



(11) **EP 3 258 393 A8**

(12) **CORRECTED EUROPEAN PATENT APPLICATION**

(15) Correction information:
Corrected version no 1 (W1 A1)
Corrections, see
Bibliography INID code(s) 72

(51) Int Cl.:
G06F 17/30 ^(2006.01)

(48) Corrigendum issued on:
19.06.2019 Bulletin 2019/25

(43) Date of publication:
20.12.2017 Bulletin 2017/51

(21) Application number: **16194936.7**

(22) Date of filing: **20.10.2016**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO
PL PT RO RS SE SI SK SM TR

(30) Priority: **13.06.2016 US 201662349548 P**
12.09.2016 US 201615262207

(71) Applicant: **Palantir Technologies Inc.**
Palo Alto, CA 94301 (US)

(72) Inventors:
• **FINK, Robert**
Palo Alto,
CA California 94301 (US)
• **CUTHRIELL, Lynn**
Palo Alto,
CA California 94301 (US)
• **ANDERSON, Adam**
Palo Alto,
CA California 94301 (US)
• **BOROCHOFF, Adam**
Palo Alto,
CA California 94301 (US)
• **LU, Catherine**
Palo Alto,
CA California 94301 (US)

- **RAFIDI, Joseph**
Palo Alto,
CA California 94301 (US)
- **MOHAN, Karanveer**
Palo Alto,
CA California 94301 (US)
- **JENNY, Matthew**
Palo Alto,
CA California 94301 (US)
- **MACLEAN, Matthew**
Palo Alto,
CA California 94301 (US)
- **GUO, Michelle**
Palo Alto,
CA California 94301 (US)
- **MENON, Parvathy**
Palo Alto,
CA California 94301 (US)
- **ROWE, Ryan**
Palo Alto,
CA California 94301 (US)

(74) Representative: **Herre, Peter**
Dendorfer & Herrmann
Patentanwälte Partnerschaft mbB
Neuhauser Straße 47
80331 München (DE)

(54) **DATA REVISION CONTROL IN LARGE-SCALE DATA ANALYTIC SYSTEMS**

(57) Computer-implemented techniques for data revision control in large-scale data analytic systems. In one embodiment, for example, the techniques encompass a method for data revision control in a large-scale data analytic system that comprises the steps of storing a first

version of a dataset that is derived by executing a first version of driver program associated with the dataset; and storing a first build catalog entry comprising an identifier of the first version of the dataset and comprising an identifier of the first version of the driver program.

EP 3 258 393 A8