

# CS 6320 Project 1

Adrianna Walls

This chatbot is intended to return information about carnivorous plants. It attempts to make conversation on the topic based on data retrieved from the websites of carnivorous plant sellers.

## Web Crawler

The web crawler that I created, Project1WebCrawler.py, starts from three urls, <https://www.californiacarnivores.com/apps/help-center>, <https://predatoryplants.com/blogs/plant-care>, and <https://www.growcarnivorousplants.com/carnivorous-plant-care/>. I created a crawl function that I started on three different threads, one for each starter url. The crawling function starts by creating a search list and adds the starter url for that thread to the search list. It then checks the first element in the search list and parses out the base url. Once it has the base website, it checks the robots.txt file to see if crawling is allowed on that webpage. If so, it loads the data on the page using beautiful soup and then finds all links to other sites present on that page, excluding links that appear to lead to social-media-type sites. It then selects the 'main' tag from the website if possible, gets the text, and replaces all newline characters with spaces. Once these modifications are made, the text is stored in a text file. The top words are then determined using the top tf-idf values from each webpage, and printed to topwords.txt. I then manually selected the top fifteen terms and placed them in modifiedtopwords.txt for the next step.

I created a second python file, Project1Knowledge.py, to finish the process of creating the knowledge base. This file reads in each of the webpage pickle files and the top terms from modifiedtopwords.txt. Each pickle file is sentence tokenized, and the sentences are checked for words and phrases from a list of "bad words" that are common in irrelevant sentences. Those sentences are then removed from the knowledge base. The sentences are then checked for their most related topic based on the wup\_similarity of the nouns and verbs to the topic after the lesk algorithm has been used to retrieve the most likely synset from each. If the sentence uses certain third person pronouns, those pronouns are replaced with the topic of the previous sentence as an attempt at word sense disambiguation. The sentences are then placed as an 's' tag under the topic tag in the xml file that we use as a knowledge base. The topic's most likely synset is included as an attribute of the topic tag.

```
1 <?xml:version="1.0" encoding="utf-8"?>
2 <kb>
3 <nepenthes:synset="nepenthes.n.01">
4 <s>
5 ...You can use undrained containers as pots, but be prepared to water frequently. One exception to the above rule is the Dewy Pine (Drosophyllum) which prefers to be in a lay
6 ...
7 </s>
8 ...Nepenthes, Heliamphora, Cephalotus and Drosophyllum want to be watered frequently but do not want to dry out.
9 </s>
10 ...
11 ...Cephalotus dislike long periods of being waterlogged, so it is best to allow the water in the tray to evaporate before adding more.
12 </s>
13 <s>
14 ...Plastic containers are best, but nepenthes also do well in glazed terracotta or glazed clay
15 </s>
16 <s>
17 ...Always use drained containers: shallow un-drained containers will rot nepenthes's roots.
18 </s>
19 ...
20 ...Cephalotus does best with moderately warm summers and cool to chilly winters.
21 </s>
22 <s>
23 ...The plants may die in long periods of very hot weather and nepenthes enjoy cool summer nights.
24 </s>
25 <s>
26 ...Nepenthes require at least three hours of direct light if grown indoors on a windowsill.
27 </s>
28 <s>
29 ...If you notice your plant has stopped making pitchers, it is often a sign that the plant needs a bit more light. In greenhouses avoid the tray method of watering entirely a
30 </s>
31 <s>
32 ...Nepenthes do not require any dormancy period. Lowlanders come from low elevations in tropical Southeast Asia.
33 </s>
34 <s>
35 ...Darlingtonia is an enigmatic plant native to Oregon and California.
36 </s>
37 <s>
38 ...Nepenthes are native to SE Asia, and these plants come in many shapes and sizes.
39 </s>
40 <s>
41 ...Most of the 170+ Nepenthes species are found in Southeast Asia - particularly Borneo, Sumatra, and the Philippines - and the genus ranges from highland montane plants to t
42 </s>
43 <s>
44 ...Nepenthes's cross-fertile nature has resulted in countless unique hybrids, both in cultivation and in the wild. These include male and female parentage (remember that Nepe
45 </s>
46 <s>
47 ...There are a nice selection of Nepenthes suitable for beginners, as well as growing supplies like compost and pots.
48 </s>
49 <s>
50 ...There are a fantastic range of Nepenthes species and hybrids suitable for growers of all abilities.
51 </s>
52 <s>
53 ...If you're not sure what species you've got, you could try putting it into the Nepenthes Species Identifier online.
54 </s>
55 <s>
56 ...Do you grow Nepenthes?
57 </s>
58 <s>
59 ...The genus Nepenthes has the only species known to have devoured whole rats!
60 </s>
61 ...
62 ...
63 ...
64 ...
65 ...
66 ...
67 ...
68 ...
69 ...
70 ...
71 ...
72 ...
73 ...
74 ...
75 ...
76 ...
77 ...
78 ...
79 ...
80 ...
81 ...
82 ...
83 ...
84 ...
85 ...
86 ...
87 ...
88 ...
89 ...
90 ...
91 ...
92 ...
93 ...
94 ...
95 ...
96 ...
97 ...
98 ...
99 ...
100 ...
101 ...
102 ...
103 ...
104 ...
105 ...
106 ...
107 ...
108 ...
109 ...
110 ...
111 ...
112 ...
113 ...
114 ...
115 ...
116 ...
117 ...
118 ...
119 ...
120 ...
121 ...
122 ...
123 ...
124 ...
125 ...
126 ...
127 ...
128 ...
129 ...
130 ...
131 ...
132 ...
133 ...
134 ...
135 ...
136 ...
137 ...
138 ...
139 ...
140 ...
141 ...
142 ...
143 ...
144 ...
145 ...
146 ...
147 ...
148 ...
149 ...
150 ...
151 ...
152 ...
153 ...
154 ...
155 ...
156 ...
157 ...
158 ...
159 ...
160 ...
161 ...
162 ...
163 ...
164 ...
165 ...
166 ...
167 ...
168 ...
169 ...
170 ...
171 ...
172 ...
173 ...
174 ...
175 ...
176 ...
177 ...
178 ...
179 ...
180 ...
181 ...
182 ...
183 ...
184 ...
185 ...
186 ...
187 ...
188 ...
189 ...
190 ...
191 ...
192 ...
193 ...
194 ...
195 ...
196 ...
197 ...
198 ...
199 ...
200 ...
201 ...
202 ...
203 ...
204 ...
205 ...
206 ...
207 ...
208 ...
209 ...
210 ...
211 ...
212 ...
213 ...
214 ...
215 ...
216 ...
217 ...
218 ...
219 ...
220 ...
221 ...
222 ...
223 ...
224 ...
225 ...
226 ...
227 ...
228 ...
229 ...
230 ...
231 ...
232 ...
233 ...
234 ...
235 ...
236 ...
237 ...
238 ...
239 ...
240 ...
241 ...
242 ...
243 ...
244 ...
245 ...
246 ...
247 ...
248 ...
249 ...
250 ...
251 ...
252 ...
253 ...
254 ...
255 ...
256 ...
257 ...
258 ...
259 ...
260 ...
261 ...
262 ...
263 ...
264 ...
265 ...
266 ...
267 ...
268 ...
269 ...
270 ...
271 ...
272 ...
273 ...
274 ...
275 ...
276 ...
277 ...
278 ...
279 ...
280 ...
281 ...
282 ...
283 ...
284 ...
285 ...
286 ...
287 ...
288 ...
289 ...
290 ...
291 ...
292 ...
293 ...
294 ...
295 ...
296 ...
297 ...
298 ...
299 ...
300 ...
301 ...
302 ...
303 ...
304 ...
305 ...
306 ...
307 ...
308 ...
309 ...
310 ...
311 ...
312 ...
313 ...
314 ...
315 ...
316 ...
317 ...
318 ...
319 ...
320 ...
321 ...
322 ...
323 ...
324 ...
325 ...
326 ...
327 ...
328 ...
329 ...
330 ...
331 ...
332 ...
333 ...
334 ...
335 ...
336 ...
337 ...
338 ...
339 ...
340 ...
341 ...
342 ...
343 ...
344 ...
345 ...
346 ...
347 ...
348 ...
349 ...
350 ...
351 ...
352 ...
353 ...
354 ...
355 ...
356 ...
357 ...
358 ...
359 ...
360 ...
361 ...
362 ...
363 ...
364 ...
365 ...
366 ...
367 ...
368 ...
369 ...
370 ...
371 ...
372 ...
373 ...
374 ...
375 ...
376 ...
377 ...
378 ...
379 ...
380 ...
381 ...
382 ...
383 ...
384 ...
385 ...
386 ...
387 ...
388 ...
389 ...
390 ...
391 ...
392 ...
393 ...
394 ...
395 ...
396 ...
397 ...
398 ...
399 ...
400 ...
401 ...
402 ...
403 ...
404 ...
405 ...
406 ...
407 ...
408 ...
409 ...
410 ...
411 ...
412 ...
413 ...
414 ...
415 ...
416 ...
417 ...
418 ...
419 ...
420 ...
421 ...
422 ...
423 ...
424 ...
425 ...
426 ...
427 ...
428 ...
429 ...
430 ...
431 ...
432 ...
433 ...
434 ...
435 ...
436 ...
437 ...
438 ...
439 ...
440 ...
441 ...
442 ...
443 ...
444 ...
445 ...
446 ...
447 ...
448 ...
449 ...
450 ...
451 ...
452 ...
453 ...
454 ...
455 ...
456 ...
457 ...
458 ...
459 ...
460 ...
461 ...
462 ...
463 ...
464 ...
465 ...
466 ...
467 ...
468 ...
469 ...
470 ...
471 ...
472 ...
473 ...
474 ...
475 ...
476 ...
477 ...
478 ...
479 ...
480 ...
481 ...
482 ...
483 ...
484 ...
485 ...
486 ...
487 ...
488 ...
489 ...
490 ...
491 ...
492 ...
493 ...
494 ...
495 ...
496 ...
497 ...
498 ...
499 ...
500 ...
501 ...
502 ...
503 ...
504 ...
505 ...
506 ...
507 ...
508 ...
509 ...
510 ...
511 ...
512 ...
513 ...
514 ...
515 ...
516 ...
517 ...
518 ...
519 ...
520 ...
521 ...
522 ...
523 ...
524 ...
525 ...
526 ...
527 ...
528 ...
529 ...
530 ...
531 ...
532 ...
533 ...
534 ...
535 ...
536 ...
537 ...
538 ...
539 ...
540 ...
541 ...
542 ...
543 ...
544 ...
545 ...
546 ...
547 ...
548 ...
549 ...
550 ...
551 ...
552 ...
553 ...
554 ...
555 ...
556 ...
557 ...
558 ...
559 ...
560 ...
561 ...
562 ...
563 ...
564 ...
565 ...
566 ...
567 ...
568 ...
569 ...
570 ...
571 ...
572 ...
573 ...
574 ...
575 ...
576 ...
577 ...
578 ...
579 ...
580 ...
581 ...
582 ...
583 ...
584 ...
585 ...
586 ...
587 ...
588 ...
589 ...
590 ...
591 ...
592 ...
593 ...
594 ...
595 ...
596 ...
597 ...
598 ...
599 ...
600 ...
601 ...
602 ...
603 ...
604 ...
605 ...
606 ...
607 ...
608 ...
609 ...
610 ...
611 ...
612 ...
613 ...
614 ...
615 ...
616 ...
617 ...
618 ...
619 ...
620 ...
621 ...
622 ...
623 ...
624 ...
625 ...
626 ...
627 ...
628 ...
629 ...
630 ...
631 ...
632 ...
633 ...
634 ...
635 ...
636 ...
637 ...
638 ...
639 ...
640 ...
641 ...
642 ...
643 ...
644 ...
645 ...
646 ...
647 ...
648 ...
649 ...
650 ...
651 ...
652 ...
653 ...
654 ...
655 ...
656 ...
657 ...
658 ...
659 ...
660 ...
661 ...
662 ...
663 ...
664 ...
665 ...
666 ...
667 ...
668 ...
669 ...
670 ...
671 ...
672 ...
673 ...
674 ...
675 ...
676 ...
677 ...
678 ...
679 ...
680 ...
681 ...
682 ...
683 ...
684 ...
685 ...
686 ...
687 ...
688 ...
689 ...
690 ...
691 ...
692 ...
693 ...
694 ...
695 ...
696 ...
697 ...
698 ...
699 ...
700 ...
701 ...
702 ...
703 ...
704 ...
705 ...
706 ...
707 ...
708 ...
709 ...
710 ...
711 ...
712 ...
713 ...
714 ...
715 ...
716 ...
717 ...
718 ...
719 ...
720 ...
721 ...
722 ...
723 ...
724 ...
725 ...
726 ...
727 ...
728 ...
729 ...
730 ...
731 ...
732 ...
733 ...
734 ...
735 ...
736 ...
737 ...
738 ...
739 ...
740 ...
741 ...
742 ...
743 ...
744 ...
745 ...
746 ...
747 ...
748 ...
749 ...
750 ...
751 ...
752 ...
753 ...
754 ...
755 ...
756 ...
757 ...
758 ...
759 ...
760 ...
761 ...
762 ...
763 ...
764 ...
765 ...
766 ...
767 ...
768 ...
769 ...
770 ...
771 ...
772 ...
773 ...
774 ...
775 ...
776 ...
777 ...
778 ...
779 ...
780 ...
781 ...
782 ...
783 ...
784 ...
785 ...
786 ...
787 ...
788 ...
789 ...
790 ...
791 ...
792 ...
793 ...
794 ...
795 ...
796 ...
797 ...
798 ...
799 ...
800 ...
801 ...
802 ...
803 ...
804 ...
805 ...
806 ...
807 ...
808 ...
809 ...
810 ...
811 ...
812 ...
813 ...
814 ...
815 ...
816 ...
817 ...
818 ...
819 ...
820 ...
821 ...
822 ...
823 ...
824 ...
825 ...
826 ...
827 ...
828 ...
829 ...
830 ...
831 ...
832 ...
833 ...
834 ...
835 ...
836 ...
837 ...
838 ...
839 ...
840 ...
841 ...
842 ...
843 ...
844 ...
845 ...
846 ...
847 ...
848 ...
849 ...
850 ...
851 ...
852 ...
853 ...
854 ...
855 ...
856 ...
857 ...
858 ...
859 ...
860 ...
861 ...
862 ...
863 ...
864 ...
865 ...
866 ...
867 ...
868 ...
869 ...
870 ...
871 ...
872 ...
873 ...
874 ...
875 ...
876 ...
877 ...
878 ...
879 ...
880 ...
881 ...
882 ...
883 ...
884 ...
885 ...
886 ...
887 ...
888 ...
889 ...
890 ...
891 ...
892 ...
893 ...
894 ...
895 ...
896 ...
897 ...
898 ...
899 ...
900 ...
901 ...
902 ...
903 ...
904 ...
905 ...
906 ...
907 ...
908 ...
909 ...
910 ...
911 ...
912 ...
913 ...
914 ...
915 ...
916 ...
917 ...
918 ...
919 ...
920 ...
921 ...
922 ...
923 ...
924 ...
925 ...
926 ...
927 ...
928 ...
929 ...
930 ...
931 ...
932 ...
933 ...
934 ...
935 ...
936 ...
937 ...
938 ...
939 ...
940 ...
941 ...
942 ...
943 ...
944 ...
945 ...
946 ...
947 ...
948 ...
949 ...
950 ...
951 ...
952 ...
953 ...
954 ...
955 ...
956 ...
957 ...
958 ...
959 ...
960 ...
961 ...
962 ...
963 ...
964 ...
965 ...
966 ...
967 ...
968 ...
969 ...
970 ...
971 ...
972 ...
973 ...
974 ...
975 ...
976 ...
977 ...
978 ...
979 ...
980 ...
981 ...
982 ...
983 ...
984 ...
985 ...
986 ...
987 ...
988 ...
989 ...
990 ...
991 ...
992 ...
993 ...
994 ...
995 ...
996 ...
997 ...
998 ...
999 ...
1000 ...
```

Only a part of the original knowledge base has been included in the screenshots, as the original file is very long. The original knowledge base file is included as kb.xml.

Before submitting to the chatbot, the knowledge base was manually cleaned to improve clarity of sentences and remove sentences that were grammatically incorrect without removing data. For example, there were many symbols that were unreadable in the xml. These were modified to be standard text symbols. A screenshot of this modified knowledge base is included in the appendices, and the file itself can be found in the project zip file.

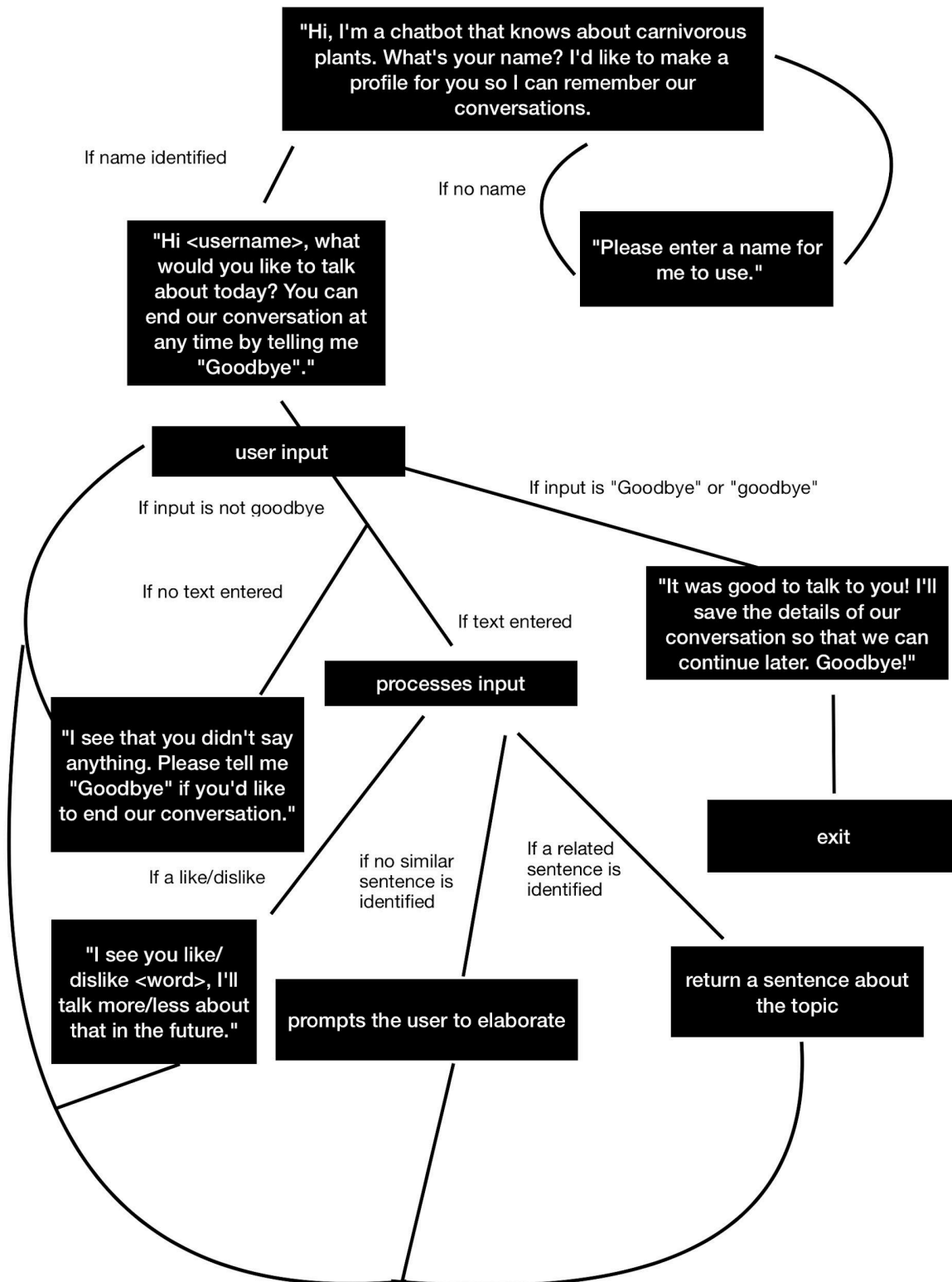
The important terms that were returned by the web crawler were < nepenthes sarracenia cart option plant view add drosera flytrap guide utricularia carnivorous seed tom carnivore care muscipula dionaea sundew sale mix pitcher by item i a maxsea fertilizer z sphagnum no read hardy california venus soil price veitchii resource review>, and these were pared down to <nepenthes sarracenia drosera flytrap utricularia seed care dionaea sundew pitcher fertilizer sphagnum soil veitchii muscipula>.

## Chatbot

The chatbot uses the file chatbot.xml as its knowledge base. It reads in the xml file and reads in the topics so it can easily retrieve the topic tags. It then enters the dialogue, asking the user to introduce themselves and recording the username. It identifies the user's username in a few different ways. If the user only enters a single word or two words, the chatbot assumes that the exact input is the username. If the user enters three words, the AI assumes that the last word is the username. If the user enters a longer string including the words "My name is," then the AI assumes that the word after "is" is the username. Otherwise, the AI reiterates its request for a name. It will then attempt to retrieve the user's profile from users.xml, and if it fails then it will create a new xml tag regarding that user.

After getting a user profile, the chatbot will ask "What would you like to talk about?" and will then begin to process the user's input to identify a starting topic from the top words. It will then loop through either selecting a sentence regarding the current topic, or it will return a random generic statement asking the user to continue. In order to identify a possible sentence from the knowledge base, it will find the topic that matches a word in the input string or that has a wup\_similarity that is closest to one of the nouns or verbs in the sentence, and then it will find the sentence within that topic with the highest cosine similarity to the user input after punctuation has been removed and the words have been changed to lowercase. If the cosine similarity is higher than 0.01, then the AI will return that sentence. Otherwise, it will ask the user to keep talking using a randomly selected sentence. If the user enters a sentence that includes the word "like" or "dislike," the chatbot will store the following word as an 'l' or 'd' tag. Any sentences that are retrieved later will have the cosine similarity artificially increased or decreased based if that word is included in the sentence. If the user enters "Goodbye" or "goodbye", the chatbot will end the conversation by notifying the user that the details of their conversation will be stored in a user profile for later.

The dialogue tree is shown below.



## Sample Dialog

Hi, I'm a chatbot that knows about carnivorous plants. What's your name? I'd like to make a profile for you so I can remember our conversations.

Adrianna

Hi Adrianna ! What would you like to talk about today? You can end our conversation at any time by telling me "Goodbye"

What can you tell me about butterworts?

You can usually tell how dry a species enjoys its winter by the size of its succulent leaves.

What types of butterworts are there?

Mexican or Tropical Butterworts are great for a sunny windowsill.

What other plants enjoy sun?

Warm Temperate Butterworts do well outdoors in part sun to full sun with a bit of protection from afternoon sun. All temperate and warm temperate species should be grown permanently wet on the tray system, with frequent overhead watering.

What about venus flytraps?

Struggling with your flytrap?

No, just curious

These are best in light-colored, preferably white, plastic or glazed ceramic pots.

How do you care for venus flytraps?

Whether you're a long-time greenhouse grower or simply curious about carnivorous plants, Tom's Carnivores has something for you.

Do they have fertilizer?

Always use fertilizers that have been recommended by experts who have tested fertilizer first. We have found MaxSea fertilizer to be the most gentle and effective.

How do I use MaxSea fertilizer?

Be careful and do not use Miracle-Gro brand peat and perlite as it has been fortified with fertilizers and will harm carnivorous plants!

Goodbye

It was good to talk to you! I'll save the details of our conversation so that we can continue later.

Goodbye!

## Evaluation

This chatbot frequently finds accurate sentences, and it stays on topic relatively well. Its weakness is when a user uses terms outside its knowledge base. It may also get confused on the more specific details of the topics, such as how in the previous dialog it confused the fertilizer names “MaxSea” and “Miracle-Gro”. It also does not distinguish between a user stating that they like something and using the subjunctive phrase “would like”, so if a user speaks more formally they may have trouble being understood.

The following are the survey results collected from users listed as the question followed by its possible answers and then each letter denotes the responses of a particular user.

1. How much did this AI seem to know about carnivorous plants? Very little, A little, Some, A lot, Quite a lot
  - a. Some
  - b. Some
  - c. A lot
2. How well did this AI identify what you wanted to talk about? Not at all, Not very well, Somewhat, Very well, Extremely well
  - a. Very well
  - b. Somewhat
  - c. Somewhat
3. How well did the responses make sense? Very little, A little, Some, A lot, Quite a lot
  - a. Some
  - b. A lot
  - c. Some
4. How likely are you to interact with this AI again? Not at all likely, Not very likely, Somewhat likely, Very likely, Extremely likely
  - a. Not very likely
  - b. Somewhat likely
  - c. Very likely

## Appendix 1: Knowledge Base

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <?xml?>
3 <?xml?>
4 <?xml?>
5 <?xml?>
6 <?xml?>
7 <?xml?>
8 <?xml?>
9 <?xml?>
10 <?xml?>
11 <?xml?>
12 <?xml?>
13 <?xml?>
14 <?xml?>
15 <?xml?>
16 <?xml?>
17 <?xml?>
18 <?xml?>
19 <?xml?>
20 <?xml?>
21 <?xml?>
22 <?xml?>
23 <?xml?>
24 <?xml?>
25 <?xml?>
26 <?xml?>
27 <?xml?>
28 <?xml?>
29 <?xml?>
30 <?xml?>
31 <?xml?>
32 <?xml?>
33 <?xml?>
34 <?xml?>
35 <?xml?>
36 <?xml?>
37 <?xml?>
38 <?xml?>
39 <?xml?>
40 <?xml?>
41 <?xml?>
42 <?xml?>
43 <?xml?>
44 <?xml?>
45 <?xml?>
46 <?xml?>
47 <?xml?>
48 <?xml?>
49 <?xml?>
50 <?xml?>
51 <?xml?>
52 <?xml?>
53 <?xml?>
54 <?xml?>
55 <?xml?>
56 <?xml?>
57 <?xml?>
58 <?xml?>
59 <?xml?>
60 <?xml?>
```



58 <5> [3] [3] ...The genus *Nepenthes* has the only species known to have devoured whole rats! [3] [3]

59 </s> [3] [3]

60 <5> [3] [3]

61 </s> [3] [3]

62 *Nepenthes* usually grow as climbing or scrambling vines. [3] [3]

63 </s> [3] [3]

64 <5> [3] [3]

65 Most species are found in Southeast Asia, *Nepenthes*'s center of distribution being the island of Borneo. [3] [3]

66 </s> [3] [3]

67 <5> [3] [3]

68 *Nepenthes* are not typically jungle plants, but prefer more open and sunny ridges, slopes and stunted forests. [3] [3]

69 </s> [3] [3]

70 <5> [3] [3]

71 Seventy percent of *Nepenthes* are tropical highland or mountain plants, growing at elevations of 3,000 to 10,000 feet above sea level. [3] [3]

72 </s> [3] [3]

73 <5> [3] [3]

74 There are both Highland *Nepenthes* and Lowland *Nepenthes*. [3] [3]

75 </s> [3] [3]

76 <5> [3] [3]

77 *Nepenthes* The Monkey Cup plant (*Nepenthes*) belongs to the *Nepenthaceae* family and is a tropical plant. [3] [3]

78 </s> [3] [3]

79 <5> [3] [3]

80 The *Nepenthes* got its name from its unique appearance and large cups. [3] [3]

81 </s> [3] [3]

82 <5> [3] [3]

83 Because *Nepenthes* eats insects, it is a carnivorous plant species. [3] [3]

84 </s> [3] [3]

85 <5> [3] [3]

86 At the bottom of the cup of the *Nepenthes* is a viscous liquid full of digestive juices (enzymes). [3] [3]

87 </s> [3] [3]

88 <5> [3] [3]

89 Some cups of *nepenthes* species can grow up to 50cm tall! [3] [3]

90 </s> [3] [3]

91 <5> [3] [3]

92 The *Nepenthes* can be found in different parts of the world and is mostly found in Southeast Asia (Borneo). In addition, the plant species can be found in Sri Lanka, Madagascar

93 </s> [3] [3]

94 <5> [3] [3]

95 Most *Nepenthes* grow in tropical slope forests. [3] [3]

96 </s> [3] [3]

97 <5> [3] [3]

98 There are *Nepenthes* plants that can be found at altitudes of up to 3,000 metres. These are highland *Nepenthes*. [3] [3]

99 </s> [3] [3]

100 <5> [3] [3]

101 A climbing *Nepenthes* in nature can grow up to 15 metres tall. [3] [3]

102 </s> [3] [3]

103 <5> [3] [3]

104 *Nepenthes* likes sun very much. [3] [3]

105 </s> [3] [3]

106 <5> [3] [3]

107 The *Nepenthes* plant can also grow beautiful yellow / brown flowers in the form of clusters. [3] [3]

108 </s> [3] [3]

109 <5> [3] [3]

110 The Cup trap is used by *Nepenthes* (Monkey Cup plant), amongst others. [3] [3]

111 </s> [3] [3]

112 <5> [3] [3]

113 *Nepenthes* is found in jungles all over the world. [3] [3]

114 </s> [3] [3]

115 <5> [3] [3]

116 Ranging in size from diminutive miniatures like *Nepenthes argentea* to gargantuan monstrosities like *Nepenthes rajah*, the genus includes the only carnivorous plants known to ...

116 ...<s>[20.63]  
117 ...</s>[20.63]  
118 ...<s>[20.63]  
119 ...The genus is hugely diverse, and some species have stricter temperature requirements than others [20.63]  
120 ...</s>[20.63]  
121 ...<s>[20.63]  
122 ...Most cultivation advice is broadly applicable to all Nepenthes. [20.63]  
123 ...</s>[20.63]  
124 ...<s>[20.63]  
125 ...There are over 170 species of Nepenthes, plus countless natural and cultivated hybrids [20.63]  
126 ...</s>[20.63]  
127 ...<s>[20.63]  
128 ...High-powered fluorescent or LED lights (such as T5 growlights) work very well with Nepenthes. [20.63]  
129 ...</s>[20.63]  
130 ...<s>[20.63]  
131 ...I have found that LEDs growlights can be particularly effective since nepenthes tend to give off less heat than fluorescent bulbs. [20.63]  
132 ...</s>[20.63]  
133 ...<s>[20.63]  
134 ...Nepenthes maxima can thrive in a terrarium. [20.63]  
135 ...</s>[20.63]  
136 ...<s>[20.63]  
137 ...Note that Nepenthes do not require a winter dormancy, but if you're growing under natural light, the shorter daylight hours will often slow your plant down significantly. [20.63]  
138 ...</s>[20.63]  
139 ...<s>[20.63]  
140 ...Each Nepenthes species grows best at the temperatures it experiences in its natural habitat. These temperatures correspond to the altitudes at which nepenthes grow in the wild. [20.63]  
141 ...</s>[20.63]  
142 ...<s>[20.63]  
143 ...Remember that Nepenthes are tropical montane plants, and so most require some kind of temperature drop at night. [20.63]  
144 ...</s>[20.63]  
145 ...<s>[20.63]  
146 ...Nepenthes villosa, grown by Simon Lumb, is an example of a true highlander. [20.63]  
147 ...</s>[20.63]  
148 ...<s>[20.63]  
149 ...Nepenthes bicalarata at Chester Zoo, is an example of a true lowlander. [20.63]  
150 ...</s>[20.63]  
151 ...<s>[20.63]  
152 ...Many Nepenthes hybrids are particularly tolerant of lower humidity. [20.63]  
153 ...</s>[20.63]  
154 ...<s>[20.63]  
155 ...Nepenthes are quite adaptable to lower levels, as long as the growing media is kept moist. [20.63]  
156 ...</s>[20.63]  
157 ...<s>[20.63]  
158 ...When you first start growing Nepenthes, space probably won't be at the top of your list. Small plants often grow slowly and may take a while to adapt to the conditions of nepenthes's new home.  
159 ...</s>[20.63]  
160 ...<s>[20.63]  
161 ...Remember that Nepenthes can get very large and will produce climbing vines -- these will need to be pruned! [20.63]  
162 ...</s>[20.63]  
163 ...<s>[20.63]  
164 ...A greenhouse heated to a minimum of 12°C and covered with shade cloth is undoubtedly the best way to grow a wide variety of highland Nepenthes species on a large scale. [20.63]  
165 ...</s>[20.63]  
166 ...<s>[20.63]  
167 ...Some growers create warm chambers within nepenthes's greenhouses for lowlanders. [20.63]  
168 ...</s>[20.63]  
169 ...<s>[20.63]  
170 ...Species and Hybrids There are over 170 known species of Nepenthes, and more are described each year. [20.63]  
171 ...</s>[20.63]  
172 ...<s>[20.63]  
173 ...All of nepenthes are cross-fertile with one another to at least some degree. [20.63]  
174 ...</s>[20.63]

These are a few screenshots of the knowledge base after cleaning. The knowledge base was modified by changing non-grammatical data to corrected sentences and by removing extraneous information and invalid characters.

## Appendix 2: User model

```
1  <?xml version="1.0" encoding="utf-8"?>CR LF
2  <users>CR LF
3  <Adrianna>CR LF
4  <l>CR LF
5  ..butterwortsCR LF
6  </l>CR LF
7  <l>CR LF
8  ..catsCR LF
9  </l>CR LF
10 <d>CR LF
11 ..pitcherCR LF
12 </d>CR LF
13 <s>CR LF
14 ..You can usually tell how dry a species enjoys its winter by the size of its succulent leaves.CR LF
15 </s>CR LF
16 </Adrianna>CR LF
17 </users>CR LF
18
```

This is an example of a user model as seen in users.xml. The username is the name of the tag under the root, <users>. Likes are listed as l tags and dislikes are listed as d tags. The last response given in the previous conversation is denoted by the s tag. The user models are relatively simple, so I only created the user Adrianna for testing this AI.