

Catálogo Padrão de Dados - Volume 2 - Anexo

Versão 1.0 – 28 Fevereiro de 2012

Introdução

Este anexo apresenta os resultados obtidos na 1ª pesquisa do CPD, que foram usados para a definição dos padrões adotados.

Houve uma boa presença de respondentes (49) e principalmente de pessoas que responderam em nome de seus órgãos (14).

Entre outros aspectos a serem avaliados, como a diferença entre as respostas gerais e as respostas oficiais dos órgãos. Por exemplo, se na pesquisa geral a proposta de CPF numérico ganhou por mais de 70%, mas nas respostas oficiais ficou em 50%.

É interessante também olhar os comentários, onde vemos algumas expectativas sobre este trabalho. Algumas talvez não possam ser atendidas, enquanto outras nos indicam o caminho para onde devemos prosseguir.

Essa pesquisa nos ajudará a definir como esse trabalho irá prosseguir. E o local de debates será a comunidade no SISP. Estão todos convidados.

Próximos Passos

Será montada uma nova pesquisa, para ampliar a lista de padrões adotados. Essa lista deve atender a solicitações que foram direcionadas a CGII por outros setores do governo. Uma dessas solicitações se refere a padrões de nomes de órgãos e datas com referência a fuso horário.

Resultado da pesquisa

Total de Fichas 49 – Pesquisa completa				
Nome				
Visualizar		Fichas Preenchidas	49	100 %
		Não responderam	0	0 %
Função/Cargo				
Visualizar		Fichas Preenchidas	49	100 %
		Não responderam	0	0 %
Email				
Visualizar		Fichas Preenchidas	49	100 %
		Não responderam	0	0 %
Nome do Órgão		Qtd	Qtd %	
Agência Nacional de Energia Elétrica - ANEEL	<div></div>	1	2.04 %	
Agência Nacional de Telecomunicações - ANATEL	<div></div>	1	2.04 %	
Agência Nacional do Petróleo - ANP	<div></div>	1	2.04 %	
Controladoria-Geral da União - CGU	<div></div>	2	4.08 %	
Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES	<div></div>	1	2.04 %	
Fundação Escola Nacional de Administração Pública - ENAP	<div></div>	1	2.04 %	
Fundação Instituto de Pesquisa Econômica Aplicada - IPEA	<div></div>	1	2.04 %	
Fundação Oswaldo Cruz	<div></div>	1	2.04 %	
Fundação Universidade Federal do Piauí	<div></div>	1	2.04 %	
Fundação Universidade Federal do Vale do São Francisco	<div></div>	1	2.04 %	
Instituto Federal de Educação Ciência e Tecnologia da Paraíba	<div></div>	1	2.04 %	
Instituto Federal de Educação Ciência e Tecnologia de Januária	<div></div>	1	2.04 %	
Instituto Federal de Educação Ciência e Tecnologia de Santa Catarina	<div></div>	2	4.08 %	
Instituto Federal de Educação Ciência e Tecnologia do Sul de Minas Gerais	<div></div>	1	2.04 %	
Instituto Nacional de Colonização e Reforma Agrária - INCRA	<div></div>	2	4.08 %	
Instituto Nacional do Seguro Social - INSS	<div></div>	1	2.04 %	
Ministério da Agricultura Pecuária e Abastecimento - MAPA	<div></div>	1	2.04 %	
Ministério da Ciência e Tecnologia - MCT	<div></div>	1	2.04 %	
Ministério da Fazenda - MF	<div></div>	3	6.12 %	
Ministério da Integração Nacional - MI	<div></div>	1	2.04 %	
Ministério da Justiça - MJ	<div></div>	1	2.04 %	
Ministério da Previdência Social - MPS	<div></div>	1	2.04 %	
Ministério do Planejamento Orçamento e Gestão - MPOG	<div></div>	13	26.53 %	
Ministério dos Transportes - MT	<div></div>	1	2.04 %	
Poder Judiciário	<div></div>	1	2.04 %	
Poder Legislativo / Câmara - Senado	<div></div>	1	2.04 %	
Universidade Federal da Bahia	<div></div>	1	2.04 %	
Universidade Federal de Campina Grande	<div></div>	1	2.04 %	
Universidade Federal de Pernambuco	<div></div>	1	2.04 %	
Universidade Federal do Pampa	<div></div>	1	2.04 %	
Universidade Federal Rural da Amazônia	<div></div>	1	2.04 %	
Universidade Federal Rural do Rio de Janeiro	<div></div>	1	2.04 %	
Fichas Preenchidas		49	100 %	

Não responderam	0	0 %
Nome da unidade onde trabalha		
Visualizar Fichas Preenchidas	49	100 %
Não responderam	0	0 %
Representante da organização	Qtd	Qtd %
Sim	14	28.57 %
Não	35	71.43 %
Fichas Preenchidas	49	100 %
Não responderam	0	0 %

Campo NOME

Tamanho	Qtd	Qtd %
50 posições	3	6.12 %
70 posições (RECOMENDADO)	30	61.22 %
100 posições	13	26.53 %
Sem limites de posições	3	6.12 %
Fichas Preenchidas	49	100 %
Não responderam	0	0 %

Acentos	Qtd	Qtd %
Nome completo conforme registro civil inclusive com acentos (RECOMENDADO)	41	83.67 %
Nome completo conforme registro civil sem acentos	6	12.24 %
Sem padronização específica	1	2.04 %
Outra padronização	1	2.04 %
Fichas Preenchidas	49	100 %
Não responderam	0	0 %

Abreviações	Qtd	Qtd %
Nome completo conforme registro civil sem abreviações (RECOMENDADO)	43	87.76 %
Sem padronização específica.	6	12.24 %
Fichas Preenchidas	49	100 %
Não responderam	0	0 %

Matriz

Parâmetros de Críticas

Deve ser maior que 5 caracteres (Recomendação: SIM)	Qtd	Qtd %
Sim	38	77.55 %
Não	11	22.45 %
Fichas Preenchidas	49	100 %
Não responderam	0	0 %

Deve atender a expressão regular abaixo (Recomendação: SIM)	Qtd	Qtd %
Sim	36	73.47 %
Não	13	26.53 %
Fichas Preenchidas	49	100 %
Não responderam	0	0 %

Deve ser obrigatório na troca de informações	Qtd	Qtd %
--	-----	-------

personais. (Recomendação: SIM)

Sim	<div><div></div></div>	42	85.71 %
Não	<div><div></div></div>	7	14.29 %
Fichas Preenchidas		49	100 %
Não responderam		0	0 %

Não deve conter três caracteres iguais - exceto III. (Recomendação: SIM)

		Qtd	Qtd %
Sim	<div><div></div></div>	27	55.10 %
Não	<div><div></div></div>	22	44.90 %
Fichas Preenchidas		49	100 %
Não responderam		0	0 %

Não deve ter palavras da língua portuguesa. (Recomendação: SIM)

		Qtd	Qtd %
Sim	<div><div></div></div>	30	61.22 %
Não	<div><div></div></div>	19	38.78 %
Fichas Preenchidas		49	100 %
Não responderam		0	0 %

Não deve ser palavras como `Ignorado`, `Inexistente`, `Desconhecido` (Recomendação: SIM)

		Qtd	Qtd %
Sim	<div><div></div></div>	36	73.47 %
Não	<div><div></div></div>	13	26.53 %
Fichas Preenchidas		49	100 %
Não responderam		0	0 %

Campo DATA

Formato		Qtd	Qtd %
Atender o formato da ISO - 8601:2004 no seu formato básico AAAAMMDD (RECOMENDADO)	<div><div></div></div>	31	63.27 %
DD/MM/AAAA	<div><div></div></div>	15	30.61 %
DD-MM-AAAA	<div><div></div></div>	1	2.04 %
Outra padronização	<div><div></div></div>	2	4.08 %
Fichas Preenchidas		49	100 %
Não responderam		0	0 %

Campo SEXO

Conteúdo		Qtd	Qtd %
ISO - 5218:2004: 0 - Não conhecido; 1 - Homem; 2 - Mulher; 9 - Não aplicável. (RECOMENDADO)	<div><div></div></div>	32	65.31 %
Seguir padrão M - Masculino e F - Feminino	<div><div></div></div>	12	24.49 %
Seguir padrão 1 - Masculino e 2 - Feminino	<div><div></div></div>	4	8.16 %
Sem padronização específica	<div><div></div></div>	1	2.04 %
Fichas Preenchidas		49	100 %
Não responderam		0	0 %

Campo CPF

Tipo		Qt d	Qtd %
Numérico (RECOMENDADO)	<div></div>	35	71.43 %
Alfanumérico sem formatação (11 caract)	<div></div>	14	28.57 %
Fichas Preenchidas		49	100 %
Não responderam		0	0 %

Matriz

Parâmetros de Críticas

Módulo 11 para validar o número do CPF (Recomendação: SIM).		Qtd	Qtd %
Sim	<div></div>	48	97.96 %
Não	<div></div>	1	2.04 %
Fichas Preenchidas		49	100 %
Não responderam		0	0 %

Padrões da Organização

Visualizar	Fichas Preenchidas	15	30.61 %
	Não responderam	34	69.39 %


Campos que desejam ser padronizados

Visualizar	Fichas Preenchidas	23	46.94 %
	Não responderam	26	53.06 %

Sugestões/Críticas



Visualizar	Fichas Preenchidas	17	34.69 %
	Não responderam	32	65.31 %



Total de Fichas 14 – Somente respostas oficiais					
Nome		Visualizar	Fichas Preenchidas	14	100 %
			Não responderam	0	0 %
Função/Cargo		Visualizar	Fichas Preenchidas	14	100 %
			Não responderam	0	0 %
Email		Visualizar	Fichas Preenchidas	14	100 %
			Não responderam	0	0 %
Nome do Orgão			Qtd	Qtd %	
Agência Nacional de Energia Elétrica - ANEEL		<div></div>	1	7.14 %	
Controladoria-Geral da União - CGU		<div></div>	1	7.14 %	
Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES		<div></div>	1	7.14 %	
Fundação Universidade Federal do Piauí		<div></div>	1	7.14 %	
Fundação Universidade Federal do Vale do São Francisco		<div></div>	1	7.14 %	
Instituto Federal de Educação Ciência e Tecnologia da Paraíba		<div></div>	1	7.14 %	
Instituto Federal de Educação Ciência e Tecnologia de Januária		<div></div>	1	7.14 %	
Instituto Federal de Educação Ciência e Tecnologia do Sul de Minas Gerais		<div></div>	1	7.14 %	
Ministério da Integração Nacional - MI		<div></div>	1	7.14 %	
Ministério do Planejamento Orçamento e Gestão - MPOG		<div></div>	1	7.14 %	
Ministério dos Transportes - MT		<div></div>	1	7.14 %	
Universidade Federal da Bahia		<div></div>	1	7.14 %	
Universidade Federal de Pernambuco		<div></div>	1	7.14 %	
Universidade Federal Rural da Amazônia		<div></div>	1	7.14 %	
		Fichas Preenchidas	14	100 %	
		Não responderam	0	0 %	
Nome da unidade onde trabalha		Visualizar	Fichas Preenchidas	14	100 %
			Não responderam	0	0 %
Representante da organização			Qtd	Qtd %	
Sim		<div></div>	14	100 %	
		Fichas Preenchidas	14	100 %	
		Não responderam	0	0 %	
Tamanho			Qtd	Qtd %	
50 posições		<div></div>	3	21.43 %	
70 posições (RECOMENDADO)		<div></div>	8	57.14 %	
100 posições		<div></div>	3	21.43 %	
		Fichas Preenchidas	14	100 %	
		Não responderam	0	0 %	
Acentos			Qtd	Qtd %	
Nome completo conforme registro civil inclusive com acentos (RECOMENDADO)		<div></div>	12	85.71 %	
Nome completo conforme registro civil sem acentos		<div></div>	2	14.29 %	
		Fichas Preenchidas	14	100 %	
		Não responderam	0	0 %	
Abreviações			Qtd	Qtd %	
Nome completo conforme registro civil sem		<div></div>	12	85.71 %	



abreviações (RECOMENDADO)			
Sem padronização específica.		2	14.29 %
Fichas Preenchidas		14	100 %
Não responderam		0	0 %



Matriz



Parâmetros de Críticas



Deve ser maior que 5 caracteres (Recomendação: SIM)		Qtd	Qtd %
Sim			71.43 %
Não		4	28.57 %
Fichas Preenchidas		14	100 %
Não responderam		0	0 %




Deve atender a expressão regular abaixo (Recomendação: SIM)		Qtd	Qtd %
Sim			71.43 %
Não		4	28.57 %
Fichas Preenchidas		14	100 %
Não responderam		0	0 %

Deve ser obrigatório na troca de informações pessoais. (Recomendação: SIM)		Qtd	Qtd %
Sim			85.71 %
Não		2	14.29 %
Fichas Preenchidas		14	100 %
Não responderam		0	0 %

Não deve conter três caracteres iguais - exceto III. (Recomendação: SIM)		Qtd	Qtd %
Sim		8	57.14 %
Não		6	42.86 %
Fichas Preenchidas		14	100 %
Não responderam		0	0 %

Não deve ter palavrões da língua portuguesa. (Recomendação: SIM)		Qtd	Qtd %
Sim		8	57.14 %
Não		6	42.86 %
Fichas Preenchidas		14	100 %
Não responderam		0	0 %

Não deve ser palavras como `Ignorado`, `Inexistente` `Desconhecido` (Recomendação: SIM)		Qtd	Qtd %
Sim			71.43 %
Não		4	28.57 %
Fichas Preenchidas		14	100 %
Não responderam		0	0 %

Formato		Qtd	Qtd %
Atender o formato da ISO - 8601:2004 no seu formato básico AAAAMMDD (RECOMENDADO)		7	50 %
DD/MM/AAAA		6	42.86 %
Outra padronização		1	7.14 %

	Fichas Preenchidas	14	100 %
	Não responderam	0	0 %
Conteúdo	Qtd	Qtd %	
ISO - 5218:2004; 0 - Não conhecido; 1 - Homem; 2 - Mulher; 9 - Não aplicável. (RECOMENDADO)	6	42.86 %	
Seguir padrão M - Masculino e F - Feminino	7	50 %	
Seguir padrão 1 - Masculino e 2 - Feminino	1	7.14 %	
	Fichas Preenchidas	14	100 %
	Não responderam	0	0 %
Tipo	Qtd	Qtd %	
Númerico (RECOMENDADO)	7	50 %	
Alfanumérico sem formatação (11 caract)	7	50 %	
	Fichas Preenchidas	14	100 %
	Não responderam	0	0 %

Matriz

Parâmetros de Críticas

Módulo 11 para validar o número do CPF (Recomendação: SIM).		Qtd	Qtd %
Sim		14	100 %
	Fichas Preenchidas	14	100 %
	Não responderam	0	0 %
Padrões da Organização			
	Visualizar Fichas Preenchidas	5	35.71 %
	Não responderam	9	64.29 %
Campos que desejam ser padronizados			
	Visualizar Fichas Preenchidas	6	42.86 %
	Não responderam	8	57.14 %
Sugestões/Críticas			
	Visualizar Fichas Preenchidas	3	21.43 %
	Não responderam	11	78.57 %

Campos de comentários

Padrões da Organização

- 1 - Padronização de número de processo
- 2 - Padronização de número de empenho
- 3 - Padronização de número de ordem bancária

A Dataprev já utiliza os padrões recomendados. Somente o domínio para o campo SEXO é diferente. Utilizamos: 1 - Masculino; 3 - Feminino; 9 - Não informado.

A Instituição já adota alguma norma do Catálogo de Padrões de Dados - CPD:

- 1- Cálculo de CPF:- Módulo 11
- 2- Formato SEXO: M-masculino, F-feminino
- 3- Campo DATA: dd/mm/aaaa

Área de conhecimento (Há um padrão acordado entre CAPES e CNPq de códigos, nomes e hierarquia);

Desconheço a existência de algum padrão.

No caso de datas o próprio campo deve prever a colocação de // ou --. A pessoa deve entrar apenas com os numeros. No Brasil o formato comum é ddmmaaaa, não é padrão internacional, mas manter assim evitara varios erros de preenchimento, pois as pessoas não costumam ler as informações que alimentam até elas acusarem erro.

NOME COM ATÉ 70 POSIÇÕES; COM PELO MENOS 2 TOKENS DE PELO MENOS 2 LETRAS CADA; SEM ABREVIACÕES; SOMENTE COM CONTEUDO ALFABETICO; SEM ACENTOS; NÃO ADMITIDOS TEXTOS "IGNORADO", "INEXISTENTE", ETC.

DATA NASC NO PADRÃO AAAAMMDD;
CPF NUMERICO E COM VALIDAÇÃO DE DV

Nome sem limite de posições e contendo acentos, data no formato dd/mm/aaaa e sexo 'M' ou 'F'

O campo NOME pode, dependendo da modelagem de dados, ser aplicável tanto a Nome de pessoa natural quanto a razão social de pessoa jurídica e deste fato utilizamos 160 posições.

O campo CPF é criado com 11 caracteres sem formatação, evitando desta forma a necessidade de formatá-lo com zeros a esquerda no momento da exibição.

O campo SEXO é utilizado com caracteres que determinam o seu significado, o que facilita o entendimento.

Padrão de nomenclatura de objetos de banco de dados: classificadores de atributos e seus possíveis datatypes, e demais objetos. Validações de domínio, como check constraint.

Padrões de identificação de e-mail.

Para o projeto Protocolo Integrado, as datas previstas para a versão do padrão de dados em fase de validação estão no formato 8 dígitos: 2 para dia, 2 para mês, 4 para ano.

TABELAS ESPECIFICAS DE UM PROJETO COM UMA SIGLA DE 3 CARACTERES NO INICIO DO NOME

Uso de prefixos dos tipos de dados com três caracteres.

Campos que desejam ser padronizados

A sugestão seria que além da validação do CPF, também seria validado o CNPJ e campos numéricos que não aceitem letras (alfabeto) com ou sem ponto decimal.

Campos de endereço, principalmente, logradouro para localidades em que o CEP não é codificado em nível de logradouro.

CEP

Cep - apenas números. Campo de endereço unico (a exceção do CEP, cidade e UF), sem partições de rua, casa.

CEP, de preferência com liberação da base de dados dos Correios com reconhecimento do interesse público na mesma.

Códigos (siglas) de país e suas subdivisões territoriais, e de línguas e dialetos, conforme padrões ISO.

CNPJ - alfanumérico de 14 posições sem formatação.

Deve existir padronização para dados do tipo boolean.

Endereço e números de documentos.

endereço, logradouro, cep, UF,

endereço,cep,município e telefone.

Formato de Campo Endereço:

ENDEREÇO. BAIRRO. CIDADE. CEP

Matrícula Siape

RG

Municípios (uso de código IBGE), UF, georreferência, endereço;

Países (uso de código ISO);

Flags de situação / status (o que é melhor: letras, números ou descrição?);
Numeração de documentos (uso do NUP);

NOME DA MÃE
NOME DO PAI
NACIONALIDADE
NATURALIDADE
ENDEREÇO

Nome de pessoa jurídica;
Coordenação, secretaria, divisão, seção, setor, serviço, etc;
Ordem das siglas, exemplo: RFB/Cotec/Digec ou Digec/Cotec/RFB;
Endereçamento;
Maiúsculo, minúsculo ou misto:

NOME E TAMANHO DE COLUNAS PARA A MESMA FINALIDADE EM TABELAS DIFERENTES.

Número de Órgãos/UGs, Funcionais programáticas, CNPJ

PAI, MAE, PAIS-DE-ORIGEM, NATURALIDADE

PAIS, UF, CEP, TELEFONE, E-MAIL(CRÍTICA), GRAU DE INSTRUÇÃO.

Podem ser padronizados os campos para endereço. Também podem ser padronizados os parâmetros de críticas para campos de data de nascimento.

telefone (0DD) XXXX-YYYY

Tipo sanguíneo, nacionalidade, Naturalidade. Presença de UF de nascimento para antigos territórios de Fernando de Noronha e estado da Guanabara. Estado Civil

Sugestões/Críticas

1. Entendo que o último padrão recomendado para nome (Não deve ser palavras como `Ignorado`, `Inexistente` `Desconhecido`) torna-se desnecessário se o padrão adotado for utilizar o nome conforme registro civil, pois teoricamente esse filtro já ocorreria no registro, e, caso não ocorra, deve ser considerado como legal.

2. Para definição de sexo, sugiro substituir os termos "Homem" e "Mulher" por "Masculino" e "Feminino"

A expressão abaixo não é inteligível por pessoas fora da área técnica; deveria ser substituída por uma linguagem coloquial.

"Deve atender a expressão regular abaixo (Recomendação: SIM): *
'^[a-zA-ZÀ-Üà-ü|'|-]+([a-zA-ZÀ-Üà-ü|'|-]+)+\$"

Abreviações (preenchido por obrigatoriedade): É comum em órgãos que lidam com educação e com desenvolvimento tecnológico utilizar o formato ABNT de referência bibliográfica, devendo os nomes estarem abreviados conforme regra da ABNT - vide currículo Lattes e COLETA CAPES;
Nomes com palavras e outros acentos: Estas mesmas instituições mantêm relações com pesquisadores estrangeiros, inclusive originários do leste europeu. Alguns palavras podem ser encontrados inclusive em sobrenome de brasileiros

Aconselho que seja feito um trabalho conjunto com o Ministério da Justiça, órgão responsável pela nova identidade civil o RIC, para que os padrões sejam igualados o que proporcionará interoperabilidade, economia de recursos e agilidade no desenvolvimento e manutenção de sw. A tendência é que o número único do RIC seja a futura chave primária a ser utilizada em todas as bases e estas estejam integradas e interoperáveis. Não vejo como definir campos sem estar em harmonia com o MER do RIC.

Atenção para o RIC - Registro de Identificação Civil. Novo código criado pelo Governo Federal.

Com este questionário, seria interessante uma padronização da entrada de dados dos órgãos participantes deste, pois facilitaria possíveis trocas de informações entre os órgãos.

Cuidado para não engessar o usuário, tivemos um caso de cadastramento onde o sistema exigia 8 caracteres, no mínimo; o nome completo só tinha 7 caracteres, não sei qual foi a solução.
No campo sexo, deveria ter o que consta no documento e um 2º campo opcional, com as variações possíveis, exemplo: a aparência, mutação cirúrgica, mutilação, próteses, hermafrodita, etc. Lembrando que o campo é opcional, para atender aos que não se sentem bem com a definição do documento.

Deveria ser possível dizer que outra padronização se poderia adotar para nomes. A gramática portuguesa, por exemplo, especifica que os nomes devem ser grafados conforme as regras ortográficas, independente dos erros consagrados no registro civil, os quais são facultativos ao portador do nome divergente da norma.

Em alguns campos é necessário uma explicação para saber em qual contexto a padronização se aplica, por exemplo, o formato do campo data se aplica para transmissão, armazenamento ou ambos? Além disso, se o nome tiver tamanho fixo, mas o tamanho do nome a ser registrado for superior a este tamanho, vai ser necessário se fazer abreviações, independente da padronização escolhida para isto.

Está bem diagramado. O tamanho da fonte deveria ser maior.

Explicar melhor se esta é uma pesquisa de opinião de como deveria ser o padrão ou se é para retratar o já utilizado no órgão.

Explicar o que quer dizer a expressão regular do item 10

Nas questões onde existe a opção de "outra padronização", deveria ser possível descrever qual seria a padronização proposta.

O CPF, embora seja um campo representado por números, não possui caráter de cálculo. Inclusive para realizar a sua verificação é preciso separar os dois últimos dígitos e esta operação é mais eficiente se feita através de um campo texto. Recomendo fortemente adotar o padrão ALFANUMÉRICO.

O campo data, dado que nosso país possui mais de um fuso horário, deve considerar a informação de fuso sempre que aplicável. Por isso recomendo a utilização do padrão ISO 8601, sem indicação de zona, no seguinte formato:

Combined date and time in UTC:
AAAA-MM-DDThh:mm±hh

Exemplo: "2011-09-05T12:05-03"

Exemplo em horário de verão: "2011-12-23T17:30-04"

OTIMA INICIATIVA. DEVE SER MAIS ABRANGENTE A FORNECIMENTO DE OUTRAS INFORMAÇÕES.

Verificar possibilidade de inclusão de padronização para dados geográficos em versões posteriores do documento.