



CETEMPúblico

Et stort korpus
av portugisisk avisspråk

av

Paulo Alexandre Rocha

SINTEF Tele og Data

Databehandling av portugisisk språk

**C
E
T
E
M**

Público

HVA ER ET KORPUS ?

corpus

1. **m.** Recopilación de los escritos de un autor, de una serie de leyes, etc.
2. Conjunto o recopilación de materiales, datos, etc., sobre una materia, doctrina, etc.
3. LING. Conjunto acabado de enunciados.

[Diccionario General de la Lengua Española VOX]

cor·pus (kôr·pʊs) *noun*

plural cor·po·ra (-pʊr-ʌ)

Abbr. cor.

1. A large collection of writings of a specific kind or on a specific subject.

[The American Heritage® Dictionary of the English Language]

corpus s.m.

compilação de documentos ou informações relativos a uma disciplina ou tema;

(Do lat. *corpus*, «corpo; conjunto; matéria»)

[Dicionário da Língua Portuguesa, Porto Editora]

II korpus m1 el. n3 (fra lat. 'legeme')

1 ...

2 innsamlet materiale

[Bokmålsordboka]

corpus n. m.

1. Recueil concernant une même matière. *Corpus d'inscriptions latines.*
2. LING Ensemble fini d'éléments, d'énoncés, réunis en vue d'une analyse linguistique.

[Dictionnaire Universel Francophone En Ligne, Hachette]

C
E
T
E
M

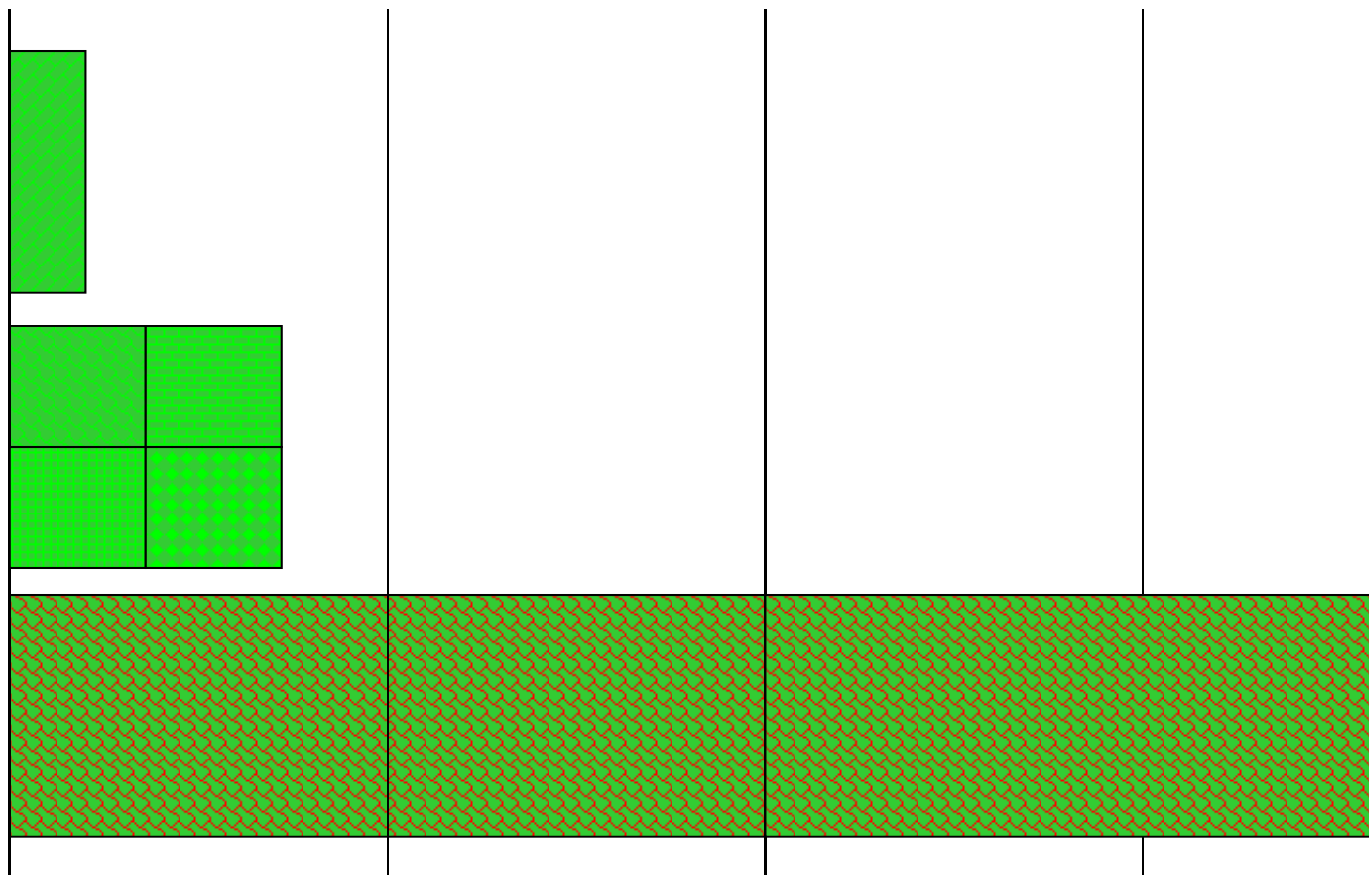
Público



Andre korpus

NILC/São Carlos

CETEMPúblico

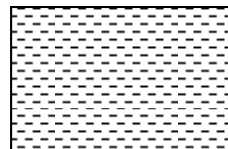


Corpus de
Extractos de
Textos
Electrónicos
MCT /
Público



X

1500



X

1.500.000

Ord

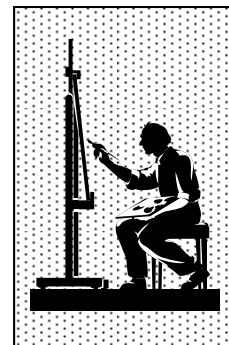
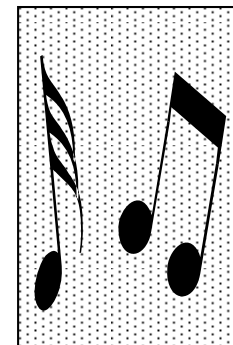
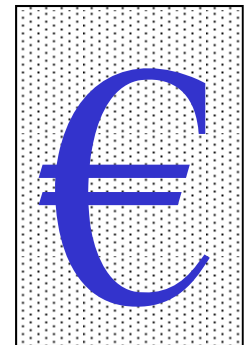
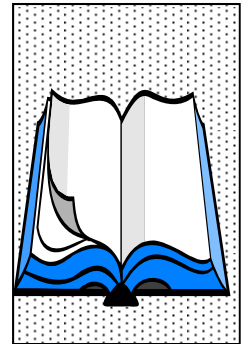
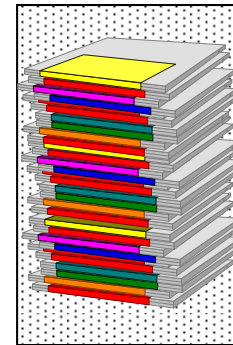
X

180.000.000



Público

1			



C
E
T
E
M

Público

Público
Lisboa



Público
Porto



Público

Guardian THE IRISH TIMES Libération

La Presse TO BHMÁ EL PAÍS

FOLHA DE S. PAULO

MLADINA

LE SOIR

der Standard

Aftenposten

יָרָאָה Независимая Газета

LA STAMPA Süddeutsche Zeitung

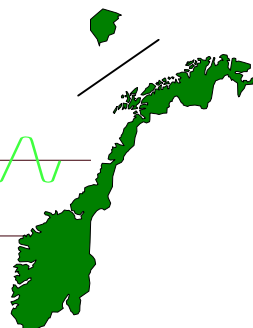
LIDOVÉ NOVINY

BerlingskeTidende



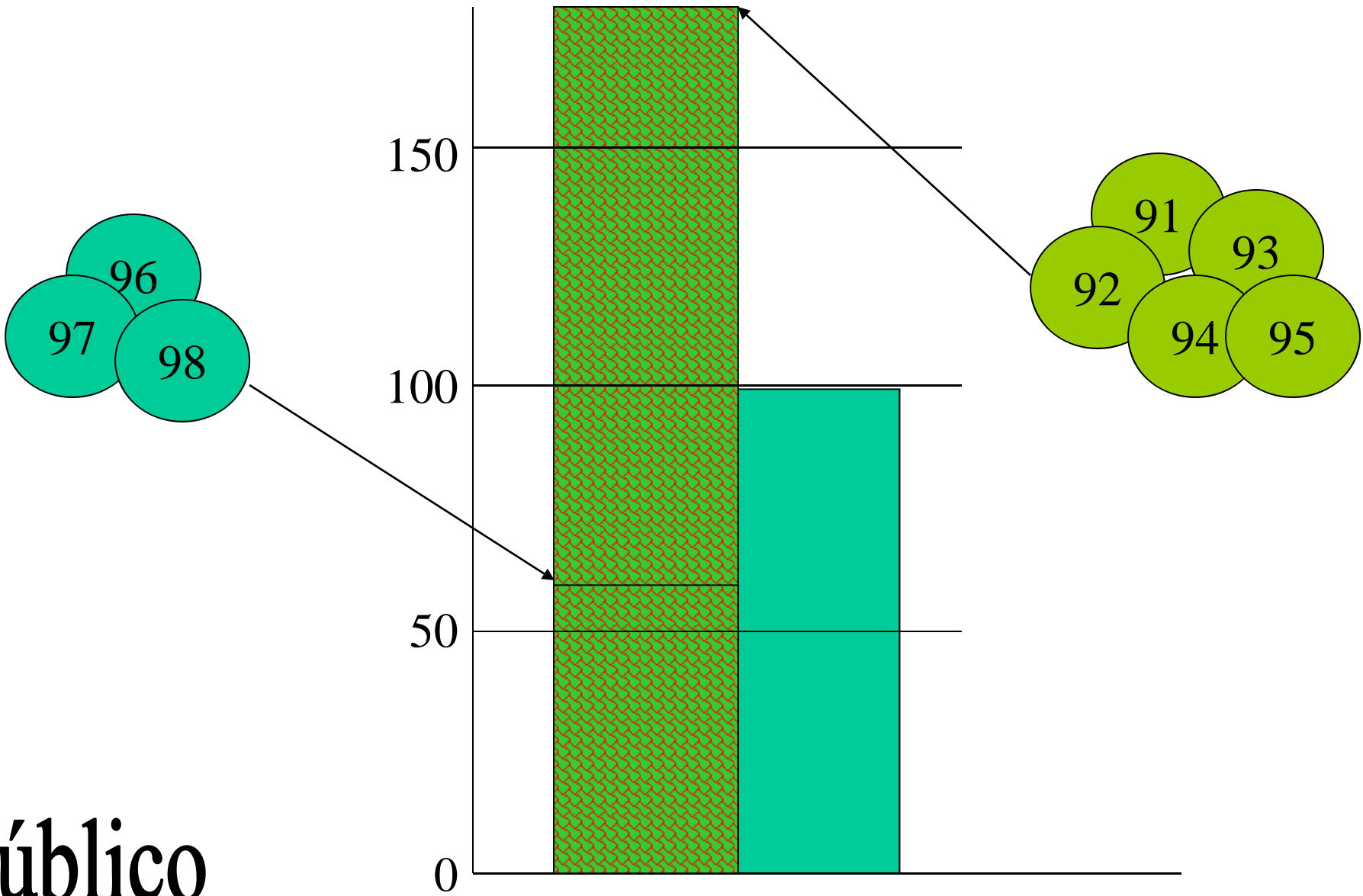
C
E
T
E
M

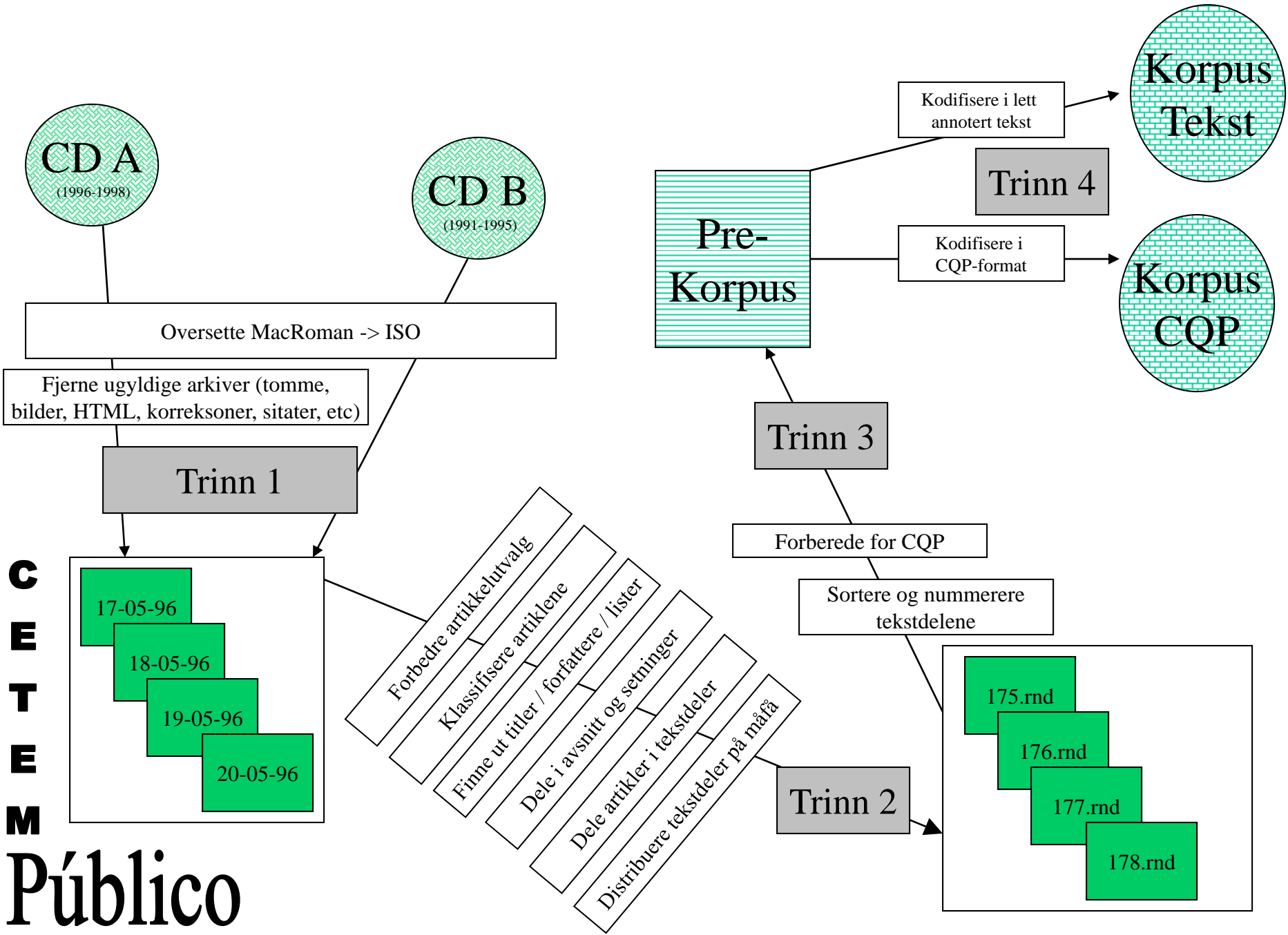
Público



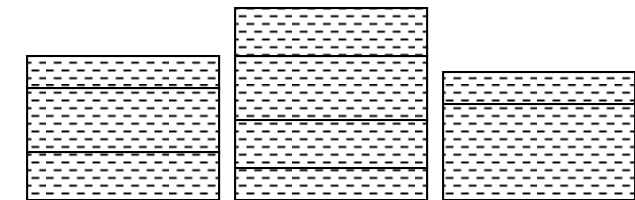


**C
E
T
E
M**
Público



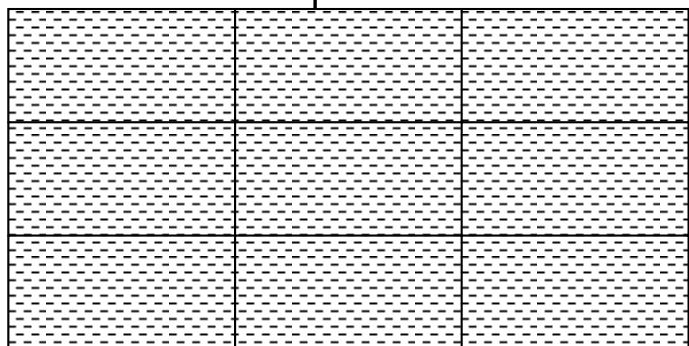


**C
E
T
E
M**
Público



	Tall	Pro artikkel	Pro tekstdele
Artikler	112.509	-	-
Tekstdele	489.033	4,35	-
Identifisert titler	269.848	0,42	0,55
Identifisert forfattere	47.763	0,42	0,10

(Tallene gjelder for tekster fra 1996 til 1998)

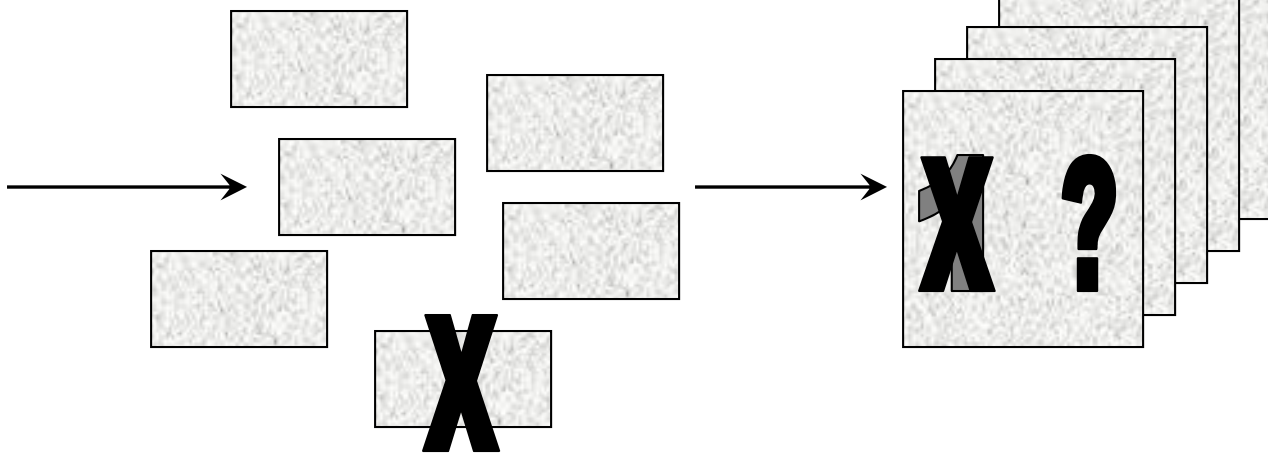
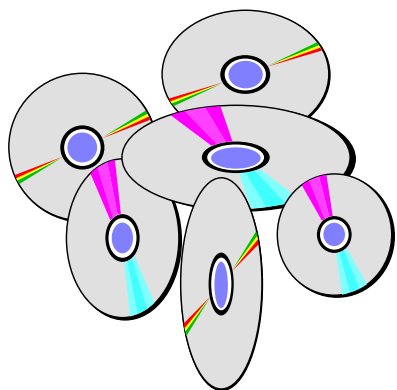


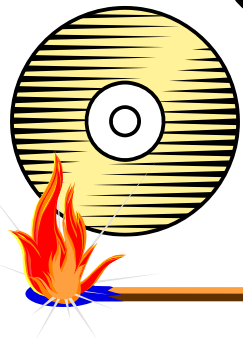
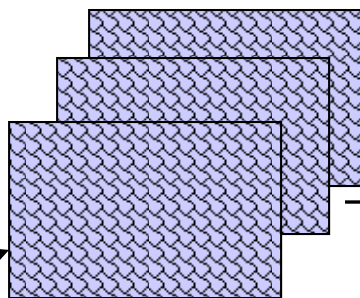
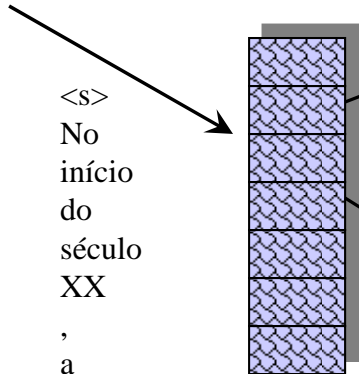
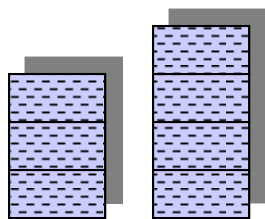
```

<t>Matpakkene og økologi </t>
<a>Gro Harlem Brundtlandt</a>
Bla bla bla bla bla. Bla bla
bla bla bla. Bla bla bla . Bla
bla bla bla bla bla
bla.....
  
```



**C
E
T
E
M**
Público



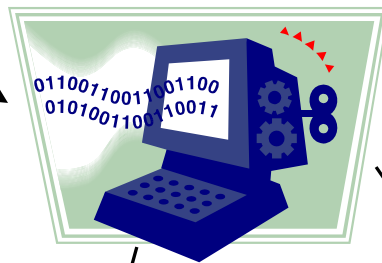


Tekst

<s>
No
início
do
século
XX
,
a
localidade
norueguesa
de

93a Hammerfest SOC

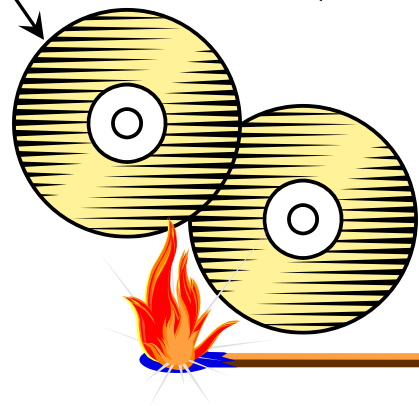
foi
o
primeiro
local
da
Europa
a
ter
luz
eléctrica
nas
ruas
.
</s>



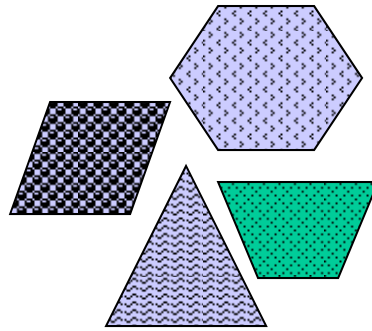
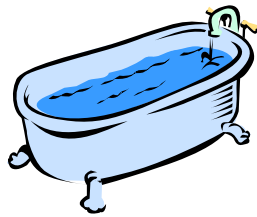
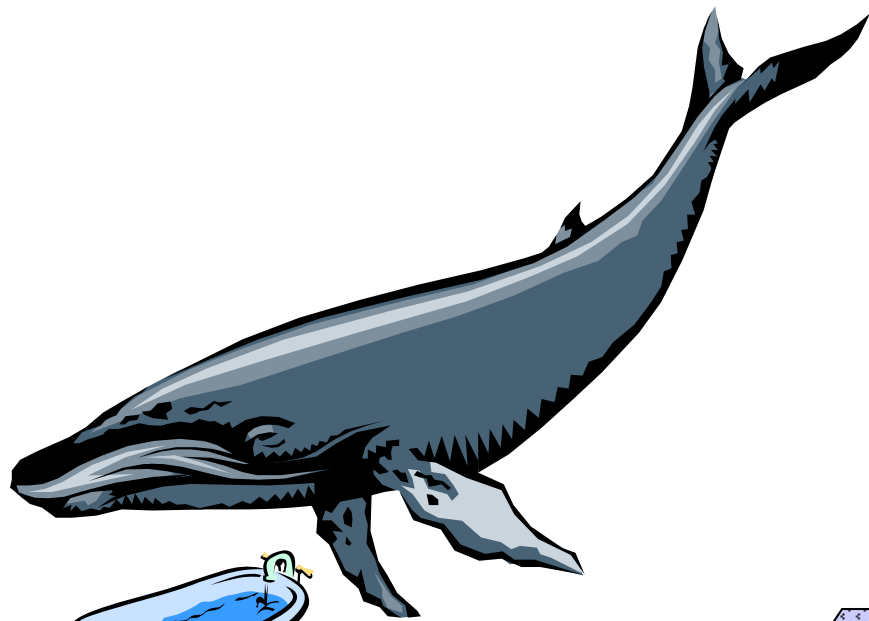
AC/DC



CQP



C E T E M Público

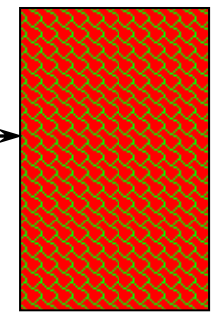
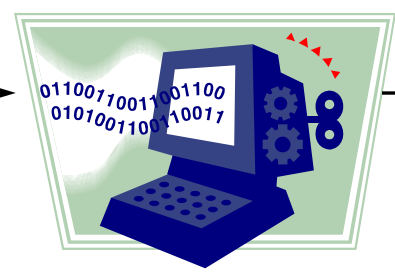
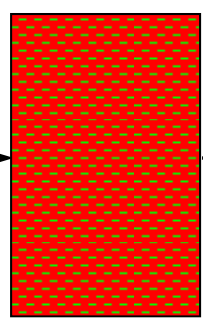


C
E
T
E
M

2 X

Público

Nobody expects the Spanish Inquisition



C
E
T
E
M
Público

Rom { CETEMPúblico
DiaCLAV

Tid { CETEMPúblico
DiaCLAV

**Resultados da procura**

Procura: "**Internet**" %c;
Distribuição de **semestre**
Corpus: CETEMPúblico 1.1

8883 casos.

Distribuição

98b	1146
98a	1931
97b	1334
97a	756
96b	409
96a	503
95b	1108
95a	1197
94b	279
94a	195
93b	3
93a	21
92a	1

Ext **332461** (nd, 96b): A oposição norueguesa criticara recentemente o ministro do Petróleo, **Jens Stoltenberg**, e o conjunto do Governo trabalhista, de não estarem suficientemente atentos à situação dos direitos humanos na Indonésia .

Ext **383637** (pol-soc, 94b): Ontem, o ministro da **Industria** e Energia, **Jens Stoltenberg**, foi obrigado a cancelar o vôo para Bruxelas porque o referendo lhe bloqueou o acesso à reunião dos ministros da Energia dos Doze .

Ext **582679** (soc, 95a): «Trata-se de encontrar um local seguro para atracar temporariamente, enquanto a Shell elabora um plano alternativo para a desactivação da Brent Spar», afirmou o ministro norueguês **Jens Stoltenberg** .

Ext **699917** (pol, 96b): Com Jagland sobe também o ministro da Indústria e Energia, **Jens Stoltenberg**, que transita para as Finanças, o que pode ser interpretado como um prémio para o seu bom desempenho na gestão dos (vastos) recursos petrolíferos do país -- que valem a este político muitos pontos nas sondagens .

Ext **1526329** (soc, 95a): Para já, no que respeita à Brent Spar, a Noruega declarou-se disposta a receber provisoriamente a plataforma, segundo declarações do ministro **Jens Stoltenberg**, que foi contactado pelo seu homólogo holandês, Hans Wijers .

Resultados da procura

Procura: "**Re.(k|que).av.***".
Pedido: Distribuição das formas
Corpus: CETEMPúblico 1.1

123 casos.

Distribuição

Reiquejavique	55
Reykjavik	46
Reikjavik	10
Reikjavic	5
Reikjavyk	3
Reykjavic	3
Reikyavik	1

Resultados da procura

Procura: "**Rosenborg**".
Distribuição de **assunto**
Corpus: CETEMPúblico 1.1

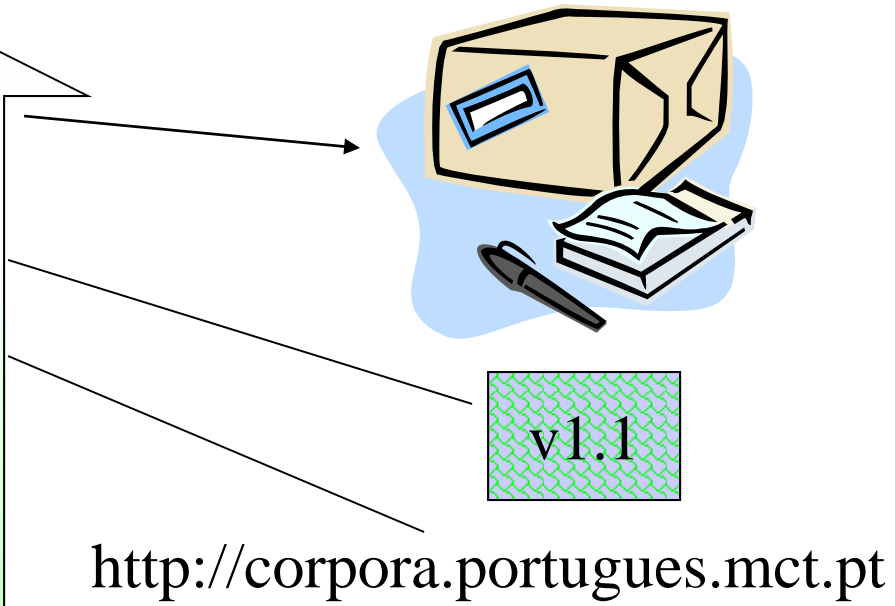
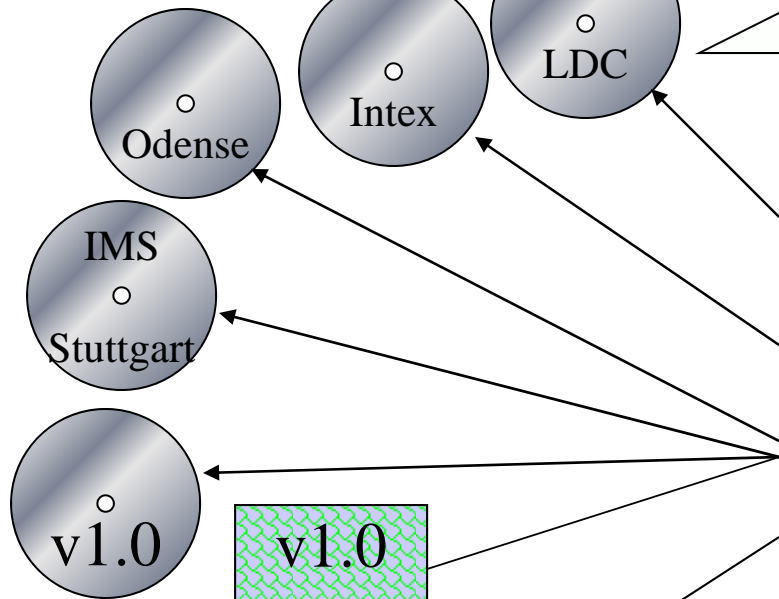
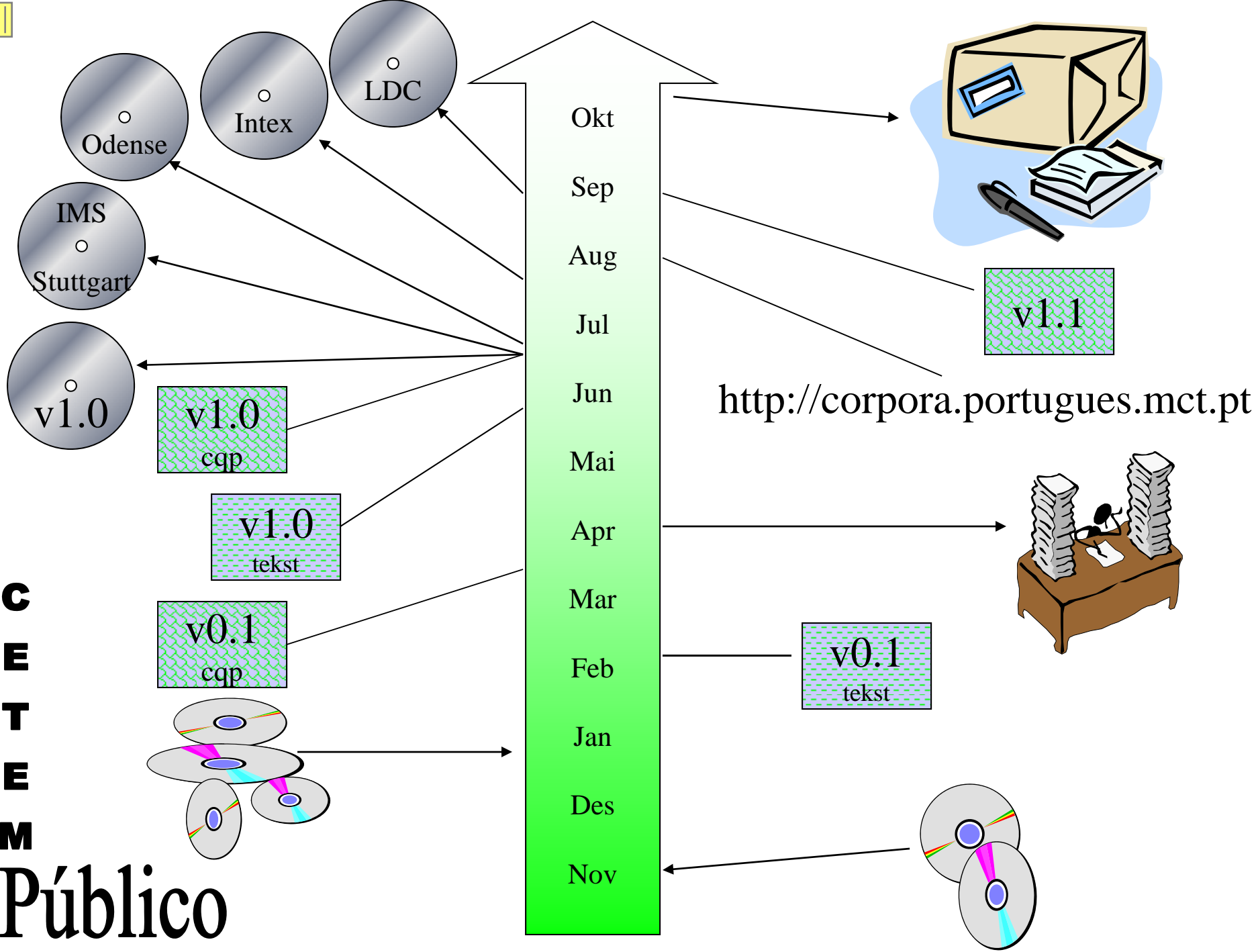
309 casos.

Distribuição

des	259
nd	42
opi	4
soc	3
clt	1



**C
E
T
E
M**
Público



v1.0
cqp

v1.0
tekst

v0.1
cqp

v0.1
tekst

v1.1

<http://corpora.portugues.mct.pt>

Okt

Sep

Aug

Jul

Jun

Mai

Apr

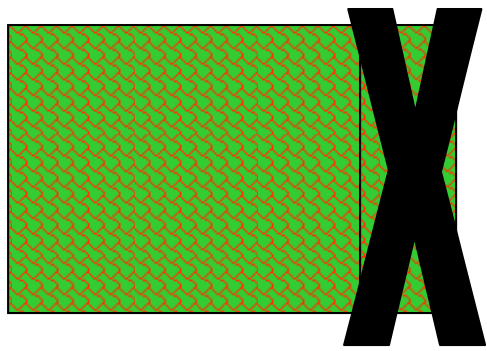
Mar

Feb

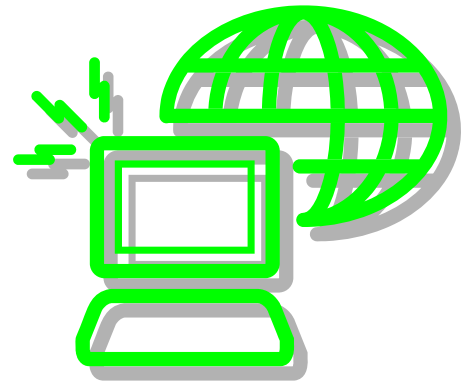
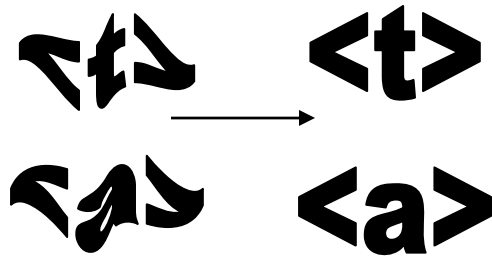
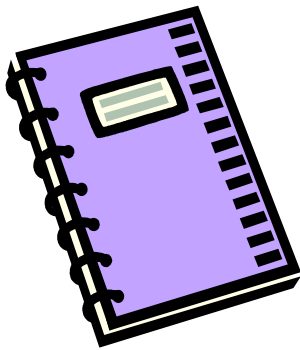
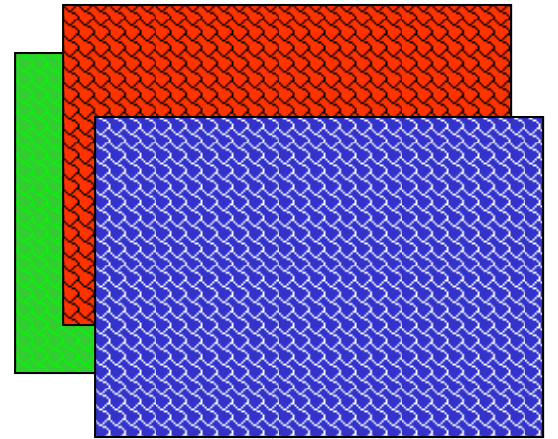
Jan

Des

Nov



Framtida



C
E
T
E
M

Público