GPT Causal Mask: Lower Triangular (No Future Peeking)

	The	cat	sat	on	the	mat
The	1	0	0	0	0	0
cat	1	1	0	0	0	0
sat	1	1	1	0	0	0
on	1	1	1	1	0	0
the	1	1	1	1	1	0
mat	1	1	1	1	1	1

Attend FROM (queries) — vertical axis

Attend TO (keys) — horizontal axis