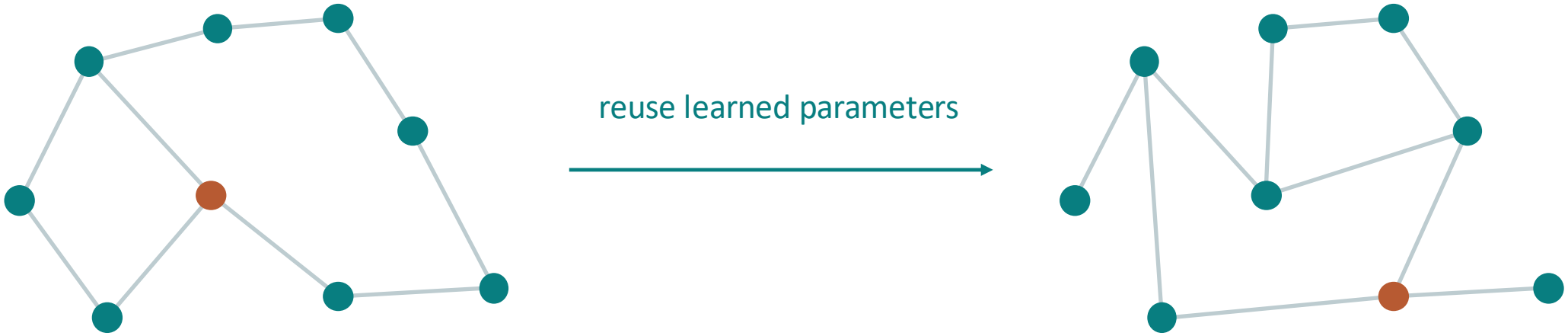
Graph foundation models

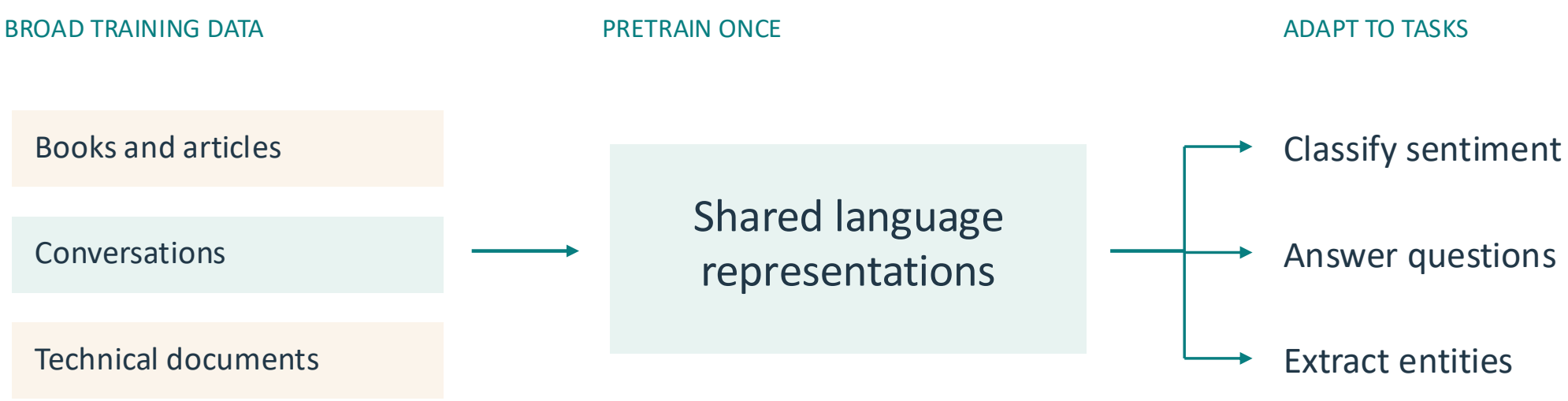
What can a model carry from one graph to another?



Source graph

New target graph

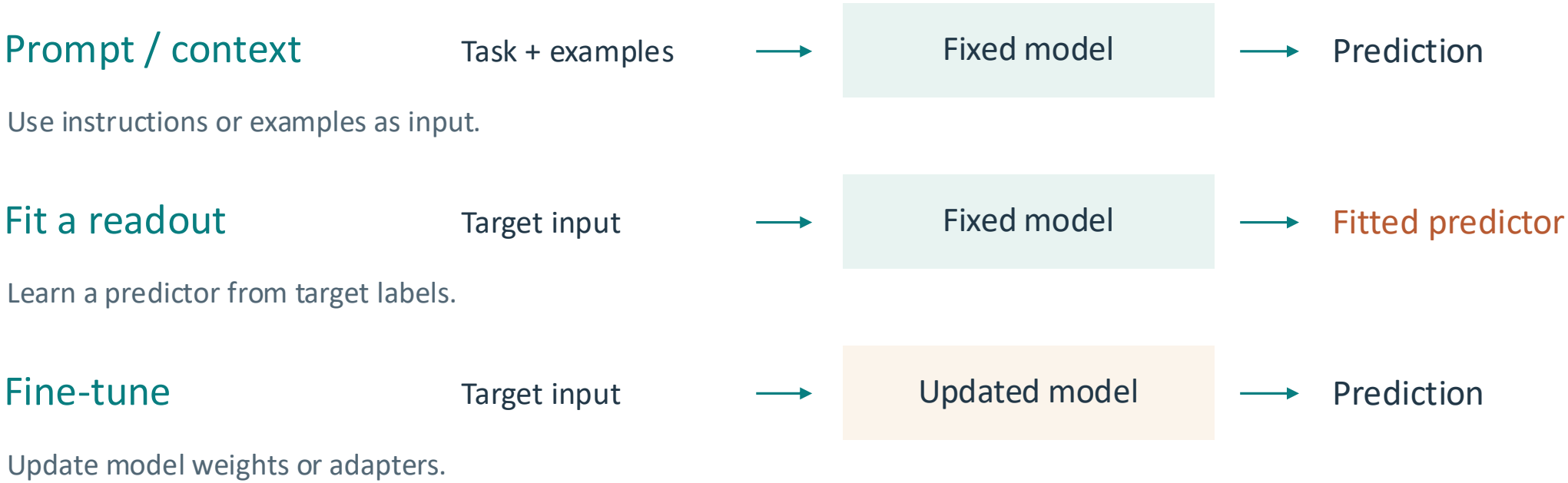
A foundation model supports many downstream tasks



Pretraining learns reusable patterns, often with self-supervised objectives.

The same idea extends to vision, audio, time series, and **graphs**.

Adaptation connects a pretrained model to a new task



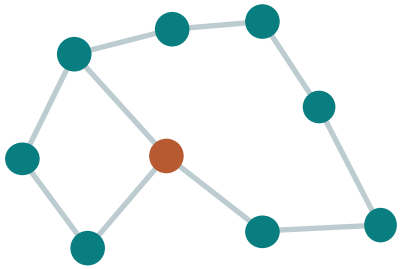
For graphs, we also need compatible features, structural information, and prediction targets.

Devlin et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL-HLT, 2019.

THE TRANSFER PROBLEM

Transfer requires a shared input and a shared prediction rule

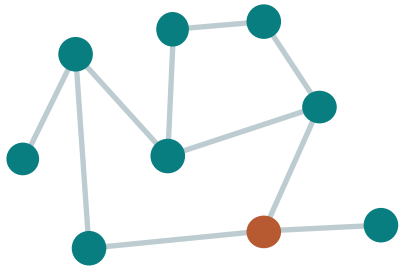
SOURCE GRAPHS



LEARNED PARAMETERS

Shared encoder

TARGET GRAPH



Train across graphs or tasks.

Choose features or structural roles with shared meaning.

Specify the query and the evidence available for adaptation.

Zero-shot transfer uses no target training labels. Few-shot transfer specifies a small target label budget.

01 / TRANSFER THROUGH TEXT

One For All gives classification tasks a shared format

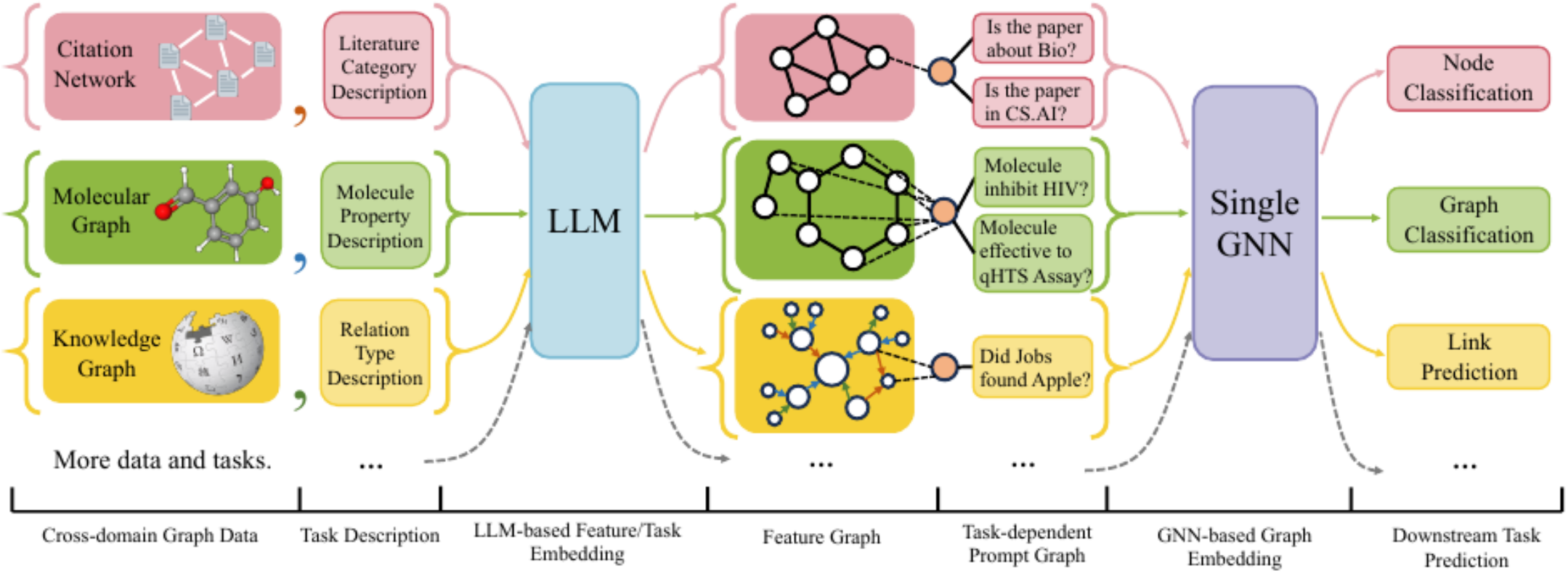
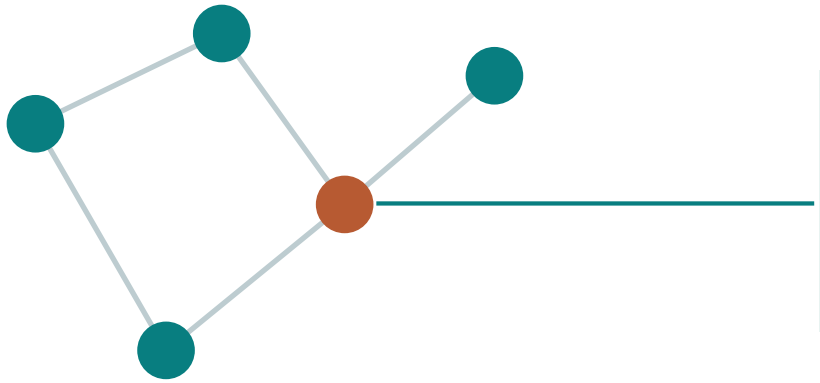


Figure 1 from Liu et al. · Follow one row from graph attributes and task text to a class prediction.

ONE FOR ALL / WORKED EXAMPLE

A class node turns a paper topic into a score

CITATION GRAPH



Target paper

TASK PROMPT

Identify the topic of this paper

CANDIDATE CLASSES

Graph learning

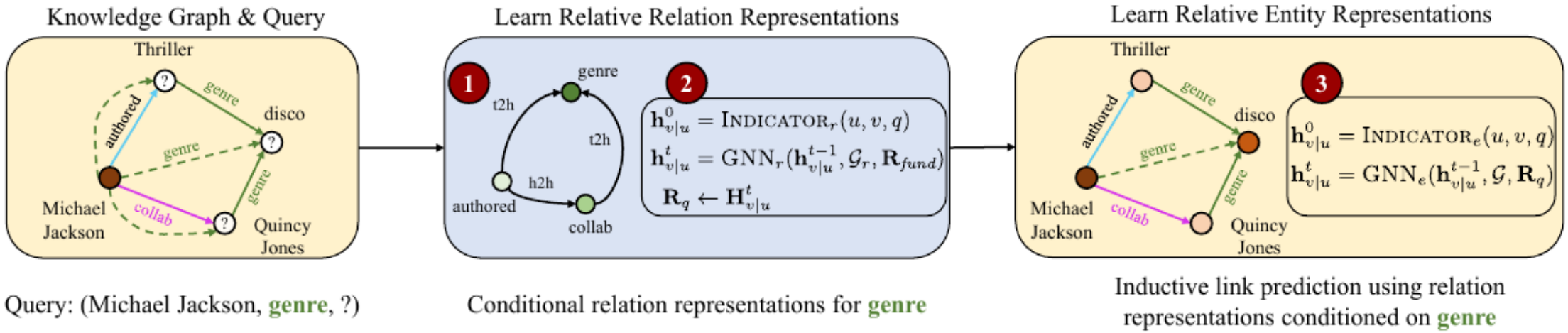
Computer vision

Each candidate receives a score from its class-node embedding after message passing.

$$P(\text{class } i \text{ applies}) = \sigma(\text{MLP}(h_i))$$

02 / TRANSFER THROUGH STRUCTURE

ULTRA represents relations by how their endpoints interact



Query: (Michael Jackson, **genre**, ?)

Conditional relation representations for **genre**

Inductive link prediction using relation representations conditioned on **genre**

1 FORM THE QUERY

Given an entity and relation, predict the missing tail.

2 ENCODE RELATIONS

Mark the query relation and propagate on the relation graph.

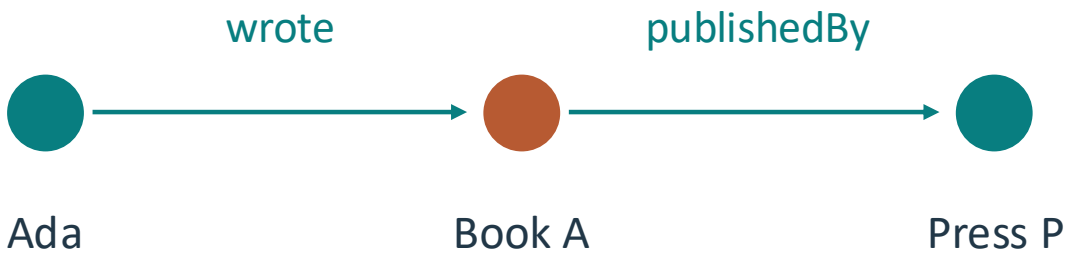
3 RANK ENTITIES

Use those vectors to propagate on the original graph.

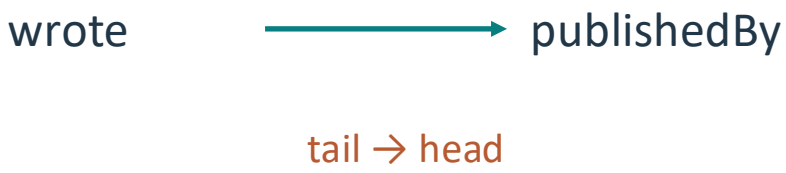
ULTRA / WORKED EXAMPLE

A shared entity connects two relation types

OBSERVED TRIPLES



RELATION GRAPH

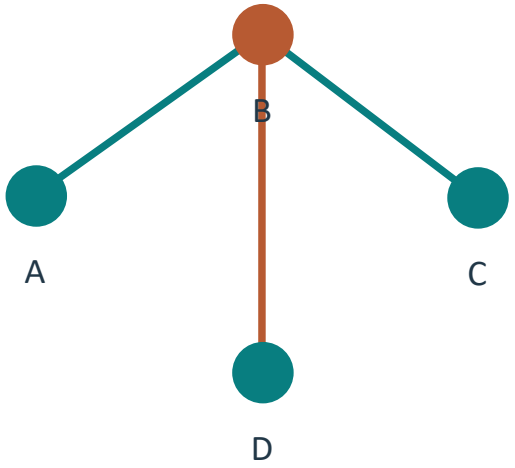


Book A is the tail of wrote and the head of publishedBy. That endpoint pattern transfers to new vocabularies.

2026 RESEARCH / FROZEN TIME-SERIES ENCODER

GraphFlowFM reads a graph as ordered paths

ILLUSTRATIVE GRAPH



TWO FLOWS COVER ITS EDGES

$A \rightarrow B \rightarrow C$

$B \rightarrow D$

Frozen TimesFM encodes each path

Nodes become sequence positions. Path order comes from graph connectivity.

Target labels fit a ridge readout over node representations. The pretrained encoder stays fixed.

Graph foundation models pursue several forms of transfer

Align meaning

Text describes attributes, tasks, and candidate labels.

OFA

Reuse graph patterns

Relations and neighborhoods provide shared structural roles.

ULTRA · GFT

Reuse other pretrained models

Convert graph evidence into tables or ordered sequences.

TFMLinker · GraphFlowFM

Assessment Promising transfer within defined tasks and domains. Broader reuse needs compatible graph semantics and structure.

Research priority Test new domains and graph formats with matched target labels, strong baselines, and reported adaptation cost.

Wang et al. Graph Foundation Models: A Comprehensive Survey. arXiv:2505.15116, 2025. · Wang et al. GFT: Graph Foundation Model with Transferable Tree Vocabulary. NeurIPS, 2024.
Yu et al. Evaluating Progress in Graph Foundation Models: A Comprehensive Benchmark and New Insights. arXiv:2603.10033, 2026.
Liao et al. TFMLinker: Universal Link Predictor by Graph In-Context Learning with Tabular Foundation Models. arXiv:2602.08592, 2026.

TAKE-HOME

A transfer claim should name what the model reuses

Shared evidence	OFA uses text, ULTRA uses relation roles, TFMLinker uses pair features, and GraphFlowFM uses ordered paths.
Target access	State which graph structure, text, examples, and labels are visible at prediction time.
A fair test	Hold out the target graph, match the supervision budget, and compare against a target-specific baseline.

A frozen model can still use target supervision through context rows or a fitted readout.