

# Загрузка АТД Индии в БД NSI (вер. 1.2)

## Оглавление

Назначение .....	1
Установка PostgreSQL и PostGIS .....	1
Восстановление базы NSI .....	2
Исходные данные .....	2
Схема базы NSI .....	3
ETL-процесс .....	4
SQL скрипты .....	5
Приложение .....	6

## Назначение

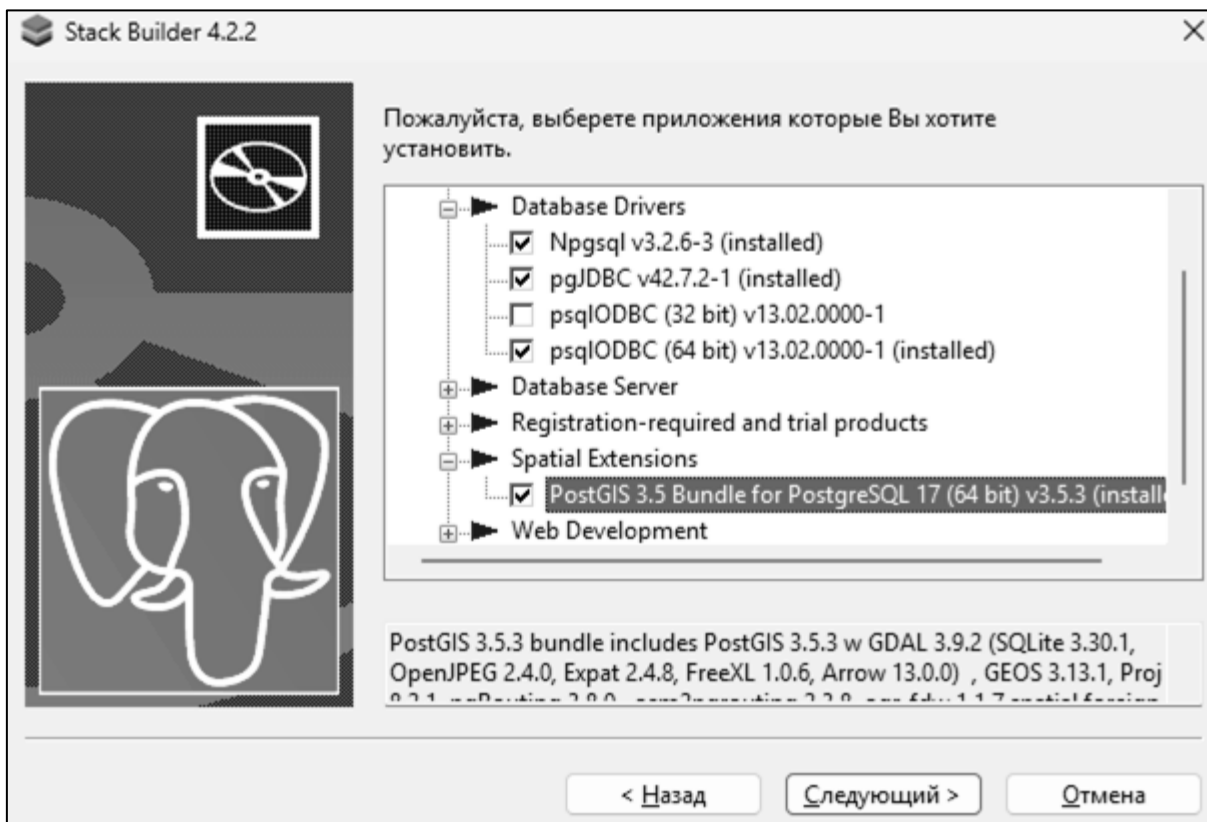
\*Загрузка данных административно-территориального деления Индии в базу данных NSI. Процесс охватывает все уровни административной иерархии от штатов до деревень с установкой соответствующих связей между единицами АТД.

## Установка PostgreSQL и PostGIS

\*Процесс установки описан в прилагаемом документе «ФПМИ МФТИ - SQL для прикладных задач - Установка Postgre.pdf». Раздел по установке DBeaver можно пропустить - перейти непосредственно к установке PostGIS. При установке PostGIS рекомендуется в Stack Builder выбрать компоненты: Spatial Extensions → PostGIS (последняя версия), Database Drivers → npgsql, pgjdbc, psqldb 64 bit. При установке PostGIS отметить опцию «Enable All GDAL Drivers» для работы с форматом .tif.

<https://www.enterprisedb.com/downloads/postgres-postgresql-downloads> -download link

Рисунок 1 -Stack Builder установка PostGIS



## Восстановление базы NSI

\*Восстановление данных из файла nsi.dump. Процесс включает создание базы данных NSI с пользователем postgres. После успешного создания базы NSI необходимо выполнить команду через Query Tools:

```
CREATE EXTENSION postgis; -- активация расширения PostGIS в базе данных PostgreSQL
```

\*\*Далее необходимо восстановить данные через Restore с форматом Custom or Tar, указанием пути к файлу nsi.dump и имени роли postgres. В разделе Data Options отметить опции несохранения (Do not save) владельца и прав (Owner and Privileges).

переходим в папку postgresql:  
cd "C:\Program Files\PostgreSQL\17\bin"

смотрим содержимое дампа:  
.\pg\_restore -l "C:\backup\nsi.dump"

создание пользователя nsi:  
.\psql -U postgres

внутри psql:  
CREATE ROLE nsi LOGIN PASSWORD 'любой пароль';  
\q

создание базы данных:  
.\createdb -U postgres -E UTF8 nsi

восстановить базу из дампа:  
.\pg\_restore -U postgres -d nsi -v "C:\backup\nsi.dump"

создание резервной копии базы после внесённых изменений:  
.\pg\_dump -U postgres -F c -b -v -f "C:\backup\nsi.dump" nsi

## Исходные данные

### Структура исходных данных

\*Структура исходных данных основана на документации «Описание полей векторной БД административно-территориального деления Республики Индия (версия 1.1)». Данные загружаются в формате SHP в системе координат WGS 84 (EPSG:4326).

**Загрузка исходных данных в схему public**

\*Для загрузки данных использовались утилиты PostGIS согласно руководству "PostGIS 2.0 postgresql2shp shp2pgsql Command Line Cheatsheet". Процесс выполняется через terminal или powershell с указанием пути к файлам SHP и схеме public.

```
cd "C:\Program Files\PostgreSQL\17\bin" -- смена директории на папку с утилитами
.\shp2pgsql -I -s 4326 -D
"C:\Путь_к_shp_файлу\India_states_20250919.shp"
public.India_states_20250919 | .\psql -U postgres -d nsi --
загрузка shp штатов в схему public с созданием пространственного
индекса в системе координат WGS 84*
/*
I - создание пространственного индекса
-s 4326 - система координат WGS 84
-D - формат дампа для быстрой загрузки
*/
```

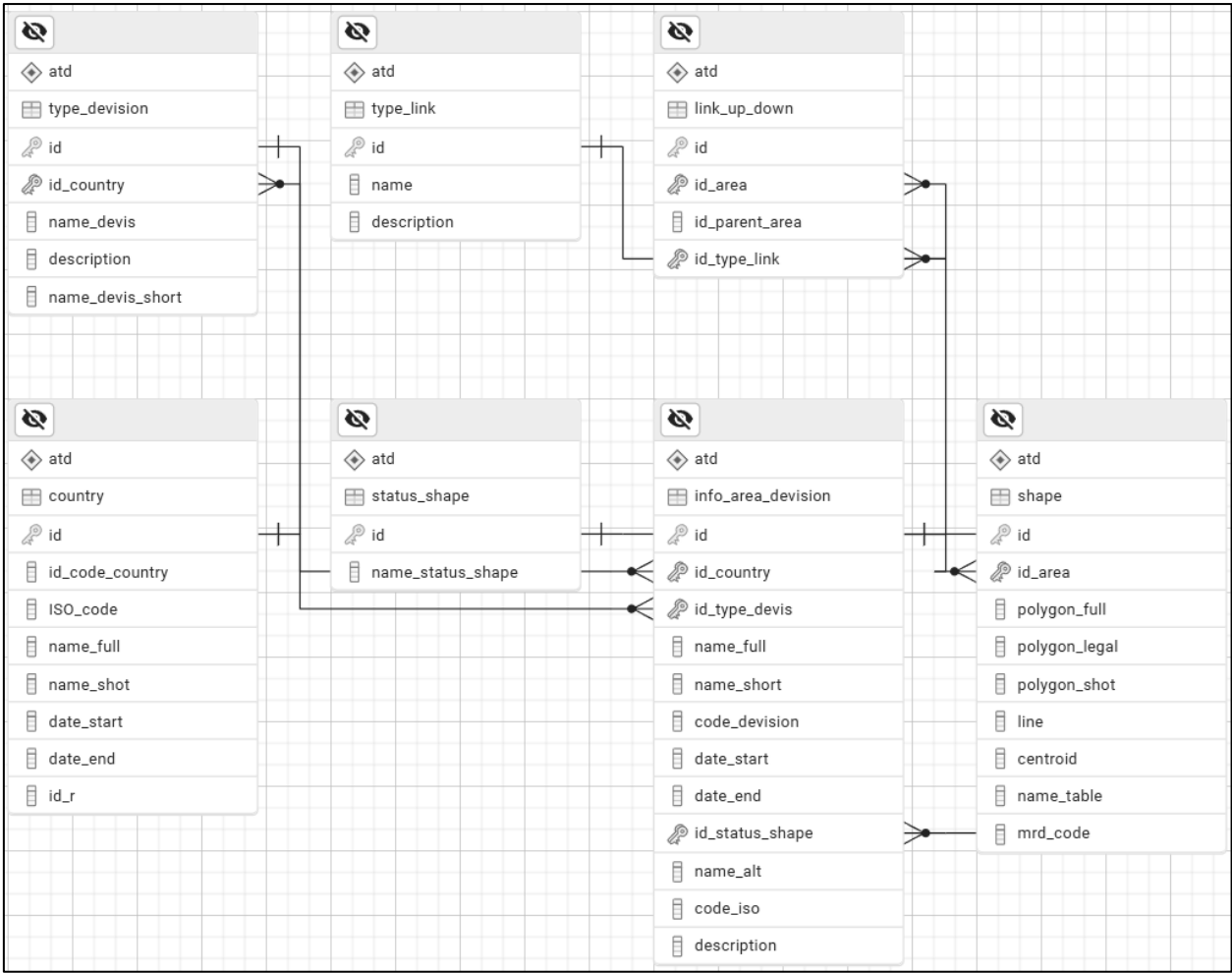
## Схема базы NSI

\*Исходные данные загружаются во временные таблицы в схеме public, затем преобразуются и связываются в схеме atd.

*Таблица 1 - Основные таблицы ATD*

Основные таблицы схемы atd	
Таблица	Назначение
info_area_devision	Административные единицы
shape	Геометрические данные
link_up_down	Иерархические связи
country, type_devision, status_shape, type_link	Дополнительная информация

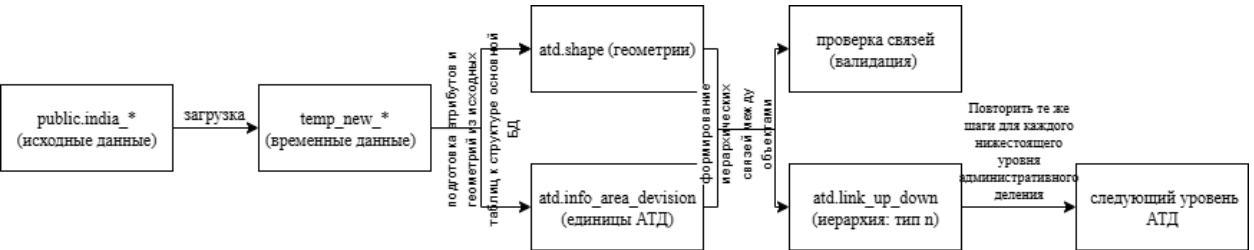
Рисунок 2 - ERD for Schema atd



## ETL-процесс

\*Преобразование данных из схемы public в схему atd. Процесс включает подготовку временных таблиц, обновление записей штатов и районов с установкой временных меток data\_start/date\_end для ведения истории изменений, обновление геометрических данных в таблице shape и установку иерархических связей между административными единицами в таблице link\_up\_down.

Рисунок 3 – ETL diagram



## SQL скрипты

\*Основные SQL запросы для преобразования данных и установки связей между административными единицами. Скрипты выполняются в строгом порядке для обеспечения корректности иерархических связей. Перед выполнением необходимо ознакомиться с особенностями загрузки городских и лесных административных единиц (см. ниже) и с приложением.

### Порядок выполнения скриптов:

1. 00\_load\_state.sql – загрузка штатов
2. 01\_load\_district.sql – загрузка районов
3. 02\_load\_tehsil.sql – загрузка техсил
4. 03\_load\_block.sql – загрузка блоков
5. 04\_load\_gp.sql – загрузка панчаятов
6. 05\_load\_village.sql – загрузка деревень
7. 06\_load\_urban.sql – загрузка городов

\*Для каждого уровня административного деления используется соответствующий id\_type\_link из таблицы type\_link. Логика обработки данных одинакова для всех уровней: подготовка временных таблиц, закрытие старых записей с установкой date\_end, вставка новых записей, обновление геометрий и установка иерархических связей.

### Особенности загрузки для городских и лесных единиц АД

\*Городские территории (City) загружаются как отдельный тип "Urban" (id\_type\_devis = 7) и связываются только с техсилами (id\_type\_link = 52) - используется существующий тип связи "Village up to tehsil", поскольку различие между Village и Urban отражено в поле id\_type\_devis.

- Если у города есть код, он используется без изменений
- Если кода нет, формируется новый по шаблону: код\_штата + код\_района + 001, 002, 003 и т. д.

\*\*Лесные территории (Forest) в систему не загружаются. На будущее в таблице создается новый тип type\_devis для "Forest and Lands" (id = 8).

```
-- добавляем тип urban для городских территорий
insert into atd.type_devis (id, id_country, name_devis,
description, name_devis_short)
values (7, 108, 'Urban', null, 'Urb');

-- добавляем тип forest and lands
insert into atd.type_devis (id, id_country, name_devis,
description, name_devis_short)
values (8, 108, 'Forest and Lands', null, 'For');
```

### Примеры применения:

- Город с кодом "123456" → 123456 + связь с техсилом (id\_type\_link = 52)
- Город без кода → 36696001 + связь с техсилом, где 36 - код штата Telangana, 696 - код района Suryapet, 001, 002, 003... - последовательные номера

\*Городские территории связываются только с техсилами. Обычные деревни связываются с панчаятами и техсилами. Деревни без кода панчаята связываются только с техсилом. Поле LGDVilCode в исходных данных не может быть пустым и не должно содержать дубликатов.

\*\*Данный подход сохраняет существующую структуру БД и обеспечивает корректную классификацию всех типов территориальных единиц.

## Приложение

\*В процессе загрузки новых данных возможны изменения или совпадения кодов административно-территориальных единиц, публикуемых в официальных источниках LGD (lgdirectory.gov.in). Такие ситуации требуют проверки актуальности информации по LGD и корректировки устаревших кодов перед финальной загрузкой. Все изменения фиксируются в реестре.

### Порядок действий:

1. Выявление совпадений кодов при загрузке новых данных
2. Проверка актуальности кодов через LGD (lgdirectory.gov.in)
3. Корректировка устаревших кодов
4. Документирование изменений в реестре
5. Повторная загрузка новых АТД в БД NSI

### Реестр изменений кодов АТД

```
-- Haryana/Bhiwani/Paluwas - обновление кода перед загрузкой  
Telangana  
update atd.info_area_devision set code_devision = '61157' where  
id = 14609 and code_devision = '576654';
```

*Реестр дополняется по мере выявления изменений в LGD.*