

# 2021 ALONZO CHURCH AWARD FOR OUTSTANDING CONTRIBUTIONS TO LOGIC AND COMPUTATION

The European Association for Theoretical Computer Science (EATCS), the ACM Special Interest Group on Logic (SIGLOG), the European Association for Computer Science Logic (EACSL), and the Kurt Goedel Society (KGS) are pleased to announce that

- **Georg Gottlob, Christoph Koch, Reinhard Pichler, Klaus U. Schulz, and Luc Segoufin**

have been selected as the winners of the *2021 Alonzo Church Award for Outstanding Contributions to Logic and Computation* for fundamental work on logic-based web data extraction and querying tree-structured data, published in:

1. Georg Gottlob and Christoph Koch. Monadic Datalog and the Expressive Power of Languages for Web Information Extraction. *Journal of the ACM* 51(1): 74–113, January 2004 (DOI: 10.1145/962446.962450),
2. Georg Gottlob, Christoph Koch, and Klaus U. Schulz. Conjunctive Queries Over Trees. *Journal of the ACM* 53(2): 238–272, March 2006 (DOI: 10.1145/1131342.1131345),
3. Georg Gottlob, Christoph Koch, and Reinhard Pichler. Efficient Algorithms for Processing XPath Queries. *ACM Transactions on Database Systems* 30(2): 444–491, June 2005 (DOI: 10.1145/1071610.1071614), and
4. Georg Gottlob, Christoph Koch, Reinhard Pichler, and Luc Segoufin. The Complexity of XPath Query Evaluation and XML Typing. *Journal of the ACM* 52(2): 284–335, March 2005 (DOI: 10.1145/1059513.1059520).

Paper (1) establishes a comprehensive logical theory of Web data extraction. At its core, this is the problem of selecting relevant nodes (subtrees) from HTML text. While the set of relevant nodes can be expressed in Monadic Second-Order logic (MSO) over finite trees, MSO has high computational complexity. The authors prove that Monadic Datalog on trees has exactly the same expressive power

as full MSO and that, surprisingly, evaluating Monadic Datalog is feasible in time linear in the size of query and input tree. These results greatly influenced theoretical and applied research, and gave rise to logic-based systems for data extraction that have been successfully used in industry.

Papers (2,3,4) present deep investigations into logical queries over tree-structured data. The complexity of evaluating XPath, a key technology in Web browsers and other systems, was unclear, and available implementations required exponential time. Paper (2) gives a full characterization of, and a dichotomy theorem for, the complexity of conjunctive queries on various representations of trees. Paper (3) shows that the full XPath standard can be evaluated in PTIME and proposes a logical core which has become seminal to research efforts at the intersection of Web data processing and (modal) logics. Finally, paper (4) establishes the precise complexity of evaluating XPath fragments.

The 2021 Alonzo Church Award Committee:

- Mariangiola Dezani,
- Thomas Eiter,
- Javier Esparza (chair),
- Radha Jagadeesan, and
- Igor Walukiewicz.

The list of the previous recipients of the Alonzo Church Award for Outstanding Contributions to Logic and Computation is available at <https://siglog.org/awards/alonzo-church-award/>.