O Impacto da Inteligência Artificial e Automação na Pesquisa Científica: Abrangência e Temores



lps COPPE/Poli/UFRJ
Laboratório de Processamento de Sinais
Inteligência Computacional. Inovação

J.M. Seixas
COPPE/Poli - UFRJ

O Impacto da Inteligência Artificial e Robótica no Futuro do Emprego e Trabalho
30 de outubro, 2017

Abrangência

- Aplicações: **Energia** (Smart-cities, Monitoração não-invasiva de carga elétrica, Descargas Parciais, manutenção preditiva, qualidade de energia, Video – Dosimetria (plantas nucleares), Classificação de Petróleos)
- Aplicações: **Defesa** (Tecnologia Sonar – PROSUB; Aplicações Duais – Meio Ambiente)
- Aplicações: **Física** Experimenal de Altas Energias (CERN – Colisão de partículas a cada 25 ns, 60 TB/s. Filtragem Online, Detecção e Estimação de Sinais)
- Aplicações: **Saúde** (Tuberculose, PAF, Engenharia Biomédica)
- **Analytics**: Esporte, Business, Política, Economia, Grandes Eventos, Web mining, Emoções, Gerência
- **Qualidade de dados/informação** (Física, Saúde, Petróleo). Online e Offline
- **IIoT** – Industrial Internet of Things
- **Detecção, Estimação, novidade, anomalia, Remoção de Ruído**
- **Telecomunicações, Transporte**
- **Jornalismo, Redes Sociais: o que está bombando**

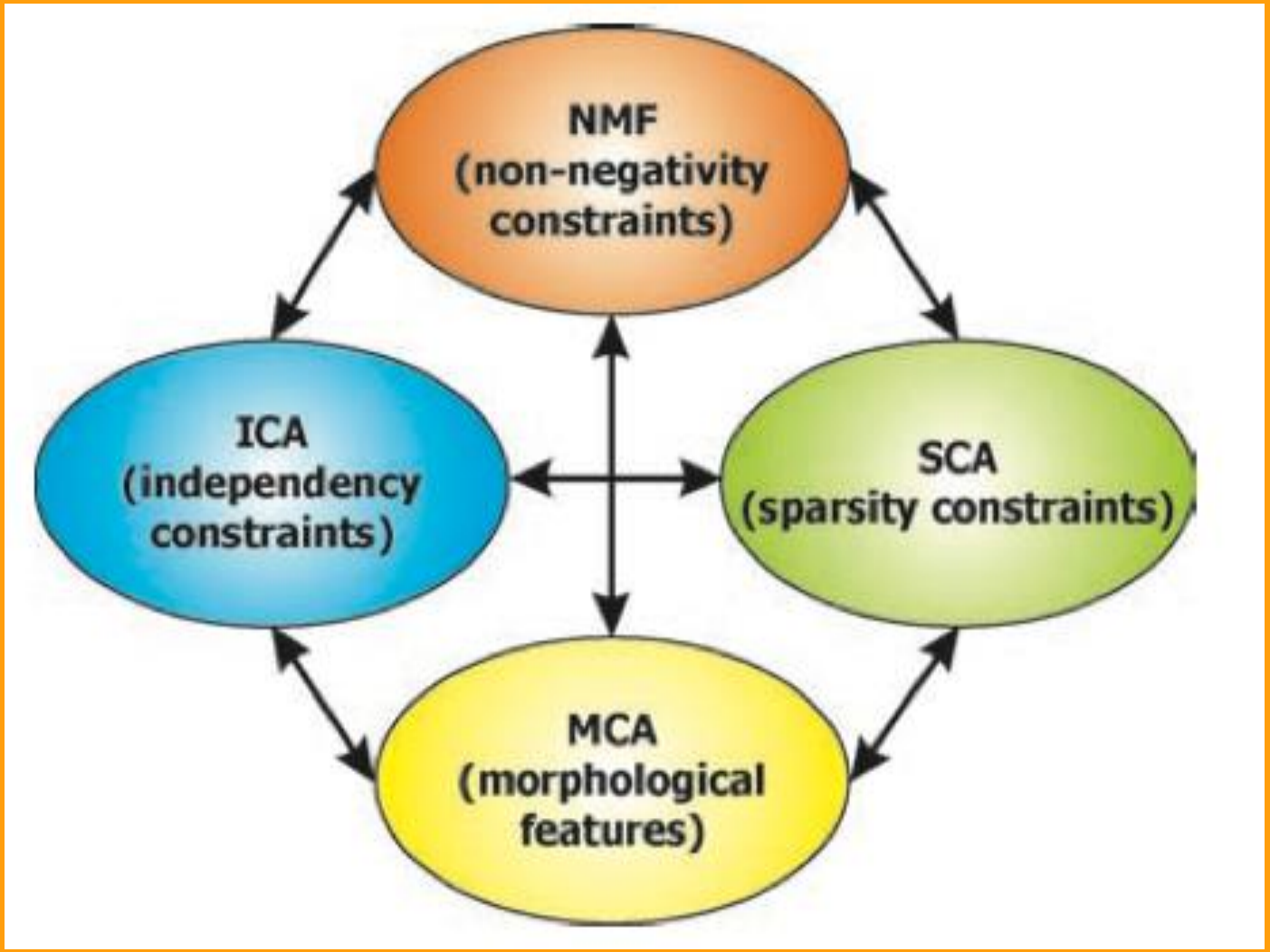
Além de Algoritmos: Representação (não)Linear de Dados Multivariáveis

- **Problema:** como achar uma boa representação para dados multivariáveis?
- **Representação**: transformação dos dados de forma que sua estrutura essencial fique mais visível (ou acessível)
- **Aprendizagem não-supervisionada**: a representação deve ser aprendida através dos próprios dados
- **Aplicações:** Uma boa representação de dados é útil para *data mining*, extração de características, separação de fontes, etc...

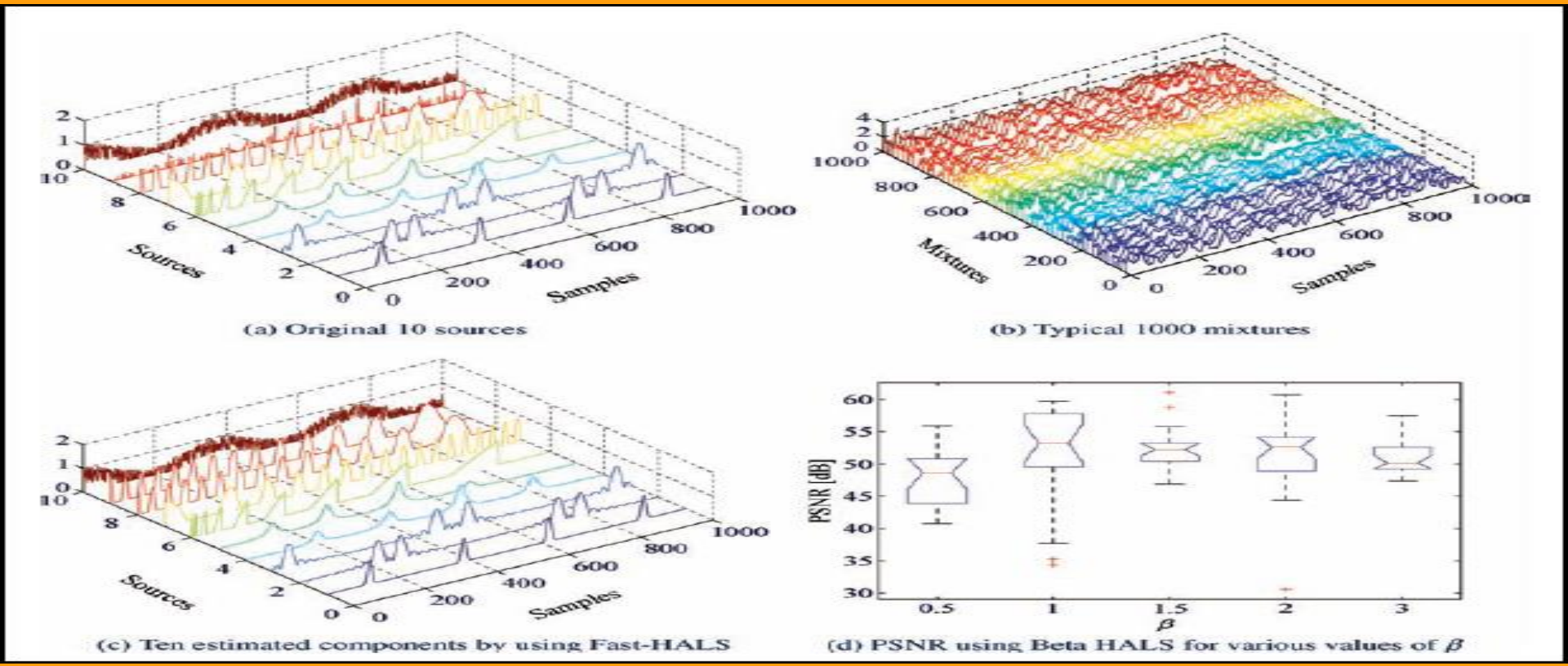
Compactação: Motivação

- Diversas aplicações com alta dimensão no espaço original de dados
  - ✓ Sonar: Informação espectral, 512 componentes
  - ✓ Calorimetria de altas energias: fina granularidade, mais de 1.000 componentes
  - ✓ Sinais de descarga parcial: rápidos, acumulados, mais de 1.000 componentes
  - ✓ Qualidade de energia: ~ 200 amostras
  - ✓ Carga elétrica: previsão, monitoração, ~ 100 amostras
  - ✓ Qualidade de dados: monitoração de séries temporais, pré-processamento, séries explicativas: ~ 100 amostras
  - ✓ Apoio ao diagnóstico médico: dezenas de variáveis; quais são relevantes?

Estrutura do Pré-Processamento



NMF: aplicação



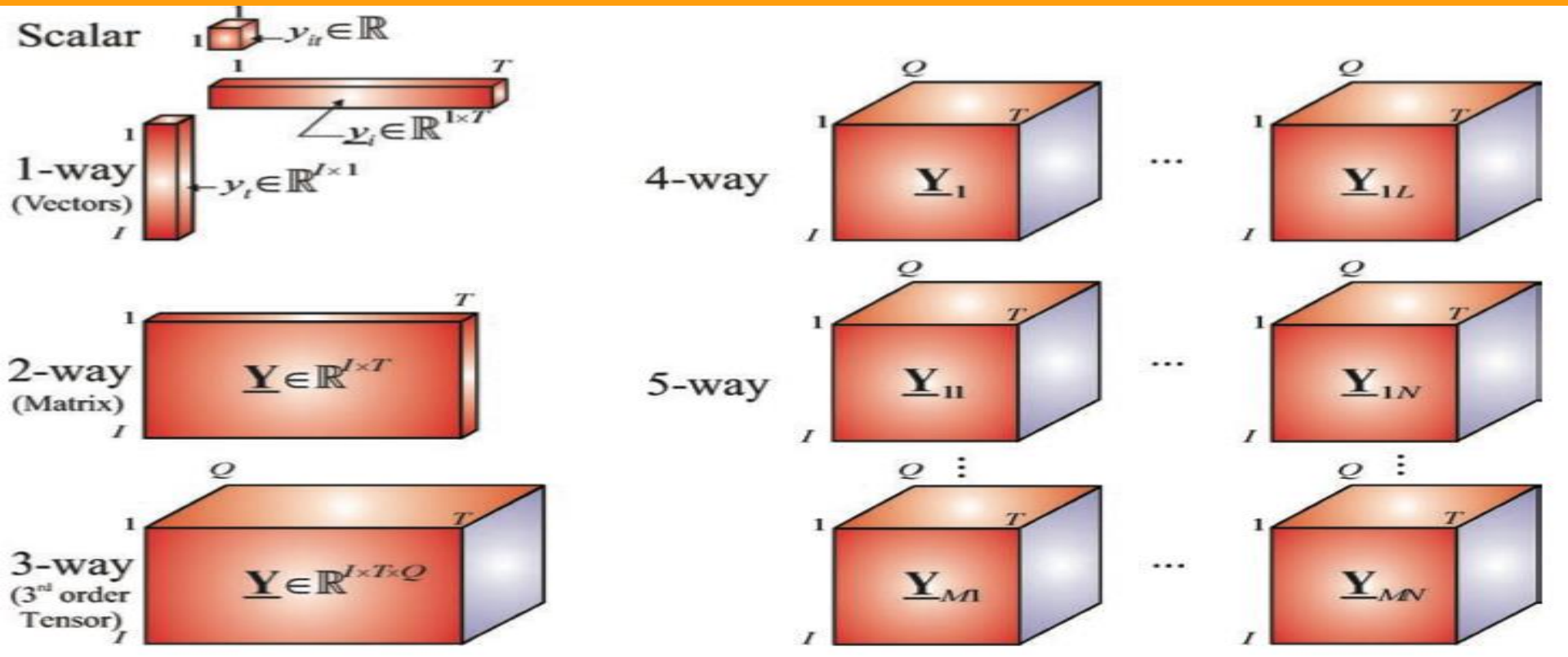
(a) Original 10 sources

(b) Typical 1000 mixtures

(c) Ten estimated components by using Fast-HALS

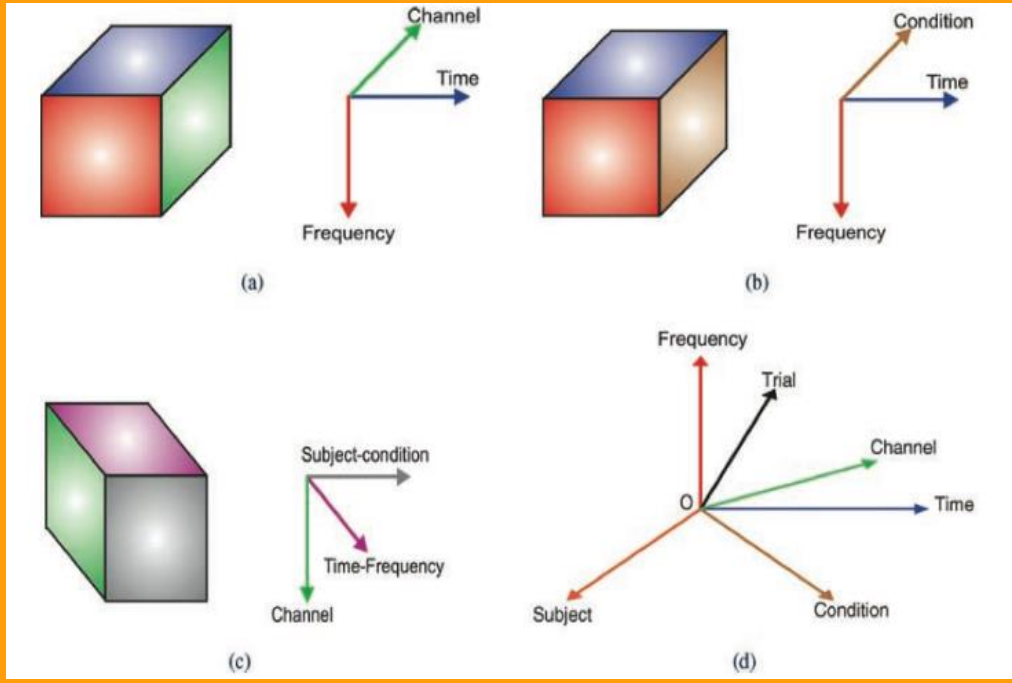
(d) PSNR using Beta HALS for various values of β

Tensores: conceito



Por que tensores?

Muitas aplicações envolvem diferentes dimensões (olhares)
Espaço, tempo-frequência, restrições, canais



Divergência

Mede a similaridade de fontes separáveis

$$D(p \parallel q) = \sum_{i=1}^{n} d(p_i, q_i) \geq 0$$

O resultado é zero apenas para o caso de p=q

Não precisam ser simétricas:

$$D(p \parallel q) = D(q \parallel p)$$

Não precisam satisfazer a inequação triangular

$$D(p \parallel q) \leq D(p \parallel z) + D(z \parallel q)$$

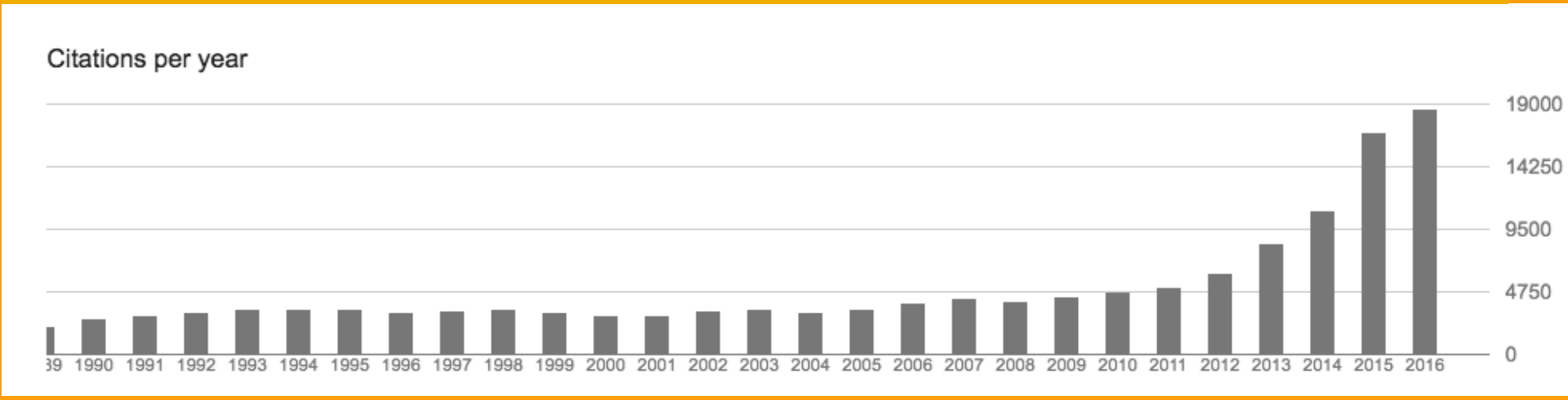
Não é métrica

Hardware/Software

- ✓ FPGA
- ✓ GPU
- ✓ Grid
- ✓ Frameworks (pre-trained)

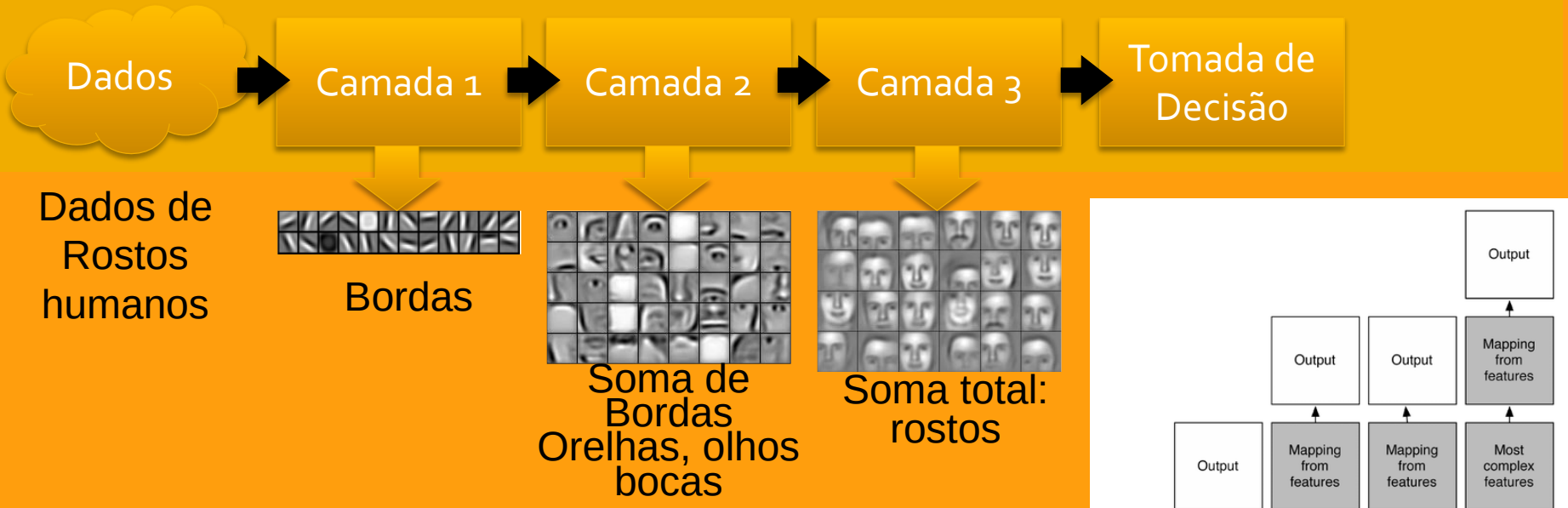
Deep Learning

Modelos baseados em níveis mais profundos de decisão
Níveis mais profundos → Mais parâmetros a serem ajustados
Mais parâmetros → Maior Probabilidade de Perda de generalização



Deep Learning

Modelos mais profundos acessam informações de mais alto nível.

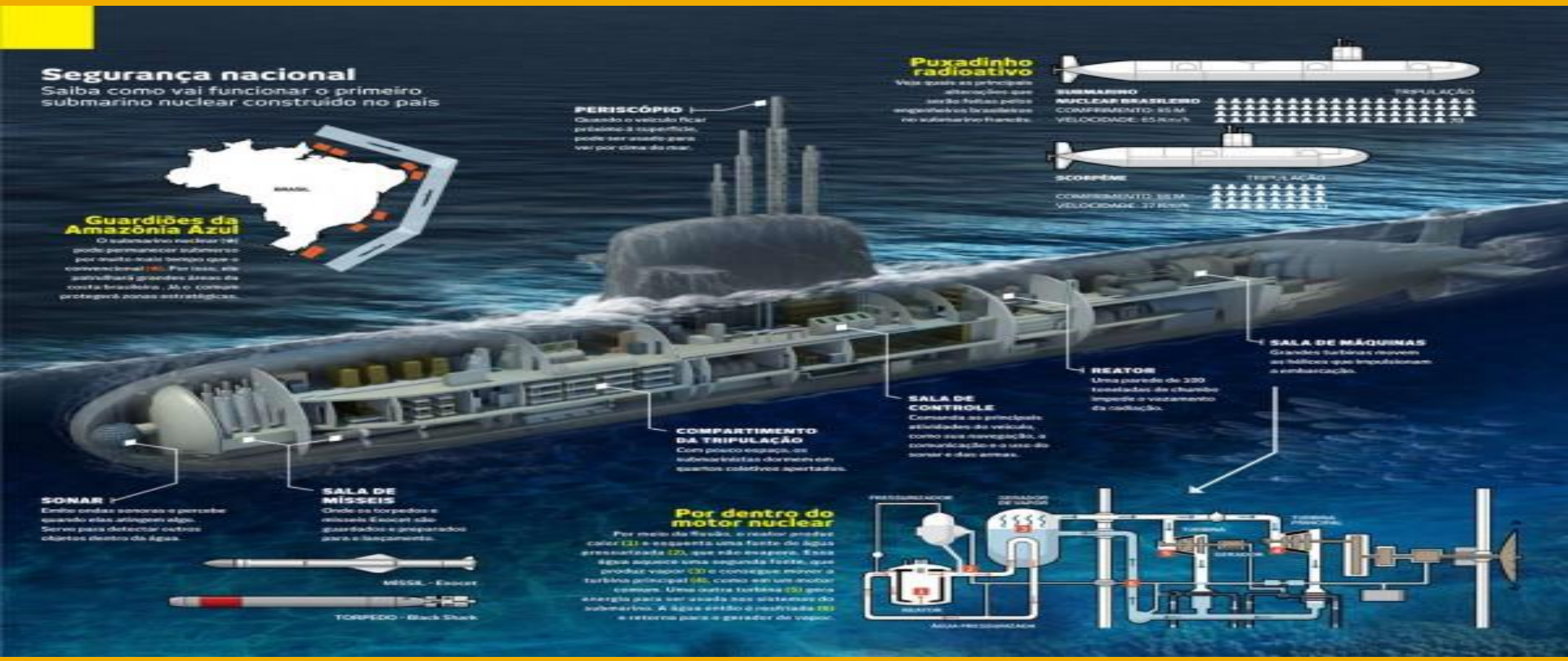


Support Vector Machine

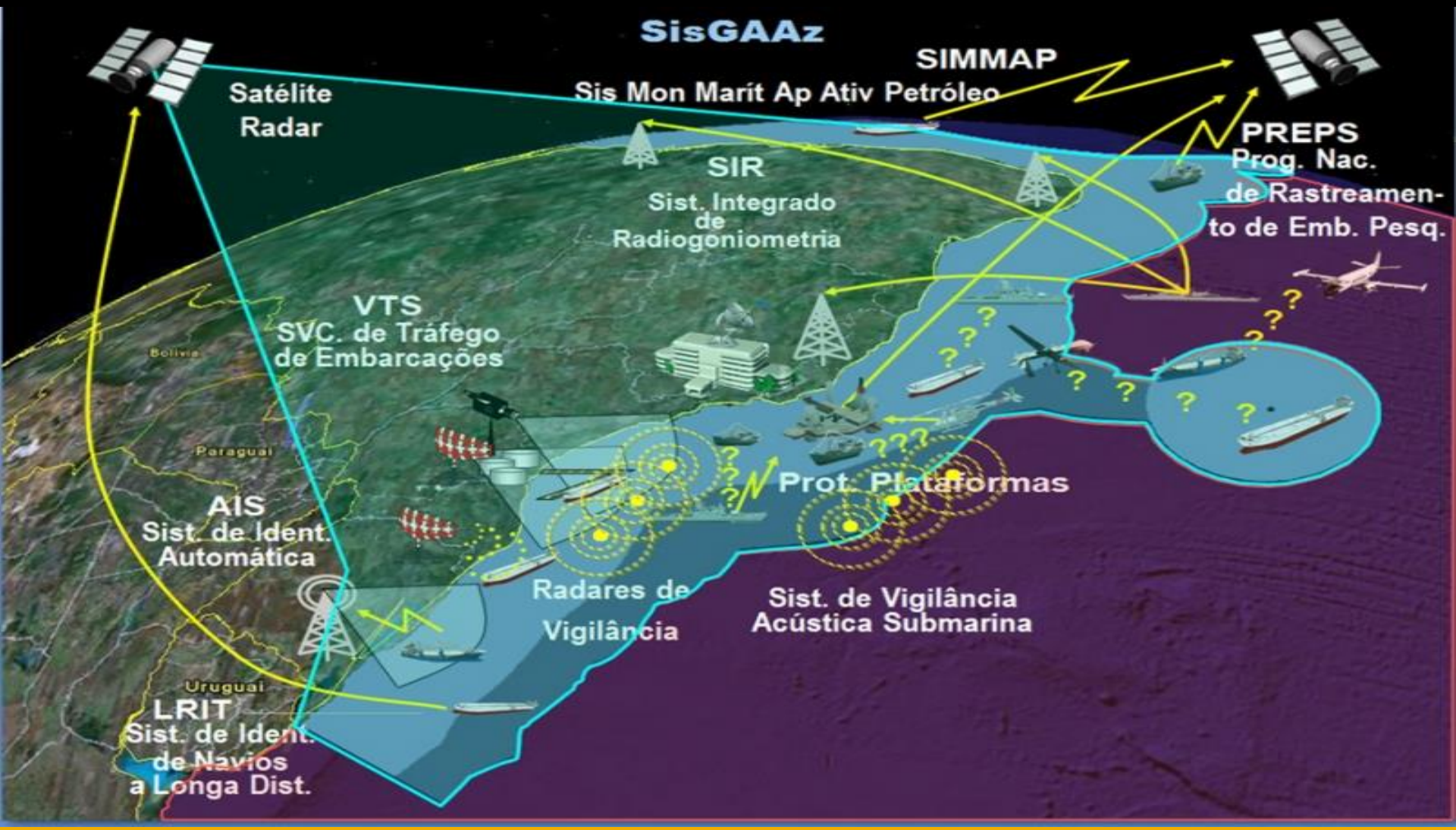
Aumentar a dimensão, para realizar um separador linear

Aproximador universal

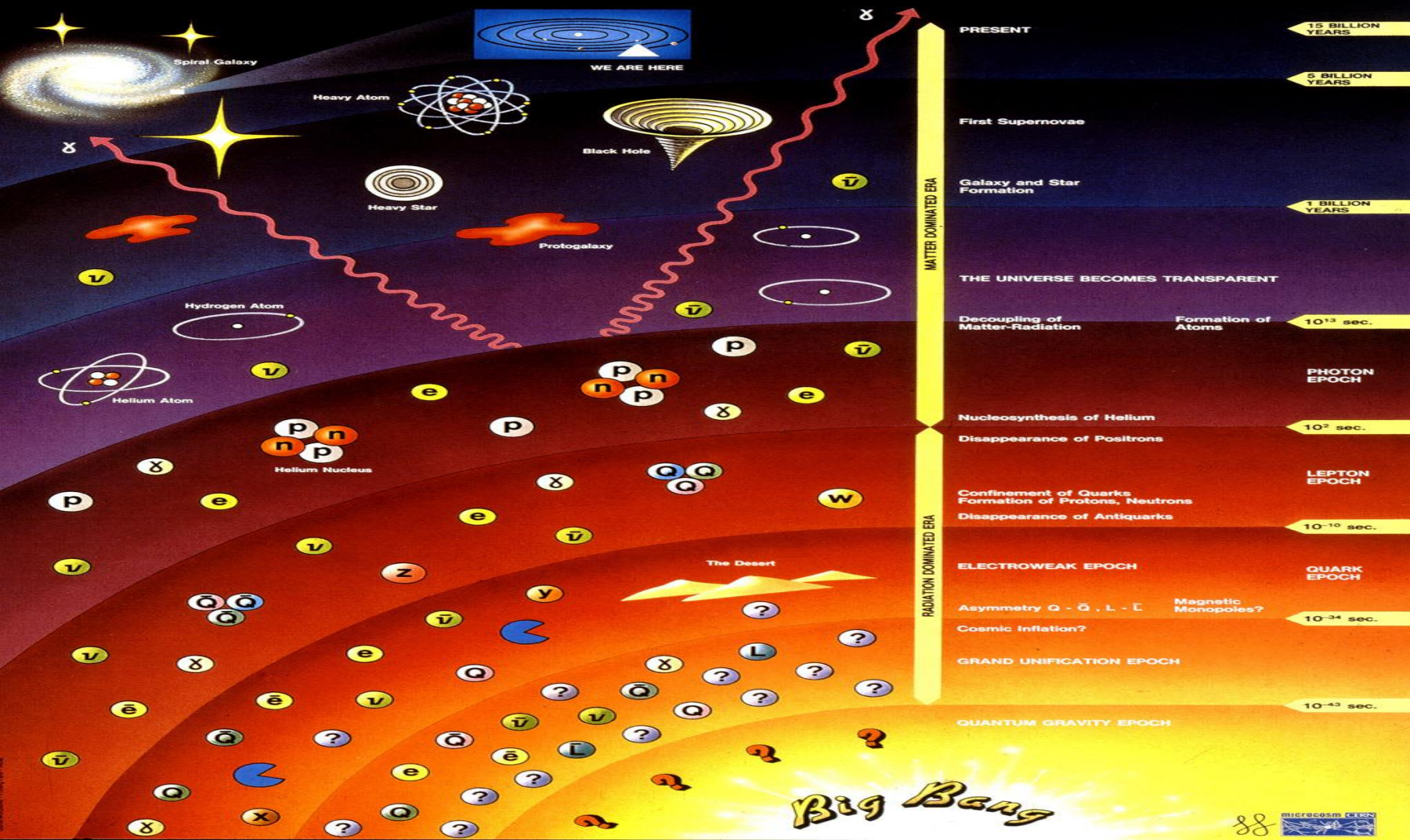
Coppe e Marinha do Brasil juntas para equipar o primeiro submarino nuclear brasileiro (imprensa)



SisGAAz


SisGAAz
Satélite
Radar
SIMMAP
Sis Mon Marít Ap Ativ Petróleo
PREPS
Prog. Nac.
de Rastreamen-
to de Emb. Pesq.
SIR
Sist. Integrado
de
Radiogoniometria
VTS
SVC. de Tráfego
de Embarcações
Bolívia
Paraguai
AIS
Sist. de Ident.
Automática
Prot. Plataformas
Radares de
Vigilância
Sist. de Vigilância
Acústica Submarina
Uruguai
LRIT
Sist. de Ident.
de Navios
a Longa Dist.

History of the Universe

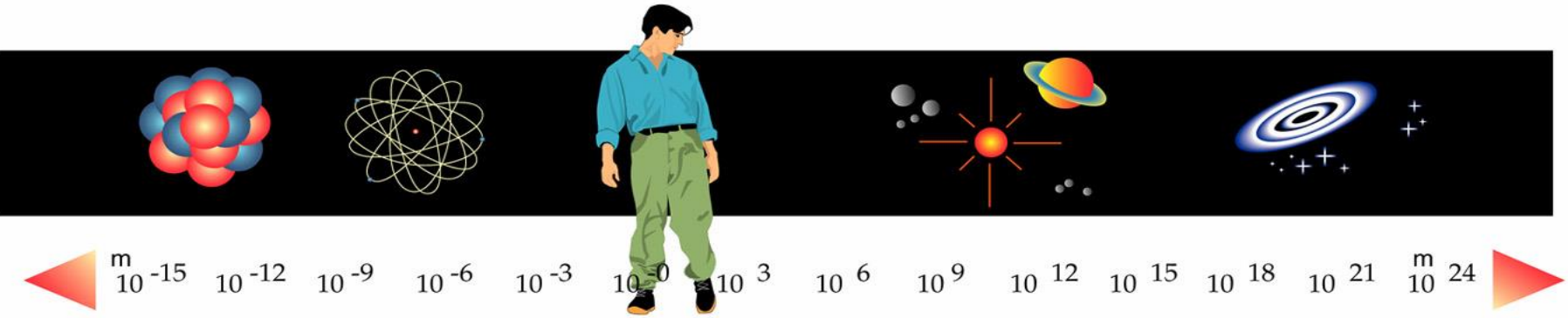


La physique des particules étudie la matière dans ses dimensions les plus petites.

Particle physics looks at matter in its smallest dimensions.

L'astrophysique étudie la matière dans ses dimensions les plus grandes.

Astrophysics looks at matter in its largest dimensions.



m
10^{-15} 10^{-12} 10^{-9} 10^{-6} 10^{-3} 10^{0} 10^{3} 10^{6} 10^{9} 10^{12} 10^{15} 10^{18} 10^{21} m 10^{24}

Microscopes
Microscopes

Jumelles
Binoculars

Telescopes optiques & radio
Optical & radio telescopes

L'oeil nu.
Naked eye

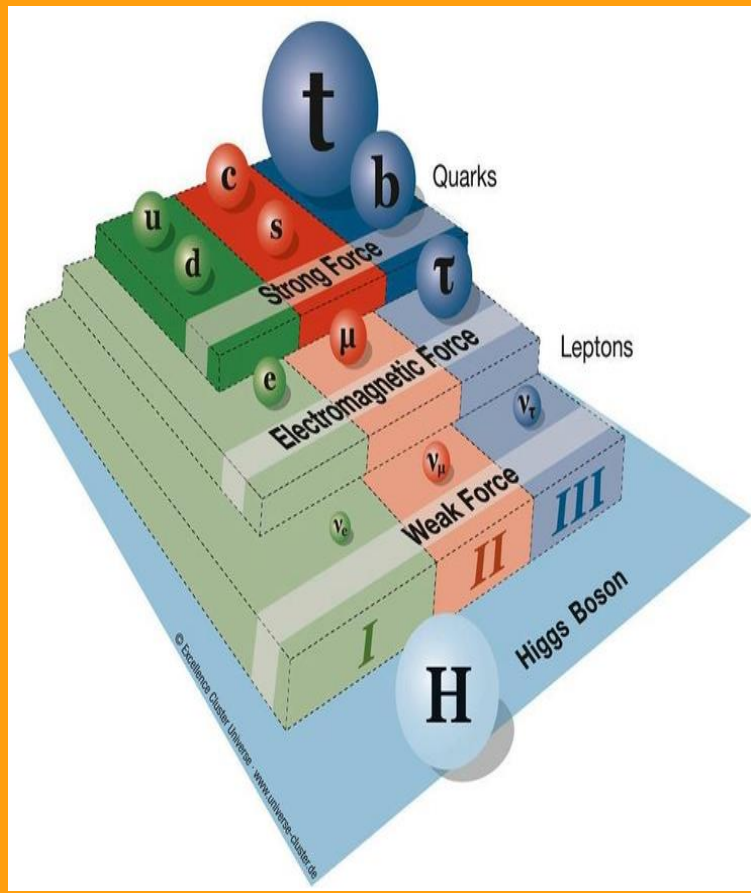
Accélérateurs et détecteurs
Accelerators and detectors

THE TWO FRONTIERS OF PHYSICS

LES DEUX FRONTIERES DE LA PHYSIQUE

O Modelo Padrão

- A matéria é compostas de partículas elementares.
  - antipartículas: mesmas propriedades, mas carga (e spin, etc.) simétricos.
  - Já foram verificadas, mas elas decaem rapidamente.
- Interações mediadas por:
  - Eletromagnéticas (fóton);
  - Nucleares Forte (glúons);
  - Nucleares Fracas ($W^{\pm}$, Z);
  - Gravitacional (gráviton)*.

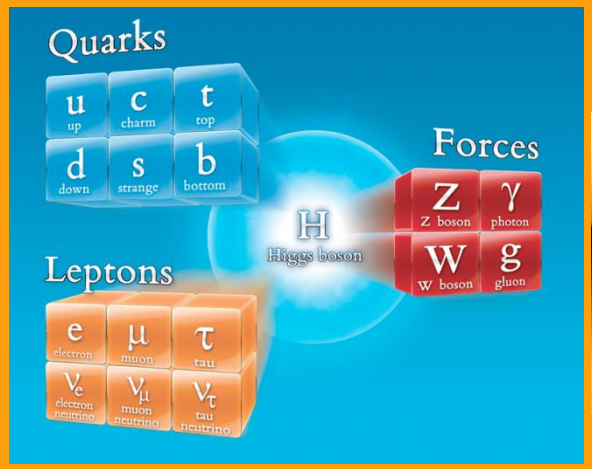


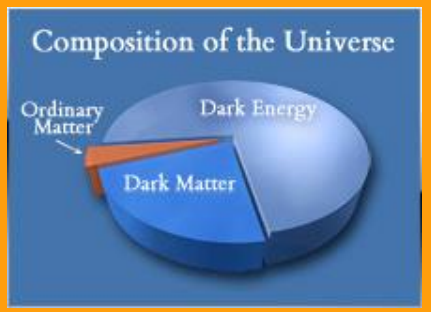
* Ainda não foi possível detectar o gráviton.

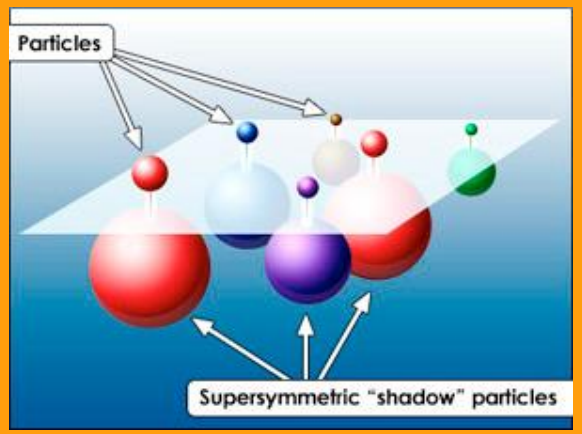
O que mais se procura?

- A Matéria e a Energia Escura : A maior parte da matéria do Universo não interage com a luz e por isso é invisível para nós. O que constitui essa matéria escura?!
- Existem companheiros “super-simétricos” para as partículas do modelo padrão?!
- Algumas teorias sugerem a existência de múltiplas dimensões. Seria essa a causa da “fraqueza” da força gravitacional?

Modelo Padrão e bóson de Higgs







Desafios

Matéria Escura:

Na cosmologia, matéria escura é uma forma postulada de matéria "invisível".

Sua presença pode ser inferida a partir de efeitos gravitacionais sobre a matéria visível.

Resultados recentes da AstroParticlePhysics indicam que o Universo é feito:

5% de matéria conhecida.

25% de Matéria Escura (Não há partícula conhecida para explicá-la)

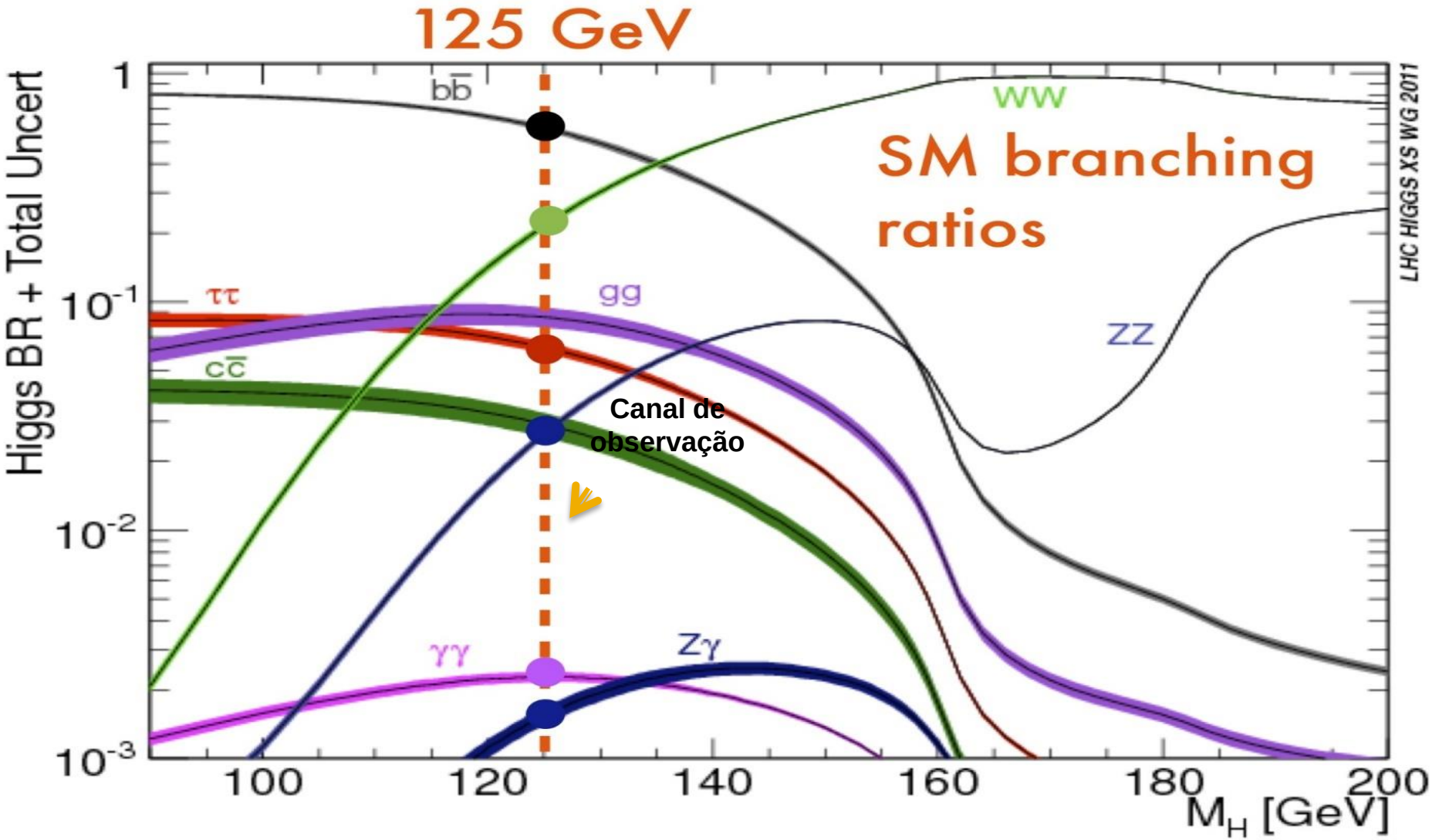
70% of "Energia Escura"

Energia Escura:

***A energia escura* é uma forma hipotética de energia que estaria distribuída por todo espaço e tende a acelerar a expansão do Universo.**

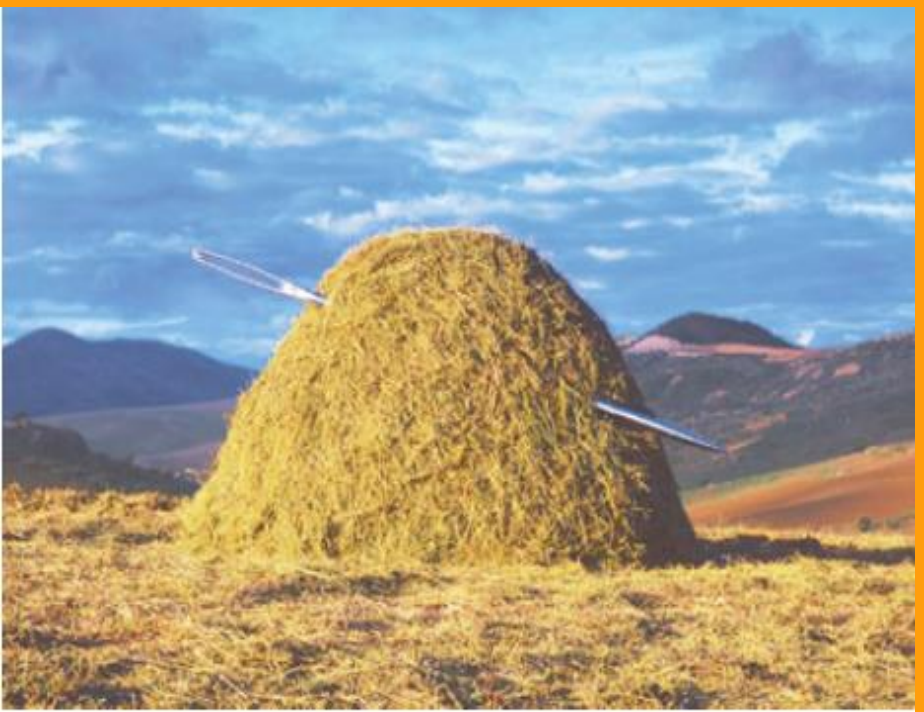
"Entendemos" somente 5% da composição do Universo.

O Bóson de Higgs



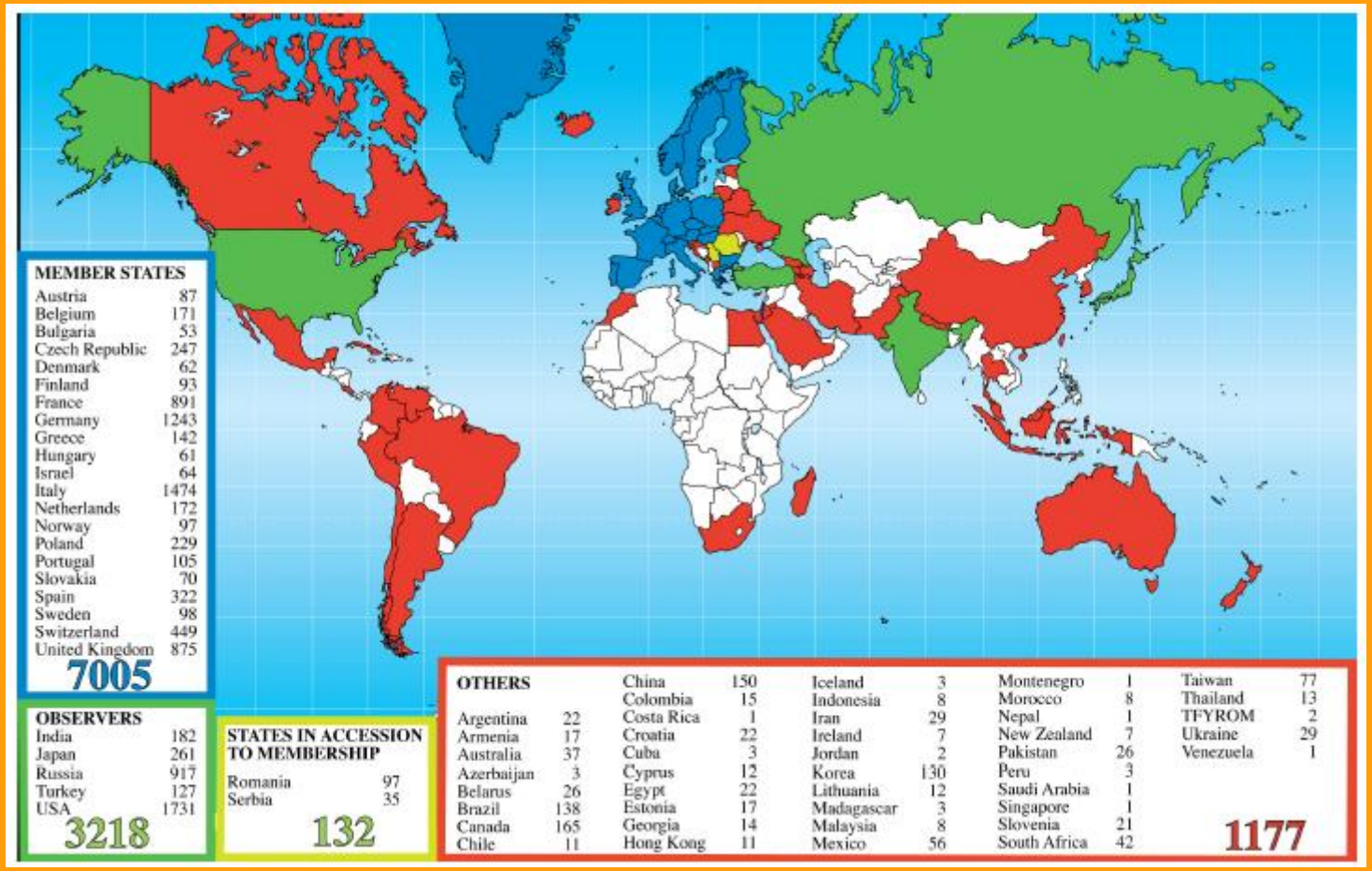
O Bóson de Higgs ou um Bóson de Higgs?

Da desoberta ao detalhamento

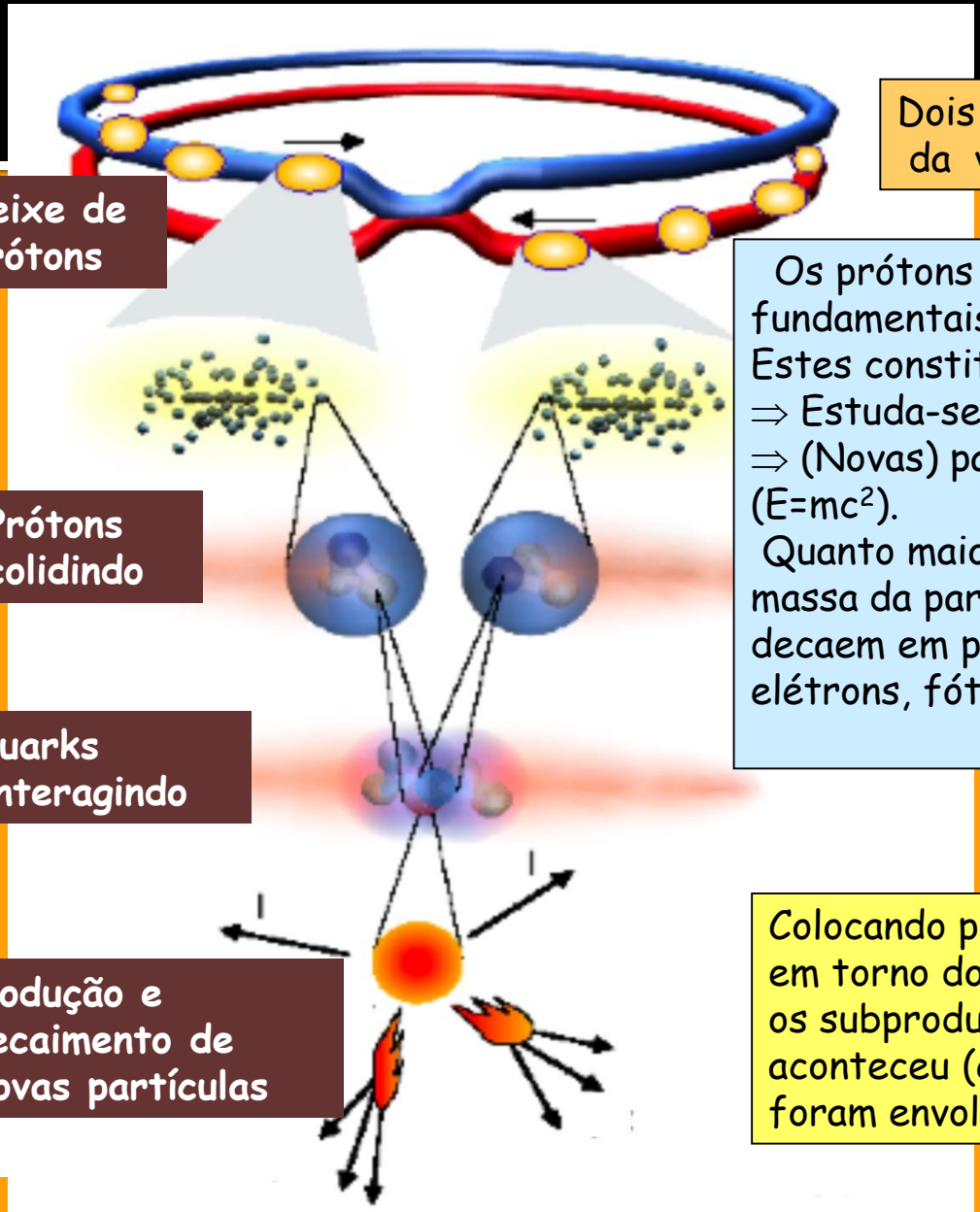


Para detectar eventos ainda mais raros: dificuldade aumenta, empilhamento de sinais.

O CERN


MEMBER STATES
Austria 87
Belgium 171
Bulgaria 53
Czech Republic 247
Denmark 62
Finland 93
France 891
Germany 1243
Greece 142
Hungary 61
Israel 64
Italy 1474
Netherlands 172
Norway 97
Poland 229
Portugal 105
Slovakia 70
Spain 322
Sweden 98
Switzerland 449
United Kingdom 875
7005
OBSERVERS
India 182
Japan 261
Russia 917
Turkey 127
USA 1731
3218
STATES IN ACCESSION TO MEMBERSHIP
Romania 97
Serbia 35
132
OTHERS
Argentina 22
Armenia 17
Australia 37
Azerbaijan 3
Belarus 26
Brazil 138
Canada 165
Chile 11
China 150
Colombia 15
Costa Rica 1
Croatia 22
Cuba 3
Cyprus 12
Egypt 22
Estonia 17
Georgia 14
Hong Kong 11
Iceland 3
Indonesia 8
Iran 29
Ireland 7
Jordan 2
Korea 130
Lithuania 12
Madagascar 3
Malaysia 8
Mexico 56
Montenegro 1
Morocco 8
Nepal 1
New Zealand 7
Pakistan 26
Peru 3
Saudi Arabia 1
Singapore 1
Slovenia 21
South Africa 42
Taiwan 77
Thailand 13
TFYROM 2
Ukraine 29
Venezuela 1
1177

Para estudar as partículas Elementares e suas interações precisamos:

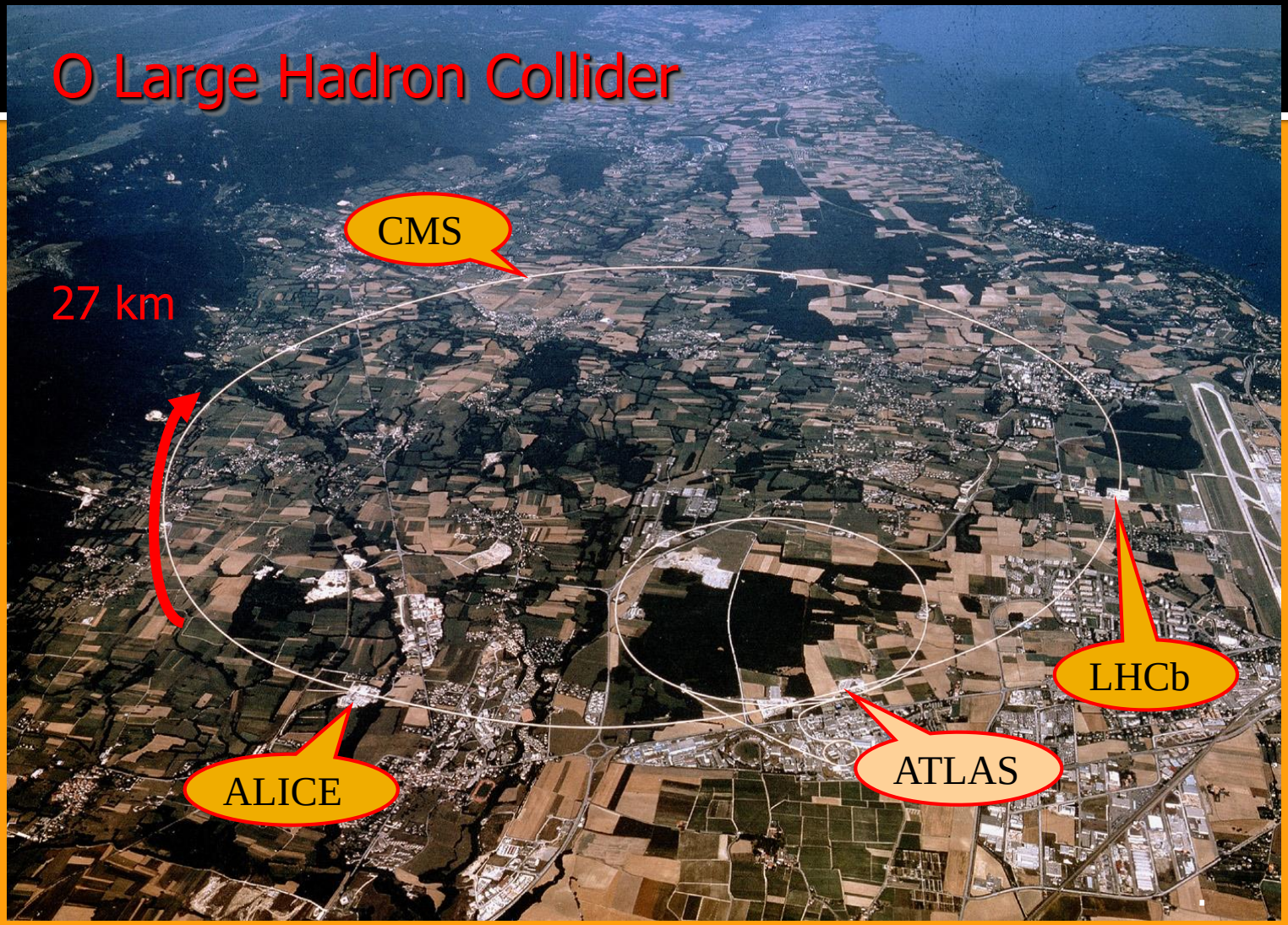


Dois feixes de particles (prótons) próximos da velocidade da luz colidem.

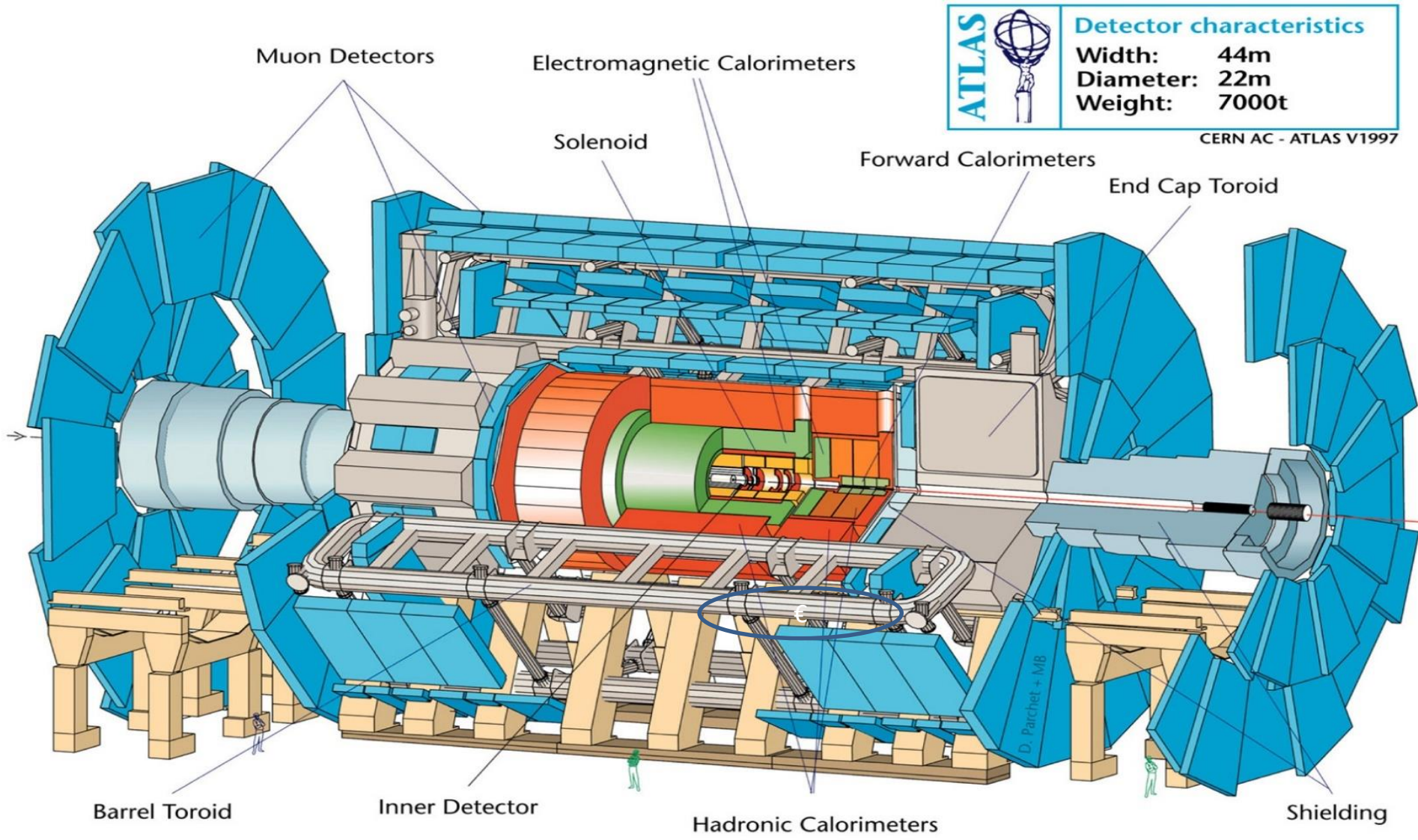
Os prótons colidindo quebram-se em seus constituintes fundamentais (p.e. quarks).
Estes constituientes interagem a altas energias:
⇒ Estuda-se como a matéria fundamental se comporta.
⇒ (Novas) partículas pesadas podem ser produzidas ($E=mc^2$).
Quanto maior a energia do acelerador, maior pode ser a massa da partícula produzida. Essas partículas então decaem em partículas (conhecidas) de menor massa: elétrons, fótons, etc.

Colocando poderosos detectores de alta tecnologia em torno dos pontos de colisão, podemos detectar os subprodutos da colisão e reconstruir o que aconteceu (quais fenomenos, partículas, forças, etc. foram envolvidas)

O Large Hadron Collider

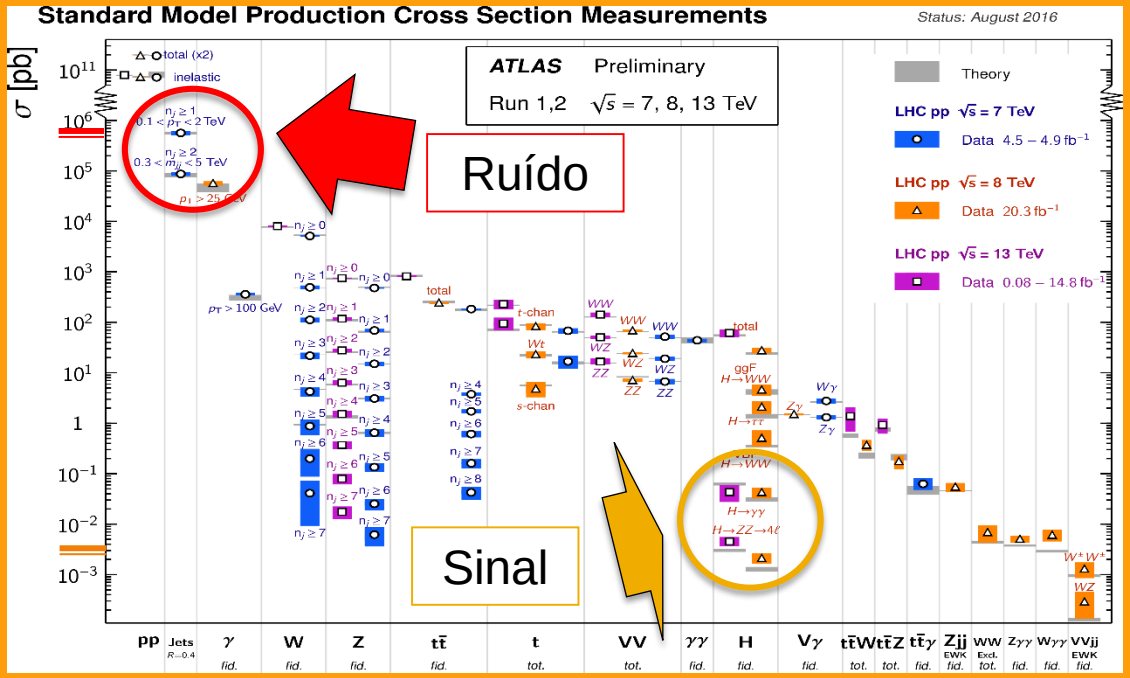


Experimento ATLAS



Aplicação de Redes Neurais: NeuralRinger

Proporcional a taxa de produção

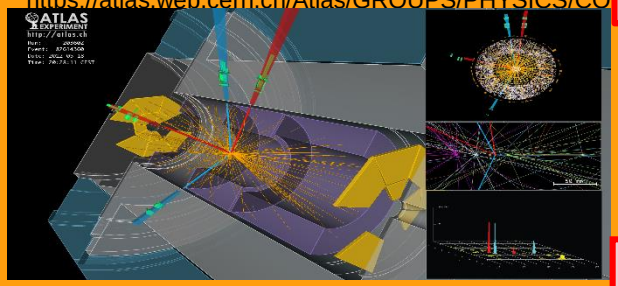


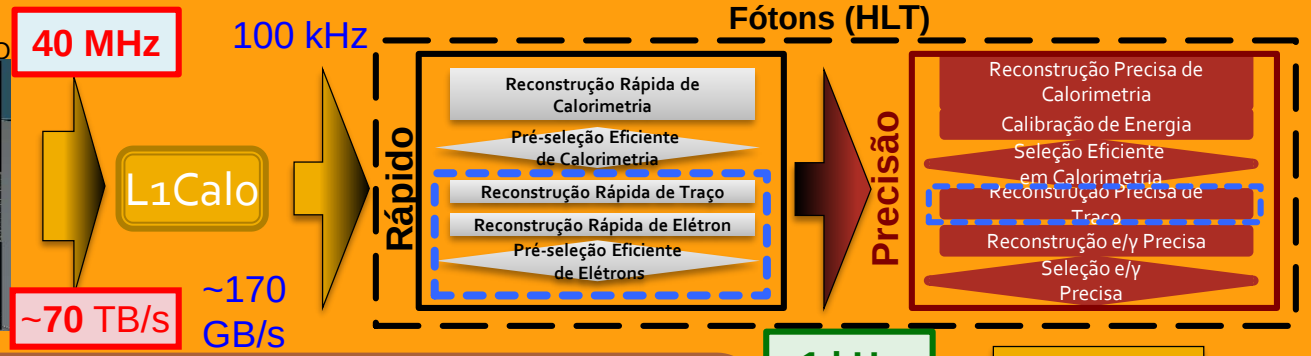
Fonte: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/ATLAS_b_SMSummary_FiducialXsect/ATLAS_b_SMSummary_FiducialXsect.png

- Um dos canais (em laranja) que levou a descoberta do Higgs em 2012;
- Para a observação do Higgs nesse canal, faz-se necessário a identificação de **elétrons (sinal)**
- Grande parte dos eventos gerados pelo LHC constitui-se de **jatos hadrônicos (ruído físico)**;
- Para 1 decaimento do Higgs citado, geram-se, em média, ~100 milhões de jatos;

Neural Ringer: Sistema de Filtragem Online do ATLAS

Fonte: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CO





O objetivo do **Sistema de Filtragem** do ATLAS é reduzir a taxa de eventos (40 MHz) para níveis viáveis de armazenamento (1 kHz), enquanto mantendo os eventos de interesse (i.e. eventos contendo elétrons como estados finais).

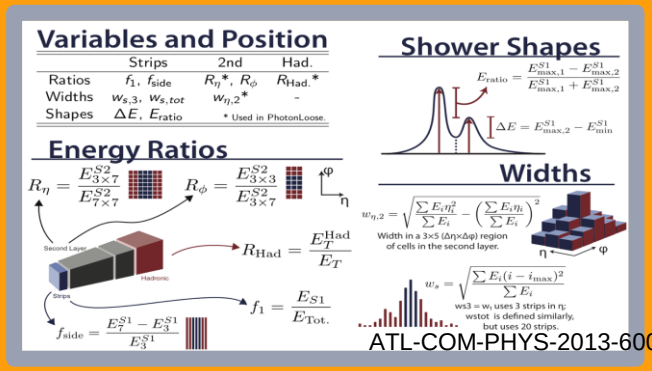


Algoritmos de Identificação de Elétrons

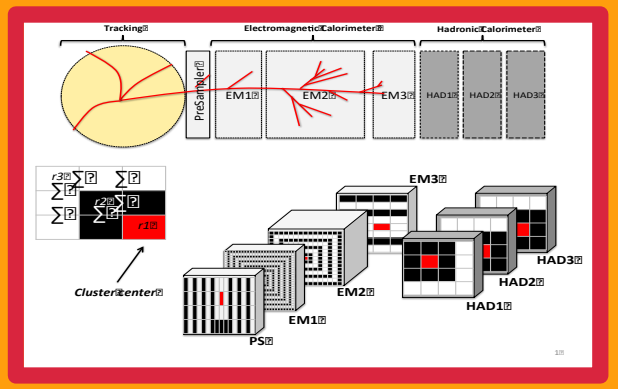
CutID e Likelihood

Ringer

Reconstrução de Calorimetria



OU



Reconstrução de Traço

Informação pura de traço
+
Combinação de traço & padrões de calorimetria

Classificadores

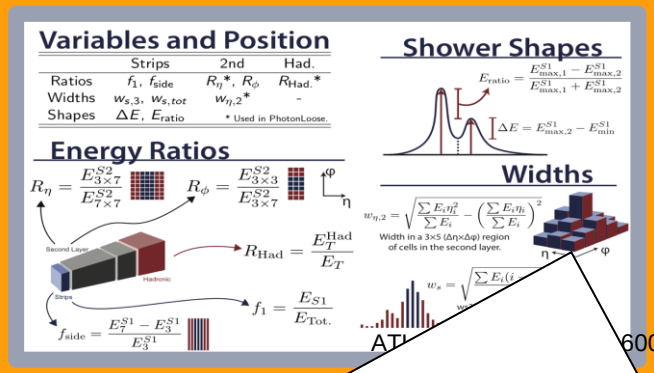
Cortes Rígidos	Likelihood	Ringer
Aplica cortes lineares sobre variáveis tradicionais	Aplica abordagem similar a Naïve Bayes sobre variáveis tradicionais	Aplica redes neurais alimentadas pelos aneis.

Algoritmos de Identificação de Elétrons

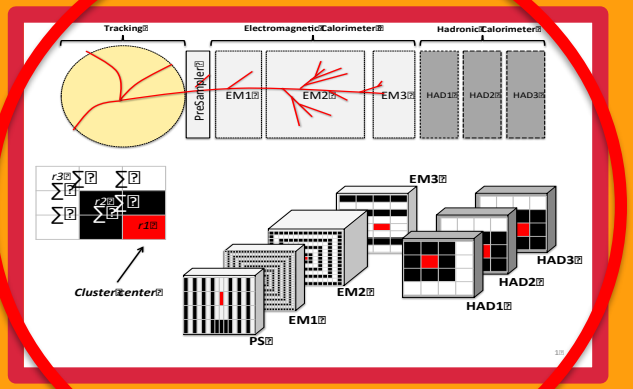
CutID e Likelihood

Ringer

Reconstrução de Calorimetria



OU



600

Reconstrução de Traço

ara de traço

+

e traço & padrões de calorimetria

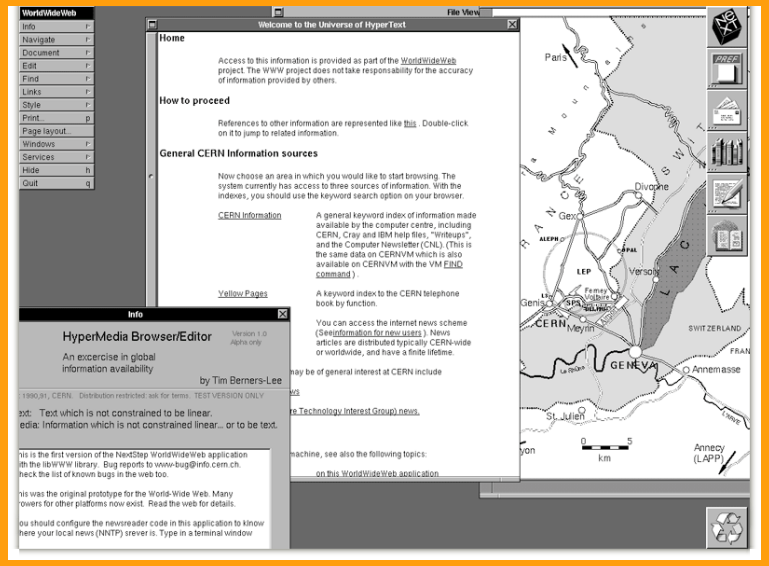
Nova abordagem

Classificadores

gidos	**Likelihood**	**Ringer**
cortes lineares sobre variáveis tradicionais	Aplica abordagem similar a Naïve Bayes sobre variáveis tradicionais	Aplica redes neurais alimentadas pelos anéis.

produtos derivados

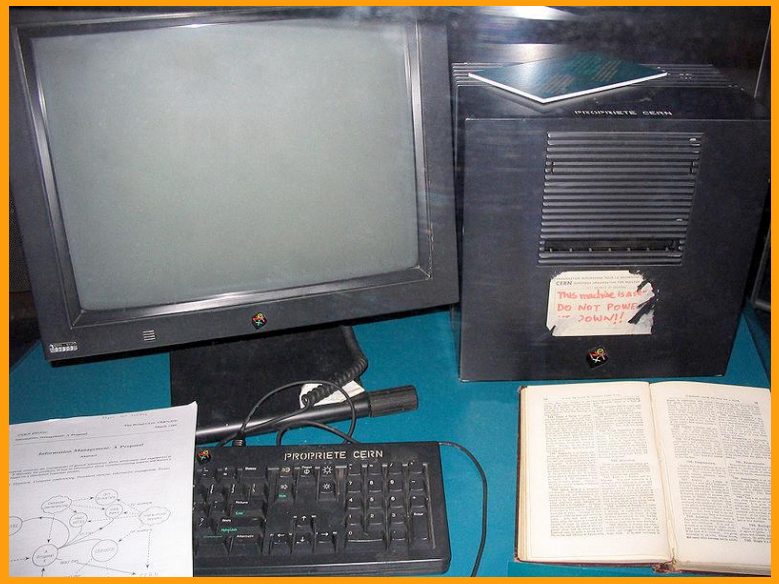
Primeira web page - conceito de hyperlink



Tim Berners-Lee inventor da world wide web



Primeiro web server!!



- Para facilitar o intercâmbio de informações, os cientistas precisam um método para misturar texto e imagens.
- Ou a possibilidade de navegar facilmente de uma página pra outra (hyperlink).
- O conceito de web page começou no CERN e agora é um instrumento comum no mundo todo produzindo uma nova era de conhecimento e acesso à informação.

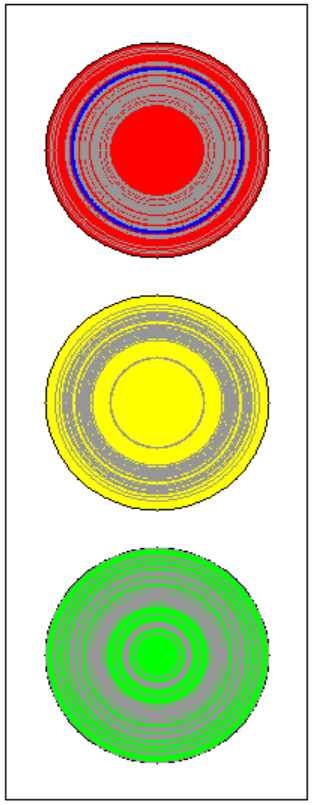
Saúde

- Amiloidose (degenerativa)
- Tuberculose
  - Pulmonar
  - Pleural
  - Gerência de leitos de isolamento
  - Co-morbidades (Aids, Diabetes)
  - Infantil
  - Multirresistente
  - Abandono de tratamento
  - Bovina

Sistema NeuralTB: Apoio ao Diagnóstico

http://www04.lps.ufrj.br/~neuraltb/cgi-bin/geraGraficoDeGrupo.cgi?grupo=2&raio=0.715501&id=0

Resultado



Marcos Gomes Silva

Saída da RN: probabilidade e grupo de risco

Legenda

- Alto risco
- Médio risco
- Baixo risco
- Densidade de pacientes
- Paciente cadastrado

Fechar

DADOS DO PACIENTE

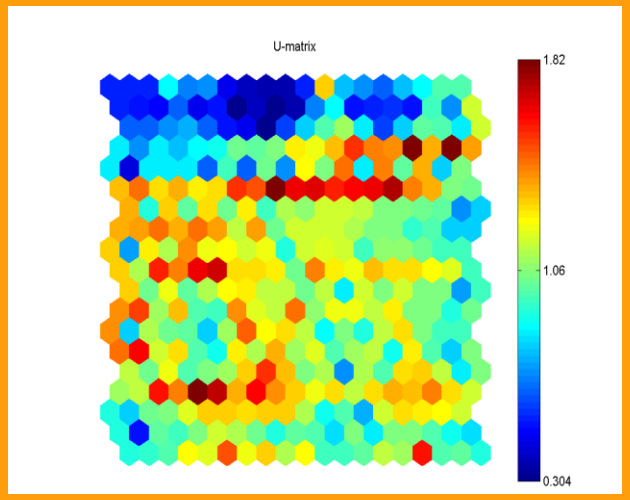
Data de nascimento: XX/XX/1992
Idade: 15
Sexo: Masculino
Tosse: Ignorado
Hemoptóico: Ignorado
Sudorese: Ignorado
Febre: Ignorado
Emagrecimento: Ignorado
Dispnéia: Ignorado
Anorexia: Ignorado
Fuma Atualmente: Ignorado
TB Extrapulmonar: Ignorado
Internação Hospitalar: Ignorado
SIDA: Ignorado
Grupo de Risco: Alto

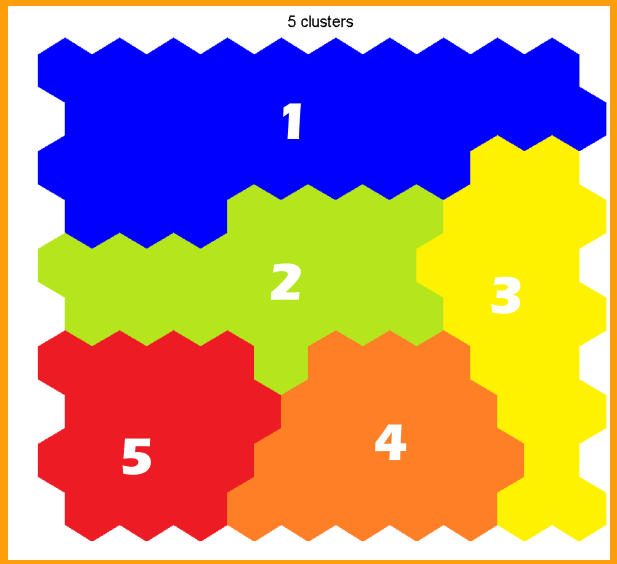
Paciente com tuberculose a nivel de 92,1%

Concluído | Internet | Modo Protegido: Ativado | 100%

Escore de gravidade para PAF usando SOM

Através da utilização de Mapas Auto-Organizáveis *(SOM)*, os dados dos pacientes com PAF *(Polineuropatia Amiloidótica Familiar)* do HU da UFRJ foram utilizados para criar um escore de gravidade para a doença.





Fatos

- China lança plano para virar líder em inteligência artificial até 2030
  - ✓ Criar indústria de $ 1.5 trilhão (crescimento do PIB em 26%)
  - ✓ China produziu mais artigos de Aprendizado Profundo em 2016 do que os EUA
  - ✓ Pesquisadores chineses estão voltando para se tornarem líderes do processo
  - ✓ Temor: desemprego em massa
- Carta de Copenhagen

Temores

- Homo Deus (Yuval Noah Harari)
  - ✓ AI God (Way of Future – this God will actually exist...)
- Rouba os empregos de mais alto nível
- Desemprego: muitos não conseguirão se realocar
  - ✓ Deslocamento já começou, mas vai se intensificar
  - ✓ Difícil escolher uma carreira sem perspectiva de automação
  - ✓ Computadores podem substituir pesquisadores (IBM Watson)
  - ✓ Menor estabilidade no emprego
  - ✓ O que acontece, quando a força de trabalho não é tão necessária?
- Educação: estamos educando para o século XX. Instituições não estão preparadas para lidar com estas mudanças
- Singularity: inteligência super-humana (máquinas superam a capacidade humana. O mundo se tornará incompreensível para cérebros racionais)
- Em 'Her', o protagonista se apaixona por um sistema operacional inteligente

Conclusões

- ✓ Impactos na pesquisa e na sociedade: de forma abrangente
- ✓ Ambientes de aplicação mais exigentes se curvam à análise multivariada e modelos inteligentes para enfrentar os desafios
- ✓ Software e hardware estão cada vez mais disponíveis
- ✓ As máquinas inteligentes podem representar um perigo?
- ✓ Data scientist x desemprego em massa
- ✓ Imigrantes?