

MANUALE UTENTE DEL PROGRAMMA EAT

NOME VARIABILE PROGRAMMA	DESCRIZIONE VALORE VARIABILE (sono tutti valori stringa)
PARAM_API_KEY_GM	codice personale per l'utilizzo dell'API google maps (opzionale)
PARAM_COLUMN_TABLE_INPUT	nome della colonna contenente gli url nella tabella di input (e.g. url)
PARAM_CREA_NUOVA_TABELLA_GEODOCUMENT	se TRUE crea una nuova tabella di geodocument nel server database
PARAM_CREA_NUOVA_TABELLA_GEODOMAIN	se TRUE crea una nuova tabella di geomaindocument nel server database
PARAM_CREA_NUOVA_TABELLA_INFODOCUMENT_ONTOLOGY	se TRUE crea una nuova tabella di infodocument nel server database
PARAM_DATABASE_INPUT_GEODOMAIN	nome del database contenente la tabella di input dei geodocument da convertire in geomaindocument (e.g. geolocationdb)
PARAM_DATABASE_OUTPUT_GEODOMAIN	nome del database contenente la tabella dei geodocument convertiti in geomaindocument (e.g. geolocationdb)
PARAM_DB_INPUT	nome database contenente la tabella di input per la realizzazione dei geodocument (e.g. urldb)
PARAM_DB_KEYWORD	nome database contenente la tabella document del sistema Ontostar (e.g. keyworddb)
PARAM_DB_OUTPUT	nome database contenente la tabella output per la realizzazione dei geodocument (e.g. geolocationdb)
PARAM_DIALECT_DATABASE	dialetto del database server utilizzato (e.g. jdbc:mysql)
PARAM_DRIVER_DATABASE	driver del database (e.g. com.mysql.jdbc.Driver)
PARAM_DS_DIR	path alla directory per la creazione e gestione di un DataStore con GATE (solo per il processo numero 3) (e.g. gate_files/Datastore)
PARAM_ERASE	se il parametro PARAM_CREA_NUOVA_TABELLA_GEODOCUMENT è settato a TRUE, il seguente parametro se anchesso a TRUE cancella la tabella PARAM_TABLE_OUTPUT con il medesimo nome se presente nel server database
PARAM_ERASE_GEODOMAIN	se il parametro PARAM_CREA_NUOVA_TABELLA_GEODOMAIN è settato a TRUE, il seguente parametro se anchesso a TRUE cancella la tabella PARAM_TABLE_OUTPUT_GEODOMAIN con il medesimo nome se presente nel server database
PARAM_ERASE_ONTOLOGY	se il parametro PARAM_CREA_NUOVA_TABELLA_INFODOCUMENT_ONTOLOGY è settato a TRUE, il seguente parametro se anchesso a TRUE cancella la tabella PARAM_TABLE_OUTPUT_ONTOLOGY con il medesimo nome se presente nel server database
PARAM_FILE_MAP_TURTLE_KARMA	file turtle (.ttl) di configurazione delle triple infodocument per il programma WEB-KARMA (e.g. R2RML_infodocument-model.ttl)

PARAM_FILE_OUTPUT_TRIPLE_KARMA	file di output delle triple generato da WEB-KARMA (e.g. triple.n3)
PARAM_FILTER	se TRUE previene l'analisi di documenti web di grosse dimensioni che farebbero andare in "out of memory for heap space " il programma d'obbligo per computer con meno di 4GB di RAM
PARAM_FREQUENZA_URL_GEODOMAIN	identifica in fase di creazione dei geodomaindocument la quantità di url per singole pagine che vogliamo impostare come limite per il medesimo dominio (e.g. 20)
PARAM_GENERATION_TRIPLE_KARMA_PROGRAMM	se TRUE fa partire il programma di creazione delle triple con WEB-KARMA
PARAM_GEODOMAIN_PROGRAMM	se TRUE fa partire il programma di creazione dei GeoDomainDocument
PARAM_HOST_DATABASE	identifica l'host del database server (e.g. localhost)
PARAM_ID_DATABASE_KARMA	codice di WEB-KARMA che identifica un databse come contenitore delle tabelle SQL da analizzare con il codice è "DB" che idientifica un database SQL generale
PARAM_INDGDOC	parametro che conserva l'indice dell'ultimo GATE Document salvato nel GATE datastore (solo per il processo numero 3) (e.g. valore di partenza zero)
PARAM_KARMA_HOME	parametro identifica la variabile di ambiente (in windows), per utilizzare la propria versione di Web-Karma invece che quella locale allegata con il progetto (e.g. KARMA_HOME)
PARAM_LIMIT	parametro che identifica la variabile LIMIT nella query SQL di estrazione degli url dalla PARAM_COLUMN_TABLE_INPUT
PARAM_LIMIT_GEODOMAIN	parametro che identifica la variabile LIMIT nella query SQL di estrazione dei geodocument dalla PARAM_TABLE_OUTPUT in pratica il numero di geodocument che vogliamo analizzare
PARAM_LOG_FILE	nome del file log in cui sono memorizzate tutte le azioni del programma è generato in automatico (e.g. LOG_20150215_181757.txt)
PARAM_NOME_DATASTORE	nome del datastore su cui vengono salvatii vari GATE Document generati dal programma (solo per il processo numero 3) (e.g. datastore_001)
PARAM_OFFSET	parametro che identifica la variabile OFFSET nella query SQL di estrazione degli url dalla PARAM_COLUMN_TABLE_INPUT
PARAM_OFFSET_GEODOMAIN	parametro che identifica la variabile OFFSET nella query SQL di estrazione dei geodocument dalla PARAM_TABLE_OUTPUT (e.g. 0)
PARAM_ONTOLOGY_PROGRAMM	se TRUE fa partire il programma di creazione degli infodocument
PARAM_OUTPUT_FORMAT_KARMA	parametro che identifica il formato delle triple generate con karma (e.g. ttl,n3,nt)
PARAM_PASS	password del database server su cui si sta lavorando (e.g. root)

PARAM_PASS_KARMA	password del database server su cui si sta lavorando (e.g. root) non l'abbiamo inglobata con la precedente variabile per motivi di sicurezza
PARAM_PATH_FILE_CFG_DB_INPUT_HIBERNATE	path al file di configurazione di Hibernate per il database PARAM_DB_INPUT (e.g. path\to\file\PARAM_DB_INPUT\hibernate.cfg.xml)
PARAM_PATH_FILE_CFG_DB_INPUT_KEYWORD	path al file di configurazione di Hibernate per il database PARAM_DB_KEYWORD (e.g. path\to\file\PARAM_DB_KEYWORD\hibernate.cfg.xml)
PARAM_PATH_FILE_CFG_DB_OUTPUT_HIBERNATE	path al file di configurazione di Hibernate per il database PARAM_DB_OUTPUT (e.g. path\to\file\PARAM_DB_OUTPUT\hibernate.cfg.xml)
PARAM_PATH_FILE_TABLE_INPUT_HIBERNATE	path al file di configurazione della tabella di input PARAM_TABLE_INPUT per hibernate (e.g. path\to\file\PARAM_DB_INPUT\website.hbm.xml)
PARAM_PATH_FILE_TABLE_OUTPUT_GEODOMAIN_HIBERNATE	path al file di configurazione della tabella di input PARAM_TABLE_INPUT_GEODOMAIN per hibernate (e.g. path\to\file\PARAM_DATABASE_OUTPUT_GEODOMAIN\geodomaindocument.hbm.xml)
PARAM_PATH_FILE_TABLE_OUTPUT_HIBERNATE	path\to\file\PARAM_DB_OUTPUT\geodocument.hbm.xml)
PARAM_PATH_FILE_TABLE_OUTPUT_KEYWORD_DOCUMENT	path al file di configurazione della tabella di output PARAM_TABLE_KEYWORD_DOCUMENT per hibernate (e.g. path\to\file\PARAM_DB_KEYWORD\document.hbm.xml)
PARAM_PATH_FILE_TABLE_OUTPUT_ONTOLOGY_HIBERNATE	path al file di configurazione della tabella di output PARAM_TABLE_OUTPUT_ONTOLOGY per hibernate (e.g. path\to\file\PARAM_DB_OUTPUT\infodocument.hbm.xml)
PARAM_PORT_DATABASE	numero della porta di comunicazione con l'host PARAM_HOST_DATABASE
PARAM_PROCESS_PROGAMM	1) ESTRAZIONE DEGLI INFODOCUMENT CON GATE E CON CORPUS PER SINGOLE PAGINE WEB 2) ESTRAZIONE DEGLI INFODOCUMENTI CON GATE E CON CORPUS PER STACK DI PAGINE WEB
PARAM_RANGE	effettivamente fatti girare dal programma è utile per sapere quanti geodocumenti sono stati salvati nel
PARAM_TABLE_INPUT	nome della tabella di input da cui andiamo a estrarre la lista degli url da analizzare (e.g. website)
PARAM_TABLE_INPUT_GEODOMAIN	nome della tabella di input per la creazione dei geodomaindocument da cui andiamo a estrarre la lista dei geodocument da analizzare (e.g. geodocument)
PARAM_TABLE_INPUT_KARMA	nome della tabella di input per il programma WEB-KARMA da cui andiamo a estrarre la lista degli infodocument da analizzare (e.g. infodocument)
PARAM_TABLE_KEYWORD_DOCUMENT	nome della tabella del database da cui andiamo a estrarre la lista dei document da analizzare per arricchire il risultato dei nostri geodocument (e.g. document)
PARAM_TABLE_OUTPUT	nome della tabella di output su cui andiamo a salvare la lista dei geodocument generati dal programma (e.g. geodocument)

PARAM_TABLE_OUTPUT_GEODOMAIN	nome della tabella di output su cui andiamo a salvare la lista dei geodomaindocument generati dal programma PARAM_GEODOMAIN_PROGRAMM (e.g. geodomaindocument)
PARAM_TABLE_OUTPUT_ONTOLOGY	nome della tabella di output su cui andiamo a salvare la lista degli infodocument generati dal programma PARAM_GEODOMAIN_PROGRAMM (e.g. geodomaindocument)
PARAM_TENTATIVI_OUT_OF_MEMORY	parametro che identifica il numero di tentativi che è possibile cercare di fare per un set di documenti salvati in un datastore prima di rinunciare (necessario solo per il processo 3)
PARAM_TYPE_DATABASE_KARMA	codice che identifica il tipo di database utilizzato per WEB-KARMA (e.g MySQL)
PARAM_TYPE_EXTRACTION	il tipo di estrazione utilizzato cioè se ci appoggiamo a HIBERNATE o al driver java MYSQL (e.g HIBERNATE oppure MYSQL)
PARAM_USER	parametro user del database server (e.g. user)
PARAM_USER_KARMA	parametro user del database server per il programma WEB-KARMA (e.g. user) non è stato inglobato con il precedente parametro per motivi di sicurezza
PARAM_DIALECT_DATABASE_HIBERNATE	parametro che identifica il tipo di driver del database utilizzato per funzionare con il particolare linguaggio di Hibernate (HQL) (e.g. org.hibernate.dialect.MySQLDialect)