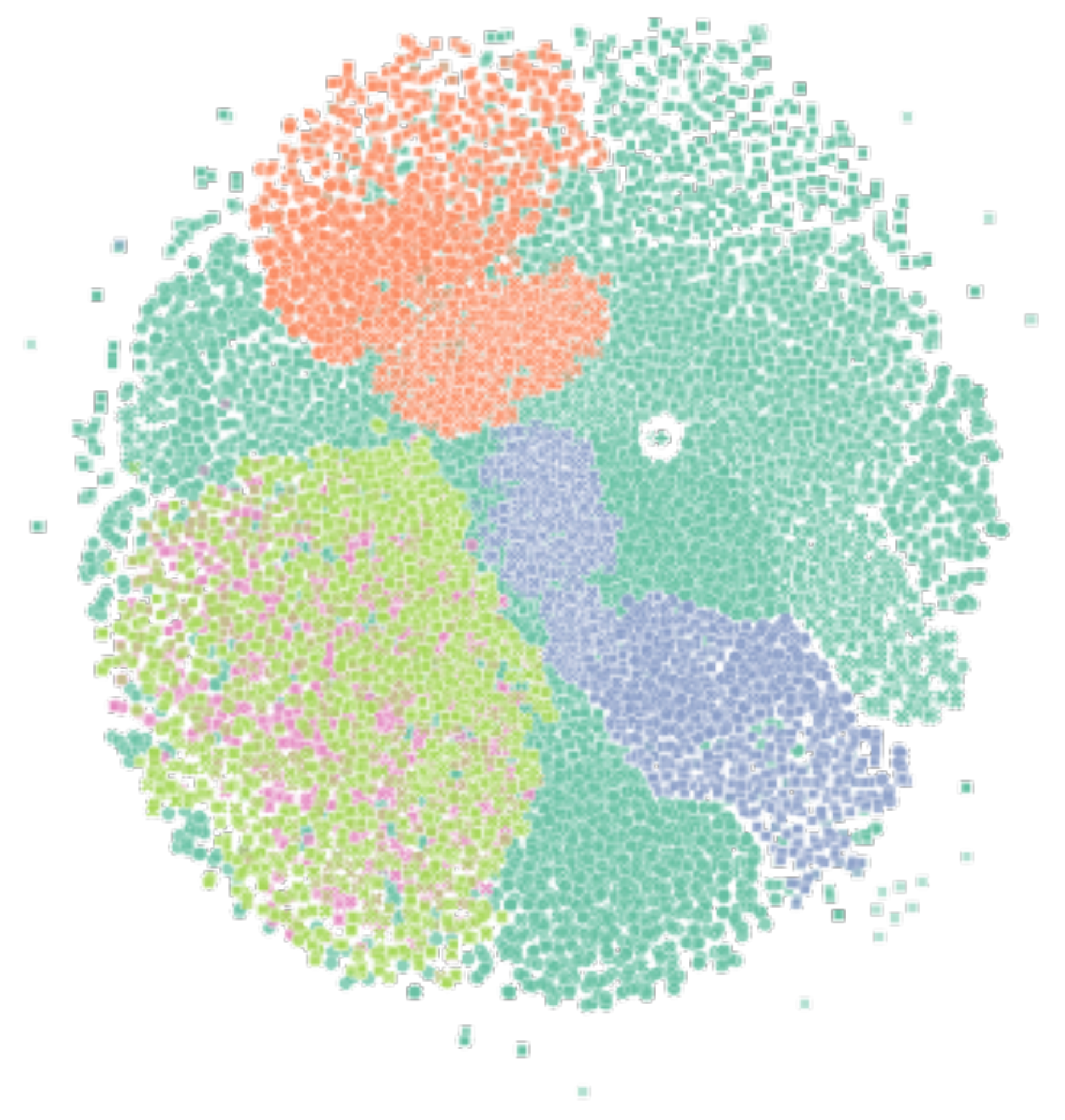
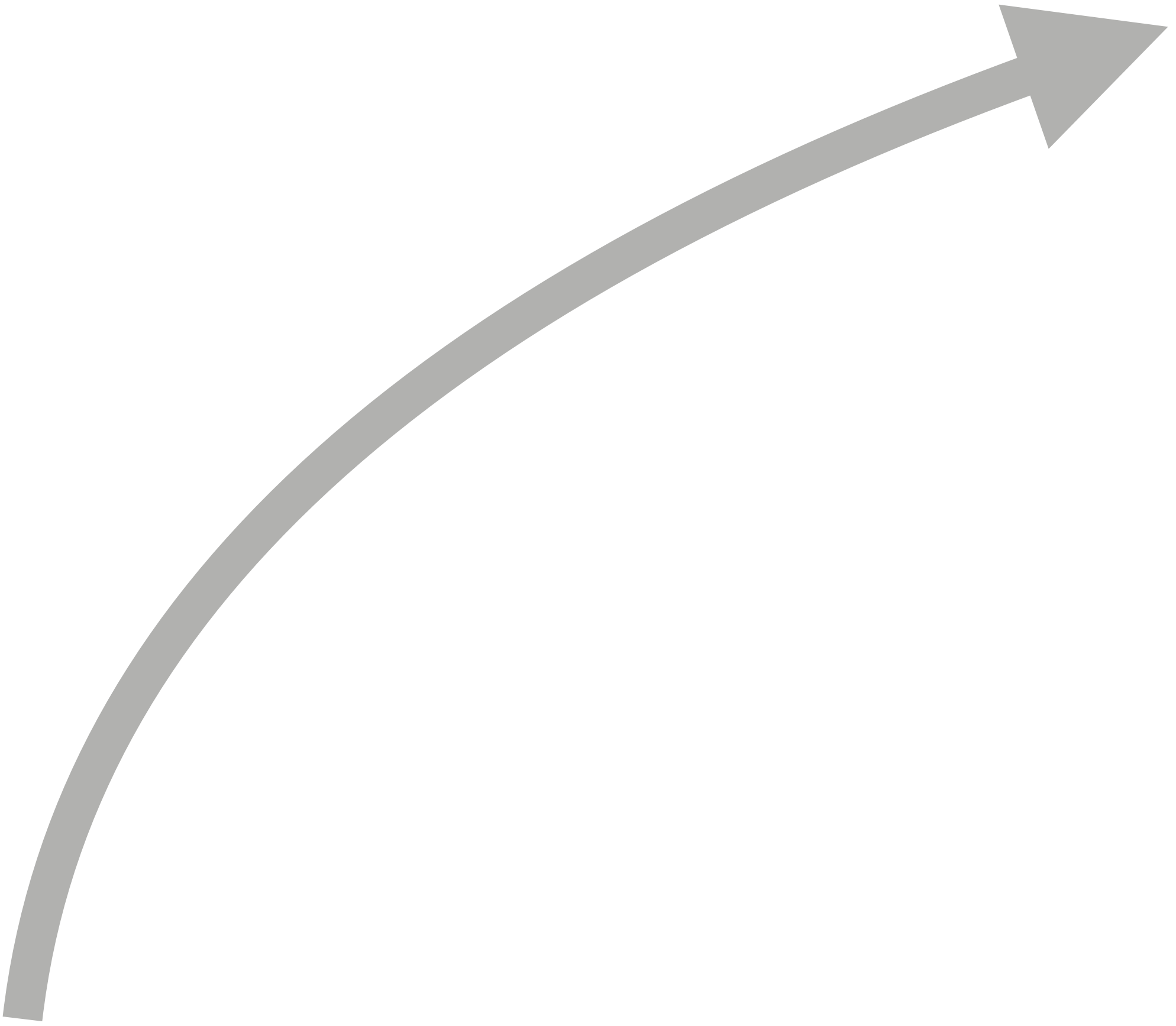
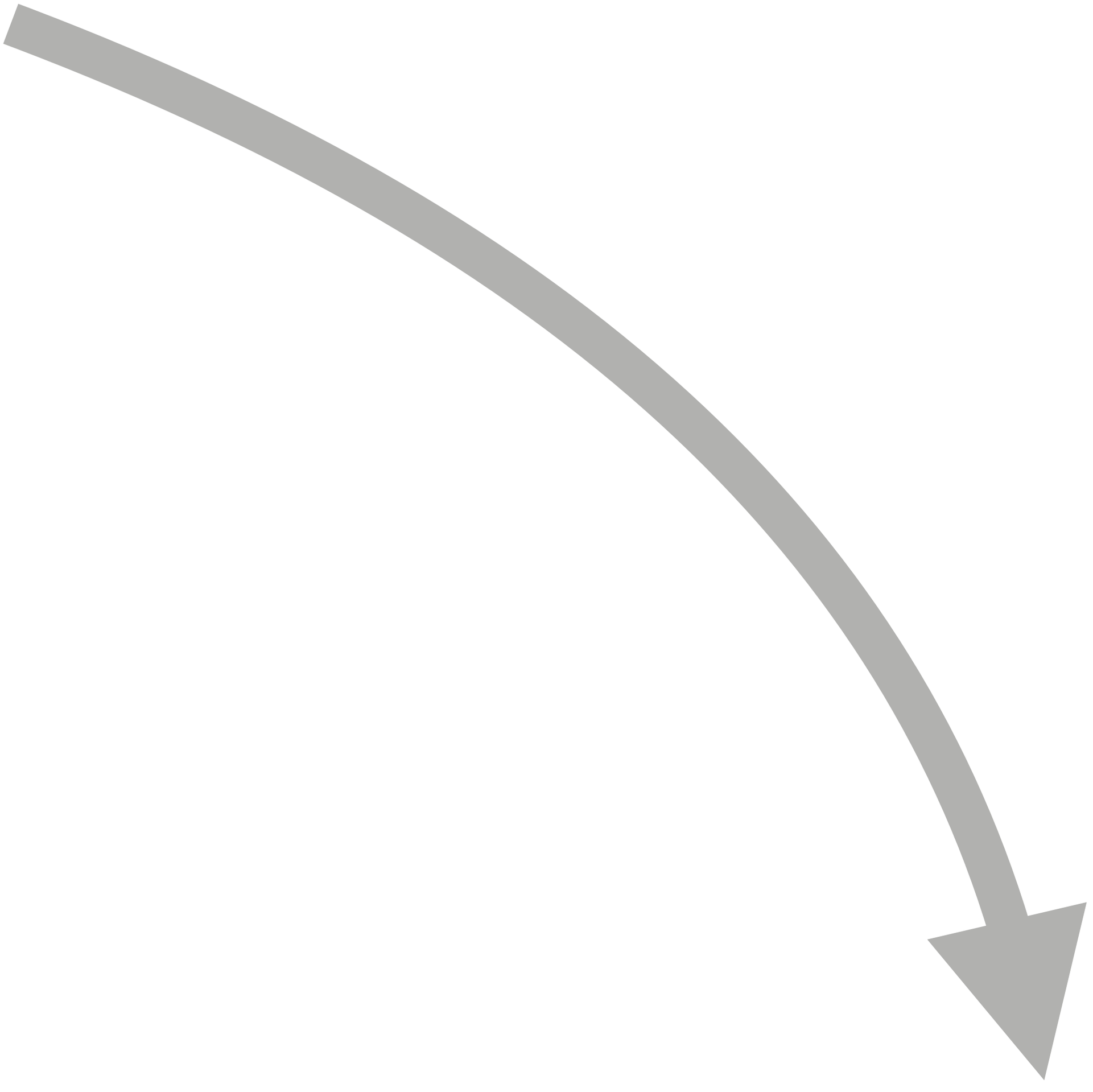
# Pretraining DFA Encoder

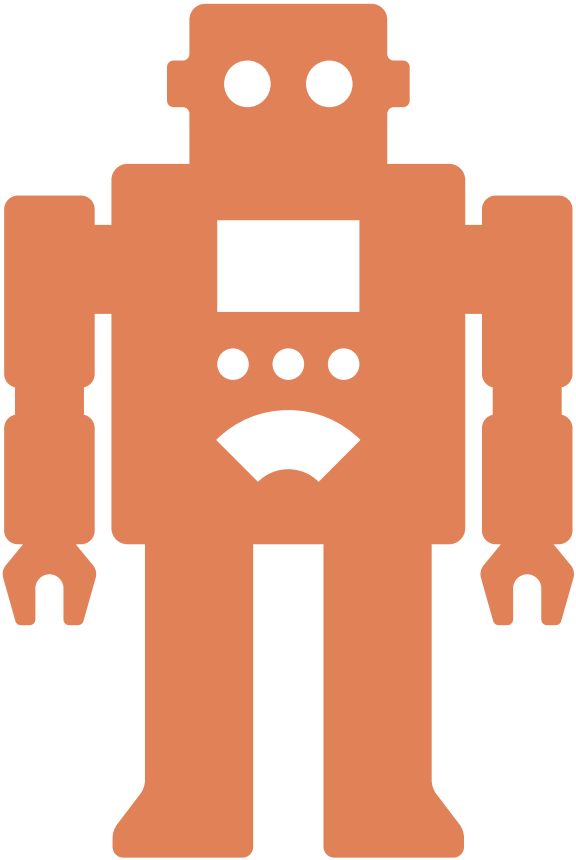
16

$a_t \sim \pi(s_t \mid g_t)$, where $a_t \in \{$ $\}$

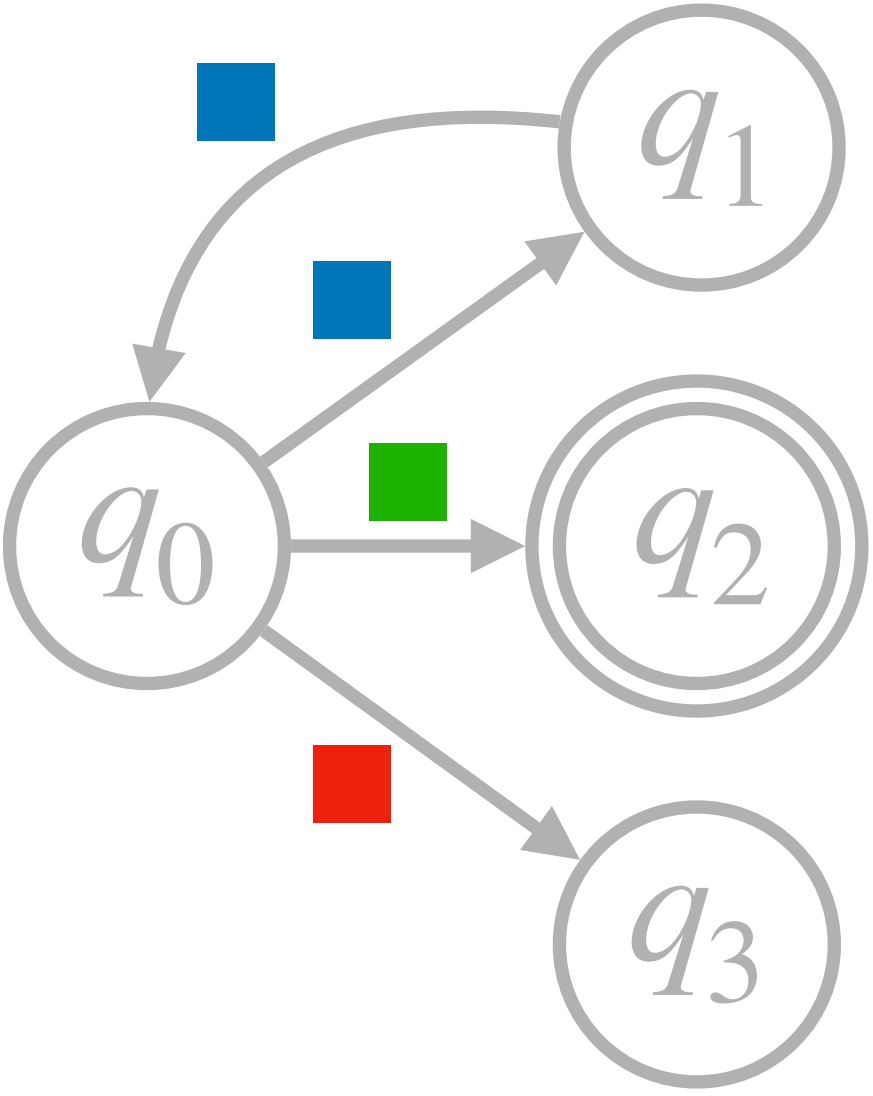








g t

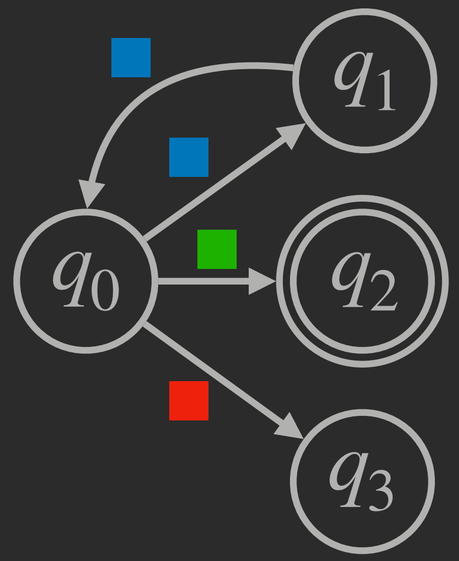

$q_1$
$q_0$
$q_2$
$q_3$

Agent learns to
“solve” a DFA

# Pretraining DFA Encoder

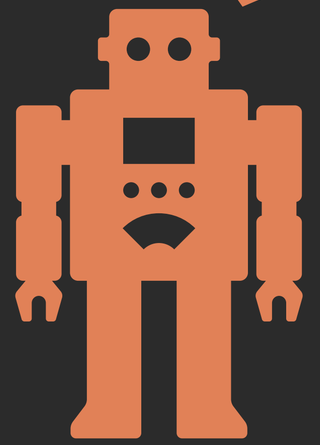
$a_t \sim \pi(s_t \mid g_t)$, where $a_t \in \{$ ■, ■, ■ $\}$

Agent learns to **“solve”** a DFA



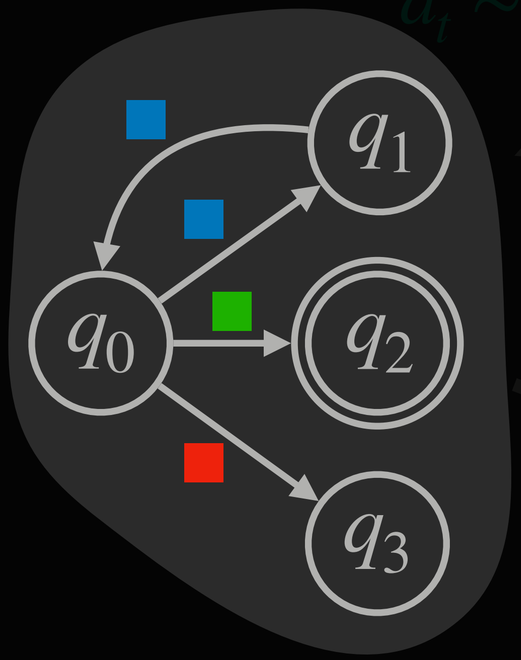

Encoder

$g_t$



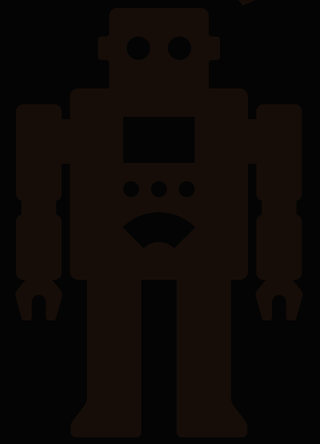
# Pretraining DFA Encoder

$a_t \sim \pi(s_t \mid g_t)$, where $a_t \in \{$ ■, ■, ■ $\}$




Encoder

$g_t$



Agent learns to "solve" a DFA

0 reward