**UTF-8 چیست؟**

UTF-8 (Unicode Transformation Format 8-bit) یکی از پرکاربردترین استانداردهای کدگذاری کاراکترها در دنیای کامپیوتر و اینترنت است. این استاندارد به عنوان یک روش کدگذاری برای نمایش تمام کاراکترهای موجود در مجموعه یونی‌کد (Unicode) استفاده می‌شود و توانایی نمایش اکثر زبان‌های دنیا را دارد. UTF-8 به دلیل ویژگی‌های منحصربه‌فردش، به استانداردی جهانی برای انتقال و ذخیره‌سازی متن تبدیل شده است.

**یونی‌کد چیست؟**

یونی‌کد یک استاندارد جهانی است که هدف آن فراهم کردن یک کد واحد برای نمایش تمام کاراکترهای زبان‌های مختلف، از جمله حروف، اعداد، نمادها و علائم، است. یونی‌کد هر کاراکتر را با یک کد عددی منحصر به فرد به نام "کدپوینت" (Code Point) مشخص می‌کند.

**ویژگی‌های UTF-8**

**فشرده و کارآمد**

UTF-8 از یک تا چهار بایت برای کدگذاری هر کاراکتر استفاده می‌کند. کاراکترهای رایج (مانند حروف لاتین) تنها یک بایت فضا می‌گیرند، در حالی که کاراکترهای پیچیده‌تر (مانند حروف زبان‌های آسیایی) از دو تا چهار بایت استفاده می‌کنند.

**سازگاری با ASCII**

کاراکترهای ASCII (کدهای ۰ تا ۱۲۷) در UTF-8 بدون تغییر کدگذاری می‌شوند، که آن را برای سیستم‌های قدیمی و مدرن به‌طور یکسان سازگار می‌سازد.

**انعطاف‌پذیری بالا**

UTF-8 می‌تواند تمام ۱,۱۱۲,۰۶۴ کدپوینت یونی‌کد را پوشش دهد، به همین دلیل برای پشتیبانی از زبان‌ها و سیستم‌های نوشتاری مختلف ایده‌آل است.

**بدون بایت‌های اضافی**

برخلاف بعضی از استانداردها، UTF-8 بایت‌های اضافی یا بی‌استفاده تولید نمی‌کند.

**UTF-8مزایای استفاده از ( 3**

**پشتیبانی جهانی**

توانایی کدگذاری هر زبان یا نماد، آن را به انتخابی ایده‌آل برای وب‌سایت‌ها، اپلیکیشن‌ها و سیستم‌های بین‌المللی تبدیل کرده است.

**سازگاری بالا**

UTF-8 با سیستم‌ها و ابزارهای قدیمی و مدرن سازگار است.

**پشتیبانی گسترده در وب**

مرورگرها و سرورهای وب به طور پیش‌فرض UTF-8 را پشتیبانی می‌کنند، و اکثر صفحات وب از این استاندارد استفاده می‌کنند.

**4) معایب احتمالی**

**پیچیدگی در پردازش**

مدیریت داده‌های UTF-8 ممکن است برای سیستم‌هایی که به صورت بایت به بایت پردازش می‌کنند، کمی پیچیده‌تر باشد.

**اندازه فایل**

برای برخی از کاراکترهای غیرلاتین، حجم فایل ممکن است نسبت به کدگذاری‌های دیگر بزرگ‌تر باشد.

**نتیجه‌گیری**

UTF-8 یکی از قدرتمندترین و پرکاربردترین استانداردهای کدگذاری کاراکتر است که امکان نمایش صحیح و بدون نقص تمامی زبان‌ها و نمادهای موجود را فراهم می‌کند. این استاندارد به دلیل سازگاری بالا، کارایی و توانایی نمایش گسترده، به یک انتخاب پیش‌فرض در برنامه‌نویسی، طراحی وب و انتقال داده‌های متنی تبدیل شده است.