Trường ĐH Bách Khoa Tp.HCM
Khoa Khoa Học và Kỹ Thuật Máy Tính

Coursework : Privacy preserving salary prediction in data mining

GV: PGS.TS ĐẶNG TRẦN KHÁNH

Chu Xuân Tình -1870583
Nguyễn Đức Huy -1870567

Coursework

- Introduction
- Data input & Model
- Data anonymization tool
- Implementation

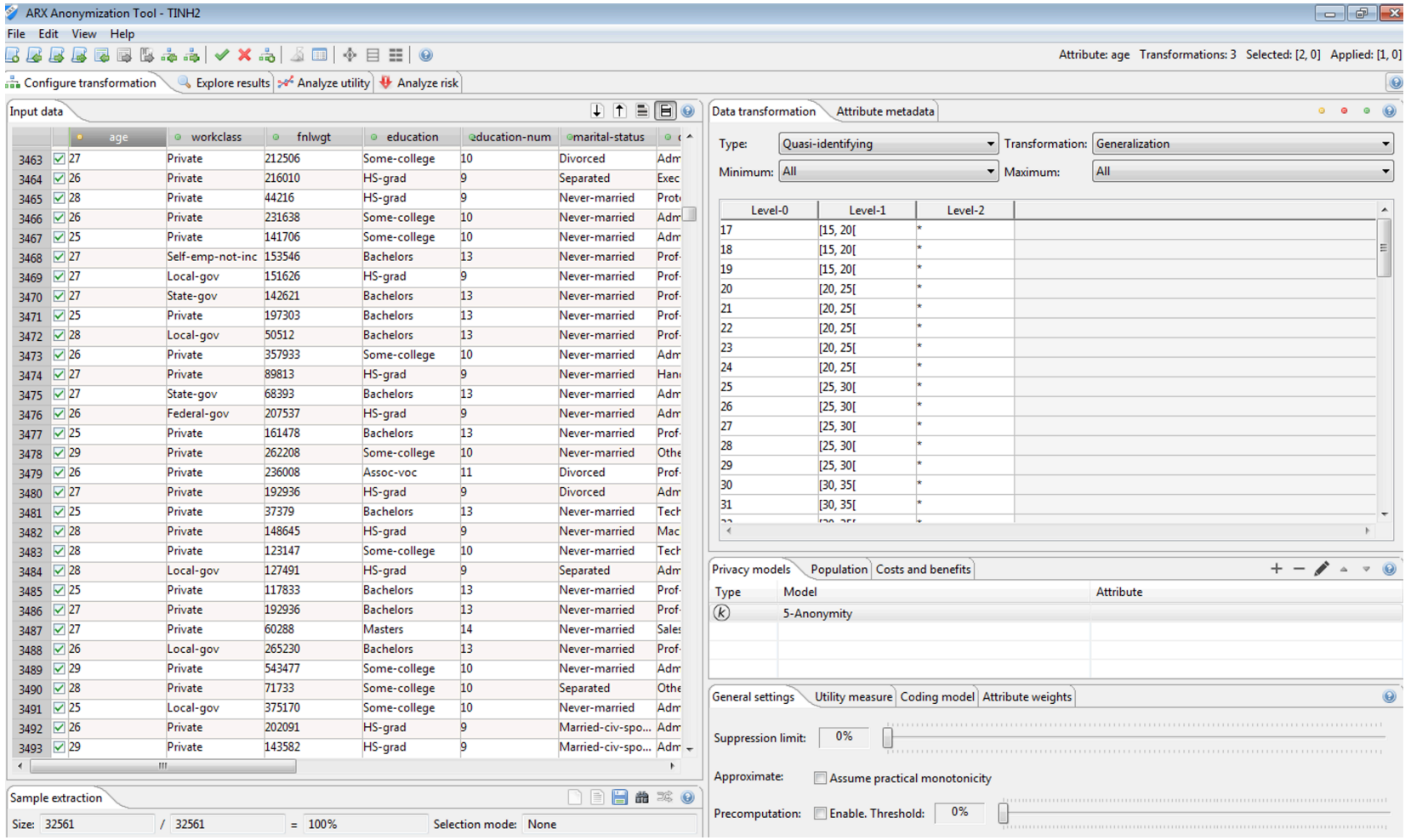
Introduction

- **Privacy preserving salary prediction in data mining**
- Data input

A	B	D	E	F	G	H	I	J	K	M	N	O
age	workclass	education	education-num	marital-status	occupation	relationship	race	sex	capital-gain	hours-per-w	native-country	class
39	State-gov	Bachelors	13	Never-married	Adm-clerica	Not-in-family	White	Male	2174	40	United-States	<=50K.
50	Self-emp-nc	Bachelors	13	Married-civ-spouse	Exec-manag	Husband	White	Male	0	13	United-States	<=50K.
38	Private	HS-grad	9	Divorced	Handlers-cle	Not-in-family	White	Male	0	40	United-States	<=50K.
53	Private	11th	7	Married-civ-spouse	Handlers-cle	Husband	Black	Male	0	40	United-States	<=50K.
28	Private	Bachelors	13	Married-civ-spouse	Prof-special	Wife	Black	Female	0	40	Cuba	<=50K.
37	Private	Masters	14	Married-civ-spouse	Exec-manag	Wife	White	Female	0	40	United-States	<=50K.
49	Private	9th	5	Married-spouse-absent	Other-servic	Not-in-family	Black	Female	0	16	Jamaica	<=50K.
52	Self-emp-nc	HS-grad	9	Married-civ-spouse	Exec-manag	Husband	White	Male	0	45	United-States	>50K.
31	Private	Masters	14	Never-married	Prof-special	Not-in-family	White	Female	14084	50	United-States	>50K.
42	Private	Bachelors	13	Married-civ-spouse	Exec-manag	Husband	White	Male	5178	40	United-States	>50K.
37	Private	Some-colleg	10	Married-civ-spouse	Exec-manag	Husband	Black	Male	0	80	United-States	>50K.
30	State-gov	Bachelors	13	Married-civ-spouse	Prof-special	Husband	Asian-Pac-Is	Male	0	40	India	>50K.
23	Private	Bachelors	13	Never-married	Adm-clerica	Own-child	White	Female	0	30	United-States	<=50K.
32	Private	Assoc-acdm	12	Never-married	Sales	Not-in-family	Black	Male	0	50	United-States	<=50K.
40	Private	Assoc-voc	11	Married-civ-spouse	Craft-repair	Husband	Asian-Pac-Is	Male	0	40	?	>50K.
34	Private	7th-8th	4	Married-civ-spouse	Transport-m	Husband	Amer-Indiar	Male	0	45	Mexico	<=50K.
25	Self-emp-nc	HS-grad	9	Never-married	Farming-fisl	Own-child	White	Male	0	35	United-States	<=50K.
32	Private	HS-grad	9	Never-married	Machine-op	Unmarried	White	Male	0	40	United-States	<=50K.
38	Private	11th	7	Married-civ-spouse	Sales	Husband	White	Male	0	50	United-States	<=50K.
43	Self-emp-nc	Masters	14	Divorced	Exec-manag	Unmarried	White	Female	0	45	United-States	>50K.

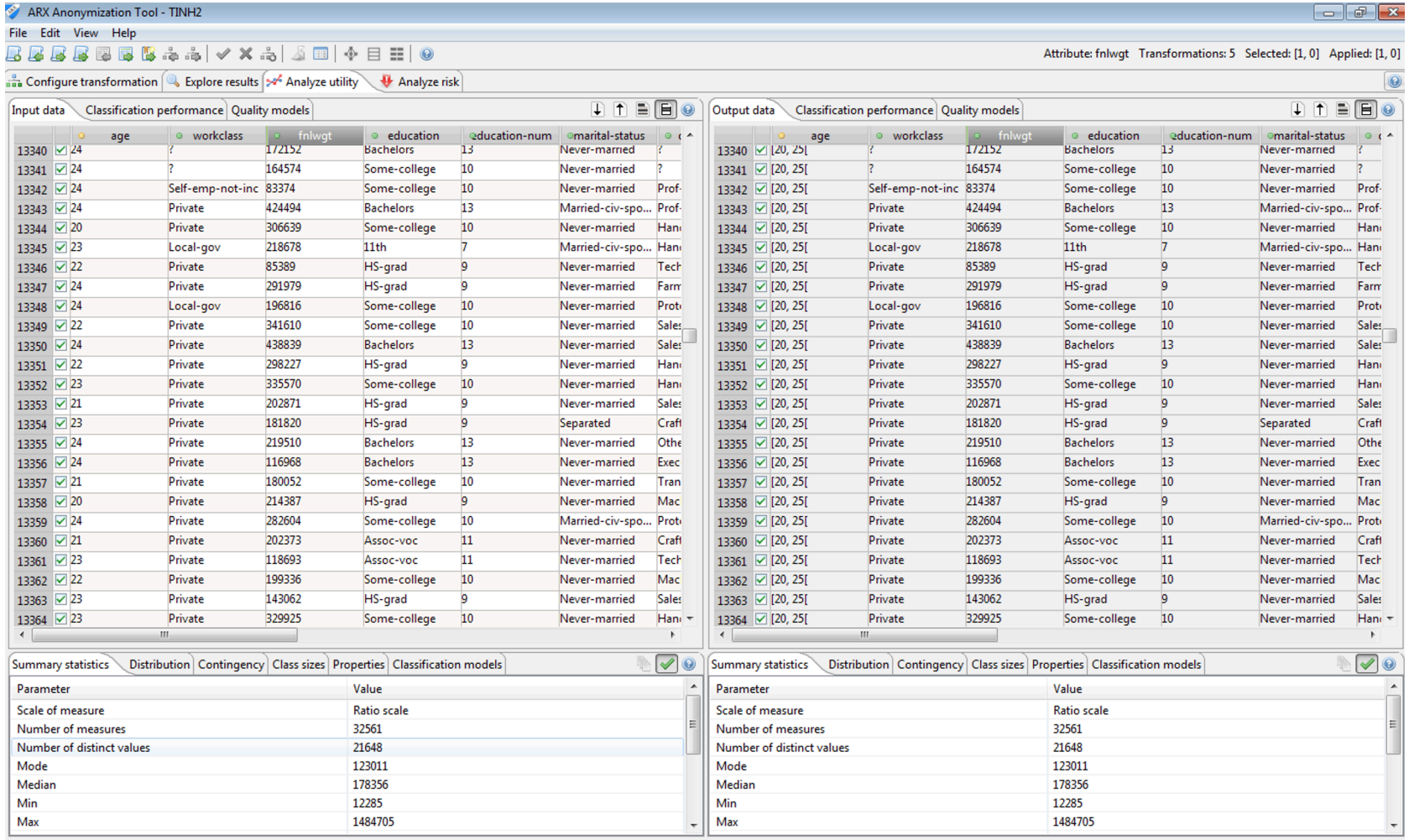
Data anonymization tool

- Data anonymization tool – ARX Tool



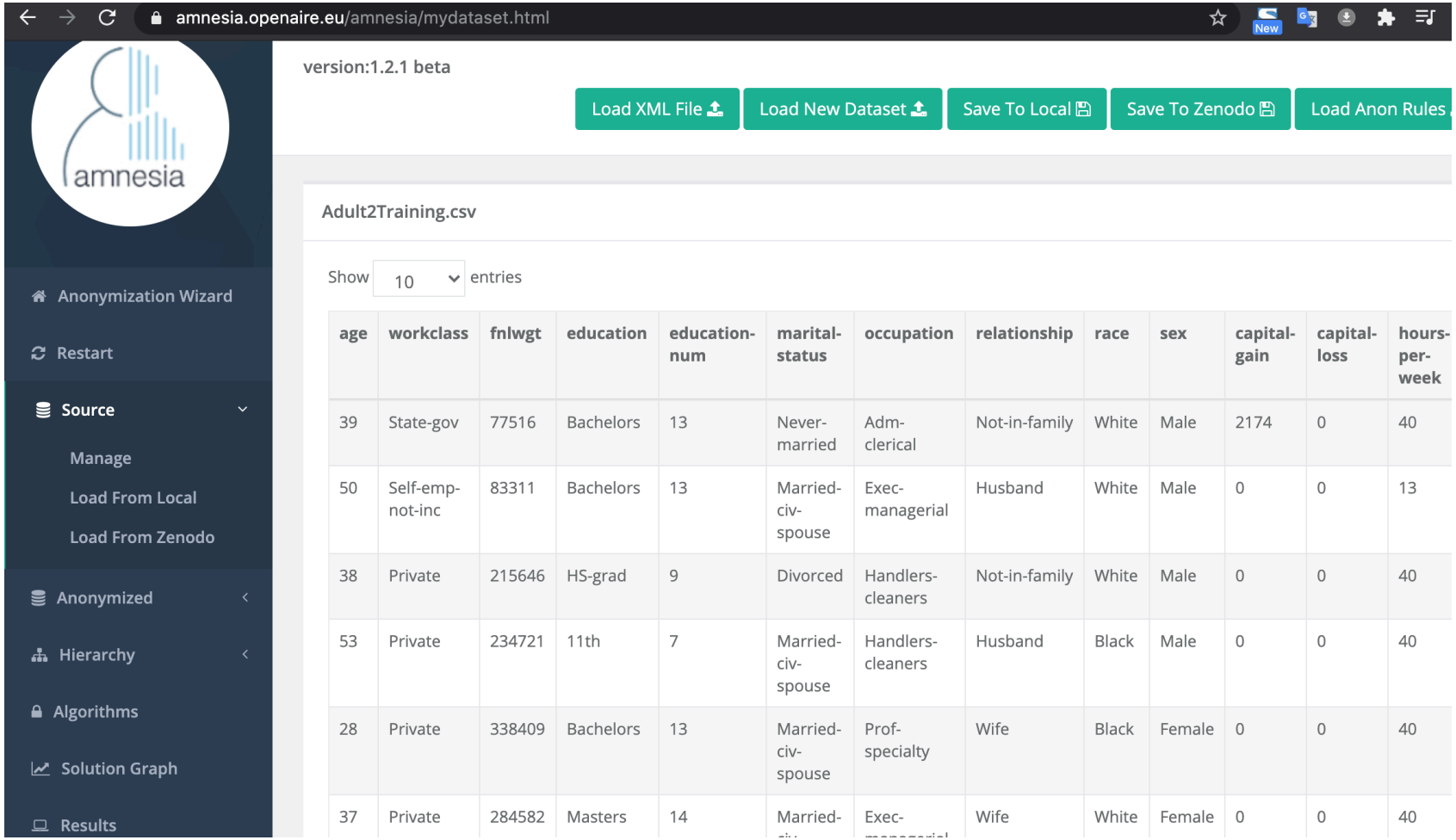
Data anonymization tool

- Data anonymization tool – ARX Tool



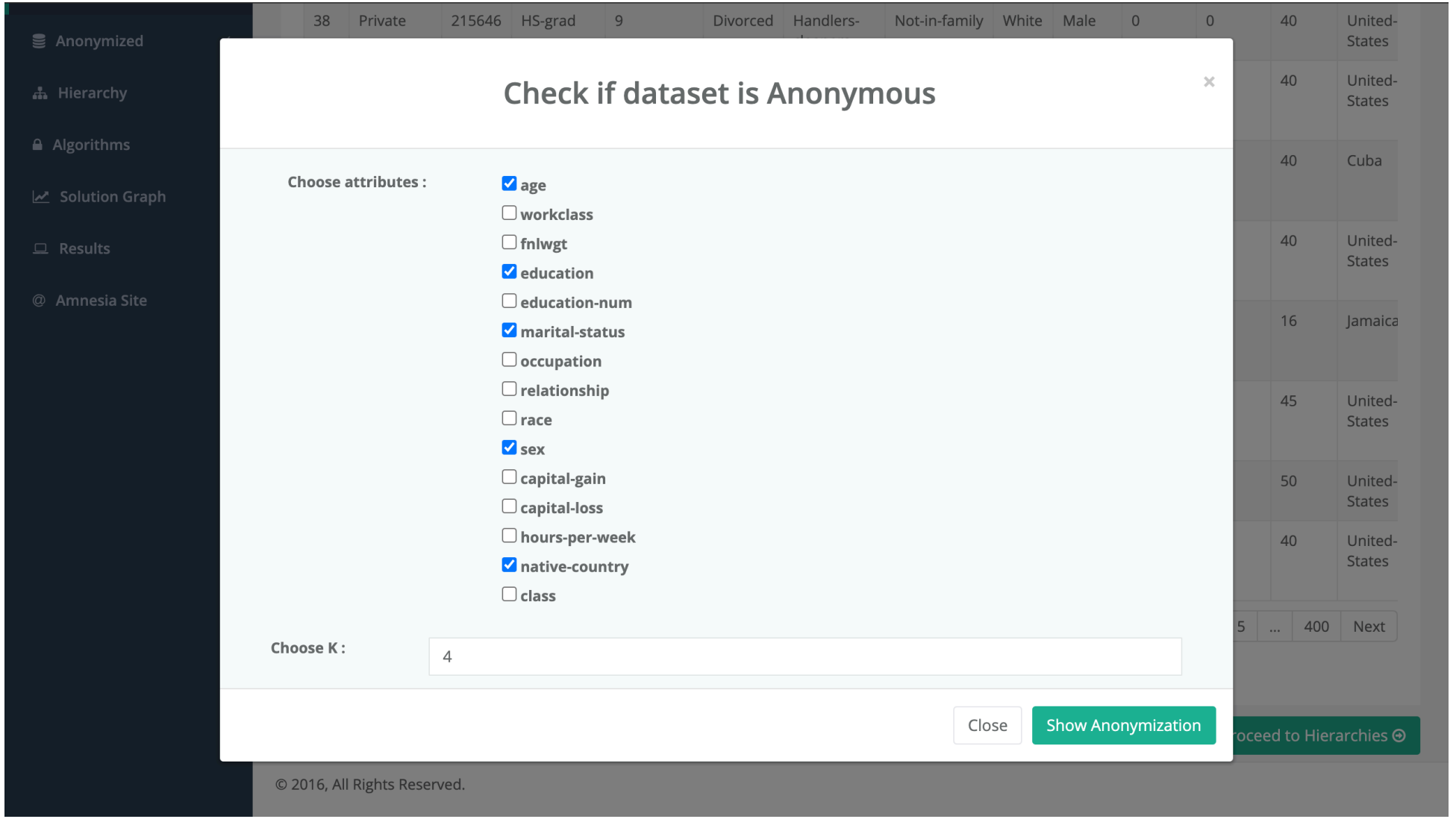
Data anonymization tool

- Data anonymization tool - https://amnesia.openaire.eu/

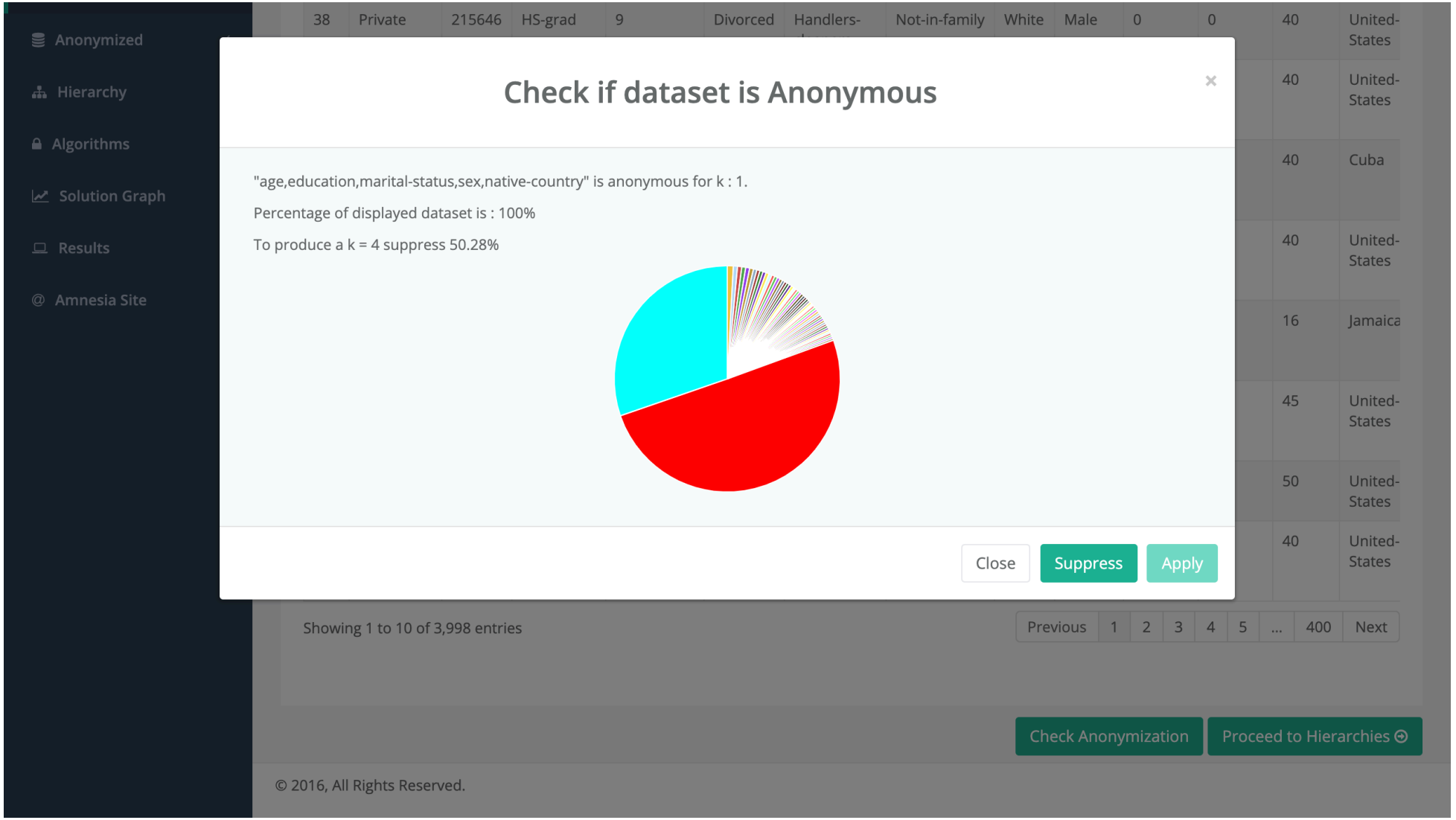


age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	sex	capital-gain	capital-loss	hours-per-week
39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40
50	Self-emp-not-inc	83311	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Male	0	0	13
38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Male	0	0	40
53	Private	234721	11th	7	Married-civ-spouse	Handlers-cleaners	Husband	Black	Male	0	0	40
28	Private	338409	Bachelors	13	Married-civ-spouse	Prof-specialty	Wife	Black	Female	0	0	40
37	Private	284582	Masters	14	Married-civ-	Exec-managerial	Wife	White	Female	0	0	40

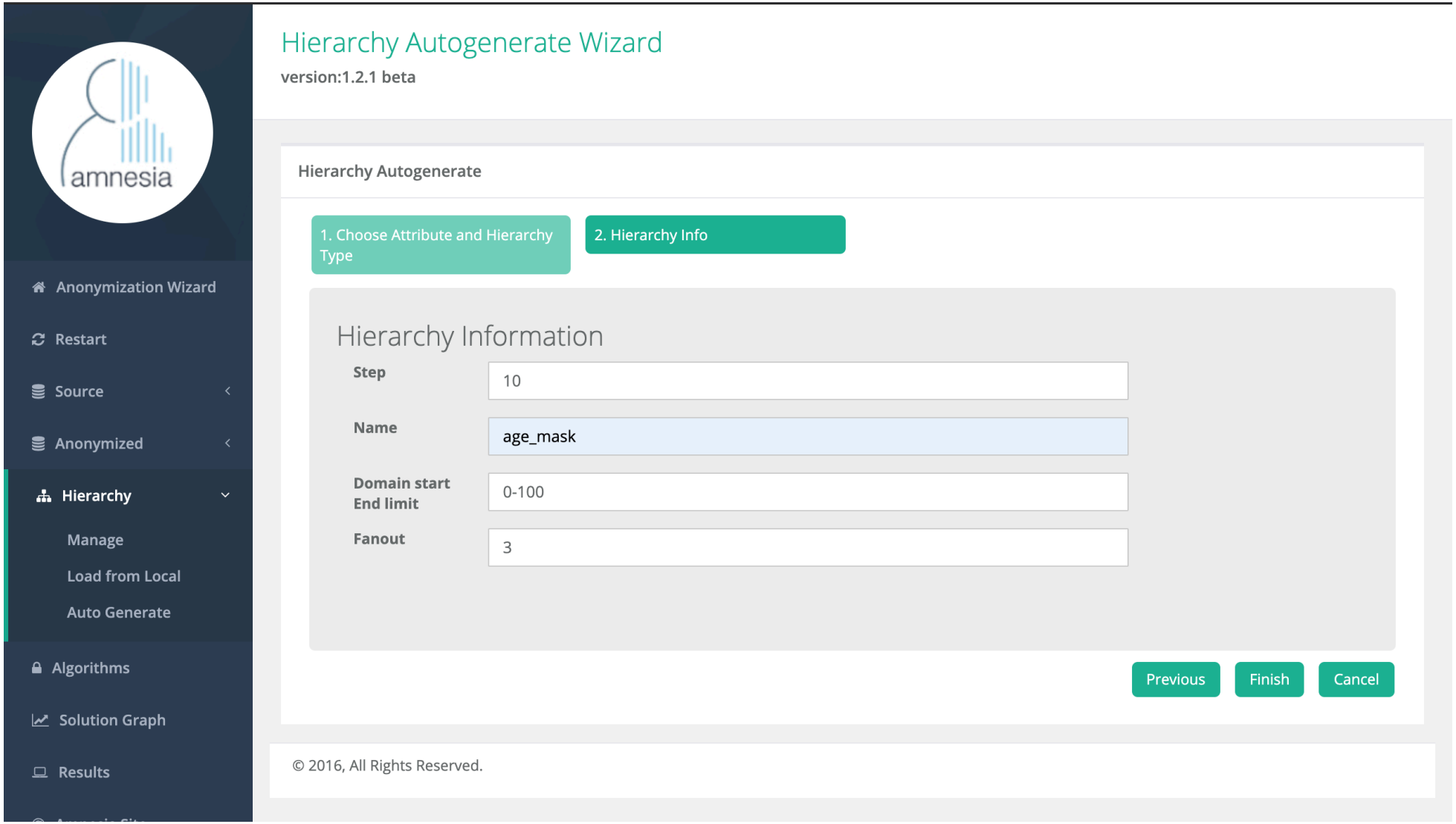
Implementation



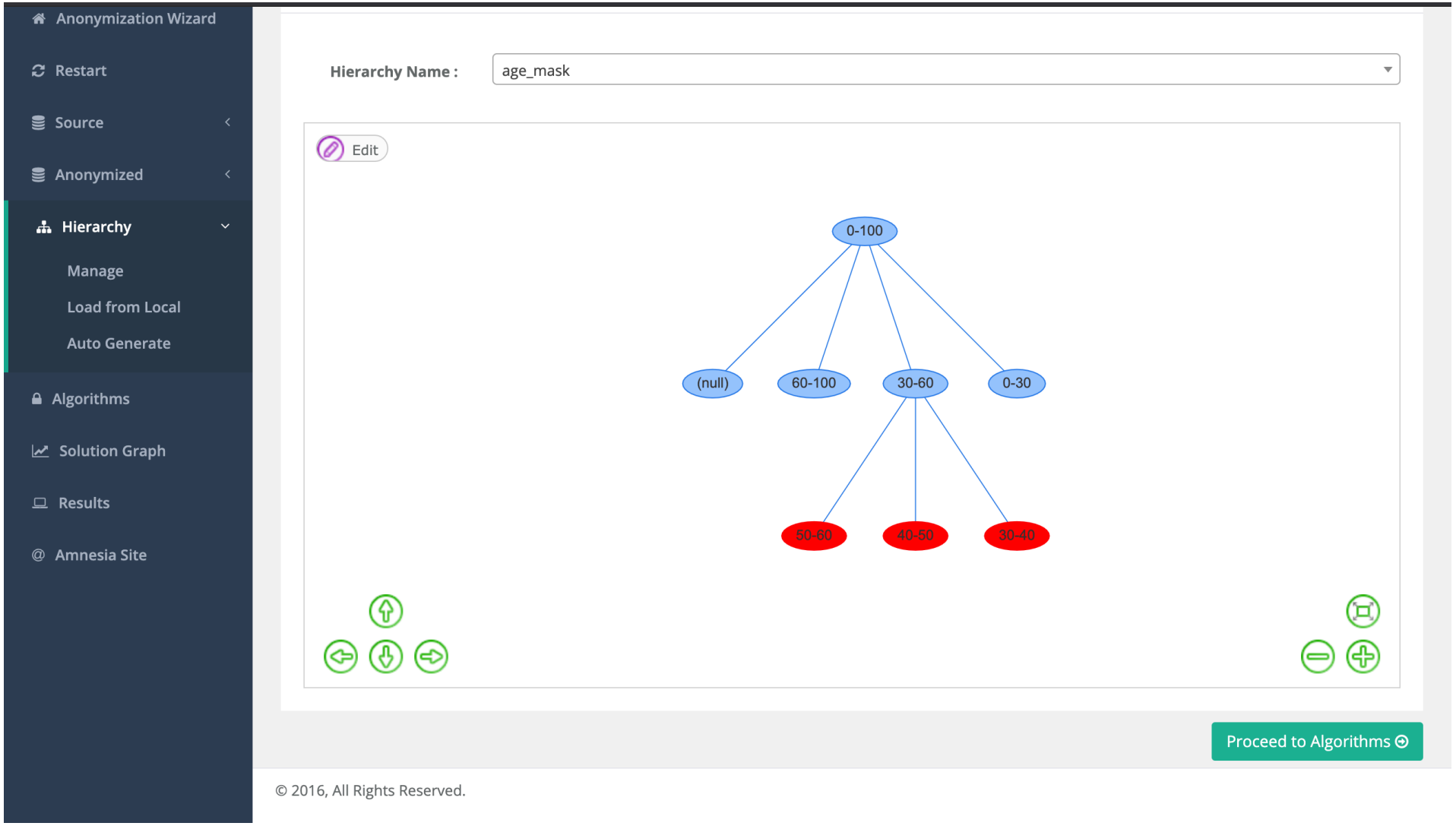
Implementation



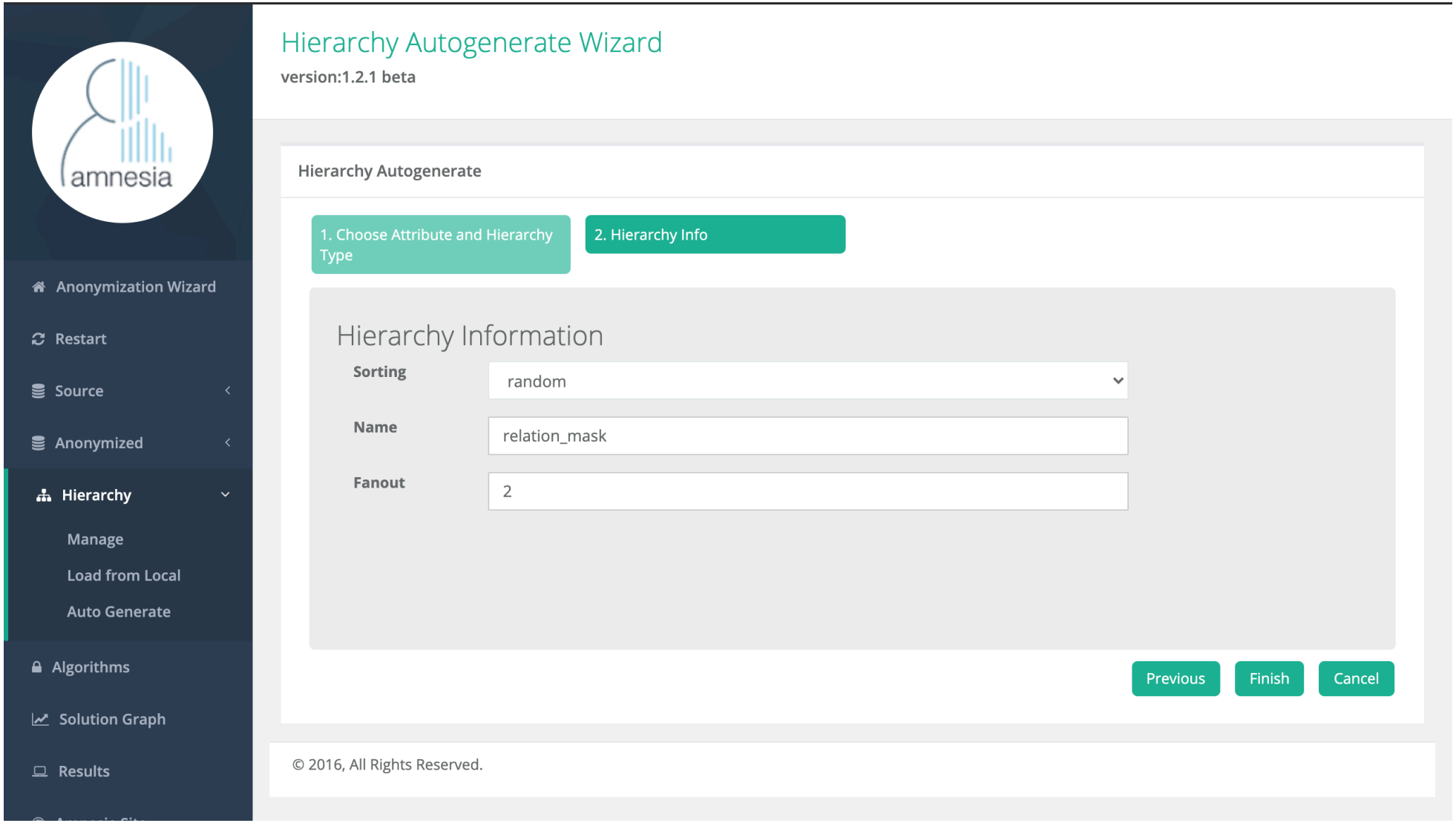
Implementation



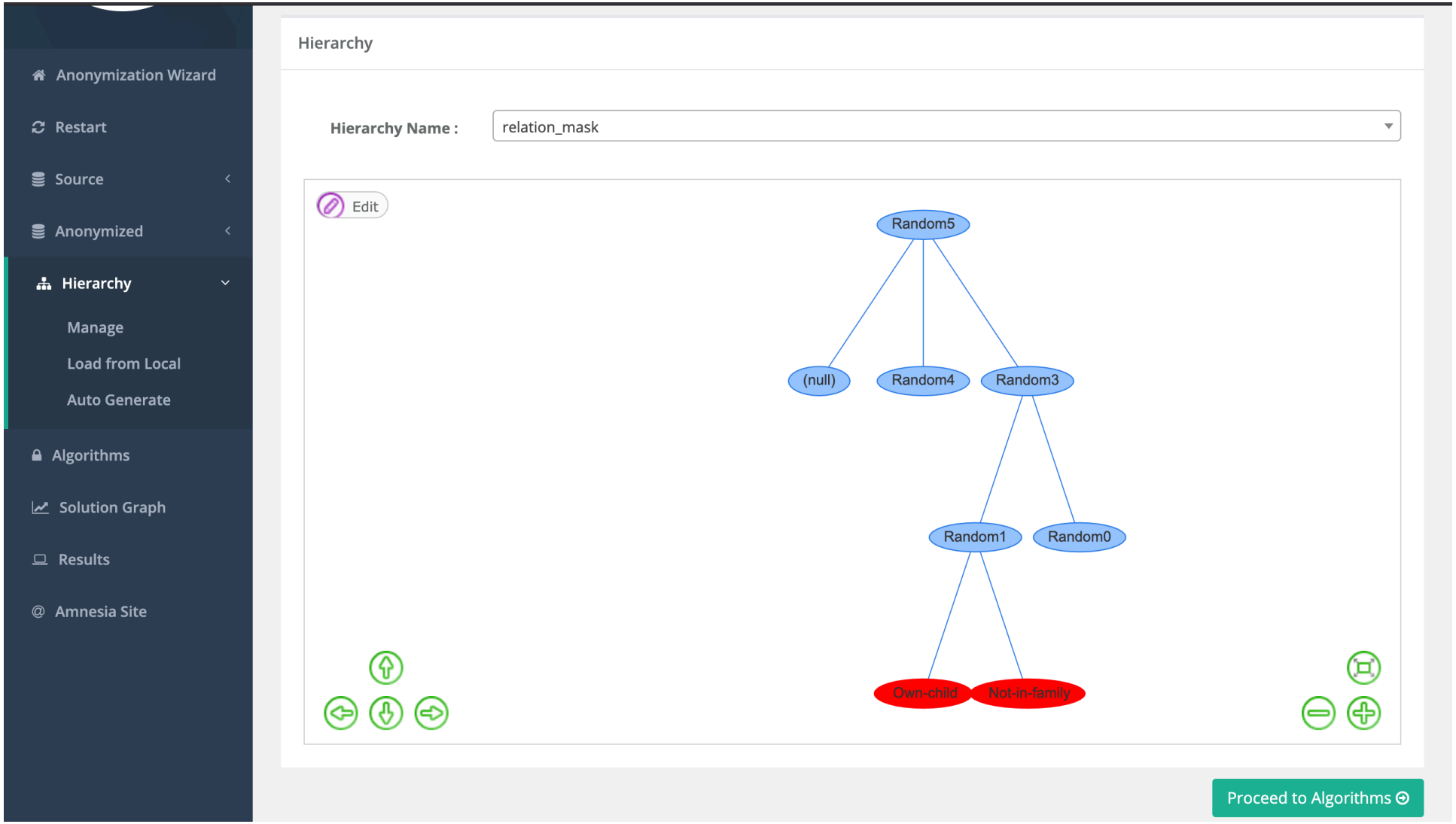
Implementation



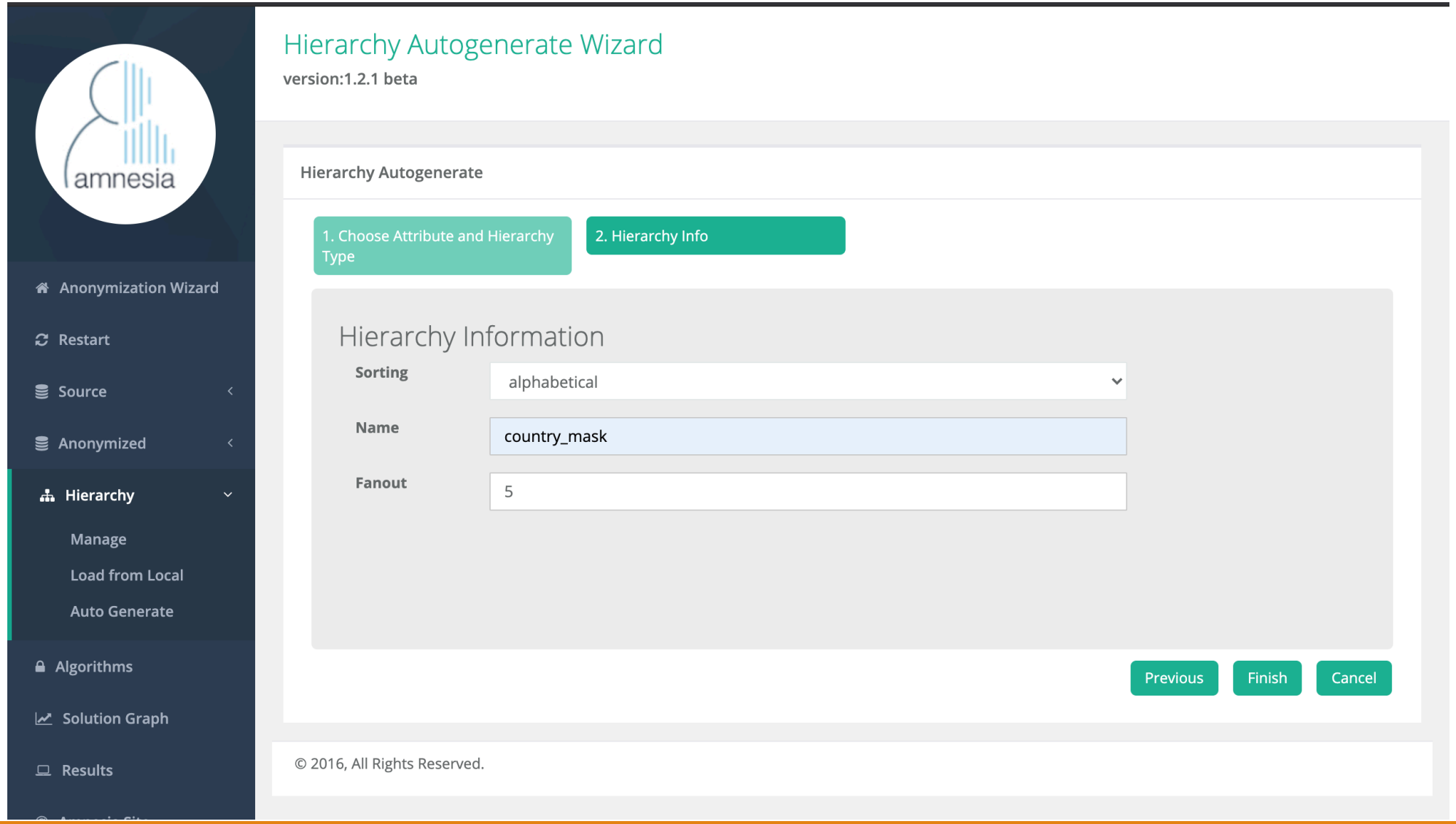
Implementation



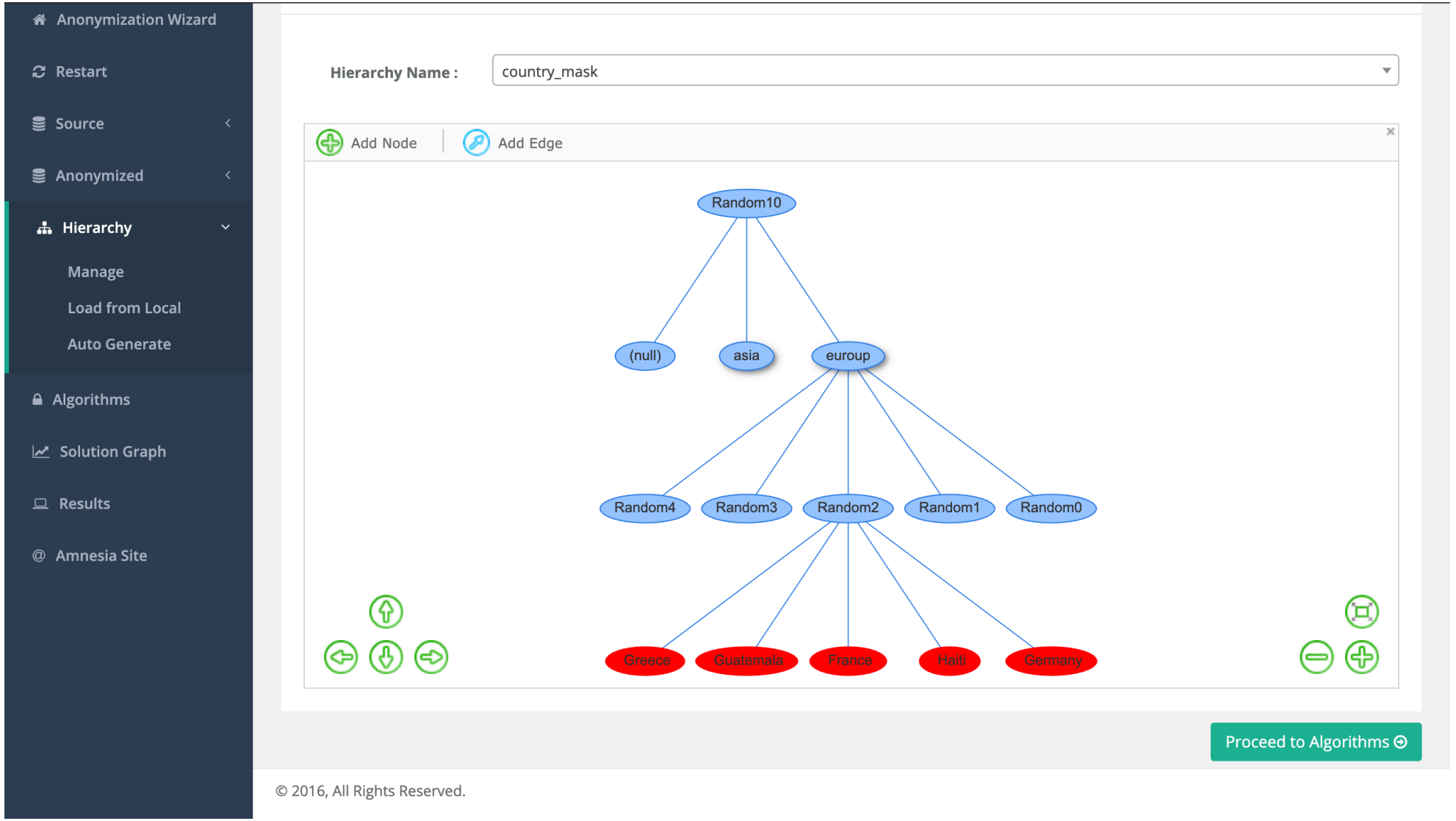
Implementation



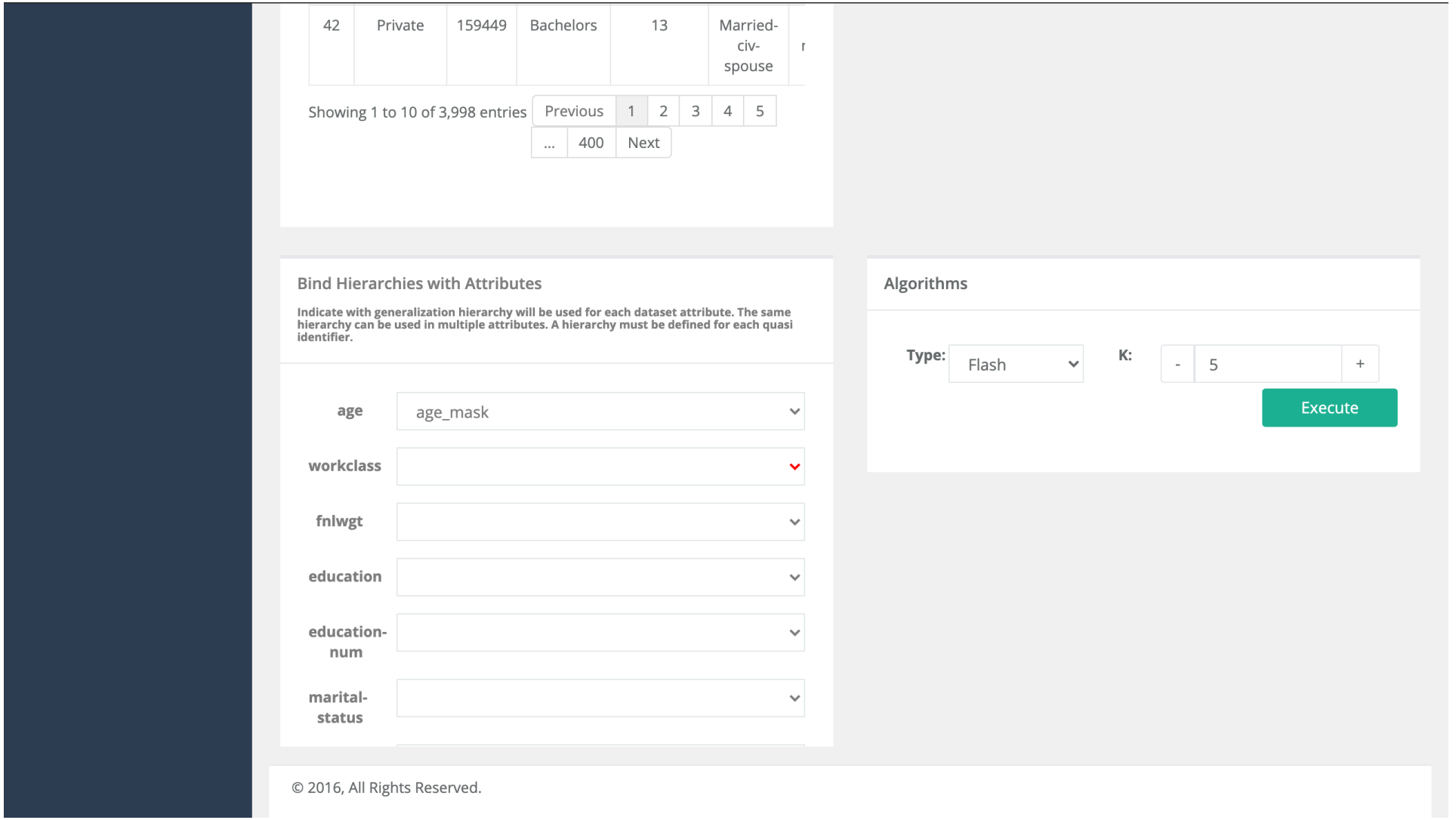
Implementation



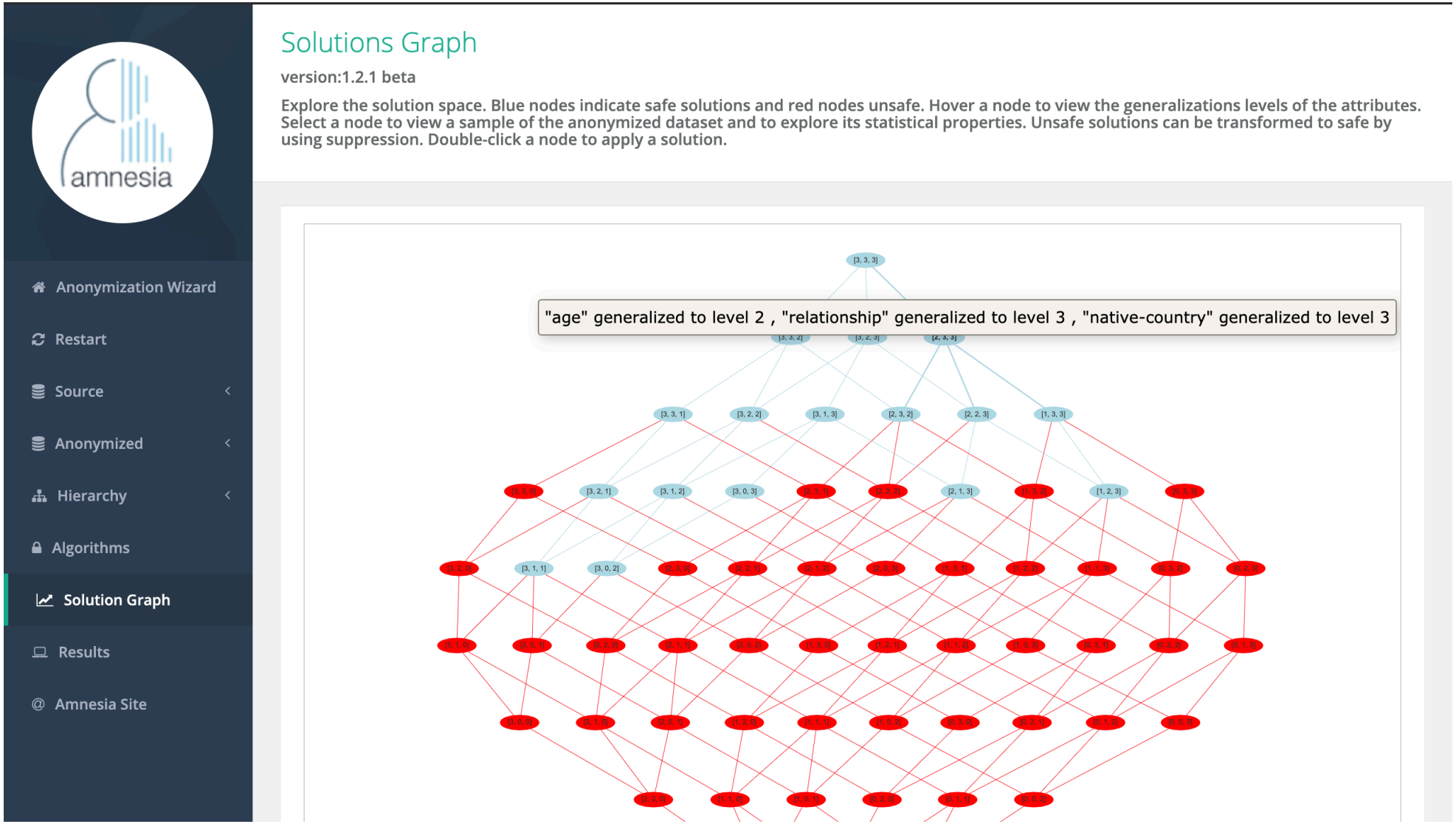
Implementation



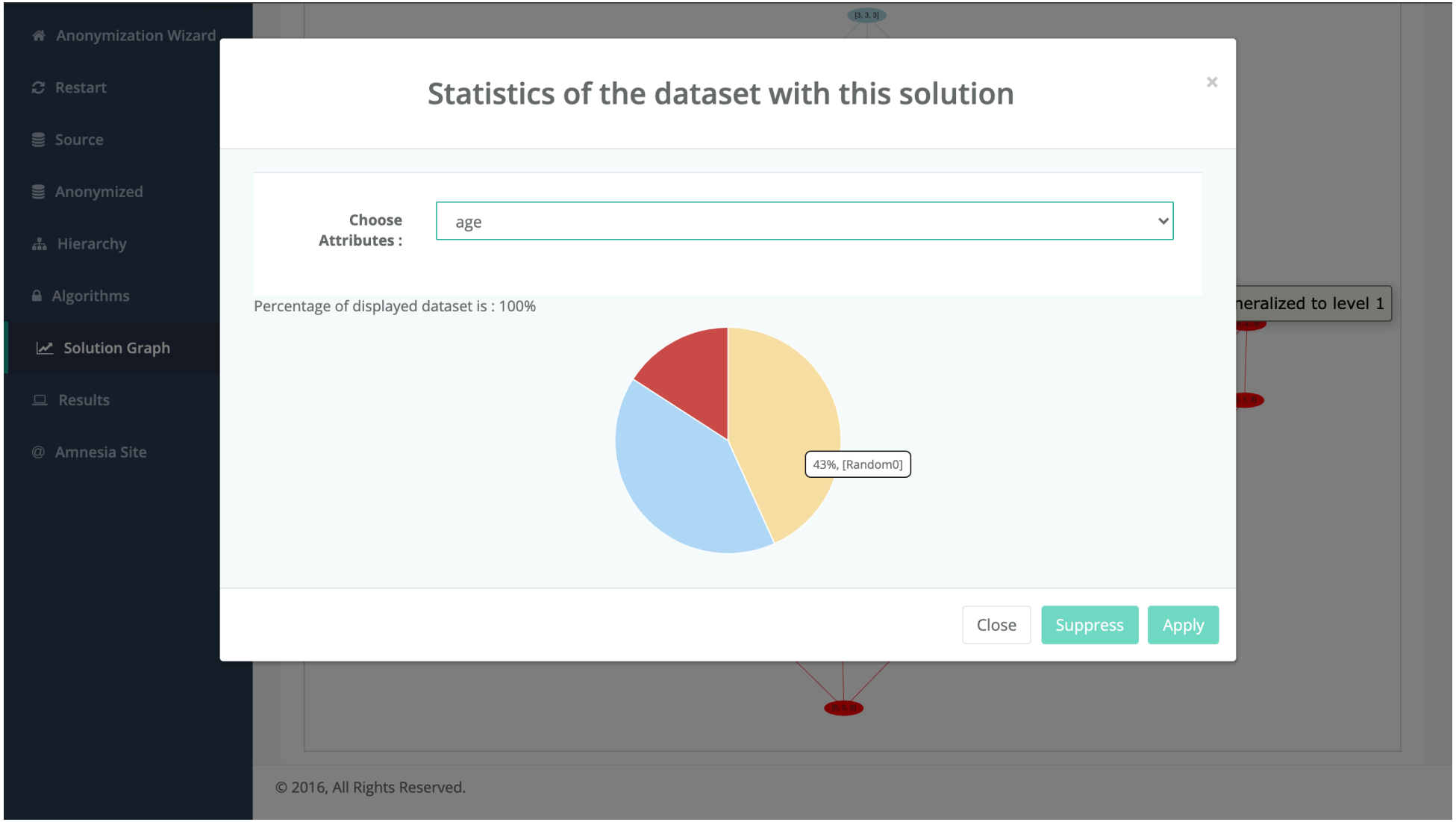
Implementation



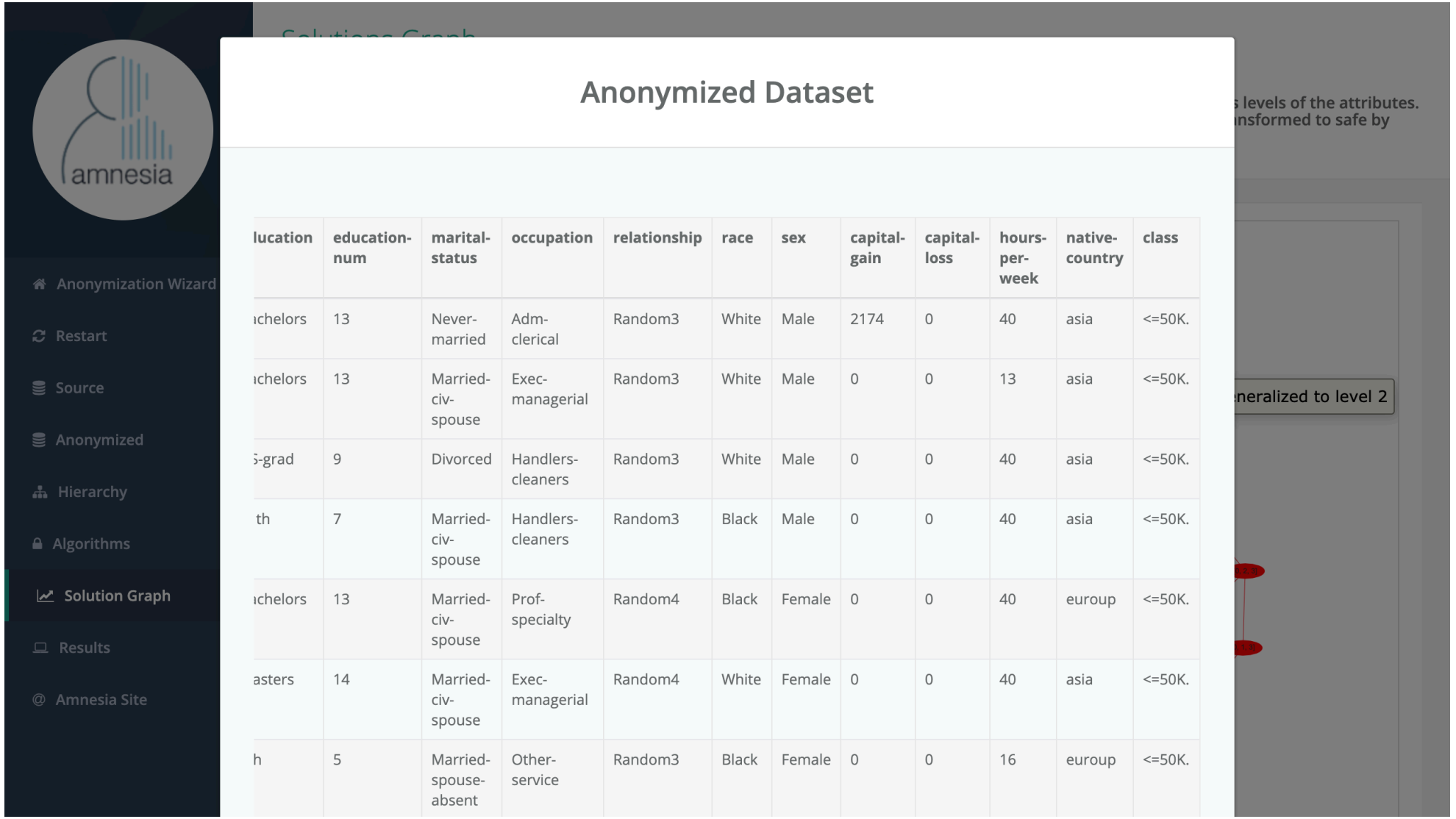
Implementation



Implementation



Implementation



THANK YOU!