

Novelty Assessment Report

Paper: Transducing Language Models

PDF URL: <https://openreview.net/pdf?id=qOyF214xmg>

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-27

Abstract

Modern language models define distributions over strings, but their outputs are not always suited to downstream tasks. For instance, a model generating byte-pair strings may not be suitable when word-level predictions are needed, and a DNA model may not fit applications requiring amino acids. In such cases, a deterministic string-to-string transformation can convert the model's output to the desired form. This is a familiar pattern in probability theory: applying a function F to a random variable $X \sim p$ yields a transformed random variable $f(X)$ with an induced distribution. While such transformations are occasionally used in language modeling, they are not treated as yielding new, fully functional language models. We formalize this perspective and introduce a general framework for language models derived from deterministic string-to-string transformations. Focusing on transformations representable as finite-state transducers—a commonly used state-machine abstraction for efficient string-to-string mappings—we develop algorithms that compose a language model with an FST to marginalize over source strings mapping to a given target. This allows us to propagate probabilities through the transducer without altering model parameters and to condition on transformed outputs. We present an exact algorithm, an efficient approximation, and a theoretical analysis. We conduct experiments in three domains: converting token-level language models to character-level language models, token-level language models to word-level models, and deriving amino-acid models from DNA models. This demonstrates inference-time adaptation of pretrained language models to match application-specific output requirements.

Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

Core Task Landscape

This paper addresses: **Transforming Language Models via Finite-State Transducers**

A total of **50 papers** were analyzed and organized into a taxonomy with **22 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Theoretical Frameworks and Algorithms for FST-Based Language Processing**
- **Speech Recognition and Acoustic Modeling with FSTs**
- **Morphological Analysis and Generation Using FSTs**
- **Machine Translation and Transliteration via FSTs**
- **Hybrid Neural-FST Architectures and Integration**
- **Language Model Combination and Adaptation with FSTs**
- **Specialized FST Applications in Language Processing**
- **Large Language Models and Dialogue Systems**

Complete Taxonomy Tree

- Transforming Language Models via Finite-State Transducers Survey Taxonomy
- Theoretical Frameworks and Algorithms for FST-Based Language Processing
 - Core FST Composition and Optimization Algorithms (2 papers)
 - [16] A generalized composition algorithm for weighted finite-state transducers. (Cyril Allauzen, 2009) [View paper](#)
 - [24] Reduction of Intermediate Alphabets in Finite-State Transducer Cascades (Kempe, 2022) [View paper](#)
 - Probabilistic Model Conversion to FST Representations (3 papers)
 - [28] Finite State Transducers Approximating Hidden Markov Models (Kempe, 2022) [View paper](#)
 - [46] Conversion of Recurrent Neural Network Language Models to Weighted Finite State Transducers for Automatic Speech Recognition. (Petr, 2012) [View paper](#)
 - [47] Look-Back and Look-Ahead in the Conversion of Hidden Markov Models into Finite State Transducers (Kempe, 2022) [View paper](#)
 - Transducer-Based Language Model Transformation Theory ★ (2 papers)
 - [0] Transducing Language Models (Anon et al., 2026) [View paper](#)
 - [32] Transformers as Transducers (Lena Strobl, 2025) [View paper](#)
- Speech Recognition and Acoustic Modeling with FSTs
 - WFST-Based Speech Recognition Architectures (3 papers)
 - [1] Speech recognition algorithms using weighted finite-state transducers (Takaaki Hori, 2022) [View paper](#)
 - [19] Continuous Punjabi speech recognition model based on Kaldi ASR toolkit (Jyoti Guglani, 2018) [View paper](#)
 - [27] Spoken language processing using weighted finite state transducers (Isabel Trancoso, 2004) [View paper](#)
 - Dynamic and Incremental FST Decoding for Speech (2 papers)
 - [18] Incremental language models for speech recognition using finite-state transducers (H.J.G.A. Dolfig, 2001) [View paper](#)
 - [29] Delay-penalized CTC implemented based on Finite State Transducer (Zengwei Yao, 2023) [View paper](#)
 - Spoken Language Understanding with FST Constraints (2 papers)
 - [8] Joint intent detection and slot filling using weighted finite state transducer and BERT (Waheed Ahmed Abro, 2022) [View paper](#)

- [13] Finstreder: Simple and fast Spoken Language Understanding with Finite State Transducers using modern Speech-to-Text models (Bermuth, 2022) [View paper](#)
- Morphological Analysis and Generation Using FSTs
 - Morphological Transducers for Low-Resource Languages (4 papers)
 - [2] Developing finite-state language technology for Maya (Robert Pugh, 2023) [View paper](#)
 - [12] Unlocking finite-state morphological transducers: Derivational networks for Inuit-Yupik languages (Haley, 2025) [View paper](#)
 - [38] Finite-state transducer based modeling of morphosyntax with applications to Hungarian LVCSR (M. Szarvas, 2003) [View paper](#)
 - [45] Finite-state transducer based hungarian LVCSR with explicit modeling of phonological changes (MÁjt Á Szarvas, 2002) [View paper](#)
 - Morphological Reinflection and Derivational Networks (1 papers)
 - [20] Morphological reinflection with weighted finite-state transducers (Alice Kwak, 2023) [View paper](#)
 - Agglutination and Dialect Processing with FSTs (2 papers)
 - [23] Tunisian Dialect Agglutination Processing with Finite Transducers (Roua Torjmen, 2022) [View paper](#)
 - [25] Automatic conversion of dialectal Tamil text to standard written Tamil text using FSTs (Marimuthu K, 2014) [View paper](#)
- Machine Translation and Transliteration via FSTs
 - FST-Based Statistical Machine Translation (4 papers)
 - [9] Machine translation with inferred stochastic finite-state transducers (Francisco Casacuberta, 2004) [View paper](#)
 - [17] Learning finite state transducers using bilingual phrases (Jorge GonzÁlez, 2008) [View paper](#)
 - [35] A weighted finite state transducer translation template model for statistical machine translation (Shankar Kumar, 2005) [View paper](#)
 - [41] Finite-state transducer-based statistical machine translation using joint probabilities (Srinivas Bangalore, 2006) [View paper](#)
 - Dialect and Script Transliteration Using FSTs (2 papers)
 - [10] Automatic transliteration of romanized dialectal Arabic (Mohamed Al-Badrashiny, 2014) [View paper](#)
 - [11] Translation from Tunisian Dialect to Modern Standard Arabic: Exploring Finite-State Transducers and Sequence-to-Sequence Transformer Approaches (Roua Torjmen, 2024) [View paper](#)
 - Mobile Input and Keyboard Transliteration with WFSTs (3 papers)
 - [15] Transliterated Mobile Keyboard Input via Weighted Finite-State Transducers (Lars Hellsten, 2017) [View paper](#)
 - [37] Mobile Keyboard Input Decoding with Finite-State Transducers (Ouyang, 2022) [View paper](#)
 - [50] Predictive Text and Error Correction Using Weighted Finite State Transducers (Staes, n.d.) [View paper](#)
- Hybrid Neural-FST Architectures and Integration
 - Neural Language Models with FST Constraints (2 papers)
 - [6] Long-form speech translation through segmentation with finite-state decoding constraints on large language models (McCarthy, 2023) [View paper](#)
 - [42] Neural-FST Class Language Model for End-to-End Speech Recognition (Antoine Bruguier, 2022) [View paper](#)
 - Shallow Fusion and Rescoring with Neural-FST Combinations (4 papers)
 - [14] Neural and FST-based approaches to grammatical error correction (Zheng Yuan, 2019) [View paper](#)
 - [31] Transformer-based language modeling and decoding for conversational speech recognition (Nassar, 2022) [View paper](#)
 - [34] Neural Grammatical Error Correction with Finite State Transducers (Stahlberg, 2022) [View paper](#)
 - [39] Shallow Fusion of Weighted Finite-State Transducer and Language Model for Text Normalization (Evelina Bakhturina, 2022) [View paper](#)
 - Neuralized Finite-State Transducers (1 papers)
 - [26] A future for universal grapheme-phoneme transduction modeling with neuralized finite-state transducers (Lin, 2023) [View paper](#)
- Language Model Combination and Adaptation with FSTs
 - Multi-Model Language Model Combination via WFSTs (3 papers)
 - [22] Fitting class-based language models into weighted finite-state transducer framework. (Pavel Irting, 2003) [View paper](#)
 - [30] Language model combination and adaptation using weighted finite state transducers (X. Liu, 2010) [View paper](#)
 - [49] Embedding grammar to N-gram Language Model Based on Weighted Finite-State Transducer (Author, 2014) [View paper](#)
 - Domain-Specific Entity Modeling with FST-Neural Mixtures (2 papers)
 - [43] Statistical Language Modeling for Historical Documents using Weighted Finite-State Transducers and Long Short-Term Memory (Azawi, 2015) [View paper](#)
 - [44] A weighted finite-state transducer (WFST)-based language model for online Indic script handwriting recognition (Suhan Chowdhury, 2011) [View paper](#)
- Specialized FST Applications in Language Processing
 - Grapheme-to-Phoneme Conversion and Pronunciation Modeling (3 papers)
 - [7] Encoding linear models as weighted finite-state transducers. (K Wu, 2014) [View paper](#)
 - [33] Finite-state super transducers for compact language resource representation in edge voice-AI (Simon DobriÁek, 2022) [View paper](#)
 - [40] Weighted finite-state transducer approach to German compound words reconstruction for Speech Recognition (Nickolay Shamraev, 2015) [View paper](#)
 - Text Normalization and Error Correction with FSTs (1 papers)
 - [5] Building Metadata Inference Using a Transducer Based Language Model (Waterworth, 2022) [View paper](#)
 - Data Transformation and Non-Linguistic FST Applications (2 papers)
 - [4] Automating the Generation of a Functional Semantic Types Ontology with Foundational Models (Sachin Konan, 2024) [View paper](#)
 - [21] Data Transformation Acceleration using Deterministic Finite-State Transducers (Marziyeh Nourian, 2022) [View paper](#)
- Large Language Models and Dialogue Systems
 - LLM-Based Intelligent Tutoring and Dialogue Systems (2 papers)
 - [3] AutoTutor meets Large Language Models: A Language Model Tutor with Rich Pedagogy and Guardrails (Sankalan Pal Chowdhury, 2024) [View paper](#)
 - [36] Combining Neural Networks with Knowledge for Spoken Dialogue Systems (Waheed, 2023) [View paper](#)
 - Comparative Evaluation of FSTs and LLMs (1 papers)
 - [48] FSTs vs ICL: Generalisation in LLMs for an under-resourced language (Ximena Gutierrez, n.d.) [View paper](#)

Narrative

Core task: transforming language models via finite-state transducers. The field of FST-based language processing spans a diverse set of branches, each addressing distinct aspects of how finite-state machinery can be applied to linguistic computation. At the highest level, the taxonomy divides into theoretical frameworks and algorithms that underpin FST operations (including composition, generalization, and transducer-based model transformation), practical applications in speech recognition and acoustic modeling (e.g., Speech Recognition FST[1], Delay CTC FST[29]), morphological analysis and generation for richly inflected languages (e.g., Inuit Morphological Networks[12], Hungarian Morphosyntax FST[38]), machine translation and transliteration tasks (e.g., Stochastic FST Translation[9], Romanized Arabic Transliteration[10]), and hybrid neural-FST architectures that integrate deep learning with classical transducer methods (e.g., Neural FST Error[14], RNN to FST[46]). Additional branches cover language model combination and adaptation, specialized applications such as text normalization and keyboard decoding, and emerging work on large language models and dialogue systems.

Within this landscape, a particularly active line of inquiry concerns the theoretical underpinnings of transducer-based language model transformation, where researchers explore how FSTs can be used to systematically modify or constrain the output distributions of probabilistic models. Transducing Language Models[0] sits squarely in this theoretical branch, focusing on the formal mechanisms by which finite-state transducers can transform language model probabilities. This work contrasts with more application-driven efforts in speech or morphology, and it shares conceptual ground with Transformers as Transducers[32], which examines the expressive power of neural architectures through a transducer lens. While many branches emphasize engineering solutions for specific tasks, the transducer-based transformation theory branch addresses foundational questions about compositionality, expressiveness, and the algebraic properties of FST operations on language models, positioning Transducing Language Models[0] as a contribution to the formal toolkit that supports the broader ecosystem of FST applications.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Transformers as Transducers

Authors: Lena Strobl, Dana Angluin, David Chiang, Jonathan Rawski, Ashish Sabharwal | **Year/Venue:** 2025 • Transactions of the Association for Computational Linguistics | **URL:** [View paper](#)

Abstract

Abstract We study the sequence-to-sequence mapping capacity of transformers by relating them to finite transducers, and find that they can express surprisingly large classes of (total functional) transductions. We do so using variants of RASP, a programming language designed to help people âthink like transformers,â as an intermediate representation. We extend the existing Boolean variant B-RASP to sequence-to-sequence transductions and show that it computes exactly the first-order rational ...

Relationship Analysis

Both papers belong to the Transducer-Based Language Model Transformation Theory category, exploring theoretical frameworks for transforming language model outputs using finite-state transducers. The original paper focuses on composing pretrained language models with FSTs to create transduced language models that marginalize over source strings, enabling inference-time adaptation without parameter modification. In contrast, the candidate paper studies the computational expressiveness of transformers themselves as transducers, analyzing what classes of sequence-to-sequence mappings transformers can represent through the RASP programming language framework, rather than transforming existing language model distributions.

Contributions Analysis

Overall novelty summary. The paper formalizes a general framework for transforming language model distributions through deterministic string-to-string mappings representable as finite-state transducers. It occupies the 'Transducer-Based Language Model Transformation Theory' leaf, which contains only two papers including this one. This is a notably sparse research direction within the broader taxonomy of 50 papers across 22 leaf nodes, suggesting the work addresses a relatively underexplored theoretical niche. The sibling paper examines neural architectures through a transducer lens, whereas this work focuses on composing arbitrary language models with FSTs to marginalize over source strings.

The taxonomy reveals that most FST research concentrates on practical applications: speech recognition (5 papers across 3 leaves), morphological analysis (9 papers across 3 leaves), and machine translation (9 papers across 3 leaves). The theoretical frameworks branch, where this paper resides, is comparatively small with only 7 papers total across 3 leaves. Neighboring work includes core FST composition algorithms and probabilistic model conversion to FST representations, but these focus on optimization techniques and HMM/RNN conversion rather than the general transformation theory this paper develops. The scope note explicitly excludes empirical applications, reinforcing the paper's foundational theoretical positioning.

Among 18 candidates examined through limited semantic search, no contributions were clearly refuted. The 'general framework for transduced language models' examined 10 candidates with none providing overlapping prior work; 'algorithms for composing language models with FSTs' examined 1 candidate; and 'prefix decomposition of the precover' examined 7 candidates. This suggests that within the search scope, the specific formalization of FST-based language model transformation and the associated marginalization algorithms appear novel. However, the limited search scale (18 candidates, not exhaustive) means substantial related work may exist outside the top-K semantic matches examined.

Based on the limited literature search, the work appears to occupy a genuinely sparse theoretical area, with minimal direct competition in its specific leaf and few closely related papers in neighboring theoretical branches. The absence of refuting candidates across all three contributions, combined with the small size of the theoretical frameworks branch, suggests the formalization may represent a meaningful conceptual advance. However, the analysis covers only top-K semantic matches and does not guarantee comprehensive coverage of all relevant prior work in FST theory or language model transformation.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: General framework for transduced language models

Description: The authors introduce a foundational framework that formalizes language models obtained by applying deterministic string-to-string transformations (encoded as finite-state transducers) to existing language models, enabling inference-time adaptation without retraining.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Finite-state transducers in language and speech processing

URL: [View paper](#)

Brief Assessment

FST Language Speech[59] focuses on finite-state transducers for string-to-string and string-to-weight transformations in language and speech processing, not on language models or inference-time adaptation without retraining.

2. Utilizing process models in the requirements engineering process through model2text transformation

[URL: View paper](#)

Brief Assessment

Model to Text[51] focuses on transforming business process models (BPMN) into textual descriptions for requirements engineering, not on general language model transformations via finite-state transducers. The technical domains and objectives are fundamentally different.

3. Text-to-text extraction and verbalization of biomedical event graphs

[URL: View paper](#)

Brief Assessment

Biomedical Event Extraction[56] focuses on extracting and verbalizing biomedical event graphs using text-to-text transformations, not on general language model transformations via finite-state transducers.

4. DecoStrat: Leveraging the capabilities of language models in D2T generation via decoding framework

[URL: View paper](#)

Brief Assessment

DecoStrat[55] focuses on a decoding framework for data-to-text generation tasks, not on transforming language models through deterministic string-to-string transformations or finite-state transducers.

5. Neural transition-based string transduction for limited-resource setting in morphology

[URL: View paper](#)

Brief Assessment

Neural String Transduction[52] focuses on morphological string transduction tasks (inflection, lemmatization) using neural transition-based systems with edit actions, not on general language model transformations via finite-state transducers for inference-time adaptation.

6. AutoTutor meets Large Language Models: A Language Model Tutor with Rich Pedagogy and Guardrails

[URL: View paper](#)

Brief Assessment

AutoTutor LLM[3] uses finite state transducers for pedagogical control in tutoring systems, not for transforming language model distributions over strings. The candidate focuses on educational applications with handcrafted pedagogy, while the original introduces a probabilistic framework for string-to-string transformations of language models.

7. Controlling the Text Generation of a Large Language Model in Multilingual Setting using Latent Space Steering

[URL: View paper](#)

Brief Assessment

Multilingual Latent Steering[57] focuses on controlling language and style attributes in multilingual LLM text generation through latent space steering vectors, not on formalizing language models derived from deterministic string-to-string transformations or finite-state transducers.

8. What languages are easy to language-model? a perspective from learning probabilistic regular languages

[URL: View paper](#)

Brief Assessment

Easy Language Modeling[53] focuses on learning probabilistic regular languages and measuring learnability of formal language models, not on deterministic string-to-string transformations or inference-time adaptation of language models through transducers.

9. Contrastive Deterministic Autoencoders For Language Modeling

[URL: View paper](#)

Brief Assessment

Contrastive Deterministic Autoencoders[58] focuses on variational autoencoders for text generation and representation learning, not on deterministic string-to-string transformations or finite-state transducers for language model adaptation.

10. De-diffusion makes text a strong cross-modal interface

[URL: View paper](#)

Brief Assessment

Dediffusion Cross Modal[54] focuses on encoding images into text using pre-trained text-to-image diffusion models as decoders, not on deterministic string-to-string transformations of language models or finite-state transducers for inference-time adaptation.

Contribution 2: Algorithms for composing language models with FSTs

Description: The authors develop exact and approximate algorithms that compose a language model with a finite-state transducer to compute probabilities over transformed strings by marginalizing over source strings, enabling sampling, scoring, and conditioning on transformed outputs.

This contribution was assessed against **1 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Fitting class-based language models into weighted finite-state transducer framework.

[URL: View paper](#)

Brief Assessment

Class Language FST[22] focuses on incorporating class-based language models (with many-to-many word-to-class mappings) into FST frameworks for speech recognition, not on composing general language models with FSTs to marginalize over source strings for transformed outputs as in the original paper.

Contribution 3: Prefix decomposition of the precover

Description: The authors introduce a prefix decomposition method that decomposes the precover into quotient and remainder sets, enabling finite-time computation of prefix probabilities for transduced language models under certain conditions on the transformation.

This contribution was assessed against **7 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. ChunkAttention: Efficient Self-Attention with Prefix-Aware KV Cache and Two-Phase Partition

[URL: View paper](#)

Brief Assessment

ChunkAttention[63] focuses on optimizing self-attention computation in LLM serving by detecting shared system prompts and structuring KV cache using prefix trees. This is an engineering optimization for attention mechanisms, not a method for computing prefix probabilities in transformed language models through mathematical decomposition.

2. Learning Tractable Distributions Of Language Model Continuations

[URL: View paper](#)

Brief Assessment

Tractable Distributions LM[64] focuses on learning tractable surrogate models (HMMs) for controlled generation with sequence-level constraints, not on prefix decomposition methods for computing prefix probabilities in transduced language models.

3. Prefix-Suffix Based Statistical Language Models of Turkish

[URL: View paper](#)

Brief Assessment

Turkish Prefix Suffix[66] focuses on statistical language models for Turkish morphology using prefix-suffix decomposition for linguistic units, not on computing prefix probabilities for transduced language models through precover decomposition into quotient and remainder sets.

4. Self-Comparison for Dataset-Level Membership Inference in Large (Vision-)Language Model

[URL: View paper](#)

Brief Assessment

Dataset Membership Inference[62] focuses on membership inference attacks for detecting training data in LLMs/VLMs using self-comparison methods. It does not address prefix decomposition methods for computing prefix probabilities in transformed language models, which is the core technical contribution of the original paper.

5. Token-wise Decomposition of Autoregressive Language Model Hidden States for Analyzing Model Predictions

[URL: View paper](#)

Brief Assessment

Token Decomposition Hidden[60] focuses on decomposing transformer hidden states into token-wise contributions for analyzing model predictions, not on prefix decomposition methods for computing prefix probabilities in transformed language models.

6. PRP: Propagating Universal Perturbations to Attack Large Language Model Guard-Rails

[URL: View paper](#)

Brief Assessment

Universal Perturbations Attack[61] focuses on adversarial attacks against LLM guard models using prefix-based perturbations, not on prefix decomposition methods for computing prefix probabilities in transformed language models. The technical domains are entirely different.

7. A Generalized Language Model in Tensor Space

[URL: View paper](#)

Brief Assessment

Tensor Space LM[65] focuses on tensor-based language modeling using tensor networks and decomposition for high-order dependencies in language sequences, not on prefix decomposition methods for computing prefix probabilities in transformed language models.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Transducing Language Models [View paper](#)
- [1] Speech recognition algorithms using weighted finite-state transducers [View paper](#)
- [2] Developing finite-state language technology for Maya [View paper](#)
- [3] AutoTutor meets Large Language Models: A Language Model Tutor with Rich Pedagogy and Guardrails [View paper](#)
- [4] Automating the Generation of a Functional Semantic Types Ontology with Foundational Models [View paper](#)
- [5] Building Metadata Inference Using a Transducer Based Language Model [View paper](#)
- [6] Long-form speech translation through segmentation with finite-state decoding constraints on large language models [View paper](#)
- [7] Encoding linear models as weighted finite-state transducers. [View paper](#)
- [8] Joint intent detection and slot filling using weighted finite state transducer and BERT [View paper](#)
- [9] Machine translation with inferred stochastic finite-state transducers [View paper](#)
- [10] Automatic transliteration of romanized dialectal Arabic [View paper](#)
- [11] Translation from Tunisian Dialect to Modern Standard Arabic: Exploring Finite-State Transducers and Sequence-to-Sequence Transformer Approaches [View paper](#)
- [12] Unlocking finite-state morphological transducers: Derivational networks for Inuit-Yupik languages [View paper](#)
- [13] Finstreder: Simple and fast Spoken Language Understanding with Finite State Transducers using modern Speech-to-Text models [View paper](#)
- [14] Neural and FST-based approaches to grammatical error correction [View paper](#)
- [15] Transliterated Mobile Keyboard Input via Weighted Finite-State Transducers [View paper](#)
- [16] A generalized composition algorithm for weighted finite-state transducers. [View paper](#)
- [17] Learning finite state transducers using bilingual phrases [View paper](#)
- [18] Incremental language models for speech recognition using finite-state transducers [View paper](#)
- [19] Continuous Punjabi speech recognition model based on Kaldi ASR toolkit [View paper](#)

- [20] Morphological reinflection with weighted finite-state transducers [View paper](#)
- [21] Data Transformation Acceleration using Deterministic Finite-State Transducers [View paper](#)
- [22] Fitting class-based language models into weighted finite-state transducer framework. [View paper](#)
- [23] Tunisian Dialect Agglutination Processing with Finite Transducers [View paper](#)
- [24] Reduction of Intermediate Alphabets in Finite-State Transducer Cascades [View paper](#)
- [25] Automatic conversion of dialectal Tamil text to standard written Tamil text using FSTs [View paper](#)
- [26] A future for universal grapheme-phoneme transduction modeling with neuralized finite-state transducers [View paper](#)
- [27] Spoken language processing using weighted finite state transducers [View paper](#)
- [28] Finite State Transducers Approximating Hidden Markov Models [View paper](#)
- [29] Delay-penalized CTC implemented based on Finite State Transducer [View paper](#)
- [30] Language model combination and adaptation using weighted finite state transducers [View paper](#)
- [31] Transformer-based language modeling and decoding for conversational speech recognition [View paper](#)
- [32] Transformers as Transducers [View paper](#)
- [33] Finite-state super transducers for compact language resource representation in edge voice-AI [View paper](#)
- [34] Neural Grammatical Error Correction with Finite State Transducers [View paper](#)
- [35] A weighted finite state transducer translation template model for statistical machine translation [View paper](#)
- [36] Combining Neural Networks with Knowledge for Spoken Dialogue Systems [View paper](#)
- [37] Mobile Keyboard Input Decoding with Finite-State Transducers [View paper](#)
- [38] Finite-state transducer based modeling of morphosyntax with applications to Hungarian LVCSR [View paper](#)
- [39] Shallow Fusion of Weighted Finite-State Transducer and Language Model for Text Normalization [View paper](#)
- [40] Weighted finite-state transducer approach to German compound words reconstruction for Speech Recognition [View paper](#)
- [41] Finite-state transducer-based statistical machine translation using joint probabilities [View paper](#)
- [42] Neural-FST Class Language Model for End-to-End Speech Recognition [View paper](#)
- [43] Statistical Language Modeling for Historical Documents using Weighted Finite-State Transducers and Long Short-Term Memory [View paper](#)
- [44] A weighted finite-state transducer (WFST)-based language model for online Indic script handwriting recognition [View paper](#)
- [45] Finite-state transducer based hungarian LVCSR with explicit modeling of phonological changes [View paper](#)
- [46] Conversion of Recurrent Neural Network Language Models to Weighted Finite State Transducers for Automatic Speech Recognition. [View paper](#)
- [47] Look-Back and Look-Ahead in the Conversion of Hidden Markov Models into Finite State Transducers [View paper](#)
- [48] FSTs vs ICL: Generalisation in LLMs for an under-resourced language [View paper](#)
- [49] Embedding grammar to N-gram Language Model Based on Weighted Finite-State Transducer [View paper](#)
- [50] Predictive Text and Error Correction Using Weighted Finite State Transducers [View paper](#)
- [51] Utilizing process models in the requirements engineering process through model2text transformation [View paper](#)
- [52] Neural transition-based string transduction for limited-resource setting in morphology [View paper](#)
- [53] What languages are easy to language-model? a perspective from learning probabilistic regular languages [View paper](#)
- [54] De-diffusion makes text a strong cross-modal interface [View paper](#)
- [55] DecoStrat: Leveraging the capabilities of language models in D2T generation via decoding framework [View paper](#)
- [56] Text-to-text extraction and verbalization of biomedical event graphs [View paper](#)
- [57] Controlling the Text Generation of a Large Language Model in Multilingual Setting using Latent Space Steering [View paper](#)
- [58] Contrastive Deterministic Autoencoders For Language Modeling [View paper](#)
- [59] Finite-state transducers in language and speech processing [View paper](#)
- [60] Token-wise Decomposition of Autoregressive Language Model Hidden States for Analyzing Model Predictions [View paper](#)
- [61] PRP: Propagating Universal Perturbations to Attack Large Language Model Guard-Rails [View paper](#)
- [62] Self-Comparison for Dataset-Level Membership Inference in Large (Vision-)Language Model [View paper](#)
- [63] ChunkAttention: Efficient Self-Attention with Prefix-Aware KV Cache and Two-Phase Partition [View paper](#)
- [64] Learning Tractable Distributions Of Language Model Continuations [View paper](#)
- [65] A Generalized Language Model in Tensor Space [View paper](#)
- [66] Prefix-Suffix Based Statistical Language Models of Turkish [View paper](#)