José Antonio Zamudio Amaya

jose.zamudio@cispa.de | www.jzamudio.com | joszamama | joszamama | zamudioamaya

Summary

3rd-year Ph.D. candidate at CISPA Helmholtz Center for Information Security, working with Prof. Andreas Zeller on automated test generation. I approach it as a search problem: evolutionary search over derivation trees, multi-objective exploration of the input space, and metaheuristic optimization, applied to language-based fuzzing, stateful protocol testing, variability-driven testing, and LLM-assisted generation.

Impact Author of Fandango, an open-source fuzzer with 135,000+ installations that has found more than 50 vulnerabilities in open-source software, including heavily fuzzed OSS-Fuzz targets, and is deployed in industry collaborations with Bosch and Volkswagen. I personally identified more than 40 of these and contributed security patches upstream.

Publications Ten peer-reviewed publications spanning fuzzing, software testing, and optimization (MATCOM, ISSTA, ASE, FSE, ICST, SSBSE, SPLC). Six manuscripts under review (ICSE, TOSEM).

Leadership Project lead, setting technical direction for collaborative work with Andreas Zeller and 7 PhD students; supervised five theses; co-authored a 353-page fuzzing tutorial; collaborated with more than 20 researchers all across the globe.

Education

Ph.D. in Computer Science - Systems Security

Saarbrücken, Germany

Saarland University · CISPA Helmholtz Center for Information Security — *Oct. 2023 - Present (expected Jan. 2028)*

- Doctoral researcher under Prof. Andreas Zeller (ACM/IEEE Fellow & Harlan D. Mills Awardee), at the CISPA Helmholtz Center and the Saarland Informatics Campus.
- Automated test generation, approached from several directions: search-based generation and multi-objective exploration of the input space, language-based fuzzing, stateful protocol testing, security testing (TLS/X.509, injection attacks), LLM-assisted test generation and constraint extraction, variability-driven testing, and high-performance input generation.

M.Sc. Software Engineering: Big Data, ML, Data Science & AI

Seville, Spain

University of Seville — *Sep. 2022 - Jul. 2023*

- Machine learning and deep learning (model training, evaluation, and the statistics of comparing them honestly), plus the engineering around them: large-scale data pipelines, distributed processing, and the infrastructure that turns a model into a system.
- Selected for a high-intensity track with 20 available seats; awarded the sole Thesis Honors Distinction, ranking as the best dissertation among all graduates that year.

B.Sc. Software Engineering · Full English Language Program

Seville, Spain

University of Seville — *Sep. 2018 - Jul. 2022*

- A heavy core, physics, calculus, linear algebra, probability & statistics, discrete mathematics, algorithms & complexity, and formal languages & automata, every subject paired with laboratory and practical sessions.
- Selected for the English Language Program: 20 places for 700+ engineering students, with 100% of technical coursework and examinations completed in English. Graduated in exactly 4 years (the minimum), a distinction achieved by 8% of the student body against an average completion time of 5.9 years; awarded Thesis Honors Distinction (top 5% of the department).

Research Experience

CISPA Helmholtz Center for Information Security

Saarbrücken, Germany

Doctoral Researcher — *Oct. 2023 - Present (expected Jan. 2028)*

- **Search-Based Test Generation.** Designed and led development of Fandango, replacing symbolic constraint solving with evolutionary search over derivation trees; achieved up to 1,355× speedup over state-of-the-art techniques at near-perfect validity (ISSTA 2024 and 2025). Produced structurally diverse valid inputs where a single fitness collapses onto whichever region of the validity space it reaches first, covering more parser code at an equal budget and surfacing 42 new defects in 40 widely used libraries.

Max Planck Institute for Security and Privacy

Bochum, Germany

Research Internship — *Sep. 2025 - Dec. 2025*

- **Statistical Fuzzing Theory.** Studied models of coverage-guided fuzzing under Dr. Marcel Böhme, specifically how input feature distributions relate to coverage saturation and the theoretical limits of grammar-based generation.
- **Feature-Focused Test Generation.** Applied those foundations to build Compás, connecting feature modeling to specification-based testing, enabling testers to target specific input features and reduce test suites by an order of magnitude while preserving aggregate coverage (ASE 2026).

Pablo de Olavide University

Seville, Spain

Student Researcher — *Feb. 2023 - Oct. 2023*

- **Metaheuristic Optimization.** Designed Iterated Greedy metaheuristics for NP-hard weighted graph problems under Prof. Alfredo García Hernández-Díaz; published in *Mathematics and Computers in Simulation* (MATCOM 2024), solving instances up to 4,000 vertices 50% faster than prior state-of-the-art.

University of Seville *Seville, Spain*

Student Researcher *Jan. 2022 - Oct. 2023*

- **Software Product Lines.** Researched variability modeling and automated artifact generation under Prof. David Benavides; built WebSPL (SISTEDES 2022) and XatKitSPL (SISTEDES 2023) open-source frameworks for generating websites and chatbots directly from feature models.
- **Feature Model Analysis.** Developed real-time analysis operations for constraint-based recommender systems (SPLC 2023).

Projects

Fandango — `https://github.com/fandango-fuzzer/fandango` *Saarbrücken, Germany*

Author & Lead Developer (12 active collaborators) *Since 2025*

- Designed and implemented an open-source language-based fuzzer that combines formal grammars with Python constraints and evolutionary search to generate semantically valid test inputs for black-box, white-box, and stateful protocol testing.
- Implemented in Python (59k+ LOC); integrates with AFL++, libFuzzer, and OSS-Fuzz; supports grammar-defined network protocols, binary formats, and structured file inputs.
- More than 135,000 installations, 120 GitHub stars; active industry collaborations with Bosch and Volkswagen for embedded systems testing, all in less than a year.
- Found more than 50 vulnerabilities in open source software (including heavily fuzzed OSS-Fuzz targets) with a total monthly download count of 2.7 billion.

Compás — `https://github.com/fandango-fuzzer/compas` *Saarbrücken, Germany*

Author & Lead Developer *Since 2025*

- Designed and implemented the first tool to apply software product line engineering to specification-based testing: it automatically extracts a feature model from a Fandango input specification and prunes it into *slices* that pin individual format features while preserving input validity by construction.
- Implemented in Python (3.5k LOC, 134 tests, CI on 3.11–3.13) as both a CLI and a library; exports extracted models to UVL for interoperability with the `flamapy` feature-model analysis ecosystem, and delegates input generation to Fandango.
- Evaluated on ten format specifications across seven parser libraries (Pillow, Mutagen, `zipfile`, `pypdf`, `pyelftools`, `tomllib`, `html.parser`): 100% mean hit rate on targeted code regions versus 26% for the unsliced baseline, feature-to-code traceability for 80 of 148 feature values, and test suites reduced to 1–2 labelled slices at equal coverage while reaching target regions in up to 18.5× fewer inputs.

Publications

Accepted

Feature-Focused Test Generation *ASE 2026 · Research Track*

J.A. Zamudio Amaya, G. Sapia, A. Turcotte, D. Benavides, M. Böhme, A. Zeller.

- Let testers aim at the input features they care about instead of the whole input space, exercising the targeted feature in 91.3% of generated inputs against 32.3% for an unfocused baseline (at 100% validity and unchanged aggregate coverage) and then shrinking suites by up to an order of magnitude with no coverage loss (PNG: 29 slices to 3; GIF: 18 to 1), by extracting a feature model from the input specification and deriving one restricted generator per feature value.

Search-Based Generation of Complex Inputs with Fandango *SSBSE 2026 · Hot off the Press*

J.A. Zamudio Amaya, M. Smytzek, A. Zeller.

- Brought language-based testing to the search-based software engineering community, arguing that evolutionary search is a practical replacement for symbolic constraint solving, by recasting valid-input generation as a search over derivation trees rather than a satisfiability problem.

How to Test Complex Systems Automatically and Systematically *FSE 2026 · Tutorial*

J.A. Zamudio Amaya, A. Liggesmeyer, M. Smytzek, A. Zeller.

- Taught practitioners and researchers to write their own generators, in a 180-minute hands-on tutorial already delivered to more than 100 participants at the RIO Summer School, by walking them from grammars through Python semantic constraints and execution feedback to fuzzing stateful network protocols.

Constraint-Driven Fuzzing at Scale with Fandango *ICST 2026 · Tool Demo*

J.A. Zamudio Amaya, M. Smytzek, A. Zeller.

- Delivered Fandango as a pip-installable tool now taught in university testing courses and adopted in industry (Volkswagen's AUTOSPEC synthesizes protocol specifications from RFCs and fuzzes them with Fandango), by unifying a grammar and arbitrary Python constraints in a single specification file driven from the command line.

FANDANGO: Evolving Language-Based Testing *ISSTA 2025 · Research Track*

J.A. Zamudio Amaya, M. Smytzek, A. Zeller.

- Replaced symbolic constraint solving with evolutionary search in language-based testing, generating valid inputs 29× to 1,355× faster than ISLa (the SMT-based state of the art) at 99.3–100% validity and with shorter, more diverse inputs, by evolving populations of derivation trees under per-constraint fitness with grammar-aware crossover and mutation that keep every offspring syntactically valid by construction.

XAVIER: Grammar-Based Testing for XML Injection Attacks

ISSTA 2025 · Short Paper

P. Kalbitzer, **J.A. Zamudio Amaya**, A. Zeller.

- Matched a premium closed-source security tool with an open-source one, generating 2,207 XML injection inputs per second at grammar coverage and diversity equal to SoapUI Pro on every subject, and reaching 65% grammar coverage on a service where SoapUI Pro reaches 0%, by deriving test suites directly from the WSDL grammar and mutating them with an SQL-injection attack grammar.

Personalized Fuzzing: A Case Study with Fandango on a GNSS Module

FUZZING 2025 · Short Paper

S. Neuhaus, **J.A. Zamudio Amaya**, A. Zeller.

- Showed that constraint-guided generation reaches safety-critical hardware where traditional fuzzers cannot, producing valid and deliberately malformed UBX messages (denormals, NaN, corrupted checksums, invalid length encodings) for a u-blox GNSS module in seconds, by specifying the binary frame as a grammar with checksum generators and Python constraints over its fields.

Shaping Test Inputs in Grammar-Based Fuzzing

ISSTA 2024 · Doctoral Symposium

J.A. Zamudio Amaya

- Argued that fuzzers should control the *distribution* of the inputs they generate, not only their validity, having shown that grammar-based fuzzers are structurally length-biased (under a standard expansion, an input of 20+ digits has a 0.0004% chance of being generated), by proposing distributional constraints as a first-class objective in a genetic search over derivation trees.

Finding the Minimum k-Weighted Dominating Sets Using Heuristic Algorithms

MATCOM 2024 · Journal

E. Barrena, S. Bermudo, A.G. Hernandez-Diaz, A.D. Lopez-Sanchez, **J.A. Zamudio Amaya.**

- Defined a new NP-hard problem (k-weighted dominating sets, previously unstudied) and made it tractable at scale, beating a Gurobi exact solver from 300 vertices upward (at 4,000 vertices: objective 16.71 in 38.5s, against 22.11 in a 1,800s timeout) and matching the state of the art on the unweighted special case in half the runtime, by pairing a greedy construction with an Iterated Greedy destruction–reconstruction search. (Author order alphabetical.)

Analysis Operations on the Run: Feature Model Analysis in Constraint-Based Recommender Systems

SPLC 2023 · Research Track

S. Lubos, A. Felfernig, V.M. Le, T.N.T. Tran, D. Benavides, **J.A. Zamudio Amaya**, D. Garber.

- Extended feature-model analysis from structural properties (dead features, void models) to the recommendation outcome itself, introducing metrics for restrictiveness, accessibility, visibility, controversy, and efficiency, by formalizing feature-model-based recommendation as a constraint satisfaction problem and enumerating the consistent recommendations with a solver.

Talks & Presentations

2026	**ASE 2026,** Feature-Focused Test Generation.	*Munich, Germany.*
2026	**FSE 2026,** How to Test Complex Systems Automatically and Systematically.	*Montreal, Canada.*
2026	**SSBSE 2026,** Search-Based Generation of Complex Inputs with FANDANGO.	*Montreal, Canada.*
2026	**ICST 2026,** Constraint-Driven Fuzzing at Scale with FANDANGO.	*Daejeon, S. Korea.*
2025	**INNOSOFT 2025,** Research abroad and Ph.D. life.	*Sevilla, Spain.*
2025	**FUZZING 2025,** Personalized Fuzzing: A Case Study with Fandango on a GNSS Module.	*Trondheim, Norway.*
2025	**ISSTA 2025,** FANDANGO: Evolving Language-Based Testing.	*Trondheim, Norway.*
2025	**ISSTA 2025,** XAVIER: Grammar-Based Testing for XML Injection Attacks.	*Trondheim, Norway.*
2024	**ISSTA 2024,** Shaping Test Inputs in Grammar-Based Fuzzing,	*Vienna, Austria.*
2023	**SISTEDES 2023,** Open-source Automatic Chatbot Generation & Training	*Ciudad Real, Spain.*
2022	**SISTEDES 2022,** Optimizing Automatic Website Generation via Feature Selection	*Galicia, Spain.*

Teaching & Supervision

Fuzzing with Fandango

Saarbrücken, Germany

Tutorial and Reference · Co-authored with M. Smytzek, A. Liggesmeyer, V. Huber, A. Zeller — *2025 - 2026*

- Co-author and co-instructor of a 353-page tutorial and reference guide on language-based fuzzing with Fandango, covering practical semantic fuzzing techniques.

Thesis Supervision Advisor

Saarbrücken, Germany

Saarland University — *2024 - 2026*

- Supervised five theses on topics including SSL/X.509 fuzzing, XML and JSON injection, property-based testing and automated program repair; one student continued to a Ph.D. at CISPA, one student became a startup founder on the topic of his thesis, and one project led to an ISSTA 2025 publication.

Teaching Assistant, Security Testing (Advanced Course)

Saarbrücken, Germany

Saarland University — *Oct. 2023 - Mar. 2024*

- Assisted in the advanced Security Testing course, designing exercises, projects, and exams for a cohort of 207 students, covering topics including fuzzing, static analysis, and vulnerability discovery.

Training

2024 **TAROT Summer School on Software Testing, Verification & Validation**, University of Bergamo
2024 **Fuzzing and Software Security Summer School 2024**, National University of Singapore
2024 **Project Management**, Impulsplus
2024 **Pitch Coaching**, Beth Susanne
2024 **Scientific Writing I & II**, CISPA Helmholtz Center for Information Security
2024 **Cybersecurity Skills**, National Institute of Cybersecurity

Service

2026 **Artifact Evaluation Track**, International Conference on Automated Software Engineering (ASE 2026)
2026 **Artifact Evaluation Track**, International Symposium on Software Testing and Analysis (ISSTA 2026)
2026 **Reviewer**, Transactions on Software Engineering and Methodology · TOSEM · ISSN: 1049-331X
2026 **Tools and Datasets Track**, International Conference on Automated Software Engineering (ASE 2026)
2025 **Website Chair**, International Workshop on Search-Based and Fuzz Testing (SBFT 2026)
2024 **Reviewer**, Journal of Systems and Software · JSS · ISSN: 0164-1212

Honors & Awards

2023 **Master's Thesis Honors Distinction (Best)**, University of Seville. Selected to present at SISTEDES'23.
2022 **Bachelor's Thesis Honors Distinction (Top 5%)**, University of Seville. Selected to present at SISTEDES'22.

Skills

Programming	Python, C/C++, Rust, Java, JavaScript, SQL.
Search-Based	Genetic algorithms, NSGA-II & multi-objective optimization, iterated greedy & local search, fitness design.
Fuzzing	AFL++, libFuzzer, OSS-Fuzz, LLVM instrumentation, ASan/UBSan/MSan/TSan, grammar-based & protocol fuzzing
Program Analysis	Z3/SMT constraint solving, symbolic execution, taint analysis
AI & LLMs	PyTorch, Hugging Face Transformers, LLM-based test generation, activation-pattern analysis
Empirical	Experimental design, statistical significance & effect size, benchmark construction, reproducible artifacts
Languages	Spanish (native), English (C2), French (B1), German (A1 — in progress)

Philosophy & Interests

Open Research	I believe publicly funded research should produce publicly available tools and reproducible experiments. Every artifact I build is open source; Fandango is the direct result of this philosophy.
Collaboration	I seek out collaborators across institutions and disciplines because the best ideas in testing come from the intersection of formal methods, software engineering, and security.
Trustworthy AI	I care about testing and securing AI systems so they behave reliably in practice, especially before deployment in high-impact domains.
Challenges	I am drawn to problems that get harder as systems grow, where naive solutions break, and principled engineering matters. I want to build the infrastructure that researchers and engineers reach for.
Real-World Impact	Every technique I develop and every tool I build is designed for researchers and practitioners alike, with a deliberate focus on practicality, modularity, and documentation.