Website Input - Bug #2462

Getting results saying that the results changed when they did not appear to

09/25/2019 08:46 PM - Luke Murphey

Status:	New	**Start date:**	09/25/2019
Priority:	Normal	**Due date:**	
Assignee:	Luke Murphey	**% Done:**	0%
Category:	Input: Web Spider	**Estimated time:**	0.00 hour
Target version:			

Description

History

#1 - 09/25/2019 09:16 PM - Luke Murphey

Questions:

- Could ordering be the problem?
- Are the results sorted anywhere?
  - update_hash() sort before generating the hash

Observations:

- You can test with this input:
  - | webscrape selector=".grid__cell-content > h2" url="https://www.nomanssky.com/release-log/" depth_limit=25 empty_matches=0
- The match count matches
- The linux host gets an MD5 of b260551723b6ded8951f3c072ba3254b when running the input. However, it gets an MD5 of a320a89e92c6f2d22699d73e496be873 from the search command and from the input on another host.
- The results look the same
- The results hash to the same results in search
  - source="web_input://no_mans_sky_updates" | table _time url match* content_md5 raw_match_count | eval n=mvjoin(match, "") | eval md5=md5(n)