

A Reciprocal Learning Paradigm: Cultivating Machine Learning and Japanese Kanji Proficiency Through System Development

Executive Summary

This report investigates the innovative proposition of concurrently developing a machine learning (ML) system for Japanese Kanji acquisition while the human creator simultaneously enhances their expertise in ML, data analysis, and data science. This approach is grounded in pedagogical principles such as "learning by building" and "machine teaching." The analysis delves into the theoretical underpinnings, examines existing technological capabilities, addresses practical development challenges, and explores the profound pedagogical implications of such a reciprocal learning framework. Current research indicates strong technical feasibility and significant potential for this interdisciplinary methodology to overcome common difficulties in Kanji learning, such as motivation and retention, while fostering advanced ML and data science competencies. The iterative and resource-intensive nature of ML development, often perceived as a burden, is reframed as an inherent component of a deeply experiential and human-centered learning journey.

1. Introduction: Synergizing Machine Learning Development with Japanese Kanji Acquisition

The journey of mastering a new skill, particularly a complex one like Japanese Kanji, often presents formidable challenges. Simultaneously, the pursuit of expertise in rapidly evolving fields such as machine learning and data science demands practical, hands-on engagement. This report explores a novel intersection of these two domains: the development of an ML-powered system designed to facilitate Japanese

Kanji acquisition, with the unique objective that the very act of building and refining this system serves as the primary mechanism for the human developer's own learning in both Kanji and ML/data science.

1.1 The Challenge of Kanji Learning: Motivation and Retention

Learning Japanese Kanji is widely recognized as a difficult endeavor, often characterized by the sheer volume of characters, their intricate stroke orders, and multiple readings and meanings.¹ A significant obstacle for learners is the inherent difficulty in remembering Kanji, which frequently leads to a pronounced lack of motivation, even among those with a strong desire for proficiency [User Query]. Traditional and many current digital methods for Kanji acquisition often rely on rote memorization or flashcard-like drills, which can feel overwhelming and fail to provide the dynamic engagement necessary for sustained effort [User Query]. This perceived difficulty and the resulting motivational decline represent a critical gap in existing learning paradigms.

The inherent motivational gap presents a unique design opportunity for ML-driven solutions. If a proposed ML project can demonstrably make Kanji learning more engaging, personalized, and effective by offering novel feedback mechanisms and progress tracking, it directly addresses this fundamental motivational problem. This transforms a personal struggle into a driving force for the project's design and a key metric for its success. The intensive ML development process, which involves significant time and resource commitment, can be viewed through a human-centered lens if it effectively solves such a deeply human learning challenge. A successful academic endeavor in this area would not only demonstrate technical prowess in ML but also a measurable improvement in the human learner's motivation, engagement, and retention of Kanji, thereby strengthening the argument for the human value embedded within the resource-intensive ML development process.

1.2 The Aspiration: Learning ML/Data Science by Building a Kanji Learning System

The central premise of this exploration is the strategic decision to leverage the process of constructing an ML program for Kanji learning as a practical pathway to

acquire and deepen skills in machine learning, data analysis, and data science [User Query]. This approach aligns seamlessly with established project-based learning (PBL) methodologies, where profound understanding and skill acquisition are cultivated through active application and real-world problem-solving.

Project-based learning emphasizes that students move beyond mere memorization to actively build, conduct original research, and collaborate across disciplines, fostering an inquisitive mindset that continually asks "why," "how," and "what next".³ The proposed project is a prime example of PBL, where the practical "material" for data science and ML learning is the complex linguistic domain of Japanese Kanji. By confronting authentic ML challenges—such as ensuring data quality, managing model complexity, and optimizing computational resources—within this specific context, the developer will acquire practical, transferable skills that extend far beyond abstract theoretical knowledge. This iterative process also inherently encourages reflection on challenges encountered and revisions made, which is fundamental for deep and lasting learning.³ The very act of designing, implementing, and debugging an ML system for Kanji will necessitate a comprehensive understanding of ML principles, data preprocessing techniques, robust model evaluation strategies, and practical deployment considerations. This engagement moves the learner beyond abstract concepts to hands-on application and troubleshooting, fostering a "T-shaped" skillset that combines broad AI knowledge with deep domain expertise.⁴

1.3 Thesis Proposition: A Reciprocal Learning Paradigm

The core hypothesis underpinning this investigation posits a reciprocal learning loop: the machine not only becomes proficient in Kanji, but, crucially, the human creator concurrently achieves proficiency in Kanji *through the very act of creating and training the machine* [User Query]. This paradigm exemplifies the pedagogical model often referred to as "learning by teaching" or, more specifically, "machine teaching."

The developer's observation that ML development "takes a long time, great resources of the physical hardware, and a lot of raw data to feed," followed by the profound question, "Isn't this process looks like the most humanist way of learning something?" [User Query], highlights a critical perspective. The intensive, iterative nature of ML development—encompassing data collection, meticulous cleaning, precise labeling, rigorous model training, detailed error analysis, continuous refinement, and ongoing monitoring⁵—strikingly mirrors the human learning process. This human learning

involves repeated exposure, deliberate practice, the assimilation of detailed feedback, and the adaptive adjustment of strategies. This interpretation transforms the perceived burdens of resource demands and time commitment into inherent features of a deep, experiential learning journey. The continuous feedback loop between human specialists and evolving ML models is recognized as essential for generating reliable and accurate outputs.⁴ Therefore, an academic exploration of this topic can argue that the resource demands and iterative nature of ML development, far from being deterrents, are integral to the reciprocal learning process. They compel the human learner into a sustained, reflective, and problem-solving engagement that cultivates both technical mastery in ML and profound domain mastery in Kanji in a uniquely integrated manner.

2. Theoretical Foundations: Machine Learning, Language Acquisition, and Cognitive Science

The proposed reciprocal learning framework draws heavily from established theories and advancements across machine learning, computational linguistics, and cognitive science. Understanding these foundational elements is crucial for appreciating the project's theoretical robustness and its potential impact.

2.1 Machine Learning in Computational Models of Language Acquisition

Machine learning research has made substantial strides in elucidating learning mechanisms, with many findings directly applicable to models of human behavior and language acquisition.⁶ Artificial Intelligence (AI) techniques, particularly Natural Language Processing (NLP), Machine Learning (ML), and Deep Learning (DL), have revolutionized linguistic research by enabling automated analysis of linguistic data and facilitating the development of predictive models for language acquisition and usage.⁷

Computational modeling and simulation techniques offer a powerful means to simulate the cognitive processes involved in language learning, allowing researchers to test theoretical frameworks against empirical data.⁷ ML algorithms have been effectively employed to predict language development trajectories and to analyze

learners' linguistic productions and errors, thereby providing valuable insights into the underlying mechanisms of language acquisition.⁷ The framework of ML tasks, which includes aggregation, clustering, characterization, and storage⁶, directly parallels the cognitive processes involved in human language acquisition. For instance, grouping utterances into component sound-sequences (aggregation), dividing objects into instances and non-instances of a concept (clustering), formulating general rules or grammars (characterization), and storing acquired knowledge (storage) are all fundamental to both human and machine learning of language. By designing an ML system that "learns" Kanji through these computational tasks, the human developer is implicitly modeling and engaging with the very cognitive mechanisms of their own learning. This offers a unique empirical avenue for investigating how humans acquire complex character systems. Therefore, an academic inquiry can explore how the architectural choices and training methodologies of the ML model reflect specific cognitive theories of Kanji learning, such as error-driven learning, spaced repetition, or contextual learning, thereby offering a novel empirical lens into human cognition and language acquisition.

2.2 Cognitive Processes in Character-Based Language Learning

The user's observation that the "most effective way to learn Kanji is by writing it, and then see it" [User Query] underscores the significance of sensorimotor engagement in character-based language learning. This traditional emphasis aligns with the importance of motor memory, visual recognition, and repetitive practice in mastering complex scripts. The inherent complexity of Kanji, with its thousands of characters and their specific, often nuanced, stroke orders, necessitates robust and multi-modal learning mechanisms for effective acquisition.¹

The user accurately identifies a limitation of many current digital methods: they are "mostly only limited to seeing, and flashcard like game" [User Query], thereby lacking the crucial "writing" component. However, research demonstrates that advanced computerized calligraphy systems can provide detailed, real-time feedback on critical aspects of writing, including stroke symmetries, pen pressure, angles, and lengths.⁸ This capability allows for the effective digitization and enhancement of the "writing" aspect, offering a richer and more precise feedback loop than traditional pen-and-paper methods or even human supervision. This directly addresses the user's desire to bridge the divide between traditional and digital learning approaches. The ML system can be designed to capture fine-grained motor data, such as stroke

velocity, pressure, sequence, and timing, from digital writing input. This data then serves as the raw material for the ML model, while simultaneously providing highly detailed, objective feedback to the human learner, thereby enriching the dataset for the machine and fostering a synergistic learning environment.

2.3 The "Learning by Building" and "Machine Teaching" Paradigms

The concept of "Machine Teaching" represents a significant paradigm shift from traditional machine learning. Instead of solely focusing on improving ML algorithms (the "learners"), it emphasizes enhancing the efficacy and productivity of the "teachers" (the human creators) who build these systems.⁹ The goal is to make the process of teaching machines easy, fast, and universally accessible.⁹ This approach places a strong emphasis on the teacher's interaction with data, the role of visualization, and the dynamic evolution of concepts during the development process.⁹

Machine teaching views the construction of ML models as a form of programming, where complex problems are systematically decomposed into smaller, manageable sub-concepts, akin to how software engineers break down functions.⁹ This framework directly addresses the growing disparity between the demand for ML practitioners and their available supply by enabling domain experts, who possess invaluable contextual knowledge, to become effective "machine teachers".⁹ Project-based learning further reinforces this by promoting deep engagement, critical thinking, and reflective practice through real-world problem-solving scenarios.³

When the user is simultaneously the ML developer and the Kanji learner, they embody the ideal "domain expert" for the system.⁴ Their personal experience with the challenges of Kanji acquisition becomes a direct and authentic feedback loop for the ML model's improvement. For instance, by correcting the system's errors or providing new, carefully curated examples, the human refines the machine's understanding. In turn, the ML model's performance and the insights it generates provide structured feedback that refines the human's own understanding of Kanji. This creates a powerful, self-reinforcing, and highly personalized learning cycle. An academic exploration can frame the user's journey as a compelling case study in "Machine Teaching," demonstrating how personal domain expertise—specifically, the challenges faced in Kanji learning—can be leveraged to build a more effective AI system. This process simultaneously deepens the developer's ML skills and their

understanding of the target domain through active, iterative engagement.

3. Current Landscape of AI-Enhanced Kanji Learning Systems

Before embarking on the development of a novel Kanji learning system, it is essential to understand the existing technological landscape. This includes reviewing current digital language learning tools, assessing advancements in Kanji handwriting recognition, and examining specialized computerized calligraphy supervision systems.

3.1 Review of Digital Language Learning Tools and Their Limitations for Kanji

Digital language learning platforms, such as Duolingo, have successfully integrated machine learning and data science to enhance language acquisition through adaptive tests and insights derived from vast datasets.¹¹ These platforms typically offer clear progress milestones and skill tracking features, which can be highly motivating for learners.¹² However, a significant limitation, as noted by the user, is that many common digital methods for Kanji learning are primarily restricted to visual exposure and flashcard-like games, often lacking the crucial "writing" component that is widely considered essential for effective Kanji acquisition [User Query].

This observation regarding the limitations of existing platforms for Kanji highlights a specific and substantial pedagogical gap. While tools like Duolingo are effective for certain aspects of general language acquisition through their data science and adaptive learning capabilities¹¹, they do not fully cater to the unique demands of character-based writing systems like Kanji. Mastering Kanji necessitates not only visual recognition but also the development of fine motor skills and precise stroke execution. This implies that current tools, despite their general effectiveness, often fall short in providing a holistic learning experience for complex character sets. The proposed project can strategically position itself to directly address this specific pedagogical and technological void. By offering a more comprehensive and traditionally-aligned digital learning experience for Kanji that integrates both visual and kinesthetic learning modalities, the system can significantly enhance its overall pedagogical value.

3.2 Advancements in Japanese Kanji Handwriting Recognition (OCR)

Significant breakthroughs have been achieved in Japanese handwritten text recognition, with researchers, notably from Waseda University, reporting an impressive 85% accuracy rate. This marks a substantial improvement over previous systems, which typically struggled to surpass 70% accuracy.¹ This progress is particularly noteworthy given the intricate nature of Japanese characters, including the vast number of Kanji symbols. Modern machine learning models are now capable of accurately handling approximately 6,000 commonly used Kanji characters.¹

Technical approaches in this domain extensively utilize advanced machine learning methods, such as Hidden Markov Models (HMMs) and Convolutional Neural Networks (CNNs), often employing deep learning techniques.¹ Crucial preprocessing techniques, including analyzing stroke width, employing projection image analysis, and incorporating symmetry, are vital for achieving higher accuracy.¹ Furthermore, automated scoring functions leveraging CNN models, such as Xception, Inception V3, and ResNet50, have demonstrated high recall (97.98-98.69%) and specificity (83.03-84.44%), with strong agreement (94.87-95.41%) with manual scoring for Japanese Kanji handwriting assessment.¹³

The high accuracy rates (85% for general recognition and over 95% for specific assessment¹) and the capability to process thousands of Kanji characters demonstrate that the core technology for recognizing handwritten Kanji is robust and sufficiently mature for practical application. This means a new project can leverage existing advancements rather than needing to prove basic feasibility. Moreover, the emphasis on analyzing structural components like symmetry and stroke order¹, and evaluating specific points within characters⁸, provides the necessary granularity for delivering detailed pedagogical feedback. This moves beyond simple correctness assessment to offering diagnostic insights into writing errors. Consequently, an academic endeavor can focus on fine-tuning and adapting existing OCR and handwriting recognition models for the specific purpose of

pedagogical feedback—for instance, identifying precise error locations and types for human learning—rather than solely for transcription. This strategic focus allows for optimized resource allocation towards the unique objectives of a reciprocal learning system.

3.3 Computerized Supervision and Feedback for Kanji Calligraphy

Beyond mere recognition, specialized systems have emerged that provide sophisticated computerized supervision for Kanji calligraphy. One such system compares a user's written calligraphic character to a model created by a human expert, precisely indicating error spots and offering explanations.⁸ This system proved significantly more efficient at reducing scripting errors, decreasing learners' error rates from 32.7% to 3.4%, in contrast to traditional practice which reduced error rates from 31.0% to 6.8%.⁸ This outcome strongly suggests that computerized supervision is more effective for calligraphy learning.

The system provides real-time reflection of pen movement, automatic evaluation results, and advice, including stroke numbers, error spots, and error types.⁸ It evaluates calligraphy across nine main categories, such as pen pressure, fades, hooks, horizontal/vertical strokes, distance between strokes, relative positions and lengths of strokes, and angles between strokes or segments, with many of these categories related to symmetry.⁸

The striking finding that computerized supervision outperformed human experts in error reduction⁸ suggests that automated systems can provide a level of precision, consistency, and instantaneousness in feedback that is challenging for human instructors to replicate. The detailed evaluation categories, such as pen pressure, angles, lengths, and symmetry⁸, imply that ML can analyze minute nuances of writing that are often overlooked or inconsistently assessed by human teachers. This capability provides highly specific, actionable insights, directly addressing the difficulty of remembering Kanji by offering targeted, objective correction and reinforcement of correct motor patterns and visual representations. Therefore, a proposed system should aim to replicate and expand upon these granular feedback mechanisms, providing highly specific, actionable insights to the learner. This detailed feedback loop is where the human learner gains a deep understanding of Kanji stroke mechanics and proper form, directly contributing to overcoming the challenge of remembering characters.

3.4 Adaptive Learning Systems for Kanji Proficiency

Adaptive learning systems for Kanji are designed to support study, for example, through mobile phone email, by dynamically adjusting content based on individual learner interests, modifying difficulty levels according to proficiency, and adapting to diverse learning styles.¹⁴ This personalized service is recognized as crucial for effective web-based learning environments.¹⁴

The broader concept of "adaptive learning" integrates rule-based methods, simple machine learning, and deep learning approaches. It optimally engages human analysts at every step, allowing them to contribute to making the machine smarter, faster, and easier to interpret.¹⁵ These systems are designed to require the least human effort over time, as they only necessitate input when it is most critical and continuously expand their knowledge base as new information is encountered.¹⁵

The principles of adaptive learning—personalization, difficulty adjustment, and learning style adaptation¹⁴—are beneficial not only for the end-user (the Kanji learner) but also for the developer. By designing and implementing these adaptive mechanisms, the developer gains profound understanding of pedagogical theory, data-driven personalization techniques, user modeling, and the complexities of human learning. The description of "adaptive learning" in the context of human analysts making machine intelligence smarter¹⁵ directly illustrates the reciprocal learning loop for the ML/data science practitioner. This demonstrates how their input refines the system, which in turn enhances their own understanding. An academic study can highlight how the iterative process of building and refining the adaptive learning components of the Kanji system directly enhances the developer's data science skills in areas such as recommendation systems, user modeling, dynamic content generation, and A/B testing for pedagogical effectiveness.

4. The Integrated Learning Framework: Building ML/Data Science Skills Through Kanji Mastery

The synthesis of machine learning development with Kanji acquisition proposes a powerful integrated learning framework. This framework aims to create a symbiotic relationship where the development process itself becomes a pedagogical tool.

4.1 Conceptualizing the ML-Powered Kanji Learning Program

The envisioned ML-powered Kanji learning program would serve as a real-world, complex project, offering a practical application for advanced ML, data analysis, and data science principles. Simultaneously, it would provide a highly personalized, engaging, and effective Kanji learning experience for the human developer. The system would extend beyond passive flashcards by incorporating interactive writing exercises, real-time diagnostic feedback, and adaptive learning paths.

The ML system can be conceptualized as a "cognitive prosthesis" and a dynamic learning partner. The user's challenges with Kanji memory and motivation [User Query] can be viewed as cognitive bottlenecks. The ML system, by providing precise, real-time, and adaptive feedback on writing and by predicting errors, functions as an external cognitive aid. It compensates for human limitations in self-correction and consistent practice. This "cognitive prosthesis" not only assists the human in learning Kanji more effectively but also provides a concrete, measurable output for the ML system's effectiveness, establishing a clear objective for the development process. The system thus evolves into a dynamic "partner" in the learning journey. An academic exploration can delve into the concept of AI as a tool for cognitive augmentation, where the machine's "intelligence" directly enhances human learning capabilities, particularly in challenging domains like Kanji, and how this partnership fosters deeper engagement and mastery.

4.2 Mechanisms for Digital Writing, Real-time Feedback, and Progress Documentation

The proposed system would necessitate robust capabilities for capturing digital handwriting input, for instance, via a tablet. This input would then be processed using advanced Optical Character Recognition (OCR) and handwriting recognition techniques.¹ The core functionality would involve providing real-time, stroke-level diagnostic feedback, similar to the advanced computerized calligraphy supervision systems discussed previously.⁸ This detailed analysis would encompass critical aspects such as stroke order, pressure, angles, lengths, and symmetries.²

Furthermore, comprehensive progress documentation would be integral to the system. This would involve meticulously tracking individual learning trajectories, identifying persistent error patterns, and visualizing the learner's mastery over time.¹⁶ To provide the detailed, diagnostic feedback on Kanji writing that is sought, the system must collect highly granular data on human stroke production, including coordinates, timestamps, and pressure, as well as the sequence of strokes. The process of designing the data capture mechanisms, storing this complex time-series data, and subsequently extracting meaningful features for ML models—such as detecting specific error types like "fades" or "hooks"⁸—represents a fundamental data science and ML engineering challenge. This process compels the developer to engage deeply with concepts of data representation, feature engineering, and real-time data processing. An academic study can detail the design of the data capture and storage pipeline, the feature engineering strategies employed for stroke analysis, and the iterative refinement of these processes, thereby demonstrating the acquisition of practical data science and ML engineering skills through tackling a complex, real-world data problem.

4.3 Leveraging Data Science for Personalized Learning Paths and Error Prediction

Data analysis would play a pivotal role in understanding individual learning styles, proficiency levels, and interests, enabling the system to personalize content delivery and dynamically adjust test difficulty.¹⁴ This level of personalization is crucial for overcoming the "lack of personalization" often cited as a challenge in AI language learning tools.¹⁸

Machine learning algorithms can be effectively employed to predict language development trajectories and to analyze learners' linguistic productions and errors.⁷ This aligns with "error-driven models of language acquisition," where the detection of prediction errors triggers updates to long-term linguistic knowledge.¹⁹ Interactive dashboards and data visualizations would be essential for tracking learning progress, identifying trends in error types, and providing actionable insights into learning patterns.¹² The human learner's challenges with Kanji, including the errors made, are not merely personal struggles; they become invaluable, authentic training data for the ML model. By analyzing these errors—a core data analysis task—the ML system can learn to predict common pitfalls. Concurrently, the human developer gains a deeper, data-backed understanding of the specific cognitive challenges associated with Kanji acquisition. This transforms a personal struggle into a rich, data-driven research

opportunity, allowing the human to "teach" the machine from their own learning experience. An academic inquiry can explore how the iterative process of collecting, implicitly labeling (through correction), and analyzing personal error data contributes to both the ML model's accuracy in error detection and the human's meta-cognitive understanding of their own learning process, ultimately leading to more effective pedagogical strategies.

4.4 The Reciprocal Benefit: How Building the System Enhances Human Kanji and ML Expertise

The process of "teaching" the machine about Kanji necessitates that the human formalize their understanding of Kanji rules, exceptions, and nuances, thereby solidifying their own knowledge. This is a direct application of "Machine Teaching" principles, where concept definitions are refined over time through continuous interaction.⁹

As the developer assumes the role of the "domain expert" ⁴, they actively participate in enriching the data—for example, through self-correction or explicit labeling—validating model outputs, and guiding the AI's learning through various techniques such as prompting and fine-tuning.¹⁰ This continuous feedback loop profoundly deepens their own domain expertise.⁴ The developer's role evolves from merely writing code to curating, reviewing, supervising, and applying critical judgment to AI-generated logic and insights.²² The concept of a "T-shaped mindset"—a broad understanding of AI tools combined with profound depth of expertise in a specific field ⁴—perfectly describes the desired outcome of this project. By building an ML system specifically for Kanji, the developer simultaneously acquires broad expertise in ML/data science (representing the horizontal bar of the 'T', encompassing model building, data management, and ethical considerations) and deep, practical expertise in Kanji learning (representing the vertical bar of the 'T', involving an understanding of its complexities and effective acquisition strategies). This interdisciplinary mastery is highly valued in the modern workforce, equipping the individual with a unique capacity to bridge technical and domain-specific challenges. An academic study can argue that this integrated approach is not merely an effective learning strategy for a specific language but also a powerful model for professional development, yielding individuals uniquely capable of leveraging AI to solve complex, real-world problems across diverse domains.

5. Technical Feasibility and Resource Considerations

The successful implementation of an ML-powered Kanji learning system hinges on its technical feasibility and the strategic management of computational resources. While machine learning projects inherently demand significant resources, existing research and methodologies offer pathways to address these requirements effectively.

5.1 Data Requirements for Kanji ML Models (Handwriting, Stroke Order, Contextual Usage)

Deep learning models generally require substantial datasets, typically ranging from 100,000 to 1,000,000 training examples, to achieve robust performance.²³ For Kanji, this encompasses various data types:

- **Handwriting Data:** Several large datasets of handwritten Kanji images are publicly available, such as Kuzushiji-Kanji, which contains 140,426 images of Kanji characters (64x64 grayscale images) with 3,999 character types, albeit with a high class imbalance.²⁴ Other valuable resources include the ETL Character Database and the KanjiVG dataset, commonly used for training handwriting recognition models.²⁶ For specific handwriting assessment purposes, researchers have utilized datasets comprising over 10,000 samples of 120 Kanji characters, including examples of correct, orthographic error, and semantic error samples.¹³
- **Stroke Order Data:** Kanji characters are composed of strokes drawn in a specific sequential order, which is critical for both aesthetic reasons and legibility.² Long Short-Term Memory (LSTM) networks are well-suited for predicting the next stroke based on previously drawn strokes, given their effectiveness with sequential data.² The Jôyô Kanji list provides structured data on stroke counts for official Kanji characters, which can be integrated into models.²⁷
- **Contextual Usage Data:** Kanji often possesses multiple meanings and readings that depend heavily on context.²⁸ Natural Language Processing (NLP) and Large Language Models (LLMs) can be leveraged to generate sentences with furigana (reading guides) and meanings, thereby facilitating contextual learning.²⁸ A wealth of curated Japanese NLP resources, including LLMs, dictionaries, and corpora,

are readily available for this purpose.²⁹

The initial concern about needing "a lot of raw data to feed" [User Query] is mitigated by the existence of several large, publicly available Kanji handwriting datasets and extensive Japanese NLP resources.²⁴ This significantly lowers the initial barrier of raw data collection. The challenge then transitions from mere acquisition to the crucial tasks of curating, preprocessing, and augmenting these datasets to align with the specific pedagogical needs of the system. This involves generating diverse handwriting styles, creating specific error examples, and handling inherent data imbalances, which are all critical data science skills.⁵ An academic study can therefore focus on novel data augmentation techniques, strategies for managing imbalanced datasets, and the development of custom datasets for specific error types (potentially derived from the developer's own learning errors), thereby demonstrating advanced data engineering and analysis capabilities.

Table 1: Essential Data Types and Available Datasets for Kanji ML Development

Data Type	Primary Purpose in System	Example Datasets/Sources	Key Characteristics/Challenges
Handwritten Kanji Images	Character Recognition, Diagnostic Feedback	Kuzushiji-Kanji, ETL Character Database, KanjiVG, User-generated data	Large scale required, Imbalanced classes, Variability in writing styles
Stroke Order Sequences	Stroke Order Correction, Motor Skill Development	Jōyō Kanji List (stroke counts), User-generated stroke data	Sequential data, Need for precise temporal capture
Kanji Contextual Usage	Contextual Meaning/Usage, Adaptive Content Delivery	Japanese NLP Corpora (e.g., Hugging Face), LLM-generated examples	Requires robust NLP models, Semantic complexity
Learner Performance Data	Adaptive Learning/Personalization, Progress Tracking, Error Prediction	User-generated data from the system itself	Privacy considerations, Requires robust logging and analysis
Specific Error	Diagnostic Feedback,	Manually created	Need for detailed

Samples	Bias Mitigation	error samples (e.g., orthographic, semantic), User-generated error data	annotation, Representativeness of error types
---------	-----------------	---	---

5.2 Computational Resources for Deep Learning in Handwriting Recognition

Deep learning for handwriting recognition, particularly when employing complex models like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, is computationally intensive.²³ Training these models can demand significant time, with examples showing 5-6 hours for a limited 50-word vocabulary even with pre-trained models.²³

Hardware requirements for efficient deep learning tasks typically include multi-core Central Processing Units (CPUs) such as Intel i7/i9 or AMD Ryzen 7/9, and CUDA-compatible NVIDIA Graphics Processing Units (GPUs), like the RTX 30-series or data center-grade A100/H100, equipped with sufficient Video RAM (VRAM)—a minimum of 8GB, with 16GB or more preferred for complex tasks.³¹ Ample system RAM (16GB minimum, 32GB or more for intensive applications) is also essential.³¹ Solid-State Drive (SSD) storage, with at least 1TB capacity, is highly recommended over Hard Disk Drives (HDD) due to significantly faster read/write speeds, which reduce data loading times.³¹

Resource scarcity can indeed pose a significant challenge.²³ To mitigate this, strategies such as distributed training across multiple GPUs or nodes can substantially reduce training time for very large models or datasets.³¹ Cloud computing platforms, including AWS, Google Cloud, and Microsoft Azure, as well as specialized platforms like Google Colab and Kaggle Kernels, offer cost-effective access to powerful GPUs without the need for extensive upfront hardware investment.³¹ The user's concern about "great resources of the physical hardware" [User Query] is validated by the high computational demands of deep learning.²³ However, the process of

optimizing resource usage—which includes selecting computationally efficient models (e.g., ResNet-18 over VGG-19²³), leveraging transfer learning from pre-trained models²³, utilizing cloud resources, implementing distributed training, and optimizing data

preprocessing and batching—becomes a crucial learning experience in ML engineering. This transforms a potential barrier into a practical skill development area, educating the developer on how to manage real-world constraints in ML projects. An academic study can detail the chosen hardware/software stack, specific optimization strategies employed (e.g., batch size tuning, early stopping, model pruning, use of lower precision data types), and the trade-offs made between model complexity, accuracy, and training time, thereby showcasing practical ML engineering expertise and resourcefulness.

5.3 Addressing Common Challenges in Machine Learning Development

Beyond hardware and data acquisition, several common challenges are inherent in machine learning development. These include issues such as poor quality of data, underfitting or overfitting of training data, the inherent complexity of the ML process itself, insufficient training data, slow implementation, and the degradation of algorithm performance as data evolves or grows (model decay).⁵ Additionally, a critical concern in AI development is the potential for AI models to perpetuate human and systemic biases present in their training data, leading to algorithmic bias.³²

Each of these challenges represents a specific, non-trivial problem that a proficient data scientist or ML engineer must learn to solve. For example, "poor data quality" necessitates rigorous data cleaning, validation, and preprocessing techniques. "Lack of training data" often requires sophisticated data augmentation strategies. "Underfitting" and "overfitting" demand careful model selection, hyperparameter tuning, and regularization methods. "Model decay" requires continuous monitoring, re-evaluation, and the establishment of robust retraining pipelines. These are not merely obstacles but structured learning opportunities that compel the developer to apply theoretical knowledge to practical problems, thereby deepening their ML/data science skills through real-world application and iterative refinement. Addressing algorithmic bias also adds an essential ethical dimension to the technical problem-solving process.³² An academic study can structure sections around how these specific ML challenges were identified and addressed within the Kanji learning project, detailing the methodologies used—such as specific data preprocessing steps, augmentation strategies, chosen model architectures, and monitoring frameworks—thereby demonstrating mastery of practical ML development workflows and robust problem-solving.

6. Pedagogical and Motivational Impact of the Integrated Approach

The proposed integrated learning framework offers a compelling solution to the long-standing challenges of Kanji acquisition, particularly those related to motivation and retention. By embedding the learning process within a hands-on ML development project, it taps into powerful pedagogical drivers.

6.1 Overcoming Kanji Learning Barriers: Engagement and Retention

The proposed system directly addresses the user's stated lack of motivation by transforming Kanji learning into an active, engaging, and personally relevant endeavor [User Query]. Incorporating gamification elements, providing real-time feedback, and offering adaptive challenges can significantly boost learner engagement.¹² The detailed and immediate feedback on writing errors⁸, coupled with adaptive difficulty adjustment¹⁴, can substantially improve retention. This is achieved by precisely targeting specific weaknesses, reinforcing correct patterns, and providing timely correction. This mechanism aligns well with error-driven learning models, where the detection of prediction errors triggers updates to long-term knowledge.¹⁹

The user's desire to build a system that helps them personally learn Kanji [User Query] creates a powerful intrinsic motivation loop. The developer is not merely a passive user but an active architect, implementer, and the primary beneficiary of the system. This deep personal investment, combined with the intellectual challenge of ML development and the tangible progress observed in both the system's performance and their own Kanji mastery, provides a far stronger motivator than traditional rote learning. The ability to visualize progress through interactive dashboards further reinforces this motivation by transforming abstract learning gains into concrete, observable achievements.¹⁶ An academic study can explore the psychological benefits of this "creator-as-learner" model, potentially quantifying its impact on motivation, persistence, and self-efficacy compared to conventional language learning methods, thereby offering a unique contribution to educational psychology and

human-computer interaction.

6.2 The Role of "Project-Based Learning" and "Learning by Teaching" in Motivation

Project-based learning (PBL) fosters an inquisitive mindset, encouraging learners to consistently ask "why," "how," and "what next." This approach promotes independent thinking, and deep reflection on the learning process.³ It provides real, meaningful experiences that learners can draw from, leading to clarity, authenticity, and purpose in their educational pursuits.³

The "learning by teaching" paradigm, or "machine teaching," requires the human to formalize their knowledge in order to instruct the machine. This act of formalization inherently deepens and solidifies their own understanding.⁹ By actively "teaching" the ML model about Kanji—through data labeling, feature engineering, defining error types, and refining feedback logic—the developer is compelled to articulate, categorize, and formalize their own understanding of Kanji rules, stroke order, radicals, and contextual nuances. This meta-cognitive process of breaking down knowledge for the machine inherently reinforces and deepens the human's own learning.

Subsequently, the machine's "performance" and the errors it makes provide direct feedback on the quality of the "teaching" (i.e., the data and model design), creating a continuous improvement cycle for both the human's Kanji knowledge and their ML system design skills. An academic study can analyze how this iterative process of formalizing domain knowledge for an AI, followed by observing and correcting the AI's "learning," leads to a more robust, transferable, and deeply internalized understanding of Kanji than passive consumption or traditional study methods.

6.3 Visualizing Progress and Gamification in Kanji Learning

Digital tools offer intuitive visualizations and advanced filters that enable seamless monitoring of learning progress.¹² Platforms like Kanshudo provide a "Mastery Map" that visually represents Kanji, vocabulary, and grammar mastery levels, tracking progress and offering personalized study recommendations.¹⁷ Similarly, Lingotrack features detailed, visualization-packed dashboards and incorporates social learning elements.¹⁶ Gamification elements, such as leaderboards and challenges, can further

enhance engagement and provide clear milestones for learners.¹²

The proposed project provides an ideal scenario for applying data visualization skills. Creating an interactive dashboard to track the developer's own Kanji learning progress—for example, characters mastered, types of errors, learning speed, and retention rates—would serve multiple purposes.²⁰ It would act as a powerful motivational tool by displaying tangible progress, a diagnostic tool for identifying persistent weaknesses in their own learning, and a practical application of data analysis and visualization skills relevant to the ML/data science aspect. This dashboard could also visualize the ML model's learning progress, such as accuracy improvements or error reduction over training epochs, providing insights into the machine's "learning curve." An academic study can include the design and implementation of such a dual-purpose dashboard as a key component, demonstrating data analysis capabilities and its direct, measurable impact on both the human learning experience and the ML model's development trajectory.

Table 2: Comparison of Traditional vs. AI-Integrated Kanji Learning Benefits

Feature/Aspect	Traditional Kanji Learning	AI-Integrated Kanji Learning (Proposed System)
Learning Motivation	Often low due to rote memorization and perceived difficulty.	High (project-driven, problem-solving, personal investment, tangible progress visualization).
Feedback Quality & Specificity	Limited, subjective, and often delayed human feedback; general correctness.	Precise, real-time, diagnostic, stroke-level feedback; identifies specific error types.
Learning Pace & Adaptivity	Fixed pace, often generic curriculum; less responsive to individual needs.	Personalized and optimized pace; adaptive (ML-driven) content and difficulty adjustment.
Skill Development Focus	Primarily Kanji knowledge (reading, writing, meaning).	Kanji mastery + ML/Data Science + Project Management + Cognitive Insight.
Progress Tracking & Insights	Manual, anecdotal tracking; limited data on learning patterns.	Automated, detailed analytics and visualization; deep insights into learning patterns.

Resource Demands	Low tech resources (paper, pen, textbooks); high human effort.	Significant computational resources (hardware, data); optimized human effort in development.
Error Identification & Correction	Delayed and general error correction; relies on human observation.	Immediate, granular, and predictive error correction; data-driven identification of common errors.

7. Ethical Considerations in AI-Enhanced Language Education

As AI becomes increasingly integrated into educational tools, particularly for language learning, it is imperative to address ethical considerations. Even for a personal development project with potential academic implications, a thoughtful approach to ethics demonstrates responsible AI practice.

7.1 Data Privacy and Security in User-Generated Content

AI-driven language learning platforms frequently collect extensive amounts of data, which can include voice recordings and other personal information.³² Ethical AI development mandates the protection of user data, ensuring transparency in how data is collected and utilized, and making privacy policies clear and accessible to users.³² While the developer is building a system primarily for personal use, framing this as an academic thesis implies its potential for broader application or as a prototype for a public tool. Therefore, considering data privacy and security from the outset, even for self-generated data, is crucial for establishing responsible AI development practices. This compels the developer to engage with concepts of data governance, potential anonymization strategies (should the data ever be shared for research), secure storage protocols, and user consent mechanisms—all of which are vital data science and ethical AI considerations.³³ This proactive approach demonstrates a mature understanding of AI's societal implications. An academic study should include a dedicated section on ethical design principles, outlining how data privacy (even for self-generated data that might be used for research) and

responsible data handling are incorporated into the system's architecture and development workflow.

7.2 Algorithmic Bias and Equitable Access to Learning Tools

Algorithmic bias can emerge if AI models are trained on imbalanced or non-representative data, potentially leading to a preference for certain accents, dialects, or writing styles.³² Such biases can perpetuate linguistic inequality and impede the learning process for affected individuals.³² Beyond individual biases, equitable access to AI-powered educational applications is a broader societal concern, particularly for students from diverse socio-economic backgrounds.³² The existence of imbalanced Kanji datasets, such as Kuzushiji-Kanji which is "highly imbalanced"²⁵, means that addressing potential biases in character recognition—for example, favoring common characters over rarer ones, or certain writing styles—is a realistic and complex ML challenge. This necessitates the application of advanced ML techniques, including data augmentation, re-sampling strategies, or specialized loss functions, to ensure fairness and robustness across the entire character set and various writing styles. This pushes the developer beyond basic model training to consider the societal impact of their algorithms. An academic study can discuss specific strategies for identifying and mitigating bias in the Kanji recognition model and the adaptive learning components, even if the primary user is the developer. This demonstrates an understanding of responsible AI development and a commitment to creating equitable and inclusive learning tools, a critical skill for future ML professionals.

7.3 Transparency and Accountability in AI-Driven Feedback

Transparency regarding how AI systems operate, including their underlying algorithms and data usage, is essential for fostering user trust.³² Large Language Models (LLMs) have the capacity to influence users based on the way information is presented.³³ Therefore, developers must ensure that AI-generated content is accurate and verifiable through other credible sources.³³ For the ML system to function as an effective

teacher for the human user, its feedback must be not only accurate but also interpretable and trustworthy. Simply stating "this stroke is wrong" is less helpful than providing a detailed explanation, such as "this stroke's angle is incorrect by X degrees compared to the model's ideal, specifically at point Y." This requirement compels the underlying ML models, such as CNNs for handwriting recognition, to not merely classify but also provide *explainable* insights into their decisions and the precise nature of the error. This pushes the developer to explore model interpretability techniques, such as saliency maps or feature importance analysis, which are crucial and evolving areas in modern ML. An academic study can detail how the system's feedback mechanisms are designed to be transparent and explainable, allowing the human learner to understand *why* an error occurred and *how* to correct it. This fosters deeper meta-learning and builds trust in the AI as a learning partner, further contributing to the "humanist" aspect by making the AI's reasoning accessible.

8. Conclusion and Future Directions for Thesis Research

8.1 Summary of Viability and Potential

The integration of machine learning and data science development with Japanese Kanji acquisition presents a highly viable and promising avenue for a reciprocal learning experience. Existing research strongly supports the technical feasibility of implementing robust Kanji handwriting recognition, detailed computerized calligraphy supervision, and adaptive learning systems. Concurrently, established pedagogical theories such as "learning by building" and "machine teaching" provide a robust theoretical framework for the profound human learning aspect. The proposed project directly addresses core challenges in Kanji learning, including difficulty, motivation, and retention, by transforming them into active problem-solving opportunities. This approach provides a rich, complex, and personally meaningful project-based environment for acquiring advanced ML and data science skills. The synergistic nature of this framework transforms traditional learning into an active, problem-solving endeavor, fostering deep engagement and comprehensive skill development.

8.2 Key Research Questions and Next Steps for a Doctoral Thesis

This exploration lays a strong foundation for a doctoral thesis, proposing a compelling research agenda centered on the reciprocal learning paradigm.

Primary Research Question: How effectively can the iterative process of designing, building, and refining an ML-powered Japanese Kanji learning system enhance both the human developer's proficiency in Kanji and their expertise in machine learning and data science, particularly through the lens of "machine teaching"?

Sub-Questions for Further Research:

- What specific machine learning and deep learning architectures, combined with tailored data preprocessing techniques, are most effective for providing real-time, granular, and diagnostic feedback on Kanji handwriting and stroke order? How does the implementation and optimization of these architectures deepen the developer's ML engineering skills and understanding of character recognition?
- How can data science methodologies, including advanced data collection, feature engineering, and predictive modeling, be applied to analyze individual Kanji learning patterns and predict common error types? What insights do these analyses provide for personalizing learning paths and how does this process enhance the developer's data analysis and modeling capabilities?
- To what extent does the "machine teaching" paradigm, wherein the human developer formalizes their Kanji knowledge to train the ML system, contribute to the human's own retention, accuracy, and overall mastery of Kanji? Can this contribution be quantitatively measured and compared against traditional learning methods?
- What are the optimal human-computer interaction (HCI) design principles for an ML-powered Kanji learning system that maximizes user engagement, provides actionable feedback, and facilitates the reciprocal learning process for the developer? How does the iterative design and testing of these interfaces refine the developer's understanding of user experience and pedagogical effectiveness?
- What are the specific computational resource requirements (hardware, data storage, processing power) for developing and deploying such a system, and what strategies for resource optimization (e.g., cloud computing, distributed training, model compression) are most effective in balancing performance with accessibility? How does navigating these resource constraints contribute to the

- developer's practical ML operations (MLOps) and resource management skills?
- How can ethical considerations, including data privacy, algorithmic bias in character recognition, and transparency in AI-driven feedback, be systematically integrated into the design and development lifecycle of the Kanji learning system? What challenges arise in mitigating these ethical concerns, and how does addressing them contribute to the developer's understanding of responsible AI?

Next Steps for Thesis Research:

1. **System Prototyping and Iterative Development:** Begin with the development of a functional prototype, focusing on core functionalities such as digital handwriting input, real-time stroke analysis, and basic character recognition. Employ an agile development methodology, allowing for continuous iteration and refinement based on self-assessment and simulated user feedback.
2. **Data Collection and Augmentation Strategy:** Systematically collect personal Kanji handwriting data, including correct forms and intentional error variations. Explore and implement advanced data augmentation techniques to expand existing public Kanji datasets, addressing class imbalances and diverse writing styles.
3. **ML Model Experimentation and Optimization:** Experiment with various CNN and LSTM architectures for handwriting recognition and stroke order prediction. Focus on optimizing models for both accuracy and interpretability, ensuring that feedback provided to the human learner is precise and actionable.
4. **Personalized Learning Algorithm Design:** Develop and test adaptive learning algorithms that adjust content difficulty, review schedules, and feedback mechanisms based on the developer's real-time performance data and identified learning patterns.
5. **Dashboard and Visualization Development:** Construct an interactive dashboard to visualize both the ML model's performance metrics and the developer's Kanji learning progress. This will serve as a critical tool for monitoring, diagnosing learning challenges, and motivating continued engagement.
6. **Ethical Framework Integration:** Document and implement a clear ethical framework for data handling, model training, and feedback generation, ensuring privacy, fairness, and transparency are prioritized throughout the project.
7. **Longitudinal Self-Study and Analysis:** Conduct a longitudinal self-study, meticulously documenting the developer's own Kanji learning journey and skill acquisition in ML/data science throughout the system's development. This will provide rich qualitative and quantitative data for the thesis.
8. **Comparative Analysis (Optional but Recommended):** If feasible, conduct a small-scale comparative study with a control group using traditional Kanji

learning methods to further validate the efficacy of the proposed reciprocal learning paradigm.

This integrated approach offers a unique and compelling thesis topic, blending cutting-edge machine learning with a deep understanding of language acquisition and pedagogical theory, all while addressing a personal learning challenge.

Karya yang dikutip

1. Understanding Japanese Kanji OCR How Modern AI Translates Complex Characters in 2024 | aitranslations.io, diakses Juli 8, 2025, https://aitranslations.io/blog/understanding_japanese_kanji_ocr_how_modern_ai_translates_co.php
2. JAPANESE CHARACTERS (KANJI) STROKE PREDICTION USING LSTM - IRJMETS, diakses Juli 8, 2025, https://www.irjmets.com/uploadedfiles/paper//issue_9_september_2023/44512/fin_al/fin_irjmets1694017232.pdf
3. Why students from project-based classrooms stand out in study abroad applications, diakses Juli 8, 2025, <https://timesofindia.indiatimes.com/education/study-abroad/why-students-from-project-based-classrooms-stand-out-in-study-abroad-applications/articleshow/122264170.cms>
4. Why deep domain knowledge can make your career AI-proof - The Times of India, diakses Juli 8, 2025, <https://timesofindia.indiatimes.com/technology/times-techies/why-deep-domain-knowledge-can-make-your-career-ai-proof/articleshow/121765874.cms>
5. 7 Major Challenges Faced By Machine Learning Professionals - GeeksforGeeks, diakses Juli 8, 2025, <https://www.geeksforgeeks.org/blogs/7-major-challenges-faced-by-machine-learning-professionals/>
6. Language Acquisition and Machine Learning. - DTIC, diakses Juli 8, 2025, <https://apps.dtic.mil/sti/tr/pdf/ADA169581.pdf>
7. (PDF) Artificial Intelligence in Linguistics Research: Applications in Language Acquisition and Analysis - ResearchGate, diakses Juli 8, 2025, https://www.researchgate.net/publication/379239839_Artificial_Intelligence_in_Linguistics_Research_Applications_in_Language_Acquisition_and_Analysis
8. (PDF) Learning System for Japanese Kanji Calligraphy with ..., diakses Juli 8, 2025, https://www.researchgate.net/publication/335336813_Learning_System_for_Japanese_Kanji_Calligraphy_with_Computerized_Supervision
9. Machine Teaching A New Paradigm for Building Machine ... - arXiv, diakses Juli 8, 2025, <https://arxiv.org/pdf/1707.06742>
10. Injecting domain expertise into your AI system | by Dr. Janna Lipenkova - Medium, diakses Juli 8, 2025, <https://medium.com/data-science/injecting-domain-expertise-into-your-ai-system-792febff48f0>

11. Research - Duolingo, diakses Juli 8, 2025, <https://research.duolingo.com/>
12. Comparing digital tools to track student progress - KeepTheScore, diakses Juli 8, 2025, <https://keepthescore.com/blog/posts/5-ways-to-track-student-progress-with-digital-scoreboards/>
13. Assessing handwriting skills in a web browser: Development and ..., diakses Juli 8, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11685258/>
14. Development of Adaptive Kanji Learning System for Mobile Phone ..., diakses Juli 8, 2025, https://www.researchgate.net/publication/344921279_Development_of_Adaptive_Kanji_Learning_System_for_Mobile_Phone
15. The fourth generation of machine learning: Adaptive learning | jungle light speed, diakses Juli 8, 2025, <https://www.junglelightspeed.com/the-fourth-generation-of-machine-learning-a-daptive-learning/>
16. Lingotrack: Landing, diakses Juli 8, 2025, <https://lingotrack.com/>
17. Japanese Mastery - Kanshudo, diakses Juli 8, 2025, <https://www.kanshudo.com/mastery>
18. elearningindustry.com, diakses Juli 8, 2025, <https://elearningindustry.com/ai-for-language-learning-benefits-and-challenges#:~:text=Challenge%20%3A%20Lack%20Of%20Personalization&text=However%2C%20those%20methods%20often%20lack,the%20progress%20of%20many%20students.>
19. The Role of Prediction Error in 4-Year-Olds' Learning of English Direct Object Datives, diakses Juli 8, 2025, <https://www.mdpi.com/2226-471X/8/4/276>
20. NLP Dashboard | Fast Data Science, diakses Juli 8, 2025, <https://fastdatascience.com/natural-language-processing/nlp-dashboard/>
21. LLM-Driven Learning Analytics Dashboard for Teachers in EFL Writing Education - arXiv, diakses Juli 8, 2025, <https://arxiv.org/html/2410.15025v1>
22. How AI Has Transformed the Role of Software Developers | Built In, diakses Juli 8, 2025, <https://builtin.com/articles/ai-transformed-role-software-developers>
23. Handwritten Text Recognition using Deep Learning - Stanford ..., diakses Juli 8, 2025, <https://cs231n.stanford.edu/reports/2017/pdfs/810.pdf>
24. Kuzushiji Kanji Classification - Kaggle, diakses Juli 8, 2025, <https://www.kaggle.com/c/kuzushiji-ml-it-kmitl-2020>
25. Kuzushiji-Kanji (KKanji) dataset, diakses Juli 8, 2025, <https://datasets.activeloop.ai/docs/ml/datasets/kuzushiji-kanji-kkanji-dataset/>
26. CaptainDario/DaKanji-Single-Kanji-Recognition: A machine learning model to recognize Japanese characters (Kanji, Katakana, Hiragana) - GitHub, diakses Juli 8, 2025, <https://github.com/CaptainDario/DaKanji-Single-Kanji-Recognition>
27. Japanese Jōyō Kanji - Kaggle, diakses Juli 8, 2025, <https://www.kaggle.com/datasets/anthaus/japanese-jy-kanji>
28. How I used NLP and LLM to supercharge my Japanese learning - DEV Community, diakses Juli 8, 2025, <https://dev.to/errietta/how-i-used-nlp-and-llm-to-supercharge-my-japanese-lear>

[ning-5bd1](#)

29. taishi-i/awesome-japanese-nlp-resources: A curated list of resources dedicated to Python libraries, LLMs, dictionaries, and corpora of NLP for Japanese - GitHub, diakses Juli 8, 2025, <https://github.com/taishi-i/awesome-japanese-nlp-resources>
30. (PDF) Handwritten recognition system using deep learning - ResearchGate, diakses Juli 8, 2025, https://www.researchgate.net/publication/385957996_Handwritten_recognition_system_using_deep_learning
31. System Requirements for Deep Learning - GeeksforGeeks, diakses Juli 8, 2025, <https://www.geeksforgeeks.org/deep-learning/system-requirements-for-deep-learning/>
32. AI Ethics And Language Learning, diakses Juli 8, 2025, https://www.meegle.com/en_us/topics/ai-ethics/ai-ethics-and-language-learning
33. Ethical AI for Teaching and Learning - Center for Teaching Innovation - Cornell University, diakses Juli 8, 2025, <https://teaching.cornell.edu/generative-artificial-intelligence/ethical-ai-teaching-and-learning>