

# Dataset Documentation (Data Sheet + Dictionary)

**\*\*مقدمه و هدف ۴/۹/۱\*\*\***

را به طور کامل مستند `nobatnou.ir` این سند مجموعه داده‌های سیستم نوبت‌دهی می‌کند. هدف، شفاف‌سازی در مورد **\*\*منبع، کیفیت، مالکیت، محدودیت‌ها و ویژگی‌های\*\*** داده‌ها است تا اطمینان حاصل شود که داده‌ها به عنوان یک «دارایی» مدیریت شده و از تبدیل شدن به «بدهی» (به دلیل مشکلات قانونی، کیفیتی یا اخلاقی) جلوگیری شود.

---

**\*\* (Data Sheet) کارت مجموعه داده ۴/۹/۲\*\*\***

**\*\*اطلاعات کلی A.\*\*\***

| فیلد | مقدار |

| ----- |

| **\*\*نام مجموعه داده\*\*** | NobatNou Appointment System Dataset |

| **\*\*نگارش (Version)\*\*** | 1.0 |

| **\*\*تاریخ ایجاد\*\*** | ۱/۰۷/۱۴۰۲ |

| **\*\*تاریخ آخرین به‌روزرسانی\*\*** | ۲۰/۰۲/۱۴۰۳ |

| **\*\*مالک (Owner)\*\*** | `nobatnou.ir` مالک شرکت نوبت نو |

مالکیت خصوصی – استفاده داخلی تنها برای ارائه | \*\* (License) مجوز استفاده \*\* |  
| سرویس نوبت‌دهی

| زبان \*\* | فارسی (با الفبای فارسی و اعداد انگلیسی) \*\* |

\*\* ترکیب مجموعه داده B. #####

| نوع داده | تعداد رکوردها (تقریبی) | درصد از کل |

| ----- | ----- | ----- |

| کاربران ثبت‌نام‌شده | ۱,۲۰۰ | ۲۰٪ |

| نوبت‌های ثبت‌شده | ۴,۸۰۰ | ۸۰٪ |

| \*\* جمع \*\* | \*\* ۶,۰۰۰ \*\* | \*\* ۱۰۰٪ \*\* |

\*\* منبع داده‌ها C. #####

منبع: \*\* داده‌ها به‌طور مستقیم از کاربران نهایی (بیماران) از طریق فرم‌های وبسایت \*\* -  
جمع‌آوری می‌شوند.

\*\* روش جمع‌آوری \*\* -

۱. فرم ثبت‌نام (نام کامل، شماره تلفن، رمز عبور).

۲. فرم درخواست نوبت (نام بیمار، شماره تلفن).

۳. تعاملات کاربر (ورود، لغو نوبت، مشاهده وضعیت).

\*\* جمع‌آوری نمی‌شود (Third-Party) هیچ داده‌ای از منابع ثالث \*\* -

## **\*\*مالکیت و مجوزها D. ###**

مالکیت: \*\* کلیه داده‌ها متعلق به \*\*کاربران\*\* (صاحبان داده) است \*\* -

(Privacy Policy) \*\*مجوز استفاده: \*\*کاربران با پذیرش \*\*سیاست حریم خصوصی\*\* -  
، به شرکت نوبت نو مجوز (Terms of Service) \*\*و\*\* شرایط استفاده (Policy) محدود و غیرانحصاری برای پردازش داده‌هایشان به منظور ارائه سرویس نوبت‌دهی را می‌دهند.

**\*\*محدودیت‌های استفاده\*\* -**

داده‌ها **\*\*فقط\*\*** برای ارائه سرویس اصلی (نوبت‌دهی، پیگیری وضعیت، بهبود تجربه -  
کاربری) استفاده می‌شوند.

فروش، اجاره یا اشتراک **\*\*داده‌ها با شرکت‌های ثالث مطلقاً ممنوع است\*\*** -

کامل و تأیید اخلاقی anonymization استفاده برای **\*\*تحقیقات دانشگاهی\*\*** تنها با -  
امکان‌پذیر است.

## **\*\*حریم خصوصی و اخلاق E. ###**

**\*\*رضایت (Consent):\*\* -**

رضایت صریح از طریق چک‌باکس «موافقت با سیاست حریم خصوصی» دریافت -  
می‌شود.

رضایت مجزا برای هر استفاده ثانویه (مثلاً ارسال پیامک تبلیغاتی) دریافت خواهد شد -

داده‌ها برای اهداف تحلیلی به صورت ناشناس (با حذف نام **\*\*Anonymization:\*\*** -  
و شماره تلفن) درمی‌آیند.

حقوق کاربران: **\*\*کاربران می‌توانند درخواست دسترسی، تصحیح یا حذف داده‌های\*\*** -  
(و قوانین ایران GDPR مطابق) خود را دهند.

\*\*\* \*\*داده ۴/۹/۳ دیکشنری (Data Dictionary) \*\*

#### \*\*جدول: `wp\_nobat\_users`\*\*

کلید	Nullable?	نام ستون	نوع داده	توضیح	مقادیر مجاز/نمونه
-----	-----	-----	-----	-----	-----
				شناسه منحصر به فرد کاربر	۱, ۲, ۳, ...   خیر
				Primary Key	
				نام کامل کاربر	"علی محمدی"   خیر
				-	
				شماره تلفن همراه (۱۱ رقم)	
				Unique Key	"۰۹۱۲۱۲۳۴۵۶۷"   خیر
				(با bcrypt) رمز عبور هش شده	
				-	
				زمان ایجاد حساب	"۱۵-۰۱-۲۰۲۴"
				-	

#### \*\*جدول: `wp\_nobat\_system`\*\*

کلید	Nullable?	نام ستون	نوع داده	توضیح	مقادیر مجاز/نمونه
------	-----------	----------	----------	-------	-------------------

شناسه منحصر به فرد نوبت	۱, ۲, ۳, ...	خیر	INT (11)	`id`	
Primary Key					
نام بیمار	"فاطمه احمدی"	خیر	VARCHAR(100)	`patient_name`	
-					
شماره تلفن بیمار	"۰۹۳۵۱۲۳۴۵۶۷"	خیر	Index	VARCHAR(20)	`phone_number`
وضعیت نوبت	pending, in-progress, completed, cancelled	خیر	Index	VARCHAR(20)	`status`
موقعیت در صف	۱, ۲, ۳, ...	خیر	Index	INT (11)	`position`
زمان ایجاد نوبت	"۲۰۲۴-۰۳-۲۰"	خیر		DATETIME	`created_at`
-					

### \*\* (Derived Features) ویژگی‌های محاسبه‌شده \*\*\*\*

نام ویژگی	فرمول/منطق	هدف
15 * (قبل از نوبت فعلی pending تعداد نوبت‌های)	`**زمان تخمینی ویزیت**	دقیقه
(completed برای نوبت‌های) مدت انتظار **	`(زمان تکمیل - زمان ایجاد) **	تحلیل عملکرد سیستم
نرخ لغو **	`(تعداد نوبت‌های لغوشده / کل نوبت‌ها) * ۱۰۰`	سنجش رضایت کاربر

---

\*\*\* \*\*۴/۹/۴ کیفیت داده (Data Quality)\*\*\*

و مقادیر مفقوده (Completeness) الف) کامل بودن \*\*\*  
(Missingness)\*\*

ستون	نرخ مقادیر مفقوده (%)	اقدام اصلاحی
-----	-----	-----
الزامی در فرم ثبت نام	۰٪	`full_name`
الزامی، اعتبارسنجی ۱۱ رقمی	۰٪	`phone_number`
الزامی در فرم نوبت	۰٪	`patient_name`
`pending` مقدار پیش فرض	۰٪	`status`
MAX(position)+1 محاسبه خودکار با	۰٪	`position`

مقادیر مفقوده: \*\* در حال حاضر هیچ مقداری مفقوده نیست (به دلیل الزامی بودن فیلدها \*\*  
(در فرم)

\*\*\* \*\*ب) صحت (Accuracy)\*\*\*

اعتبارسنجی می شود `^([0-9]{11})\$` regex شماره تلفن: \*\* با \*\* -

- نام: \*\* فقط حروف فارسی و فاصله مجاز است (اعتبارسنجی سمت سرور) \*\* -
- محاسبه می‌شود تا از  $\text{MAX}(\text{position}) + 1$  موقعیت نوبت: \*\* با کوئری \*\* -
- تکراری بودن جلوگیری شود

#### #### \*\* (Consistency) ج سازگاری \*\*

- `YYYY-MM-DD HH:MM:SS` ها در قالب timestamp قالب زمانی: \*\* همه \*\* -
- ذخیره می‌شوند (IRST) و زمان محلی ایران
- شماره تلفن: \*\* همه شماره‌ها با حذف `+98` یا `0` اول به صورت ۱۱ رقمی ذخیره \*\* -
- می‌شوند (مانند `09121234567`)

#### #### \*\* (Uniqueness) د یکتایی \*\*

- در جدول کاربران یکتا است (یک شماره، یک حساب) `phone\_number` -
- در یک لحظه می‌تواند تکراری باشد (اگر دو نوبت همزمان ثبت شوند) - `position` -
- این یک \*\* نقطه ضعف \*\* است که باید با تراکنش دیتابیس حل شود

---

#### #### \*\* در داده‌ها (Bias) اریبی ۴/۹/۵ \*\*

#### #### \*\*:منابع اریبی شناسایی شده\*\*

| نوع اریبی | منبع احتمالی | تأثیر بر سیستم | راهکار کاهش |

|-----|-----|-----|-----|

کاربران فقط از **\*\*شیراز\*\*** هستند (محدوده | **\*\* (Selection Bias) گزینشی\*\*** |  
جغرافیایی) | مدل‌های تحلیلی برای شهرهای دیگر قابل تعمیم نیستند | گسترش خدمات به  
| شهرهای دیگر

داده‌ها در ساعات کاری (۸ صبح تا ۴ عصر) | **\*\* (Temporal Bias) زمانی\*\*** |  
بیشتر جمع‌آوری می‌شوند | پیش‌بینی‌ها برای شب‌ها و آخر هفته نادرست است | جمع‌آوری  
| داده در همه ساعات

فقط کاربران با دسترسی به اینترنت و تلفن | **\*\* (Technology Bias) فناوری\*\*** |  
همراه | محرومیت از سالمندان یا افراد با دسترسی محدود | ارائه امکان نوبت‌گیری تلفنی  
|

فقط فارسی‌زبانان | عدم پشتیبانی از سایر | **\*\* (Language Bias) زبانی\*\*** |  
| زبان‌ها | افزودن پشتیبانی چندزبانه

**\*\*اقدامات کاهش اریبی\*\*** #####

ثبت **\*\*متاداده\*\*** (مانند شهر، زمان ثبت) برای تحلیل اریبی -

متنوع‌سازی روش‌های ثبت نوبت (تلفنی، حضوری) -

نظرسنجی دوره‌ای از کاربران برای شناسایی گروه‌های محروم -

---

**\*\* (Monitoring) و پایش (Data Drift) تغییرات داده ۴/۹/۶\*\*** ###



**\*\*قابل رخداد Drift انواع\*\*** #####

| تعریف | شاخص پایش | Drift نوع |

| ----- | ----- | ----- |

| **\*\*Concept Drift\*\*** | تغییر در رابطه بین ویژگی‌ها و نتیجه (مثلاً تغییر میانگین  
| زمان ویزیت) | میانگین زمان انتظار در هفته‌های مختلف

| **\*\*Data Drift\*\*** | تغییر در توزیع داده‌های ورودی (مثلاً افزایش ناگهانی نوبت‌های  
| شبانه) | توزیع ساعت ثبت نوبت

| **\*\*Label Drift\*\*** | تغییر در مقادیر خروجی (مثلاً افزایش نرخ لغو نوبت) | نرخ  
| لغو ماهانه

**\*\*Drift: ابزارهای پایش\*\*** #####

- در بازه‌های زمانی `status` و `created\_at` داشبورد داخلی: **\*\*نمودارهای\*\***

- هشدار خودکار: **\*\*اگر نرخ لغو از ۲۰٪ عبور کند، ایمیل به تیم فنی ارسال شود\*\*** -

- از تغییرات داده quarterly report تحلیل دوره‌ی: **\*\*بررسی\*\*** -

---

**\*\*امنیت و دسترسی به داده ۴/۹/۷\*\*** ###

**\*\*سطوح دسترسی\*\*** #####

| نقش | دسترسی به داده‌ها | محدودیت‌ها |

|-----|-----|-----|  
| کاربر عادی\*\* | فقط داده‌های خود | از طریق داشبورد کاربری\*\* |  
| ثابت، لاگ‌گیری IP | پشتیبانی\*\* | داده‌های همه کاربران (فقط خواندن)\*\* |  
| توسعه‌دهنده\*\* | دسترسی کامل (برای عیب‌یابی) | تنها در محیط تست با داده‌های\*\* |  
| ساختگی  
| کوتاه session timeout، FA ادمین\*\* | همه داده‌ها + توان حذف/تغییر | \*\*۲ |

\*\*رمزنگاری####

TLS 1.3:در انتقال\*\* -

bcrypt. در ذخیره:\*\* رمز عبور هش شده با\*\* -

AES-256 با `phone\_number` برنامه آینده:\*\* رمزنگاری ستون\*\* -

---

\*\* (Backup & Recovery) پشتیبان‌گیری و بازیابی ۴/۹/۸\*\* ####

\*\*سیاست پشتیبان‌گیری####

| نوع | فرکانس | محل نگهداری | مدت نگهداری |

|-----|-----|-----|-----|

روزانه (ساعت ۲ بامداد) | سرور جداگانه + کلود | **\*\*Full Backup\*\*** |  
| روز ۳۰ | (encrypted)  
| هر ۶ ساعت | سرور جداگانه | ۷ روز | **\*\*Incremental Backup\*\*** |  
| همان سرور | ۱۴ روز | Real-time | **\*\*لاگ‌های تراکنش\*\*** |

**\*\*تست بازیابی\*\*** #####

بازیابی کامل داده‌ها **\*\*هر فصل یکبار تست می‌شود -**

RPO (Recovery Point Objective) < ۴ ساعت،  
(Recovery Point Objective) < ۱۵ دقیقه. هدف -

---

**\*\*مستندات وابسته ۴/۹/۹\*\*** ###

سند	توضیح	لینک/مرجع
سیاست حریم خصوصی **\*\***	توضیح نحوه جمع‌آوری و استفاده از داده **\*\***	
[https://nobatnou.ir/privacy`](https://nobatnou.ir/privacy)		
قوانین استفاده از سرویس	**\*\* (ToS) شرایط استفاده\*\***	
[https://nobatnou.ir/terms`](https://nobatnou.ir/terms)		
(در صورت وجود)	داخلی endpoints مستندات	**\*\*API راهنمای\*\***

| گزارش کیفیت داده\*\* | گزارش ماهانه کیفیت داده‌ها | داخلی\*\* |

---

\*\*جمع‌بندی و اقدامات آینده ۴/۹/۱۰\*\* ###

\*\*نقاط قوت داده‌های فعلی\*\* ####

.(missing value هیچ) کامل بودن -

.صحت بالا (اعتبارسنجی سمت سرور) -

.مالکیت شفاف (تمامی داده‌ها از کاربران با رضایت گرفته شده‌اند) -

\*\*نقاط ضعف و بهبودهای آینده\*\* ####

| ضعف | اقدام بهبود | مهلت |

| ----- | ----- | ----- |

در `phone\_number` عدم رمزنگاری داده‌های حساس\*\* | رمزنگاری\*\* |

| دیتابیس | ۶ ماه

| اریبی جغرافیایی\*\* | گسترش خدمات به شهرهای دیگر | ۱۲ ماه\*\* |

| نداشتن داده‌های جمعیتی\*\* | جمع‌آوری اختیاری سن و جنسیت (با رضایت) | ۳ ماه\*\* |

| Grafana با dashboard راه‌اندازی | \*\*Data Drift\*\* عدم پایش خودکار\*\* |

| ۴ ماه |

---

**\*\*نتیجه‌گیری ۴/۹/۱۱\*\*** ###

با کیفیت بالا، شفافیت در مالکیت و رعایت حریم `nobatnou.ir` مجموعه داده‌های خصوصی جمع‌آوری شده‌اند. با این حال، **\*\*پایش مستمر کیفیت، اریبی و تغییرات داده\*\*** ضروری است. این سند به‌عنوان مرجع اصلی برای همه ذینفعان (توسعه‌دهندگان، مدیران، ناظران) عمل کرده و تضمین می‌کند که داده‌های سیستم به عنوان یک **\*\*دارایی مطمئن و پایدار\*\*** مدیریت شوند.