

Gökhan ERCAN

Software Engineer / Researcher

[in](#) linkedin.com/in/gokhanercan [g](#) github.com/gokhanercan [t](#) twitter.com/gokhanercan
[+90 533 5211430](#) [@](#) mail@gokhanercan.com [gokhanercan.com](#)
[Levent, Istanbul / TURKEY](#) [i](#) Birth:1983



SKILLS AND INTERESTS

Professional Expertise Object Oriented Design & Architecture, Web Development, Unit Testing, TDD, DDD, Agile, Product Development, Software Process Development, Project Management

Primary Languages C# (21 years), Java (12 years), Python (7 years), Javascript (21 years), Node (2 years)

Research Interests Information Retrieval, Natural Language Processing, Machine Learning, Semantics

PhD thesis (final stage) Semantic Relation Extraction by Enriching Word Embeddings Exploiting Turkish Morphology

Personal Interests Guitar, Keyboard, Rock, Programming, Philosophy, Science, Evolution, AI, Linguistics, Sports

CAREER SUMMARY

Sep 2022 - ... **Software Engineering Leader** - [BLUE.CLOUD](#) , U.S.A. (Istanbul Remote)
> Leading software developers and cloud-based development projects.

(Sep 2022 - Mar 2023) **Principal Software Engineer**
> Consulting, analyzing, designing, and developing cloud migration projects.

2002 - 2022 (2017-2022) **Director of Software Development** - [HOSTAMEDIA](#) , İstanbul
> Designing and developing corporate development frameworks and products.
> Hiring, training, and managing back-end/front-end software development teams.
> Developing software development lifecycles and business process workflows.

4 Products	150+ Web Applications	70 Developers Hired
Company received 27 awards	200+ Plugins	On 1 Company Framework

Previous roles:

(2012 - 2017) Principal Architect A (2006 - 2012) Lead Developer A (2002 - 2006) Software Developer

2010 - 2022 **External Lecturer/Consultant** - [BT AKADEMI](#) , İstanbul
Recent Courses:
> Test Driven Development & Unit Testing (8 years)
> Enterprise Design Patterns & Architectures (11 years)
> Object Oriented Programming Principles & Design Patterns (12 years)

75+ lectures	1200+ hours of teaching	1000+ developers trained
---------------------	--------------------------------	---------------------------------

2018 - 2019 **External Lecturer** - [IŞIK UNIVERSITY](#) , COMPUTER SCIENCE DEPARTMENT , İstanbul
Lecture:
> CSE 112 - Object Oriented Programming

EDUCATION

PhD. 2014 - 2022* [Işık University](#) , Istanbul/Turkey - **Computer Engineering** (Full scholarship)
[\[Best paper awarded\]](#) *Final paper submitted. Expecting graduation in early 2023.

MSc. 2010 - 2014 [Bogazici University](#) , Istanbul/Turkey - **Software Engineering**

B.A. 2006 - 2010 [Anadolu University](#) , Eskisehir/Turkey - Business Administration

2000 - 2006 [Yıldız Technical University](#) , Istanbul/Turkey - Urban and Regional Planning (foundation degree)

1997 - 2000 [Akdeniz College](#) , Antalya/Turkey - Science High School

2022 - ...

BLUE.CLOUD, U.S.A. - SOFTWARE ENGINEERING LEADER, - ISTANBUL (REMOTE)

- › Leading software developers and cloud-based development projects.

(Sep 2022 - Mar 2023) Principal Software Engineer

- › Consulting, analyzing, designing, and developing cloud migration projects.

2002 - 2022

HOSTAMEDIA, İSTANBUL - [4 ROLES]

ROLES

(2017 - 2022) Director of Software Development: Responsible for planning, designing, developing, and managing the company's core framework, products, and development processes. Also responsible for recruitment, training, and appraisals of software developers. The role was %70 hands-on.

(2012 - 2017) Principal Architect: Responsible for designing, developing, and maintaining the company's core framework and products. Developed E-Commerce and Information Retrieval products of the company. The role was %100 hands-on.

(2006 - 2012) Development Team Lead: Developed the company's core Framework, CMS, and Hotel Reservation products. The role was %100 hands-on.

(2002 - 2006) Software Developer: Developed software solutions for customers and wrote the foundation of the company's first product, CMS. The role was %100 hands-on.

SIGNIFICANT CONTRIBUTIONS

Development Framework (2009 - 2022): It is a core software development framework that provides high extensibility and maintainability to the products and solutions, along with the opinionated business and cross-cutting features.

- › Cross-cutting: auth, session, security, logging, exceptions, data access, scheduler, etc.
- › Dependency injection, aspect-oriented support, dynamic proxies.
- › The plugin system facilitates the advent of 200+ plugins over the years.
- › Meta-programming: DSL, declarative behaviors, scripting.
- › Multi-layer caching system (aspect-oriented, auto invalidation, trickle- down effect)
- › Multi-lingual database support: All providers and components to support ML.
- › Web composite UI components and theming
- › Declarative scaffolding UI. Meta screens.

Product 1: Information Retrieval System: The system facilitated entity matching and price grabbing solutions for Turkey's leading e-commerce companies such as Vatan Bilgisayar and Hepsi Burada. NLP enhanced infrastructure: Turkish morphology enriched subword-level semantic modeling, entity matching and named entity recognition tasks. Distributed durable job processing. Adaptive web wrappers. Focused crawler. Rule-based intelligence.

Product 2: Content Management System (CMS): Web CMS system allows developers to focus on domains by reducing boilerplate artifacts regarding dynamic website development and rich content. Meta Search, Markdown live editing, Dynamic Image Sizing/Rendering, Meta Url Management, Dynamic entities. Structured UI theming system with Material Design.

Product 3: E-commerce Product: Commerce domain modeling. Order state machine. Payment and POS provider abstractions. Multi-tenant database and shopping system.

Product 4: Hotel Reservation System: Hotel domain modeling. Multi-property and multi-tenancy architecture. Booking provider integrations.

Tech. Stack: .Net, .Net Core, C#, Asp.Net, WebForms, Blazor, JS, TS, MVC, WCF, Redis, LLBLGen, Enterprise Library, Unity DI, Quartz, NUnit, Specflow (BDD), Cruise Control, TeamCity, SQL Server, Lucene, Node.

2010 - 2022

BT AKADEMİ, İSTANBUL - EXTERNAL LECTURER / CONSULTANT BT Akademi is one of the best software consultancy/education companies in Turkey that has been working with the most prominent enterprises from various sectors such as Banking, Telecommunication, and Technology. I was responsible for teaching and consulting architectural software design and quality concepts to software architecture teams of such organizations.

75+ lectures

1200+ hours of teaching

1000+ developers trained

Lectures:

› [Test Driven Development & Unit Testing \(2014 - 2022\)](#)

Outline: Delivering high-quality production code requires well-organized unit tests with enough coverage. However, paradoxically, proper unit testing requires well-designed software that flawlessly handles dependencies of inner components. TDD methodology addresses the problem by offering development teams to follow a straightforward process called red-green-refactor. When developers clearly define requirements before implementing them, they can isolate test cases easily and introduce lesser bugs. This course focuses on patterns, practices, and processes for developing high-quality software with unit tests while making in-depth analyses of mocking/stubbing, dependency injection, and testing tools.

› [Enterprise Design Patterns & Architectures \(2011 - 2022\)](#)

Topics: Framework Development, Fowler's Patterns of Enterprise Architecture Book, DDD, BDD, Anemic Model, Domain Model, Entity/Data Layer Tradeoffs, DTO, POCO, Mocking/Stubbing, IOC, Dependency Injection, Coupling, Crosscutting Concerns, Repository, Versioning, MVVM, ...

› [Object Oriented Programming Principles & Design Patterns \(2010 - 2022\)](#)

Outline: This course covers advanced object-oriented programming techniques and design patterns essential in developing reusable and extensible libraries. In a modern development world, every team should agree on the fundamentals of the object-oriented theory regardless of the platforms they use. The course iterates through every GOF design patterns after making an in-depth analysis of popular SOLID principles.

› [Advanced ASP.NET Development \(2010 - 2014\)](#)

› [Developing .NET Applications with C# \(2010 - 2012\)](#)

Recently lectured/gave consultancy:

Banks: Garanti BBVA Technology, Akbank, SoftTech, Ziraat Technology, Kuveyt Turk, Albaraka Turk, ...

Technology / Telecommunication: Ericsson, Siemens, Turk Telekom, Innova, Çiçek Sepeti, Intertech, ...

Other: Honda, NCR, TÜBİTAK Bilgem, Belbim, Enerjisa, Intertech, Amerikan Hospital, Anadolu University

Microsoft Certified Trainer (MCT, MCPD, MCTS, MCP) (2007-2011). Transcript

2018 - 2019

ISIK UNIVERSITY, İSTANBUL - EXTERNAL LECTURER

Lecture:

- › [CSE 112 - Object Oriented Programming:](#) The objective of this course is to introduce the concepts of object oriented thinking, design, and programming on Java language platform.

2001 - 2002

TEKTONIK GROUP EURASIA, ISTANBUL - SOFTWARE DEVELOPER

- › Development of XML integration services and user interfaces for the Eskidji and KolayTedarik.com projects.

2014 - 2022

PHD RELATED NLP WORK - SEMANTIC AND MORPHOLOGICAL MODELING

Thesis: "Semantic Relation Extraction by Enriching Word Embeddings Exploiting Turkish Morphology"
[[Best paper awarded](#)] My research focuses on leveraging morphological resources to improve unsupervised semantic modeling techniques. I have been developing my own NLP framework to jointly model morphology and semantics to build complex pipelines for [my papers](#). The framework is built on Java and Python stack communicating through micro-services and shared file systems along with wrappers of commonly used libraries and frameworks such as Gensim, Word2vec, FastText, etc. I have never worked as a full-time academician, and I always see academic research as a complimentary part of my software development journey.

Fields: Semantic modeling, Language modeling, Morphological processing, Data wrangling and pre-processing, Knowledge-bases, Dataset modeling, Automatic dataset generation.

Tasks: Word similarity/relatedness, Word analogy, Morphological analysis, Morphological disambiguation, Word segmentation, Root detection, Textual/String similarity, Correlations, Classifiers.

STACK

Java: Java, Guava, Commons, JUnit, Spring, Play, Vaadin, Jetty, Tomcat, Jersey, Maven, Kotlin, IntelliJ.

Python: Cython, pybind, pandas, numpy, sklearn, scipy, sympy, networkx, matplotlib, PyCharm, Docker.

NLP / ML: word2vec, fasttext, dl4j, gensim, spacy, polyglot, lucene, nltk, morfeessor, tsne, C, C++.

Public resources: WordNet, Wikipedia, Polyglot, CommonCrawl, Morpholex.

Some open-source contributions that I made:

- [StarlangSoftware/TurkishMorphologicalAnalysis](#)
- [MudBlazor/MudBlazor](#)
- [luozhouyang/python-string-similarity](#)
- [gokhanercan/WSQuest](#)
- [gokhanercan/nlptoolkit.ui](#)

2010 - 2014

Bogazici University, Istanbul/Turkey - Software Engineering

Masters project: "Adaptive Web Wrappers with Genetic Algorithms"

By leveraging generic algorithms, the system is able to generate redundant web wrapper scripts (XPath) that can parse the given target value of HTML pages. The system uses various textual similarity metrics as a fitness function to determine the best parsers in the population. With numerous readily-available redundant parsers with different structures, the system can update itself (adding and killing some genes) by relying on majority voting mechanisms, which provides high robustness and adaptability against possible structural changes in web pages.



My paper has been awarded as the "Best resource paper" at the COLING conference, which is one of the top NLP conferences.

Gökhan Ercan and Olcay Taner Yıldız. "AnlamVer: Semantic Model Evaluation Dataset for Turkish - Word Similarity and Relatedness". In: *Proceedings of the 27th International Conference on Computational Linguistics*. 2018, pp. 3819–3836

- REVIEW... 2022*** | **G. Ercan, Yildiz, O. T.** - THE ROLE OF MORPHOLOGY ON DISTINGUISHING SEMANTICALLY UNRELATED WORDS, *Submitted to journal. Under review. Full-text available upon request.
N-grams based segmentation introduces *the noise* to semantic spaces which makes models incapable of distinguishing two unrelated concepts from each other (<2 accuracy), especially when the words are textually similar. We presented TSimUnr dataset that consists of nearly one million English and Turkish word-pairs that are automatically generated by a complex pipeline relying on graph-based WordNet relatedness approximations. We showed morphological segmentation overcomes the problem and outperforms the n-gram segmentation with a very high margin (en:65%, tr:68% accuracy) without sacrificing performance on conventional datasets of word-similarity task such as RareWords, MTurk771 and MEN.
- RANLP 2019** | **Yildiz, O. T., B. Avar, G. Ercan** - AN OPEN, EXTENDIBLE, AND FAST TURKISH MORPHOLOGICAL ANALYZER , [Recent Advances in Natural Language Processing 2019](#)
We presented a two-level morphological analyzer for Turkish which consists of five main components: finite state transducer, rule engine for suffixation, lexicon, trie data structure, and LRU cache. We use Java language to implement finite state machine logic and rule engine, Xml language to describe the finite state transducer rules of the Turkish language, which makes the morphological analyzer both easily extendible and easily applicable to other languages.
- COLING 2018 (**)** | **G. Ercan, Yildiz, O. T.** - SEMANTIC MODEL EVALUATION DATASET FOR TURKISH - WORD SIMILARITY AND RELATEDNESS , [The 27th International Conference on Computational Linguistics](#) (**Best resource paper awarded)
Presented AnlamVer dataset, which is a semantic model evaluation dataset for Turkish designed to evaluate word similarity and word relatedness tasks while discriminating those two relations from each other. The dataset word-pairs are balanced by their frequencies to evaluate the robustness of semantic models concerning out-of-vocabulary and rare words problems, which are caused by the rich derivational and inflectional morphology of the Turkish language.

[See all publications on Google Scholar](#)

 ACADEMIC REVIEWER

- SIU 2020 | 28th Signal Processing and Communications Applications Conference
AI Review 2019 | Artificial Intelligence Review - An International Science and Engineering Journal
SIU 2019 | 27th Signal Processing and Communications Applications Conference

 LANGUAGES

- > Turkish (native)
- > English (Professional working proficiency)

 ORGANIZATIONS AND GROUPS

- > Philosophical Dialogues - Meetup Group [2020 - present]
- > T90 Turkey Blockchain Consortium (Board Member) [2018]