



University of Glasgow | School of
Computing Science

THE AWARDS
2020

UNIVERSITY
OF THE YEAR

MIMIC Critical Care Dataset: The Impact

Dr. Fani Deligianni,

fani.deligianni@glasgow.ac.uk

Lecturer (Assistant Professor)

Lead of the Computing Technologies for Healthcare Theme

<https://www.gla.ac.uk/schools/computing/staff/fanideligianni>

WORLD
CHANGING
GLASGOW



Basic Information

- All of the material in this course is based on de-identified EHR provided via the Physionet database
- Physionet: Research Resource for Complex Physiologic Signals
- MIMIC-III Database – the only publicly available database of EHR, which links data from ICU and the hospital.



Motivation

Researchers at the Laboratory of Computational Physiology at MIT in 2000 recognised that a lot of rich data was being collected (“petabytes of data are captured daily during care delivery in the country’s ICU”) but none of it was being used to generate new knowledge



Motivation

- Freely accessible deidentified critical care dataset under a data user agreement
- Available both for academic, industrial research and higher education
- Dataset spam over a decade
- Analysis is relatively unrestricted
- Interoperability – data integration
- Reproducibility



MIMIC-III

Update to MIMIC-II, adds four more years of ICU data (2008-2012)

With other data updates (e.g., adding a new monitoring system company)

First published in 2016



Publications using MIMIC

According to Web of Science:

- MIMIC II: A massive temporal ICU patient database to support research in intelligent patient monitoring (2002)
- Multiparameter Intelligent Monitoring in Intensive Care II: A public-access intensive care unit database (2009)
- MIMIC-III, a freely accessible critical care database (2016)



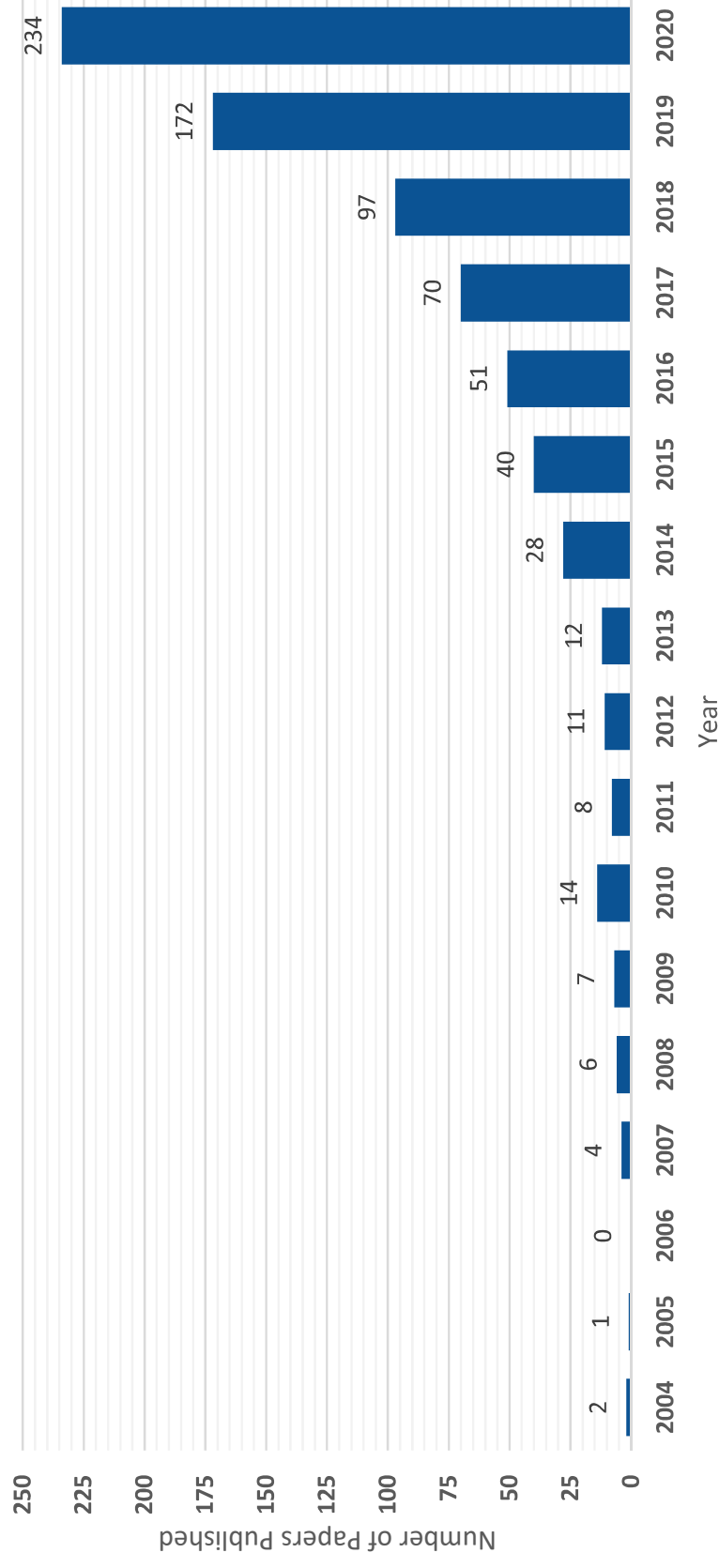
Fields of Research

365 COMPUTER SCIENCE	210 ENGINEERING	174 CRITICAL CARE MEDICINE	96 GERIATRICS GERONTOLOGY	88 PATHOLOGY	84 NEUROSCIENCE NEUROLOGY
278 HEALTH CARE SCIENCES SERVICE	206 SCIENCE TECHNOLOGY OTHER TOPICS	173 CARDIOVASCULAR SYSTEM CARDIOLOGY	69 BIOPHYSICS	59 MEDICAL LABORATORY TECHNOLOGY	53 RESEARCH EXPERIMENTAL MEDICINE
270 MATHEMATICAL COMPUTATIONAL BIOLOGY	193 GENERAL INTERNAL MEDICINE	134 COMMUNICATION	69 INFECTIOUS DISEASES	43 RESPIRATORY SYSTEM	39 HEMATOLOGY
250 MATHEMATICS	190 MEDICAL INFORMATICS	122 RADIOLOGY NUCLEAR MEDICINE MEDICAL IMAGING	69 INFORMATION SCIENCE LIBRARY SCIENCE	39 TELECOMMUNICATIONS	37 ANESTHESIOLOGY

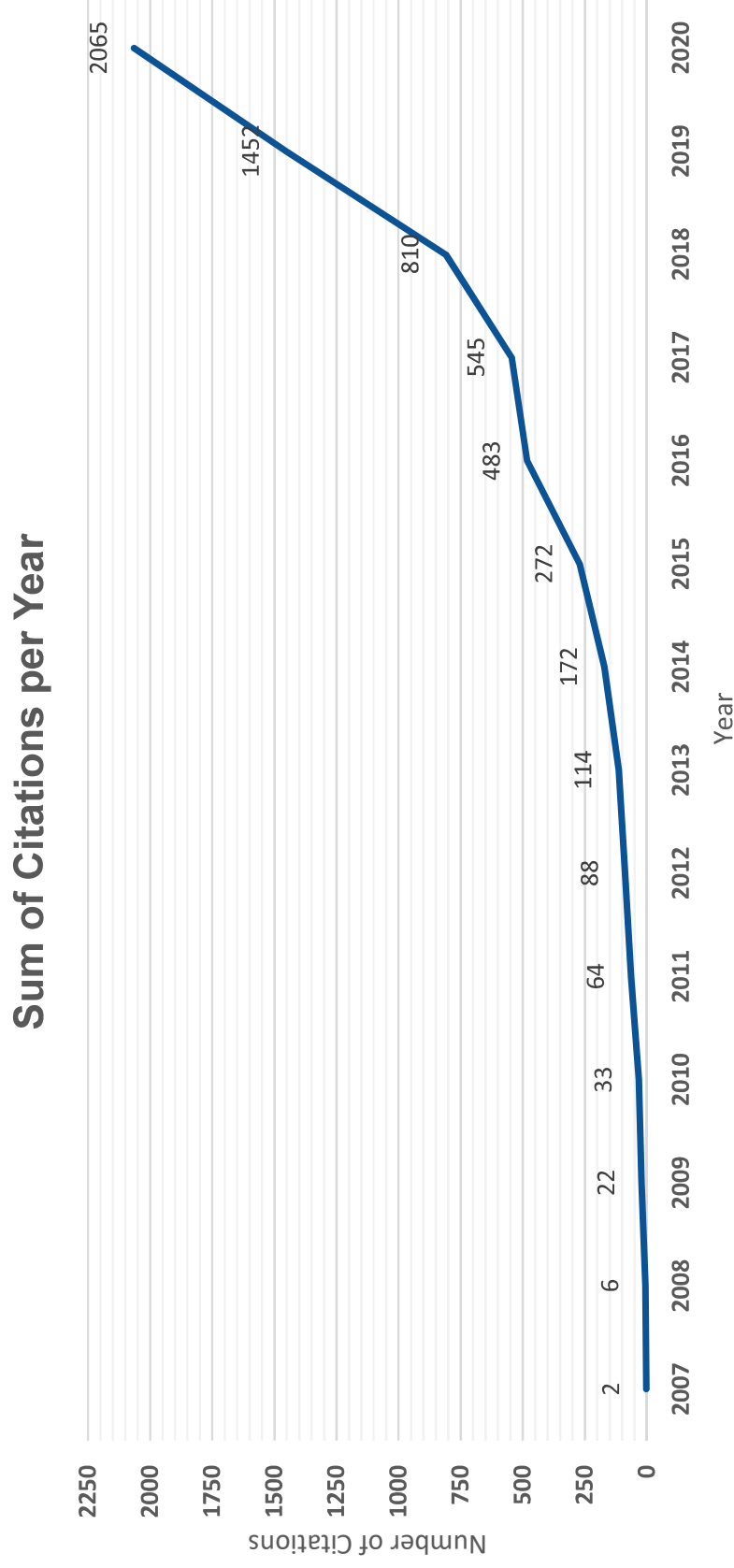


Publications by Year

Papers Published Per Year



Citations from Papers using MIMIC



Papers which use MIMIC...

Recurrent Neural Networks for Multivariate Time Series with Missing Values (2018)

- Cited 235 times

Automated de-identification of free-text medical records (2008)

- Cited 158 times

Big Data Analytics in Healthcare (2015)

- Cited 146 times



Summary

- Availability of de-identified EHR
- MIMIC III Dataset:
 - Reproducibility of clinical studies
 - Development of new algorithms
 - Development of new technologies
 - Unique clinical database with impactful research

