

АГриЧака – Ваше решение, когда чаты в кармане

Сбор, обработка, аналитика, статистика общения и подготовка для
обучения, нейросети и возможное предсказание поведения

Шуруповёрт вместо отвёртки: собери свои беседы, чаты, переписки и исследуй их вместе с
аналитикой и ИИ.





Краткое содержание:

Вводная часть

Общая архитектура, основные процессы

Основные составляющие проекта

Реализованный модуль

Patroni

PostgreSQL

ETCD

Backup

Схема БД

Основные функции и возможности

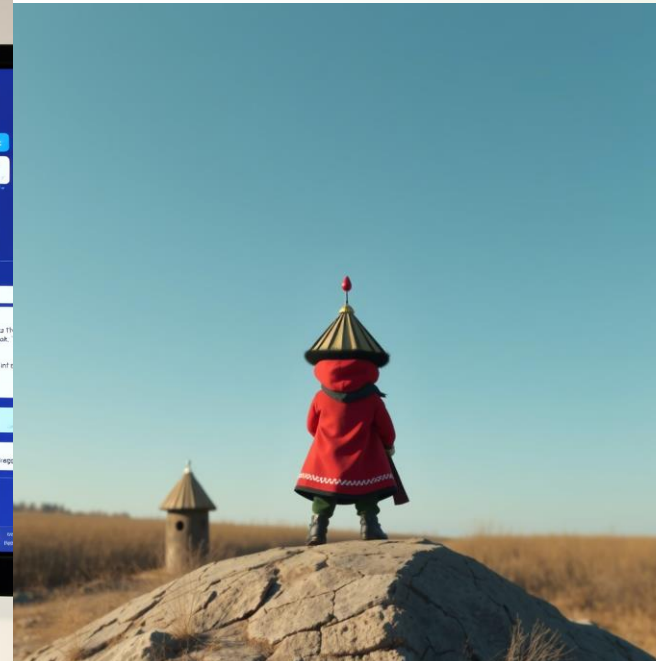
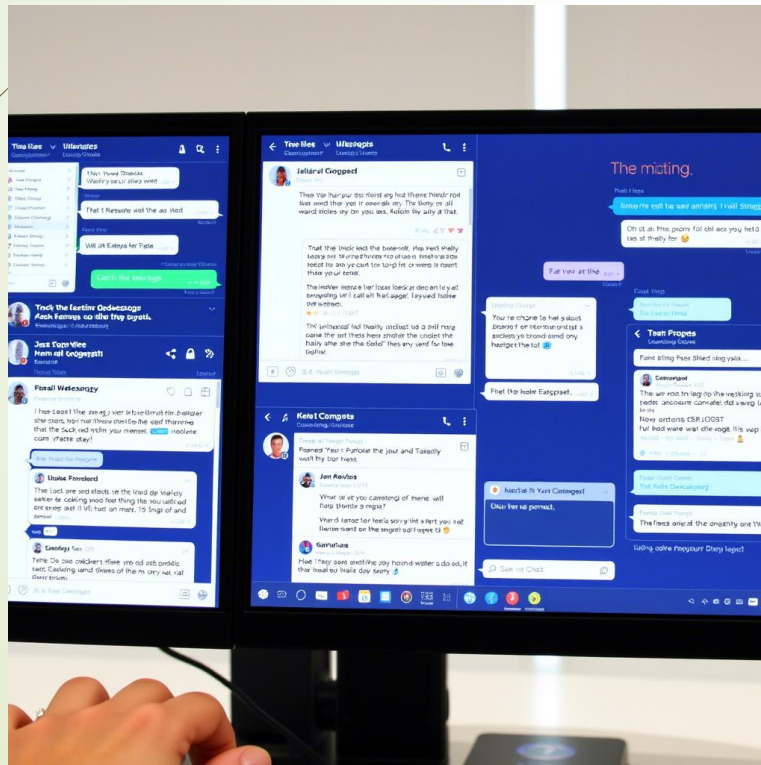
Перспективы развития, гипотезы и возможные решения

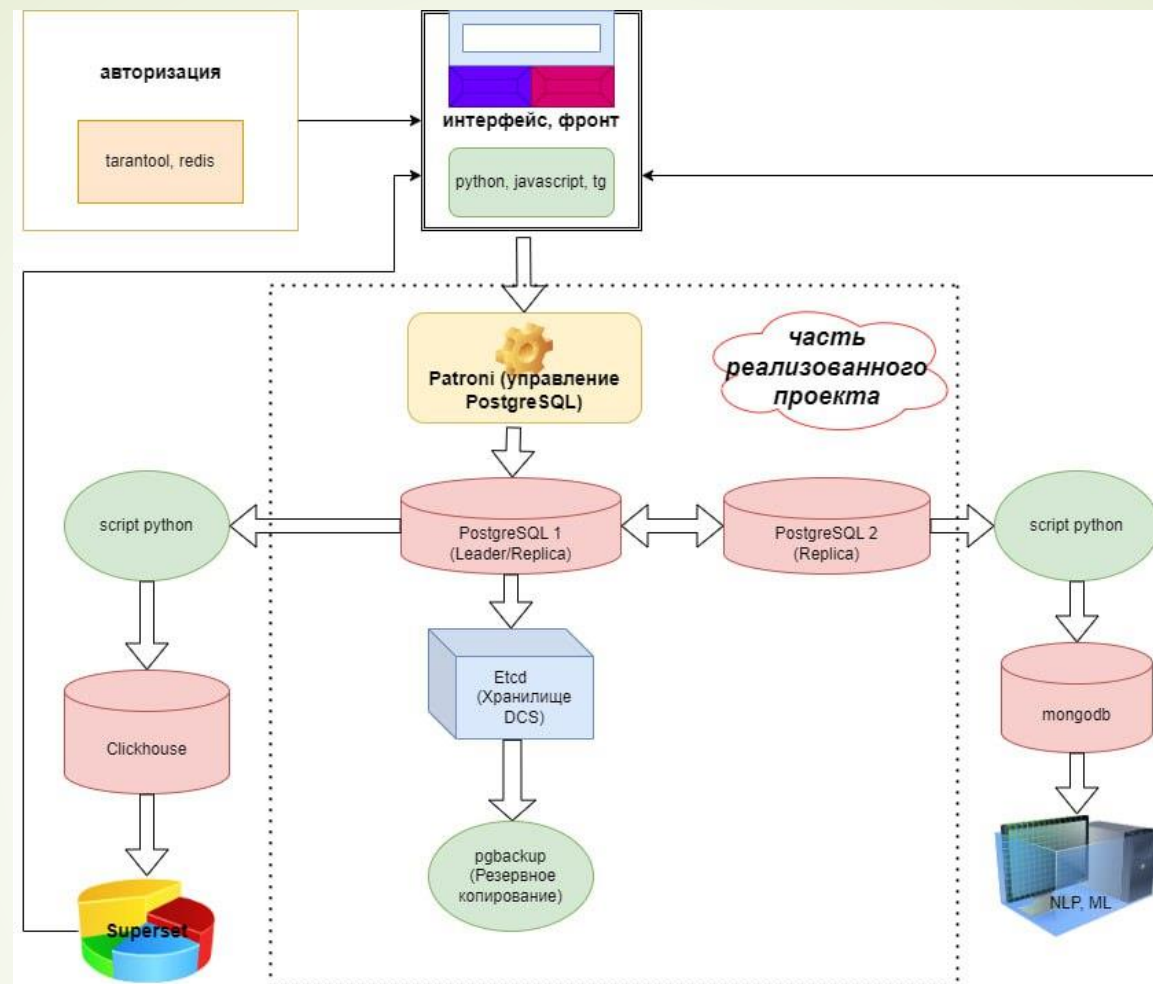
Рефлексия

Вводная

Сегодня мы окружены тысячами чатов, сообщений, переписок, бесед. Точно знаем, что в этом ворохе повседневных деловых и не очень деловых бесед, под серостью рабочих будней скрыто золото информации, алмазы решений, изумруды алгоритмов. Надо всего лишь отыскать, раскопать. Много ли времени есть.

С помощью инструмента «АГриЧака» Вы можете увереннее ориентироваться в ворохе чатов и груде переписок выделяя для себя главное и отсеивая ненужное, ускоряя работу в несколько раз.





Основные составляющие проекта, модуль загрузки

Прототип прототипа





Реализованный модуль

High-Availability PostgreSQL кластер с Patroni

Архитектура решения

[Docker Host]

- etcd (хранилище конфигурации)
- Patroni 1 (PostgreSQL Master)
- Patroni 2 (PostgreSQL Replica)
- pgbackup (сервис бэкапов)

Компоненты:

etcd: Координация кластера

Patroni: Управление PostgreSQL (+авто-фейловер)

pgbackup: Ежедневные бэкапы

Настройка Patroni

```
scope: my_cluster
restapi:
  listen: 0.0.0.0:8008
etcd:
  host: etcd:2379
postgresql:
  pg_hba:
    - host all all 0.0.0.0/0 md5
```

Преимущества:

Автоматическое переключение при сбоях
Централизованное управление через etcd

Система бэкапов

Схема работы:
Скрипт backup-script.sh запускается по cron
pg_dump создает бэкап
Файлы сохраняются в ./backups

Пример команды:

```
pg_dump -h patroni1 -p 5432 -U postgres -F c -f /backups/backup_$(date +"%Y%m%d").sql
```



Docker-реализация

services:

patroni1:

image: registry.opensource.zalan.do/acid/spilo-14:2.1-p7

environment:

PATRONI_POSTGRESQL_DATA_DIR: /var/lib/postgresql/data/pgdata

pgbackup:

image: postgres:14

volumes:

- ./backups:/backups

Преимущества:

Изоляция сервисов

Простота развертывания

Проверка работы

Команды для мониторинга:

Статус кластера

curl http://localhost:8008

Проверка репликации

docker exec -it patroni1 psql -c "SELECT * FROM pg_stat_replication;"

Просмотр бэкапов

ls -lh ./backups



Дополнительные возможности

Расширения системы:

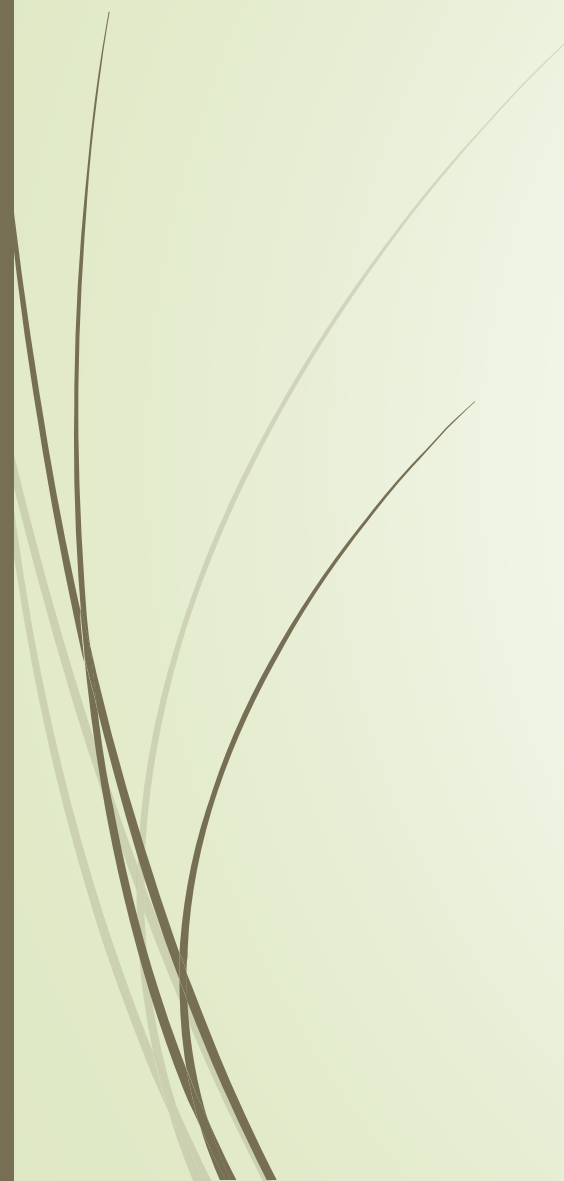
- Добавьте мониторинг (Prometheus + Grafana)
- Настройте алертинг при сбоях
- Интегрируйте с S3 (для хранения бэкапов)

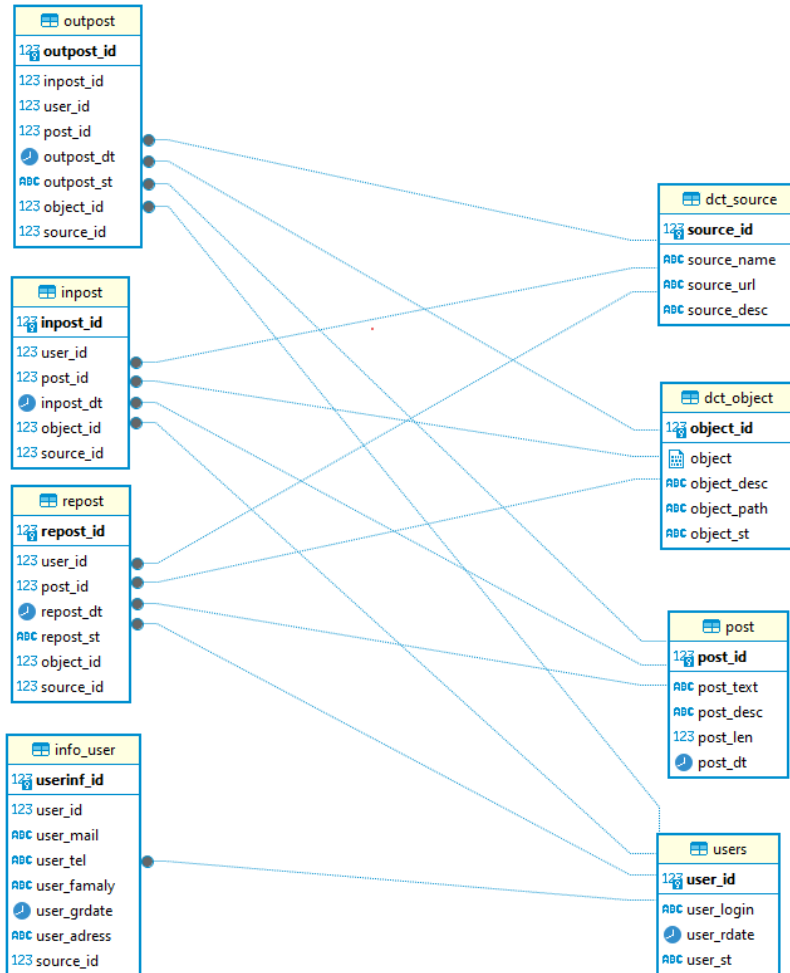
Статистика:

- Доступность 99.99%
- Время восстановления < 30 сек

Итоги

Что достигнуто:

- Отказоустойчивый PostgreSQL-кластер
 - Автоматические бэкапы
 - Простое управление через Docker
- 



33

--создание информационной схемы

34

--DROP schema IF EXISTS info;

35

CREATE SCHEMA info AUTHORIZATION user_info;

36

--создание целевой схемы

37

DROP schema IF EXISTS calc;

38

CREATE SCHEMA calc AUTHORIZATION user_admin;

39

--создание целевой схемы

40

DROP schema IF EXISTS sapu;

41

CREATE SCHEMA sapu AUTHORIZATION user_sapu;

42

--создание схемы расширений

43

CREATE SCHEMA extensions;

44

45

46

47

--создание таблицы в нужном нам табличном пространстве

48

DROP TABLE IF EXISTS info.info_users ;

49

CREATE TABLE info.info_users(users_id serial PRIMARY KEY, users_text jsonb) ;--TABLESPACE info_space;

50

51

--создание последовательности для таблицы user

52

CREATE SEQUENCE sapu.seq_user_id

53

START WITH 1

54

INCREMENT BY 1

55

NO MINVALUE

56

NO MAXVALUE

57

CACHE 1;

58

--создание таблицы пользователей

59

CREATE TABLE sapu.users (

60

user_id bigint DEFAULT nextval('sapu.seq_user_id'::regclass) NOT NULL PRIMARY key,

61

user_login character varying(50) NOT NULL,

62

user_rdate timestamp without time zone DEFAULT now() NOT NULL,

63

user_st varchar(50) DEFAULT 'ACTIVE' NOT NULL

64

);

65

66

--создание последовательности для идентификатора таблиц информации о пользователях

67

CREATE SEQUENCE sapu.seq_userinf_id

68

START WITH 1

69

INCREMENT BY 1

70

NO MINVALUE

71

NO MAXVALUE

72

CACHE 1;

73

--создание таблицы информации о пользователях

74

CREATE TABLE sapu.info_user(

75

Сохраняет данные в базу данных

76

connection = psycopg2.connect(**DB_CONFIG)

77

cursor = connection.cursor()

78

79

insert_query = """

80

call sapu.create_inpost(%s, %s, 1, %s::timestamp);

81

"""

82

83

for message in messages:

84

cursor.execute(insert_query, (message['author'], message['text'],message

85

86

87

connection.commit()

88

cursor.close()

89

connection.close()

90

91

if __name__ == '__main__':

92

Путь к папке с HTML-файлами

93

folder_path = r'C:\Python\messages'

94

95

Получаем список всех HTML-файлов в папке

96

html_files = [f for f in os.listdir(folder_path) if f.endswith('.html')]

97

98

total_messages_saved = 0

99

for html_file in html_files:

100

file_path = os.path.join(folder_path, html_file)

101

102

Парсим файл

103

messages = parse_html(file_path)

104

105

Сохраняем данные в базу данных

106

save_to_db(messages)

107

108

Подсчитываем общее количество сохраненных сообщений

109

total_messages_saved += len(messages)

110

print(f'Файл {html_file} обработан. Сохранено {len(messages)} сообщений.

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

Язык: [Русский]

PostgreSQL » postgres » otus » sapu » Выбрать: users

Adminer 5.0.6 5.1.0

DB: [otus]

Схема: [sapu]

SQL-запрос

Импорт

Экспорт

Создать таблицу

выбрать dct_object

выбрать dct_source

выбрать info_user

выбрать inpost

выбрать outpost

выбрать post

выбрать repost

выбрать users

Выбрать: users

Выбрать

Показать структуру

Изменить таблицу

Новая запись

Выбрать

Поиск

Сортировать

Лимит

Длина текста

Действие

SELECT * FROM "users" LIMIT 50 OFFSET 750 (0.001 s) Редактировать

Изменить

user_id

user_login

user_rdate

user_st

редактировать

751

Natasha

2025-03-26 08:44:07.544089

ACTIVE

редактировать

752

Vitali Efimov

2025-03-26 08:44:07.544089

ACTIVE

редактировать

753

Rumila Sinelnikova

2025-03-26 08:44:07.544089

ACTIVE

редактировать

754

R.

2025-03-26 08:44:07.544089

ACTIVE

редактировать

755

Святослав Смирнов

2025-03-26 08:44:07.544089

ACTIVE

редактировать

756

Roman Solomatn

2025-03-26 08:44:07.544089

ACTIVE

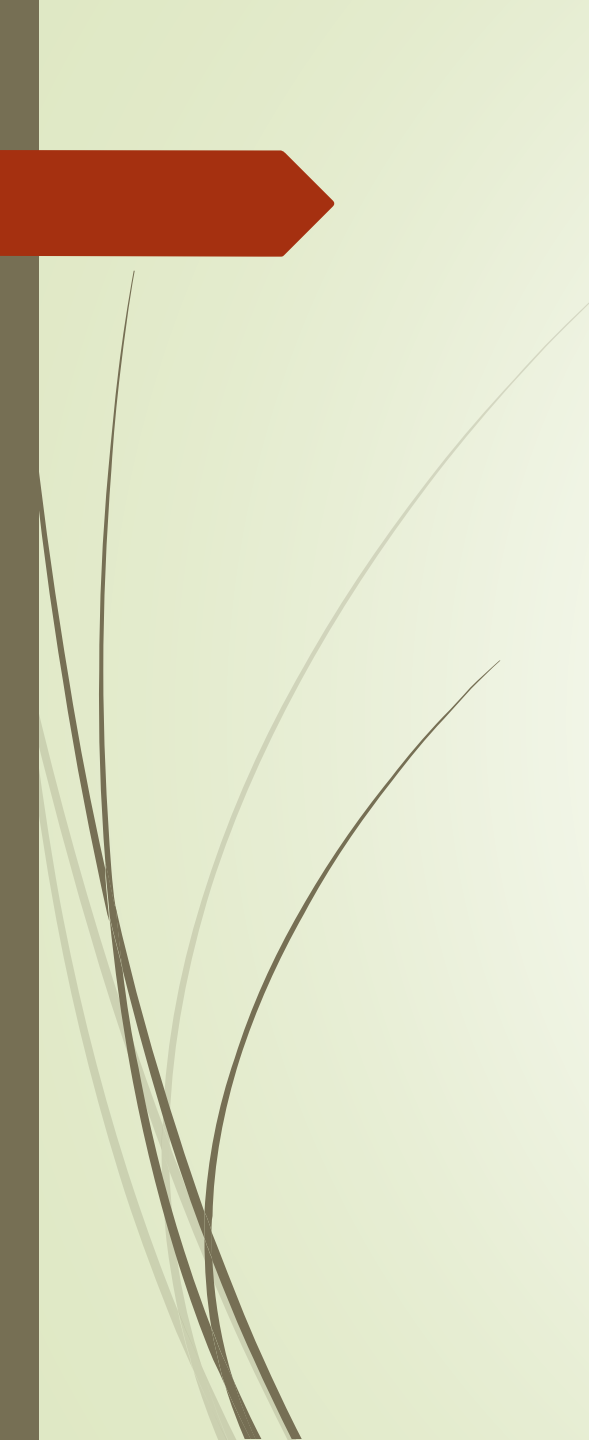
редактировать

757

Alexander Polyakov

2025-03-26 08:44:07.544089

ACTIVE



Данный программный комплекс возможно создать на микросервисной архитектуре, что по сути и преследуется в данном проекте.

В дальнейшем планируется рассмотреть возможность использования Kubernetes, подключение средств мониторинга работы, добавление скриптов для взаимодействия (API) между модулями (сервисами), обучение в коробочном варианте AI-моделей, разработке модели анализа и прогнозирования поведения собеседника, сбора и предоставления статистики, гипотетически получать так называемый «почерк автора».

К примеру, в процессе создания проекта возникло желание местами использовать MySQL, как бы я не любил PostgreSQL. Интересные решения приходили в голову с использованием Cassandra и MongoDB, ClickHouse как таковой включен в архитектуру, но можно было бы «поиграться», GreenPlum хотелось применить, но на другом уровне.

Проект является прототипом, но даже на текущем уровне решает часть затрат во всяком случае в моём списке задач. Курс позволил изучить помимо тем курса часть возможностей докер, с которым был всего лишь знаком на словах, bash-скрипт, Linux, Markdown и некоторые возможности инструментов программных средств.