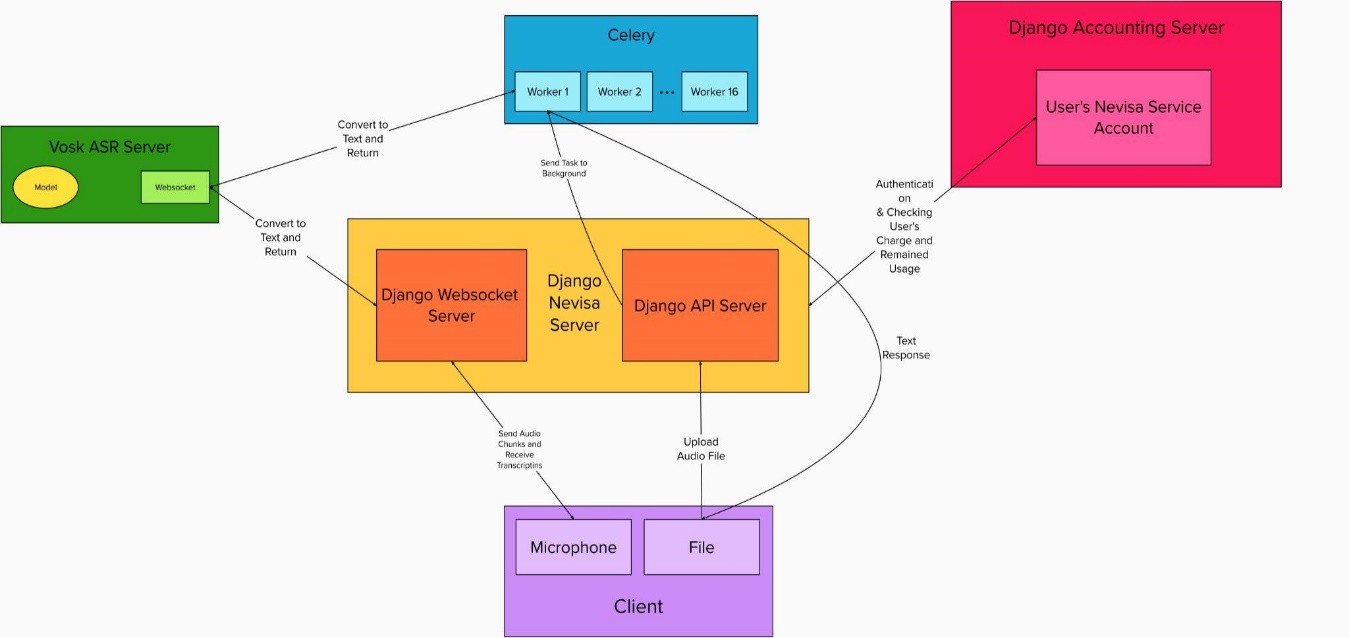
# معماری نویسا

معماری سامانه نویسا به صورت سرویس‌دهنده-سرویس‌گیرنده طراحی شده است. سرویس‌دهنده‌ها شامل موارد زیر است:

1. VOSK-ASR Server (هسته پردازشی تبدیل گفتار به متن)
2. Django Accounting Server (سرویس مدیریت حساب کاربران)
3. Django Nevisa Server (سرویس دهنده نویسا)
4. Celery (برای انجام تسک‌های بک‌گراند)
5. Database Server (سرویس دهنده پایگاه داده)

این سرویس دهنده‌ها روی یک سرور در قالب ماشین فیزیکی یا مجازی ارائه می‌شود. دیاگرام نحوه ارتباط سرویس‌دهنده‌ها در شکل 1 نشان داده شده است.



شکل 1- ارتباط سرویس‌دهنده‌ها

سرویس‌گیرنده نویسا به دو صورت ارسال داده میکروفونی (از طریق وب سوکت) و ارسال فایل (از طریق API نویسا) با سرویس‌دهنده ارتباط برقرار می‌کند:

1. **ارتباط از طریق میکروفون:**

* در این قسمت صدای کاربر از میکروفون ضبط شده و از طریق ارتباط وب‌سوکت به سرور نویسا ارسال می‌شوند.
* سرور نویسا از طریق ارتباط با سرور اکانتینگ احراز هویت کاربر را انجام داده و در صورتی که کاربر اعتبار کافی برای استفاده از سرویس نویسا داشته باشد، شروع به انجام عملیات تبدیل صوت به متن می‌کند.
* در این روش قطعات صوتی در وب‌سوکت نویسا از کلاینت دریافت می‌شوند و مجددا از طریق یک ارتباط وب‌سوکتی دیگر برای سرور ASR ارسال شده و در پاسخ، متن معادل هر قطعه صوتی برای کاربر ارسال می‌شود. سرور ASR مدل Vosk را در حافظه‌ی خود بارگذاری کرده و با استفاده از آن تبدیل صوت به متن را انجام می‌دهد.

1. **ارتباط از طریق API نویسا**

* در این بخش، فایل صوتی کاربر (فارغ از اینکه در چه فرمت فایل صوتی قرار دارد) در قالب یک درخواست POST برای API تبدیل فایل ارسال می‌شود.
* پس از اینکه سرور نویسا (دقیقا مانند حالت میکروفون) احراز هویت کاربر را با اکانتینگ انجام داد، عملیات تبدیل فایل به متن آغاز می‌شود.
* از آنجایی که تبدیل فایل‌های طولانی می‌تواند زمان نسبتا زیادی ببرد، ما این عملیات را در background و به کمک Celery انجام می‌دهیم که ابزاری است برای اجرای تسک‌ها در background و به خوبی با Django هماهنگ می‌شود. به تعداد هسته‌های پردازشی موجود می‌توانیم Celery worker داشته باشیم که این worker ها به صورت موازی کار تبدیل فایل‌ها را انجام می‌دهند.
* پس از اینکه تسک مربوط به هر فایل به صورت خودکار به worker خاصی داده شد، مانند حالت ارتباط میکروفونی به کمک ارتباط وب‌سوکتی با سرور ASR و مدل Vosk کار تبدیل فایل صوتی به متن انجام داده می‌شود.
* در زمان تبدیل فایل ارسالی، کاربر می‌تواند از طریق URL ای که در پاسخ درخواستش دریافت کرده میزان پیشرفت تسک خود را در هر لحظه مشاهده کند و در صورت اتمام تسک، فایل متنی نهایی را دانلود کند.
* کاربر با فعالسازی گزینه‌ی "ذخیره در تاریخچه" امکان ذخیره‌سازی فایل‌های متنی تبدیل یافته‌ی خود را در سرور دارد تا بتواند بعدا به آنها رجوع کند. با فعالسازی این گزینه کاربر می‌تواند پس از آپلود فایل صوتی از سامانه خارج شود و بعدا فایل متنی تبدیل شده را از قسمت تاریخچه دریافت کند.

# سرویس دهنده نویسا

برای استفاده از API های سرور نویسا هر API به انتهای آدرس IP سرور اضافه می‌شود. به عنوان مثال اگر آدرس IP سرور 192.168.1.100 باشد، URL برای API ورود به صورت زیر خواهد بود.

https://192.168.1.100/account/login

## ورود

POST: /account/login

با استفاده از این API می‌توانید با ارسال نام‌کاربری خود به همراه رمز عبور حساب کاربری خود، token و api-key خود را دریافت کنید. شما برای استفاده از API های دیگر به این دو مقدار نیاز خواهید داشت. در نسخه سازمانی شماره تلفن و ایمیل می‌تواند خالی باشد.

Request:

{  
 'username\_or\_phone\_or\_email': "uuupppeee",   
 'password': "xxxxxx",  
}

Response (200):

{  
 'message': "Successfully logged in.",  
 'user': {  
 'id': "user's UUID",  
 'token': "user's auth-token",  
 'username': "username",  
 'phone': "989123456789",  
 'email': "xxxxxx@yyyy.zzz",  
 'is\_staff': bool,  
 'is\_telegram\_user': bool,  
   
 'nevisa\_service\_account': {  
 'current\_service\_record': {  
 'key': "your api-key",  
 ...  
 },  
 ...  
 }  
 }  
}

**auth-token:**

response.data['user']['token']

**api-key:**

response.data['user']['nevisa\_service\_account']['current\_service\_record']['key']

**توجه:** مقدار **authentication token** با هر بار لاگین کردن تغییر می‌کند.

**توجه:** در صورتی که بسته‌ی فعال نداشته باشید، مقدار **current\_service\_record** برابر **null** خواهد بود.

## تشخیص فایل

POST: /recognize-file

حال با استفاده از auth-token و api-key خود که از Login دریافت کردید، می‌توانید فایل صوتی خود را (بدون نیاز به تبدیل فرمت فایل) به این API فرستاده و در پاسخ متن حاصل از تبدیل فایل صوتی را دریافت کنید.

**توجه:** فرمت request شما برای این API باید حتما به صورت **multipart/form-data** باشد. در غیر این صورت با خطا مواجه خواهید شد.

Request:

{  
 'auth\_token': "user's authentication token", # from login API  
 'api\_key': "user's api-key", # from login API  
 'file': <file>,  
}

Response (200):

{  
 'progress\_url': "/celery-progress/<task\_id>/",  
 'task\_id': "task\_id",  
}

**نکته:** می‌توانید با ارسال یک درخواست GET به progress\_url میزان پیشرفت یا progress عملیات تبدیل صوت به متن فایل ارسال شده‌ی خود را مشاهده کنید. برای جزئیات بیشتر به API میزان پیشرفت (Task Progress) رجوع کنید.

## میزان پیشرفت

GET: /celery-progress/<task\_id>/

با این API می‌توانید میزان پیشرفت یا progress عملیات تبدیل صوت به متن فایل ارسال شده را مشاهده کنید. برای این منظور از progress\_url مربوط به task مورد نظر که از API تشخیص فایل دریافت می‌شود، استفاده کنید.

همچنین هرگاه مقدار state در پاسخ برابر SUCCESS شد، می‌توانید متن حاصل از تبدیل فایل صوتی خود را در کلید result پیدا کنید.

Response [**Before the Final Result**]:

{  
 'state': "PROGRESS",  
 'complete': false,  
 'success': null,  
 'progress': {  
 'pending": false,  
 'current': 65, # less than total  
 'total': 100,  
 'percent': 65  
 },   
}

Response [**The Final Result**]:

{  
 'state': "SUCCESS",  
 'complete': true,  
 'success': true,  
 'progress': {  
 'pending": false,  
 'current': 100,   
 'total': 100,  
 'percent': 100  
 },   
 'result': {  
 'transcription': {  
 'result': [  
 {'conf': 1.0, 'end': 2.61, 'start': 0.0, 'word': "word"},  
 {...},  
 ...  
 ],  
 'text': "final and complete text"  
 },  
 'final': true  
 },  
}

**result:**

response.data['result']['transcription']['result'] = لیستی از تمام کلمات تشخیص داده شده به همراه زمان دقیق شروع و پایان بیان هر کلمه در فایل صوتی و میزان اطمینان مدل از تشخیص هر کلمه  
  
response.data['result']['transcription']['text'] = متن نهایی

## متوقف کردن پردازش

POST: /file/terminate-task

با استفاده از این API می‌توانید عملیات تبدیل صوتی خود را لغو کنید. برای این کار نیاز به مقدار task\_id دارید که از پاسخ File Recognition API دریافت کردید.

همچنین، برای این API باید مقدار auth-token خود را به شکل زیر در **header** درخواست خود قرار دهید:

{'Authorization': "Token <auth-token>"}

Request:

{  
 'auth\_token': "user's authentication token", # from login API  
 'api\_key': "user's api-key", # from login API  
 'task\_id': "task\_id",  
}

در صورتی که تسک با موفقیت لغو شود، پاسخ شامل کد status = 200 خواهد بود.