flying about 10 miles amid overcast conditions. Friday and Saturday Moray spent most of her time around Torness. During this time she was seen mainly on farmland while making occasional journeys to heather.
	Will Moray continue exploring the same area next week as well?

Evaluation Scores

Comp I E F	H5 I E F	Preferred	Comments
7 5 6	4 4 2	Comp	The two versions clearly put emphasis on different aspects.Both versions tend to speak about the red kite as if it were human, e.g. stating that Moray is "enjoying" the sunny day. Version one [Comp] gives much more detailed information about when and where the red kite was spotted.
4 2 3	3 3 4	Neither	Each sentences has almost the same content. I think both versions are not very good. They describe the behavior of the kite badly.
4 3 2	2 1 2	Comp	that she flew 12 miles [Comp]
5 5 5	3 3 2	Comp	Version 1 [Comp] was extremely informative and interested the reader. It was also well written and caught the readers attention. Version 2 [Human] lacked much information and didnt explain the information on the graph well. It [Human] was also written in a very broken and incoherent, non-flowing manner. Version 1 [Comp] had high level of background information. It explained in depth the movements of the birds and possible reasons why in a logical and coherent manner.
6 5 6	3 3 2	Comp	Version 1 [Human] was good to start off with but by the end the sentences did not really make any sense. There seems to be quite a few assumptions about the kites movements rather than factual reasons. Version 2 [Comp] had much more information about where the kite flew to and what kind of land was present at each site. This maybe had an effect on what type of food the kite was looking for? The weather information is also useful in understanding the flying pattern of the kite.
4 4 3	5 5 5	Human	Version 2 [Human] is written much better than version 1. It is also a lot more informative and interesting. Version 1 [Comp] is too long for the amount of information it has included. Version 2 [Human] includes information about changes in habitats and why the red kite may change her habitat.