Search String Development

In this auxiliary document, we described how we created, refined and validated the keywords which were used to search the news database. Upon any publication, we will include a link to this document. We hope this could be a tutorial for future HCI research that leverage news search databases. The following steps are mainly adapted from Stryker et al. [6] and Lacy et al. [4].

1. Creating Comprehensive Search Terms

Following this guidance, we composed our keywords with the following four steps:

Step 1: identify topical terms. Stryker et al. [6] suggested to define the story relevance as the first step. We approached this by identifying the key topics of target new stories. Given the scope of this paper, we identified "GPS", "Failure" and "Catastrophic" as three topics that define the story relevance.

Step 2: extract keywords from existing literature. We surveyed the existing literature that analyze interaction issues of in-car GPS to add keywords. Specifically, we leveraged the keywords that describe the interaction problems of GPS in [1–3]. The extracted key words include "traffic", "deviate", "timing", "attention", "route", "destination", "map" and "detour".

Step 3: extract keywords from real news stories. We augmented the communication studies best practices with this additional common HCI approach. We considered the keywords from existing literature not sufficient since 1) they focused on issues in everyday scenarios, not in catastrophes, and 2) they are high-level summative keywords generated by researchers, which is foreign to the words used in the news. As such, we incorporated the idea behind Grounded Theory Method to discover insights grounded in actual data [5]. Four researchers collected 40 well known GPS failure stories and extracted the keywords along the three topics that we identified in **Step 1**.

Step 4: add synonyms to all keywords. Finally, Table 1 lists the search terms generated by the previous steps and the combined search string that is applied to the news databases.

Topics	Synonyms
GPS	GPS, satnav
Failure	wrong, fail, incorrect, bad, gaffe, lost, problem, mistake, destination, error, astray, misled, blame, glitch, route, map, timing, deviate, detour, traffic, attention,
Catastrophic	accident, damage, stuck, strand, crash, rescue, police, illegal, hit, kill, death, charged, die, struck, strike
Combined search string	(GPS OR satnav) AND (wrong! OR fail! OR incorrect! OR bad! OR lost! OR problem! OR mistake! OR astray! OR accident! OR damag! OR error! OR stuck! OR strand! OR crash! OR rescue! OR destination! OR route! OR map! OR timing! OR deviat! OR detour! OR traffic! OR attention! OR police! OR turn! OR direction!)

Table 1. Constitutions of the search terms and the combined initial search string

* Some words are modified to account for tense variations in the combined search string.

2. Refine and Validate the Search String

We refined and validated the search string by investigating the retrieved samples on LexisNexis and conducting a formal test of search string's recall and precision. In our context, while recall measures the fraction of relevant articles that are retrieved by the search string, precision measures the fraction of retrieved articles that are relevant. We calculate precision and recall using the following steps:

Step 1: retrieve valid samples of news using the initial search string. We applied the search string from Table 1 in LexisNexis to retrieve a dataset. Since we focused on comprehensiveness when we constructed the search string, it produced many irrelevant stories. For example, the retrieved articles contain lots of news about general practitioners ("GPs"). From this dataset, the researchers identified the first 50 relevant GPS catastrophic incidents news as the **valid samples set**.

Step 2: investigate the sampled dataset to refine the search string. We analyzed the stories in **valid samples set** to refine the initial search string. We removed search terms as they didn't appear in the **valid samples set** and found new terms that are related to the three topical words. In addition, since the retrieved dataset contained large amount of false positives, we also black-listed keywords have high frequencies in the irrelevant articles.

Step 3: validate the new search string with recall and precision. The new search string need to be validated by computing the recall and precision. The recall is computed as the proportion of the articles in the valid sample set that can be retrieved by the new search string. The precision is computed as the proportion of the relevant articles among all retrieved articles. Since assessing the precision requires reading new articles at each time of assessment, we compute the precision of the proportion of the relevant articles among the first 150 articles, a number also used in [6].

Step 4: iterate the Step 2 and Step 3 until satisfied.

Table 2 shows the final search string we used to search LexisNexis. The recall of this search string is 94% and the precision is 40.8%, resulting 408 relevant articles retrieved[1], which are desirable according to the best practices [6]. As a point of reference for future HCI studies that leverage similar method, we also included all the search strings that we iterated in the attachment.

Final search string
(GPS OR satnav) AND (wrong! OR fail! OR incorrect! OR bad! OR gaffe! OR lost! OR problem! OR mistake! OR error! OR detour! OR deviat! OR astray! OR mislead! OR misled! OR blam! OR glitch! OR accident! OR damag! OR stuck! OR strand! OR crash! OR rescu! OR police! OR illegal! OR hit! OR kill! OR death! OR die! OR charg! OR struck! OR strik!) AND NOT (patient! OR doctor! OR ship OR sailor! OR airplane! OR plane! OR bracelet! OR stole! OR pilot! OR boat! OR tracker! OR monitor!)

Table 2. Constitutions of the search terms and the combined initial search string

Reference

1. Abdullah Al Mahmud, Omar Mubin, and Suleman Shahid. 2009. User experience with in-car GPS navigation systems: comparing the young and elderly drivers. 1. http://doi.org/10.1145/1613858.1613962
2. Barry Brown and Eric Laurier. 2012. The Normal Natural Troubles of Driving with GPS. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '12), 1621–1630. http://doi.org/10.1145/2207676.2208285
3. Markus Hipp, Florian Schaub, Frank Kargl, and Michael Weber. 2010. Interaction Weaknesses of Personal Navigation Devices. In *Proceedings of the 2Nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (AutomotiveUI '10), 129–136. http://doi.org/10.1145/1969773.1969796
4. Stephen Lacy, Brendan R. Watson, Daniel Riffe, and Jennette Lovejoy. 2015. Issues and Best Practices in Content Analysis. *Journalism & Mass Communication Quarterly*: 1077699015607338. http://doi.org/10.1177/1077699015607338
5. Michael Muller. 2014. Curiosity, Creativity, and Surprise as Analytic Tools: Grounded Theory Method. In *Ways of Knowing in HCI*, Judith S. Olson and Wendy A. Kellogg (eds.). Springer New York, 25–48. http://doi.org/10.1007/978-1-4939-0378-8_2
6. Jo Ellen Stryker, Ricardo J. Wray, Robert C. Hornik, and Itzik Yanovitzky. 2006. Validation of Database Search Terms for Content Analysis: The Case of Cancer News Coverage. *Journalism & Mass Communication Quarterly* 83, 2: 413–430. http://doi.org/10.1177/107769900608300212

[1] Due to the limitation set by LexisNexis, we can only retrieve 1000 articles.

Appendix A: Iteration of the Search String

Iteration	search_string
1	(GPS OR satnav) AND (wrong! OR fail! OR incorrect! OR bad! OR lost! OR problem! OR mistake! OR astray! OR accident! OR damag! OR error! OR stuck! OR strand! OR crash! OR rescue!) AND NOT (patient! OR doctor!) AND (destination! OR route! OR map! OR timing! OR deviat! OR detour! OR traffic! OR attention! OR police! OR turn! OR direction!)
2	(GPS OR satnav) AND (wrong! OR fail! OR incorrect! OR bad! OR lost! OR problem! OR mistake! OR astray! OR accident! OR damag! OR error! OR stuck! OR strand! OR crash! OR rescu!) AND NOT (patient! OR doctor!)
3	(GPS OR satnav) AND (wrong! OR fail! OR incorrect! OR bad! OR gaffe! OR lost! OR problem! OR mistake! OR error! OR detour! deviate! OR astray! OR accident! OR damag! OR stuck! OR strand! OR crash! OR rescu! OR police! OR illegal! OR hit! OR kill!) AND NOT (patient! OR doctor!)
4	(GPS OR satnav OR navi) AND (wrong! OR fail! OR incorrect! OR bad! OR gaffe! OR lost! OR problem! OR mistake! OR error! OR detour! OR deviat! OR astray! OR mislead! OR misled! OR blam! OR glitch! OR accident! OR damag! OR stuck! OR strand! OR crash! OR rescu! OR police! OR illegal! OR hit! OR kill! OR death! OR die! OR charg! OR struck! OR strik!) AND NOT (patient! OR doctor!)
5	(GPS OR satnav) AND (wrong! OR fail! OR incorrect! OR bad! OR gaffe! OR lost! OR problem! OR mistake! OR error! OR detour! OR deviat! OR astray! OR mislead! OR misled! OR blam! OR glitch! OR accident! OR damag! OR stuck! OR strand! OR crash! OR rescu! OR police! OR illegal! OR hit! OR kill! OR death! OR die! OR charg! OR struck! OR strik!) AND NOT (patient! OR doctor! OR ship OR sailor! OR airplane! OR plane! OR bracelet! OR stolen!)
7	(GPS OR satnav) AND (wrong! OR fail! OR incorrect! OR bad! OR gaffe! OR lost! OR problem! OR mistake! OR error! OR detour! OR deviat! OR astray! OR mislead! OR misled! OR blam! OR glitch! OR accident! OR damag! OR stuck! OR strand! OR crash! OR rescu! OR police! OR illegal! OR hit! OR kill! OR death! OR die! OR charg! OR struck! OR strik!) AND NOT (patient! OR doctor! OR ship OR sailor! OR airplane! OR plane! OR bracelet! OR stole! OR pilot! OR boat! OR tracker! OR monitor!)