Supplementary Material

In this supplementary material section, we provide detailed benchmark results for all web agents and all GPUs as benchmarked in our paper. These include the energy expendend for the whole benchmark (Mind2Web), as well as for individual splits of the benchmark, listing energy per token. We further extend our CO_2 calculations, providing a complete overview for all agents, all GPUs and the three energy mixes mentioned in our paper.

Total Energy Consumption on Mind2Web Benchmark

Table 1 shows the expended energy in kWh and time in minutes for all web agents on all GPUs on the Mind2Web benchmark (Deng et al. 2024), extending the table in the paper (only showing results for the H100-NVL).

Agent	GPU	Energy(kWh)	Time (min)
AutoWebGLM	A100-SXM4	0.46 ± 0.02	88.3 ± 3.3
AutoWebGLM	A100-PCIe	0.48 ± 0.01	88.4 ± 1.26
AutoWebGLM	H100-SXM5	0.36 ± 0.01	44.9 ± 1.06
AutoWebGLM	H100-NVL	0.33 ± 0.01	57.0 ± 0.8
AutoWebGLM	H200-SXM5	0.35 ± 0.02	44.5 ± 1.25
AutoWebGLM	L40S	0.58 ± 0.01	81.1 ± 3.15
AutoWebGLM	RTX 3090	0.80 ± 0.05	91.6 ± 2.35
AutoWebGLM	RTX A6000	0.76 ± 0.02	103.0 ± 2.83
MindAct	A100-SXM4	1.90 ± 0.40	468.0 ± 92.6
MindAct	A100-PCIe	1.39 ± 0.34	459.0 ± 89.9
MindAct	H100-SXM5	1.58 ± 0.26	326.0 ± 55.4
MindAct	H100-NVL	1.22 ± 0.29	301.9 ± 94.3
MindAct	H200-SXM5	1.51 ± 0.31	300.0 ± 67.4
MindAct	L40S	2.01 ± 0.30	402.0 ± 55.3
MindAct	RTX 3090	1.28 ± 0.04	202.0 ± 3.58
MindAct	RTX A6000	2.57 ± 0.64	570.0 ± 126.0
MultiUI	A100-PCIe	1.18 ± 0.02	218.0 ± 1.79
MultiUI	H100-SXM5	0.93 ± 0.02	113.0 ± 4.53
MultiUI	H100-NVL	0.822 ± 0.01	129.9 ± 1.20
MultiUI	H200-SXM5	0.89 ± 0.01	107.0 ± 1.52
MultiUI	L40S	1.52 ± 0.03	209.0 ± 0.335
MultiUI	RTX 3090	2.00 ± 0.03	223.0 ± 3.16
MultiUI	RTX A6000	2.00 ± 0.05	269.0 ± 4.06
Synapse	A100-SXM4	2.62 ± 0.05	592.0 ± 5.00
Synapse	A100-PCIe	2.48 ± 0.06	606.0 ± 13.9
Synapse	H100-SXM5	1.90 ± 0.01	277.0 ± 1.52
Synapse	H100-NVL	1.74 ± 0.02	356.3 ± 2.80
Synapse	H200-SXM5	1.83 ± 0.03	276.0 ± 2.47
Synapse	L40S	2.40 ± 0.28	339.0 ± 38.2
Synapse	RTX A6000	3.70 ± 0.20	574.0 ± 19.4
Synatra	A100-SXM4	4.53 ± 0.11	668.0 ± 2.27
Synatra	A100-PCIe	4.36 ± 0.06	683.0 ± 2.15
Synatra	H100-SXM5	3.89 ± 0.06	409.0 ± 0.737
Synatra	H100-NVL	3.31 ± 0.04	426.5 ± 1.40
Synatra	H200-SXM5	3.74 ± 0.07	369.0 ± 1.71
Synatra	L40S	6.33 ± 0.10	906.0 ± 0.551
Synatra	RTX A6000	8.72 ± 0.07	1143.0 ± 2.04

Table 1: Energy and completion times per web agent and GPU

Energy Consumption per Benchmark Split and Energy per Token

We fruther compiled the detailed results for energy consumption per token for each web agent, the benchmark split and GPU. In Table 2, the complete results for AutoWebGLM (Lai et al. 2024) are shown. Table 3 shows the complete results for MindAct (Deng et al. 2024), Table 4 for MultiUI (Liu et al. 2024), Table 5 for Synatra (Ou et al. 2024), and Table 6 for Synapse (Zheng et al. 2023).

Agent	GPU	Split	Energy(kWh)	Energy/Token (10^{-9}kWh)
AutoWebGLM	A100-SXM4	domain	0.308 ± 0.011	1245.35 ± 44.48
AutoWebGLM	A100-SXM4	task	0.088 ± 0.008	1021.18 ± 92.83
AutoWebGLM	A100-SXM4	website	0.066 ± 0.005	1143.15 ± 86.6
AutoWebGLM	A100-PCIe	domain	0.314 ± 0.009	1269.61 ± 36.39
AutoWebGLM	A100-PCIe	task	0.094 ± 0.005	1090.8 ± 58.02
AutoWebGLM	A100-PCIe	website	0.068 ± 0.004	1177.8 ± 69.28
AutoWebGLM	H100-SXM5	domain	0.242 ± 0.011	978.49 ± 44.48
AutoWebGLM	H100-SXM5	task	0.072 ± 0.008	835.51 ± 92.83
AutoWebGLM	H100-SXM5	website	0.048 ± 0.004	831.38 ± 69.28
AutoWebGLM	H100-NVL	domain	0.22 ± 0.007	889.54 ± 28.3
AutoWebGLM	H100-NVL	task	0.062 ± 0.004	719.47 ± 46.42
AutoWebGLM	H100-NVL	website	0.048 ± 0.004	831.38 ± 69.28
AutoWebGLM	H200-SXM5	domain	0.232 ± 0.011	938.06 ± 44.48
AutoWebGLM	H200-SXM5	task	0.066 ± 0.005	765.88 ± 58.02
AutoWebGLM	H200-SXM5	website	0.05 ± 0.0	866.03 ± 0.0
AutoWebGLM	L40S	domain	0.388 ± 0.008	1568.82 ± 32.35
AutoWebGLM	L40S	task	0.112 ± 0.008	1299.68 ± 92.83
AutoWebGLM	L40S	website	0.084 ± 0.005	1454.92 ± 86.6
AutoWebGLM	RTX 3090	domain	0.532 ± 0.043	2151.06 ± 173.86
AutoWebGLM	RTX 3090	task	0.154 ± 0.005	1787.06 ± 58.02
AutoWebGLM	RTX 3090	website	0.114 ± 0.018	1974.54 ± 311.77
AutoWebGLM	RTX A6000	domain	0.516 ± 0.026	2086.37 ± 105.13
AutoWebGLM	RTX A6000	task	0.142 ± 0.008	1647.81 ± 92.83
AutoWebGLM	RTX A6000	website	0.102 ± 0.004	1766.69 ± 69.28

Table 2: Energy cost per Token for each Split on all GPUs for the AutoWebGLM web agent

Agent	GPU	Split	Energy(kWh)	Energy/Token (10^{-9}kWh)
MindAct	A100-SXM4	domain	1.0 ± 0.162	6.63 ± 1.07
MindAct	A100-SXM4	task	0.569 ± 0.145	8.41 ± 2.14
MindAct	A100-SXM4	website	0.328 ± 0.097	7.78 ± 2.3
MindAct	A100-PCIe	domain	0.73 ± 0.132	4.84 ± 0.87
MindAct	A100-PCIe	task	0.418 ± 0.129	6.18 ± 1.91
MindAct	A100-PCIe	website	0.245 ± 0.083	5.81 ± 1.97
MindAct	H100-SXM5	domain	0.823 ± 0.078	5.45 ± 0.52
MindAct	H100-SXM5	task	0.483 ± 0.117	7.14 ± 1.73
MindAct	H100-SXM5	website	0.271 ± 0.073	6.43 ± 1.73
MindAct	H100-NVL	domain	0.656 ± 0.152	4.47 ± 1.10
MindAct	H100-NVL	task	0.349 ± 0.086	4.23 ± 1.33
MindAct	H100-NVL	website	0.21 ± 0.066	4.98 ± 1.57
MindAct	H200-SXM5	domain	0.806 ± 0.103	5.34 ± 0.68
MindAct	H200-SXM5	task	0.454 ± 0.124	6.71 ± 1.83
MindAct	H200-SXM5	website	0.274 ± 0.068	6.5 ± 1.61
MindAct	L40S	domain	1.09 ± 0.113	7.22 ± 0.75
MindAct	L40S	task	0.575 ± 0.115	8.5 ± 1.7
MindAct	L40S	website	0.341 ± 0.075	8.09 ± 1.78
MindAct	RTX 3090	domain	0.47 ± 0.01	3.11 ± 0.07
MindAct	RTX 3090	task	0.498 ± 0.024	7.36 ± 0.35
MindAct	RTX 3090	website	0.308 ± 0.013	7.31 ± 0.31
MindAct	RTX A6000	domain	1.4 ± 0.356	9.28 ± 2.36
MindAct	RTX A6000	task	0.734 ± 0.206	10.85 ± 3.04
MindAct	RTX A6000	website	0.428 ± 0.124	10.15 ± 2.94

Table 3: Energy cost per Token for each Split on all GPUs for the MindAct web agent

Agent	GPU	Split	Energy(kWh)	Energy/Token (10^{-9}kWh)
MultiUI	A100-PCIe	domain	0.744 ± 0.011	468.49 ± 6.93
MultiUI	A100-PCIe	task	0.268 ± 0.008	412.35 ± 12.31
MultiUI	A100-PCIe	website	0.172 ± 0.008	432.53 ± 20.12
MultiUI	H100-SXM5	domain	0.588 ± 0.022	370.26 ± 13.85
MultiUI	H100-SXM5	task	0.206 ± 0.009	316.96 ± 13.85
MultiUI	H100-SXM5	website	0.136 ± 0.005	342.0 ± 12.57
MultiUI	H100-NVL	domain	0.518 ± 0.008	326.18 ± 5.04
MultiUI	H100-NVL	task	0.182 ± 0.004	280.03 ± 6.15
MultiUI	H100-NVL	website	0.122 ± 0.004	306.8 ± 10.06
MultiUI	H200-SXM5	domain	0.558 ± 0.008	351.37 ± 5.04
MultiUI	H200-SXM5	task	0.2 ± 0.0	307.73 ± 0.0
MultiUI	H200-SXM5	website	0.132 ± 0.004	331.94 ± 10.06
MultiUI	L40S	domain	0.966 ± 0.025	608.28 ± 15.74
MultiUI	L40S	task	0.336 ± 0.005	516.98 ± 7.69
MultiUI	L40S	website	0.222 ± 0.004	558.27 ± 10.06
MultiUI	RTX 3090	domain	1.26 ± 0.019	793.41 ± 11.96
MultiUI	RTX 3090	task	0.448 ± 0.016	689.31 ± 24.62
MultiUI	RTX 3090	website	0.29 ± 0.017	729.27 ± 42.75
MultiUI	RTX A6000	domain	1.24 ± 0.046	780.81 ± 28.97
MultiUI	RTX A6000	task	0.46 ± 0.0	707.77 ± 0.0
MultiUI	RTX A6000	website	0.3 ± 0.0	754.42 ± 0.0

Table 4: Energy cost per Token for each Split on all GPUs for the MultiUI web agent

Agent	GPU	Split	Energy(kWh)	Energy/Token (10^{-9}kWh)
Synatra	A100-SXM4	domain	2.85 ± 0.101	117.11 ± 4.15
Synatra	A100-SXM4	task	1.0 ± 0.015	112.05 ± 1.68
Synatra	A100-SXM4	website	0.676 ± 0.009	122.9 ± 1.64
Synatra	A100-PCIe	domain	2.77 ± 0.052	113.82 ± 2.14
Synatra	A100-PCIe	task	0.946 ± 0.023	106.0 ± 2.58
Synatra	A100-PCIe	website	0.638 ± 0.004	115.99 ± 0.73
Synatra	H100-SXM5	domain	2.48 ± 0.062	101.9 ± 2.55
Synatra	H100-SXM5	task	0.852 ± 0.016	95.47 ± 1.79
Synatra	H100-SXM5	website	0.558 ± 0.004	101.44 ± 0.73
Synatra	H100-NVL	domain	2.11 ± 0.027	86.7 ± 1.11
Synatra	H100-NVL	task	0.722 ± 0.008	80.9 ± 0.9
Synatra	H100-NVL	website	0.478 ± 0.004	86.9 ± 0.73
Synatra	H200-SXM5	domain	2.38 ± 0.056	97.79 ± 2.3
Synatra	H200-SXM5	task	0.814 ± 0.011	91.21 ± 1.23
Synatra	H200-SXM5	website	0.54 ± 0.019	98.17 ± 3.45
Synatra	L40S	domain	4.02 ± 0.094	165.18 ± 3.86
Synatra	L40S	task	1.38 ± 0.053	154.64 ± 5.94
Synatra	L40S	website	0.932 ± 0.029	169.44 ± 5.27
Synatra	RTX A6000	domain	5.61 ± 0.05	230.51 ± 2.05
Synatra	RTX A6000	task	1.87 ± 0.09	209.54 ± 10.08
Synatra	RTX A6000	website	1.25 ± 0.048	227.25 ± 8.73

Table 5: Energy cost per Token for each Split on all GPUs for the Synatra web agent

Agent	GPU	Split	Energy(kWh)	Energy/Token (10^{-9}kWh)
Synapse	A100-SXM4	domain	1.61 ± 0.05	234.05 ± 7.27
Synapse	A100-SXM4	task	0.64 ± 0.014	215.49 ± 4.71
Synapse	A100-SXM4	website	0.374 ± 0.011	193.82 ± 5.7
Synapse	A100-PCIe	domain	1.54 ± 0.055	223.88 ± 8.0
Synapse	A100-PCIe	task	0.59 ± 0.021	198.65 ± 7.07
Synapse	A100-PCIe	website	0.354 ± 0.005	183.46 ± 2.59
Synapse	H100-SXM5	domain	1.17 ± 0.022	170.09 ± 3.2
Synapse	H100-SXM5	task	0.458 ± 0.008	154.21 ± 2.69
Synapse	H100-SXM5	website	0.272 ± 0.004	140.96 ± 2.07
Synapse	H100-NVL	domain	1.07 ± 0.018	155.55 ± 2.62
Synapse	H100-NVL	task	0.422 ± 0.004	142.09 ± 1.35
Synapse	H100-NVL	website	0.252 ± 0.004	130.6 ± 2.07
Synapse	H200-SXM5	domain	1.13 ± 0.023	164.27 ± 3.34
Synapse	H200-SXM5	task	0.438 ± 0.011	147.47 ± 3.7
Synapse	H200-SXM5	website	0.262 ± 0.004	135.78 ± 2.07
Synapse	L40S	domain	1.56 ± 0.044	226.79 ± 6.4
Synapse	L40S	task	0.61 ± 0.008	205.38 ± 2.69
Synapse	L40S	website	0.354 ± 0.005	183.46 ± 2.59
Synapse	RTX A6000	domain	2.28 ± 0.183	331.46 ± 26.6
Synapse	RTX A6000	task	0.922 ± 0.105	310.43 ± 35.35
Synapse	RTX A6000	website	0.5 ± 0.047	259.12 ± 24.36

Table 6: Energy cost per Token for each Split on all GPUs for the Synapse web agent

Carbon Dioxide Emissions Equivalents

To provide a more complete overview of the CO_2 emission of each web agent, based on their energy consumption, we provide CO_2 equivalents for all conducted benchmark tests. Table 7 expands our CO_2 emission table in the paper, which only included the equivalents for the H100-NVL GPU. For comparison, our theoretical estimations for MindAct and LASER (Ma et al. 2023) are also attached, showcasing that even the most inefficient GPU coupled with the most inefficient open-source web agent in our lineup uses far less energy than our energy cost estimation for LASER.

Agent	GPU	Energy (kWh)	Norway (gCO_2)	US (gCO_2)	Australia (gCO_2)
AutoWebGLM	A100-SXM4	0.46	9	208	368
AutoWebGLM	A100-PCIe	0.48	9	217	384
AutoWebGLM	H100-SXM5	0.36	7	163	288
AutoWebGLM	H100-NVL	0.33	6	149	264
AutoWebGLM	H200-SXM5	0.35	7	158	280
AutoWebGLM	L40S	0.58	11	262	463
AutoWebGLM	RTX 3090	0.8	16	362	640
AutoWebGLM	RTX A6000	0.76	15	344	608
MindAct	A100-SXM4	1.9	38	860	1520
MindAct	A100-PCIe	1.39	27	629	1112
MindAct	H100-SXM5	1.58	31	715	1264
MindAct	H100-NVL	1.22	24	552	976
MindAct	H200-SXM5	1.51	30	684	1208
MindAct	L40S	2.01	40	910	1607
MindAct	RTX 3090	1.28	25	579	1024
MindAct	RTX A6000	2.57	51	1164	2056
MultiUI	A100-PCIe	1.18	23	534	944
MultiUI	H100-SXM5	0.93	18	421	744
MultiUI	H100-NVL	0.822	16	372	657
MultiUI	H200-SXM5	0.89	17	403	712
MultiUI	L40S	1.52	30	688	1216
MultiUI	RTX 3090	2.0	40	906	1600
MultiUI	RTX A6000	2.0	40	906	1600
Synapse	A100-SXM4	2.62	52	1186	2096
Synapse	A100-PCIe	2.48	49	1123	1984
Synapse	H100-SXM5	1.9	38	860	1520
Synapse	H100-NVL	1.74	34	788	1392
Synapse	H200-SXM5	1.83	36	828	1464
Synapse	L40S	2.4	48	1087	1920
Synapse	RTX A6000	3.7	74	1676	2960
Synatra	A100-SXM4	4.53	90	2052	3624
Synatra	A100-PCIe	4.36	87	1975	3488
Synatra	H100-SXM5	3.89	77	1762	3112
Synatra	H100-NVL	3.31	66	1499	2648
Synatra	H200-SXM5	3.74	74	1694	2992
Synatra	L40S	6.33	126	2867	5064
Synatra	RTX A6000	8.72	174	3950	6976
Estimation: MindAct		9.01	180	4081	7208
Estimation: LASER		99.21	1984	44942	79368

Table 7: The CO_2 consumed by running the web agents on different GPUs for the Mind2Web benchmark evaluated on the energy mix of Norway, US and Australia. Additional theoretical estimation for MindAct and LASER.

References

Deng, X.; Gu, Y.; Zheng, B.; Chen, S.; Stevens, S.; Wang, B.; Sun, H.; and Su, Y. 2024. Mind2web: Towards a generalist agent for the web. *Advances in Neural Information Processing Systems*, 36.

Lai, H.; Liu, X.; Iong, I. L.; Yao, S.; Chen, Y.; Shen, P.; Yu, H.; Zhang, H.; Zhang, X.; Dong, Y.; et al. 2024. Autowebglm: A large language model-based web navigating agent. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 5295–5306.

Liu, J.; Ou, T.; Song, Y.; Qu, Y.; Lam, W.; Xiong, C.; Chen, W.; Neubig, G.; and Yue, X. 2024. Harnessing webpage uis for text-rich visual understanding. *arXiv preprint arXiv:2410.13824*.

Ma, K.; Zhang, H.; Wang, H.; Pan, X.; and Yu, D. 2023. Laser: Llm agent with state-space exploration for web navigation. *arXiv preprint arXiv:2309.08172*.

Ou, T.; Xu, F. F.; Madaan, A.; Liu, J.; Lo, R.; Sridhar, A.; Sengupta, S.; Roth, D.; Neubig, G.; and Zhou, S. 2024. Synatra: Turning indirect knowledge into direct demonstrations for digital agents at scale. *Advances in Neural Information Processing Systems*, 37: 91618–91652.

Zheng, L.; Wang, R.; Wang, X.; and An, B. 2023. Synapse: Trajectory-as-exemplar prompting with memory for computer control. *arXiv preprint arXiv:2306.07863*.