



25 ANOS
1991-2016

eurolab
Portugal

Simpósio

Desafios da Inovação

Lisboa, 9 de maio de 2017

LNEG - Alfragide

Informação e/ou Desinformação

Uma mentira dá uma volta inteira ao mundo antes
mesmo de a verdade ter oportunidade de se vestir.



BIG DATA

Embora o termo "big data" seja relativamente novo, o ato de recolher e armazenar grandes quantidades de informações para eventual análise de dados é bem antigo. O conceito ganhou força no início dos anos 2000, quando um analista famoso deste setor, Doug Laney, articulou a definição de big data como os três Vs:

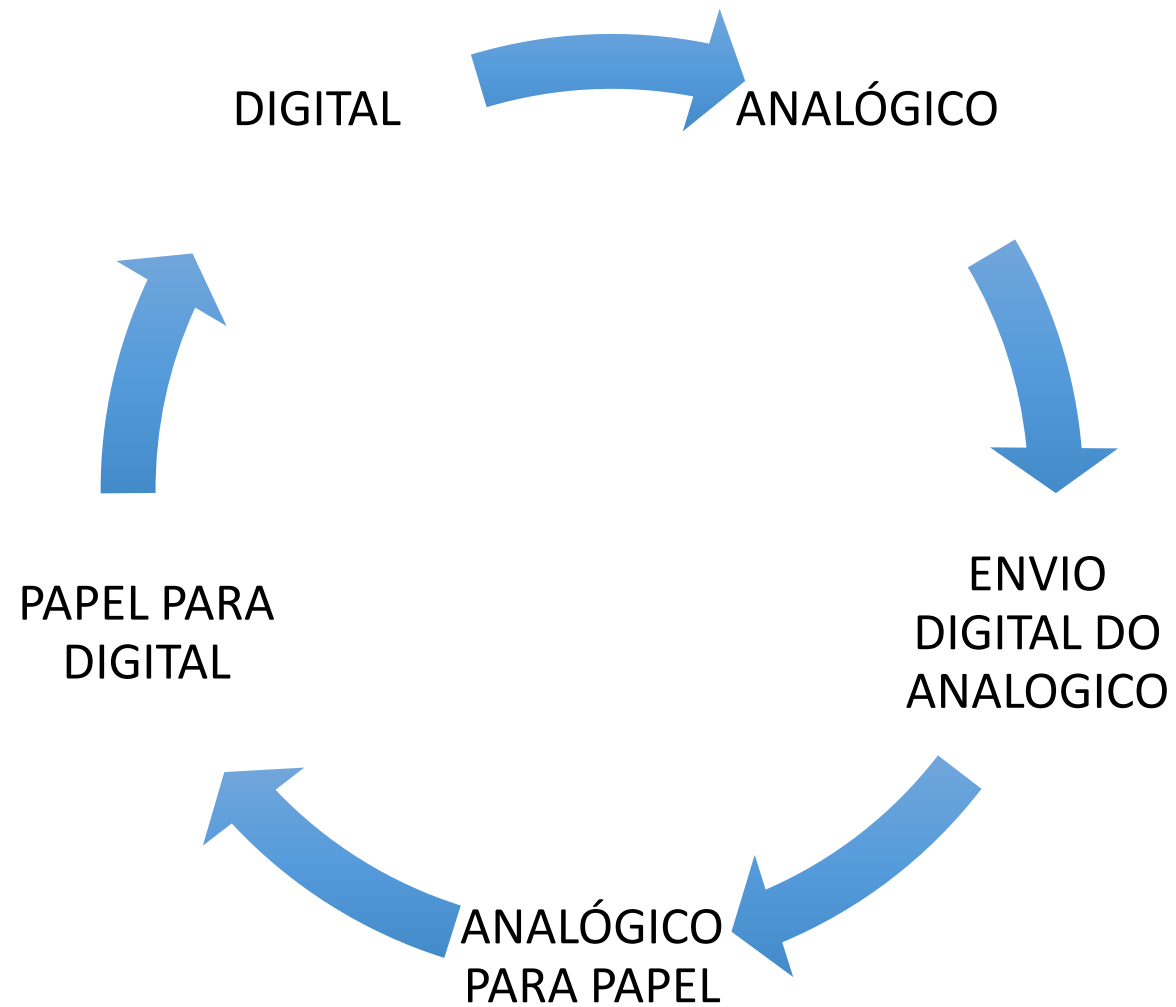
Volume. Organizações coletam dados de uma grande variedade de fontes, incluindo transações comerciais, redes sociais e informações de sensores ou dados transmitidos de máquina a máquina. No passado, armazenar tamanha quantidade de informações teria sido um problema.

Velocidade. Os dados fluem em uma velocidade sem precedentes e devem ser tratados em tempo hábil. Tags de RFID, sensores, celulares e contadores inteligentes estão impulsionando a necessidade de lidar com imensas quantidades de dados em tempo real, ou quase real.

Variedade. Os dados são gerados em todos os tipos de formatos - de dados estruturados, dados numéricos em bancos de dados tradicionais, até documentos de texto não estruturados, e-mail, vídeo, áudio, dados de cotações da bolsa e transações financeiras.



1. 3261 RELATÓRIOS ANO 2016
2. CADA RELATÓRIO 8 PARAMETROS
3. 26088 PARAMETROS ANUAIS
4. ENVIADOS PARA A APA E PARA AS CCDRS COMPETENTES
5. 30 LABORATÓRIOS EM PORTUGAL



LAB A

- nINFORMAÇÃO

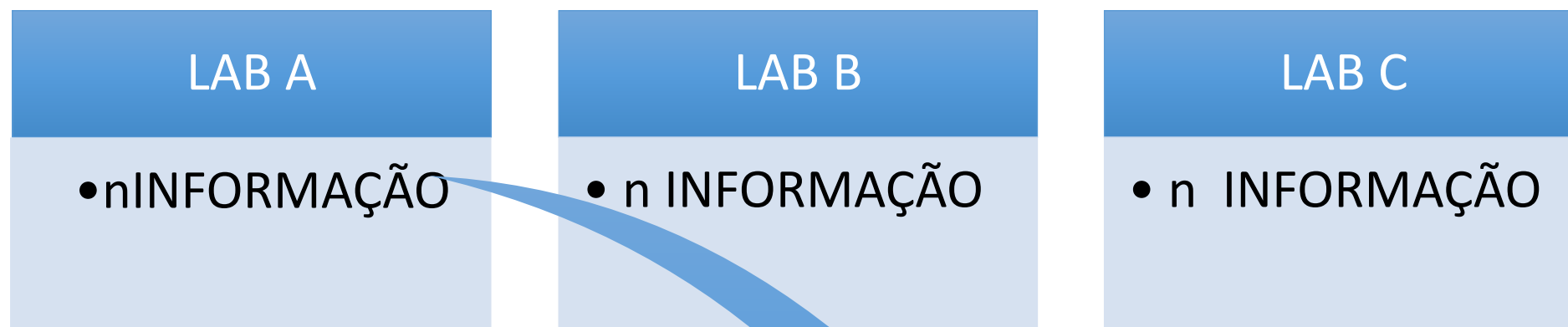
LAB B

- n INFORMAÇÃO

LAB C

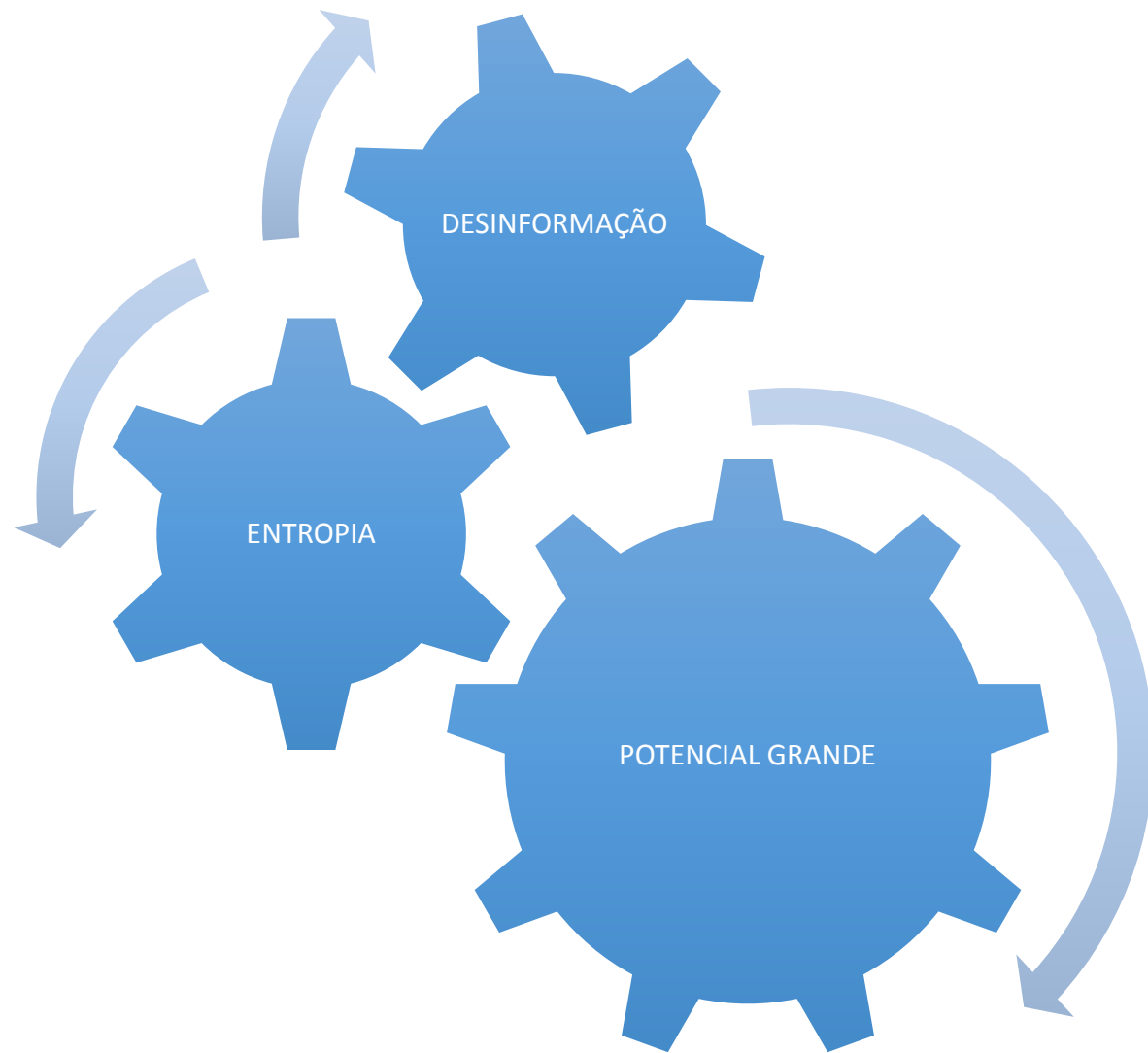
- n INFORMAÇÃO

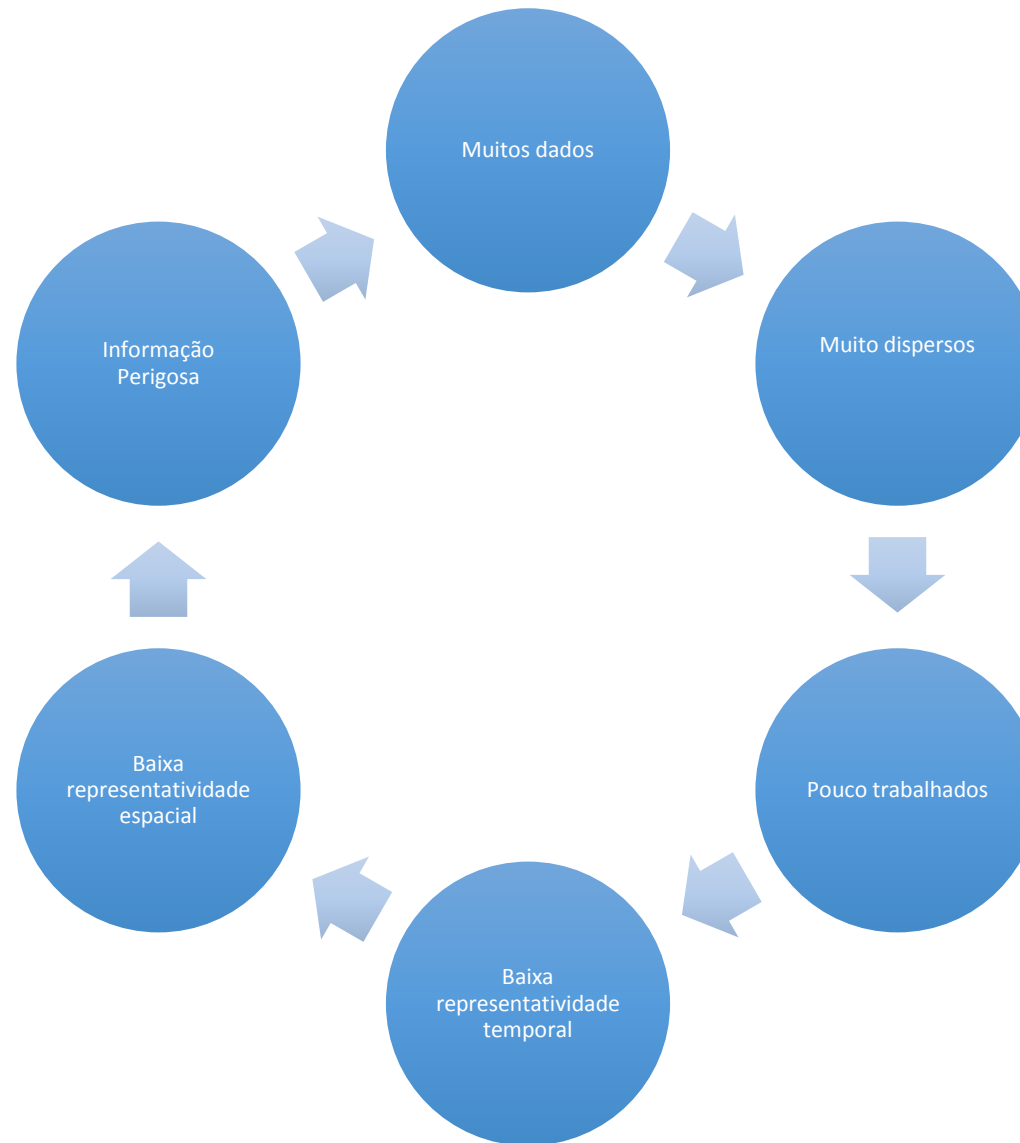
GERAÇÃO DE DADOS

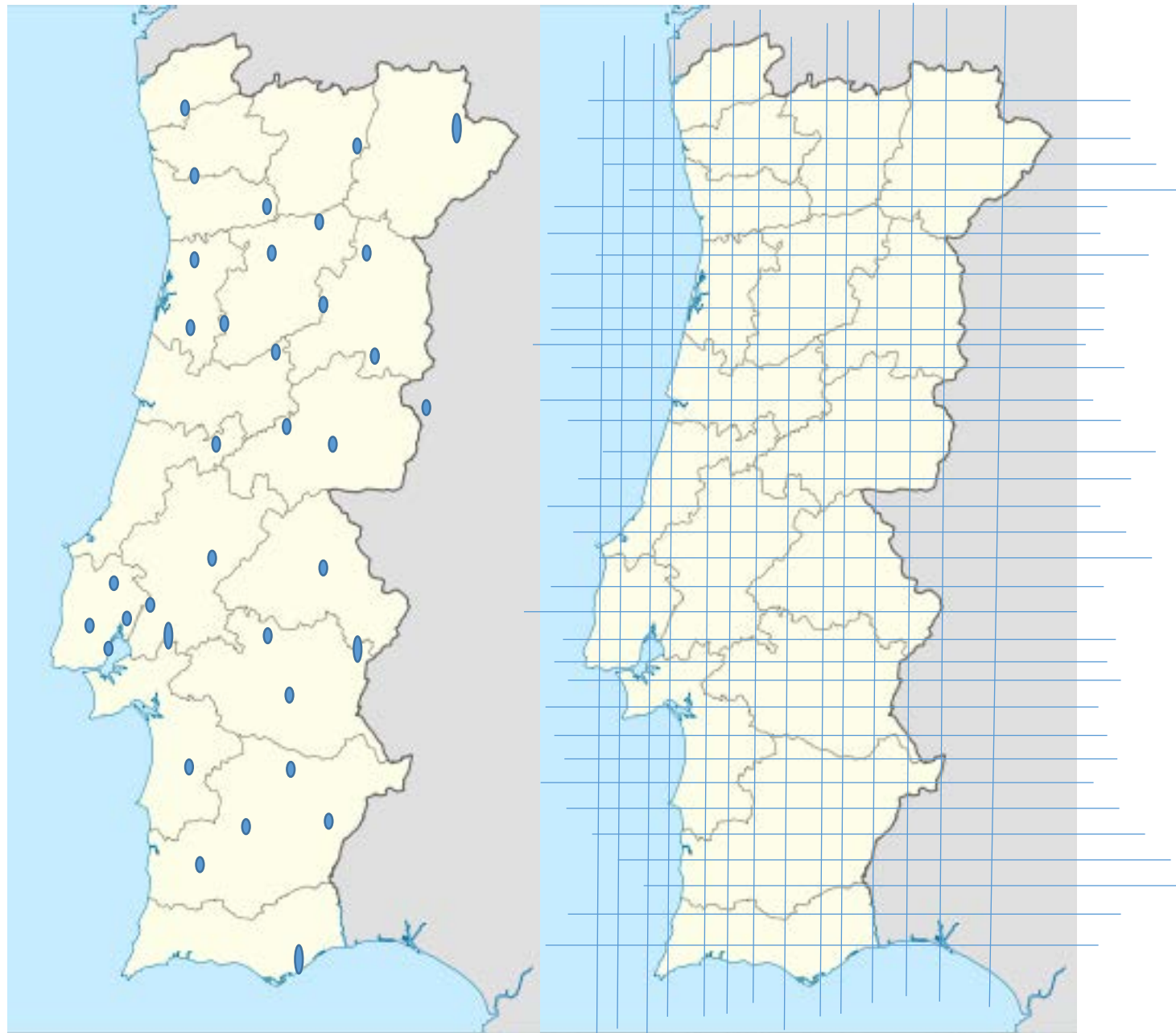


inútil









Funcionamento e/ou garante do cumprimento de requisitos legais em Portugal (e na Europa..)

Qualidade do Ar Exterior

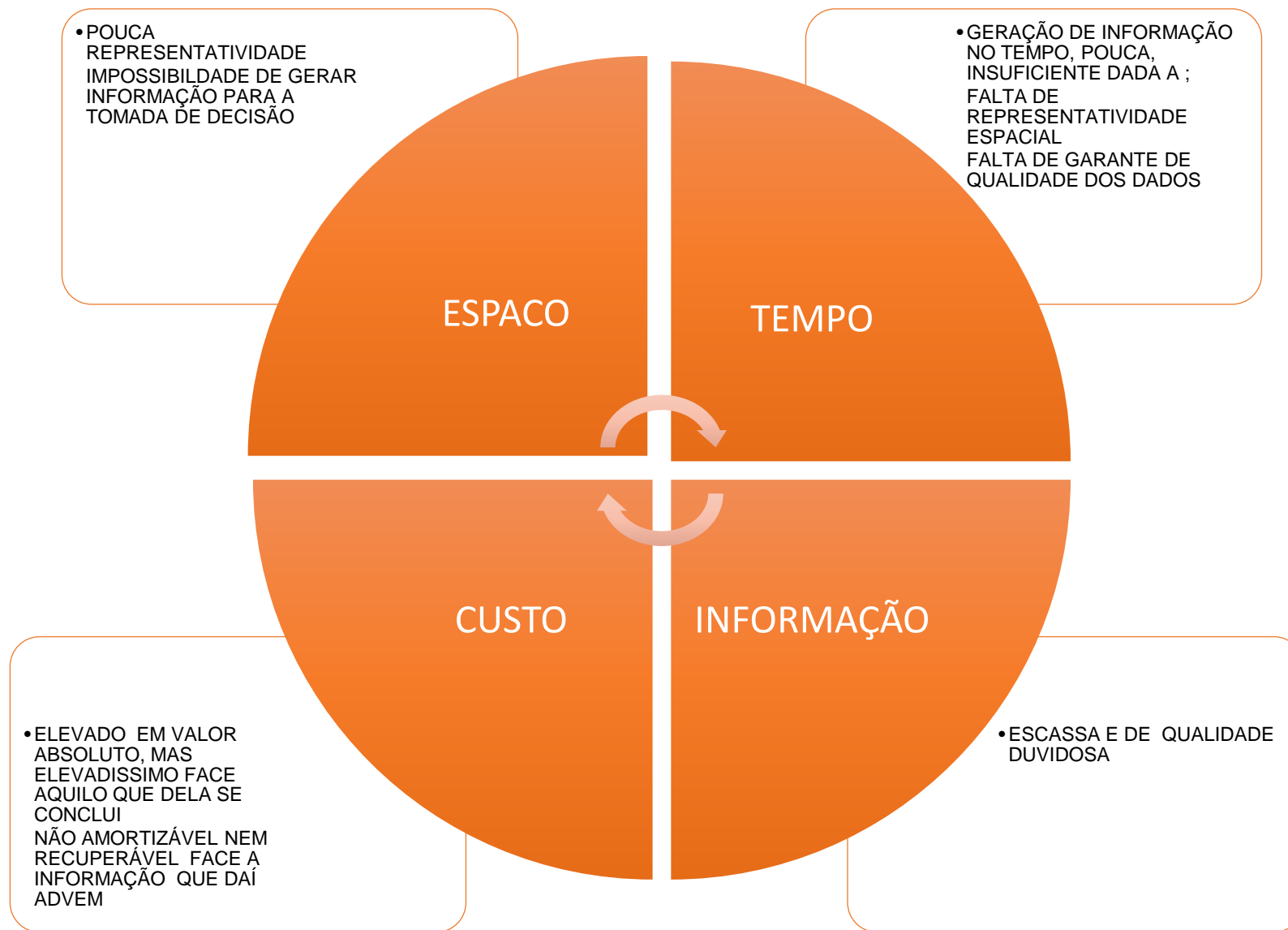
Rede nacional com cerca de 70 Estações
para todo o território

Emissões gasosas para a Atmosfera

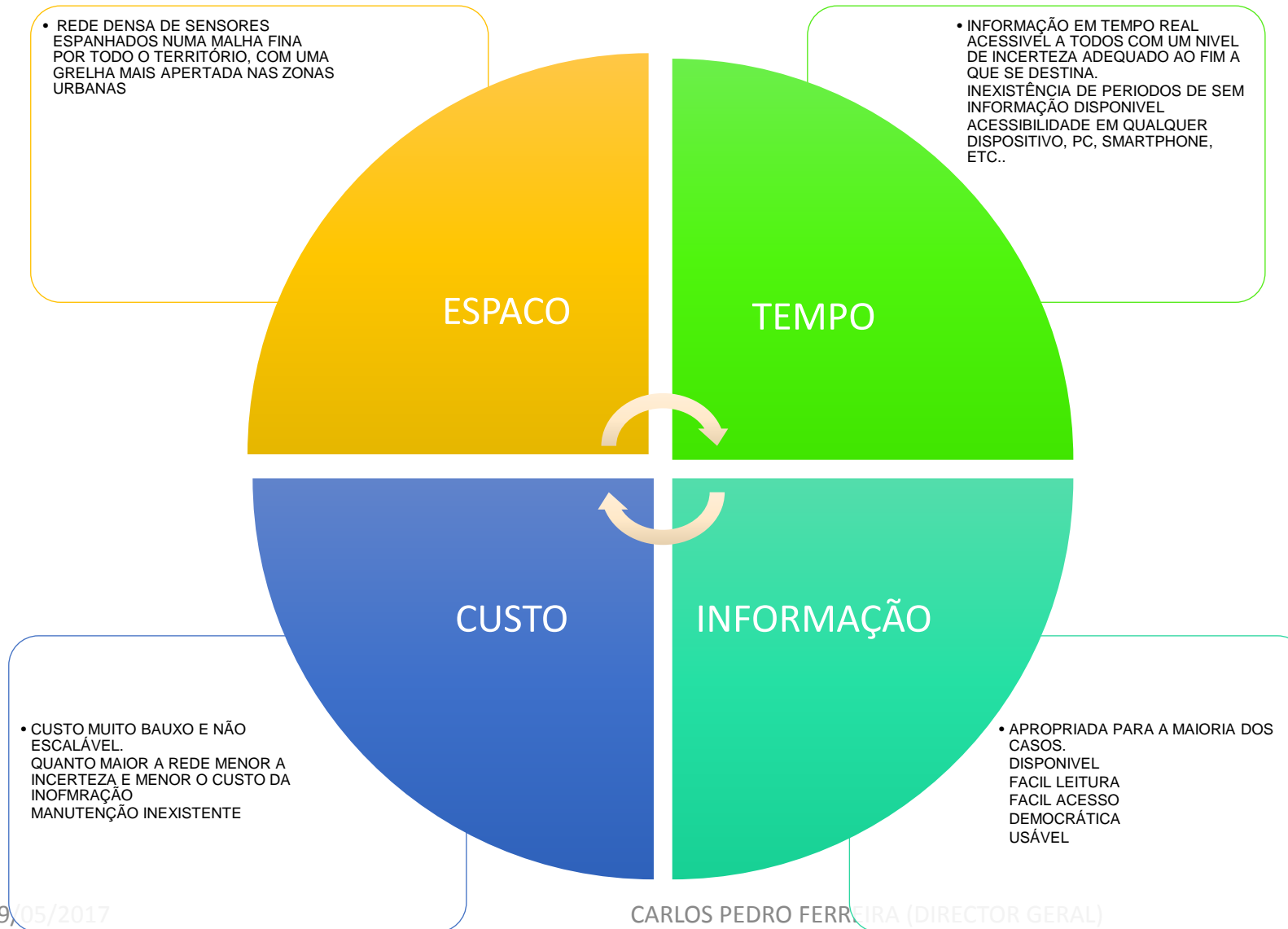
Uma medição de 3 em 3 anos
Duas medições por ano
Medições em continuo

O território português tem uma **área** total de 92 090 km²

- 70 Estações de Qualidade do Ar
- 1 estação por cada 1316 km²
- Sendo que grande parte delas estão nos aglomerados Urbanos Porto/Lisboa
- A maior parte do território não está coberta
- A parte coberta não merece confiança nos dados fornecidos
- Uma parte das estações em funcionamento não medem todos os parâmetros
- Reportamos estimativas com incertezas desconhecidas
- Reportamos dados com incertezas duvidosas



ALTERNATIVA ACERTIVA E REALÍSTICA. REAL TIME MONITORING LOW COST



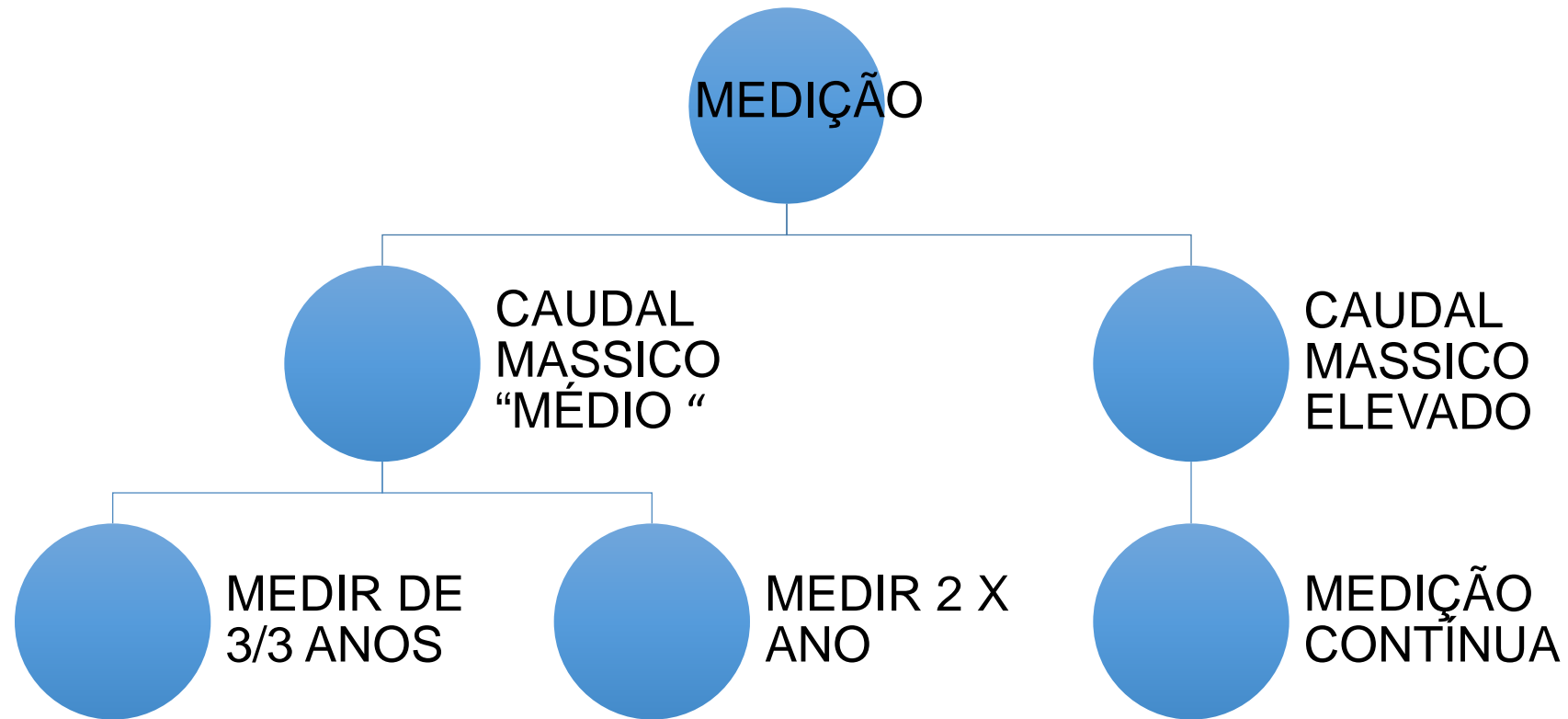
EMISSÕES ATMOSFERICAS

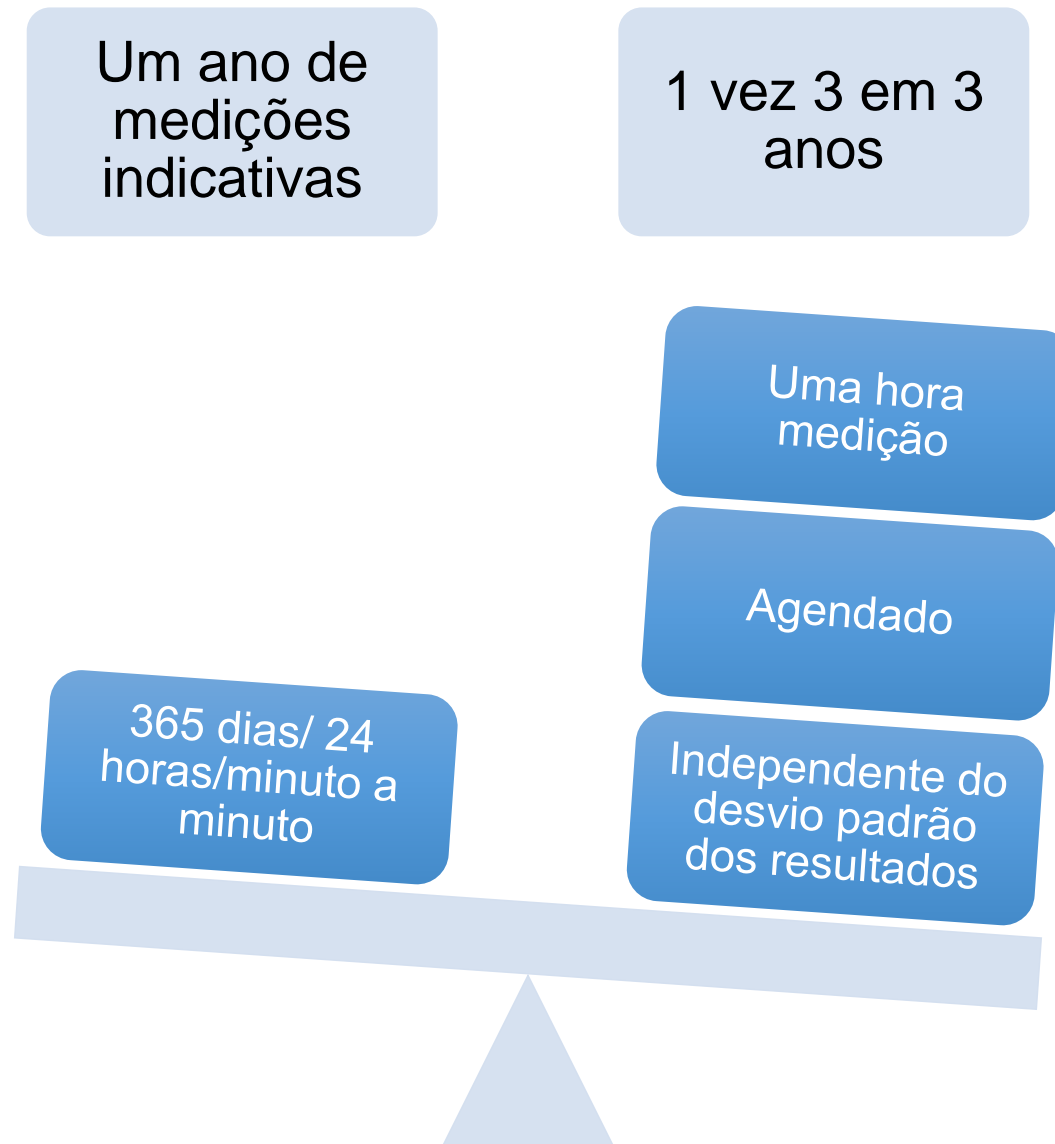
3 EM 3
ANOS

2XANO

CONTINUO

CRITÉRIOS PARA A DEFINIÇÃO DA PERIODICIDADE

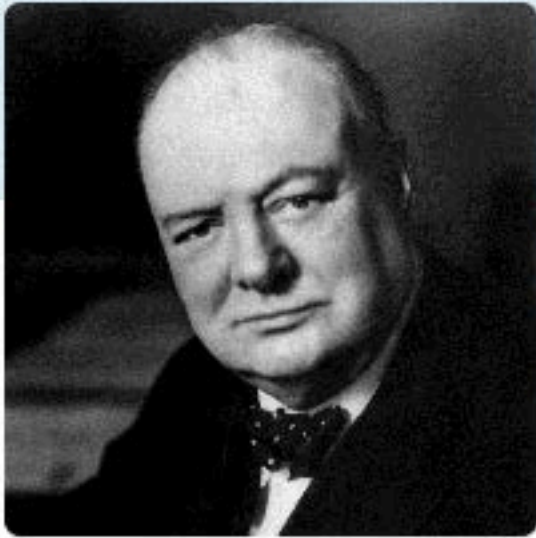




Com uma hora de
informação por fonte

Independentemente
da variabilidade dos
resultados nessa hora
e independentemente
de sabermos o que
essa hora representa

Afirmamos a
conformidade legal da
fonte e introduzimos
essa informação no
bolo da BIG DATA



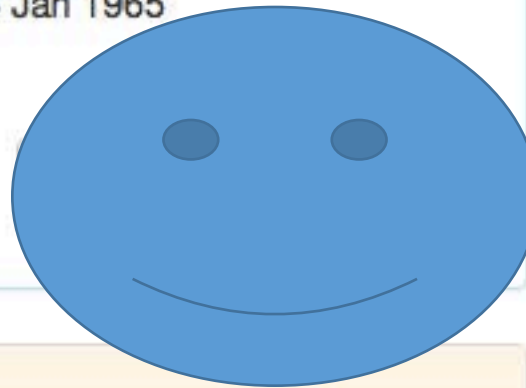
Winston Churchill

Inglaterra

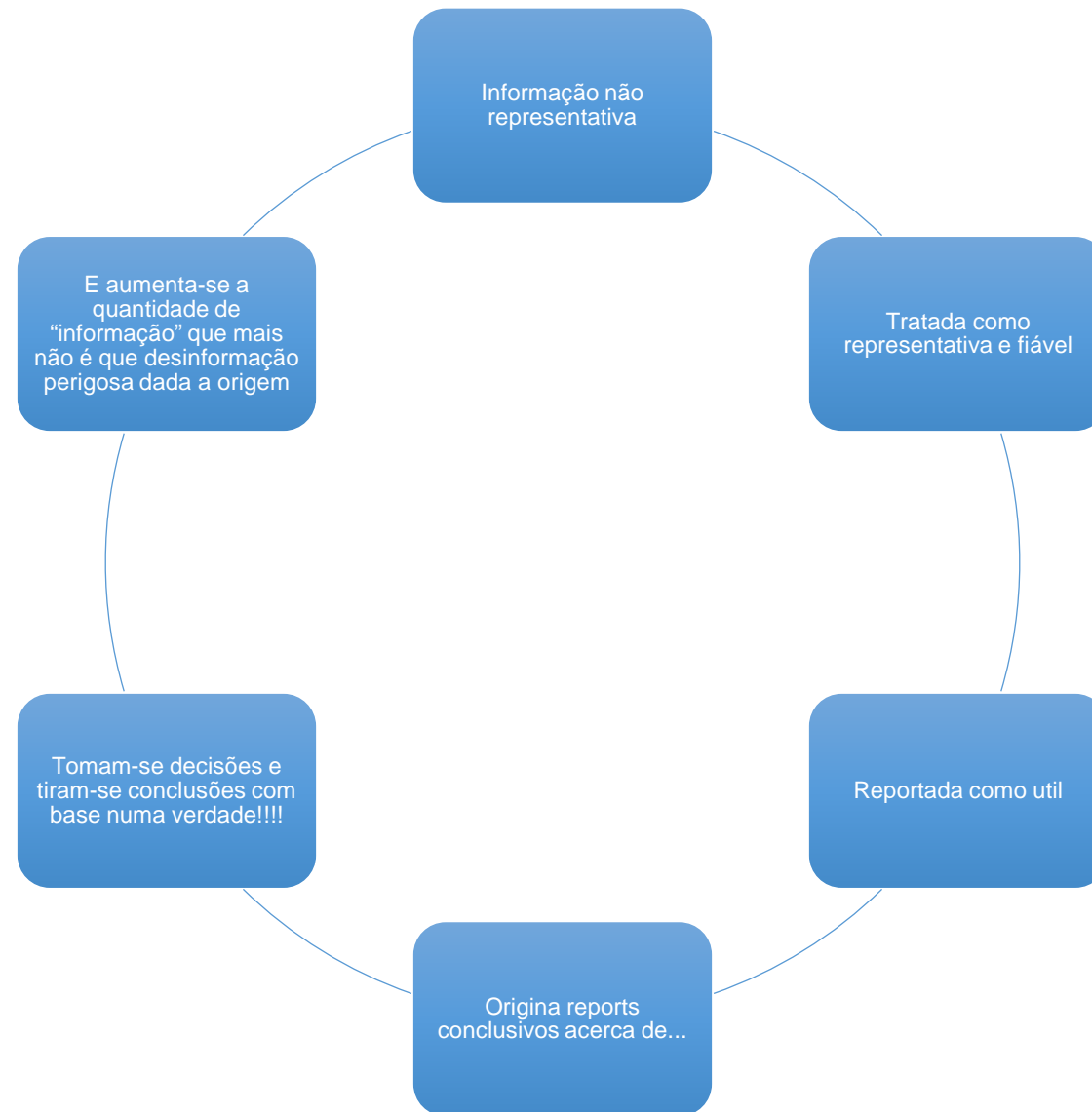
30 Nov 1874 // 24 Jan 1965

Político

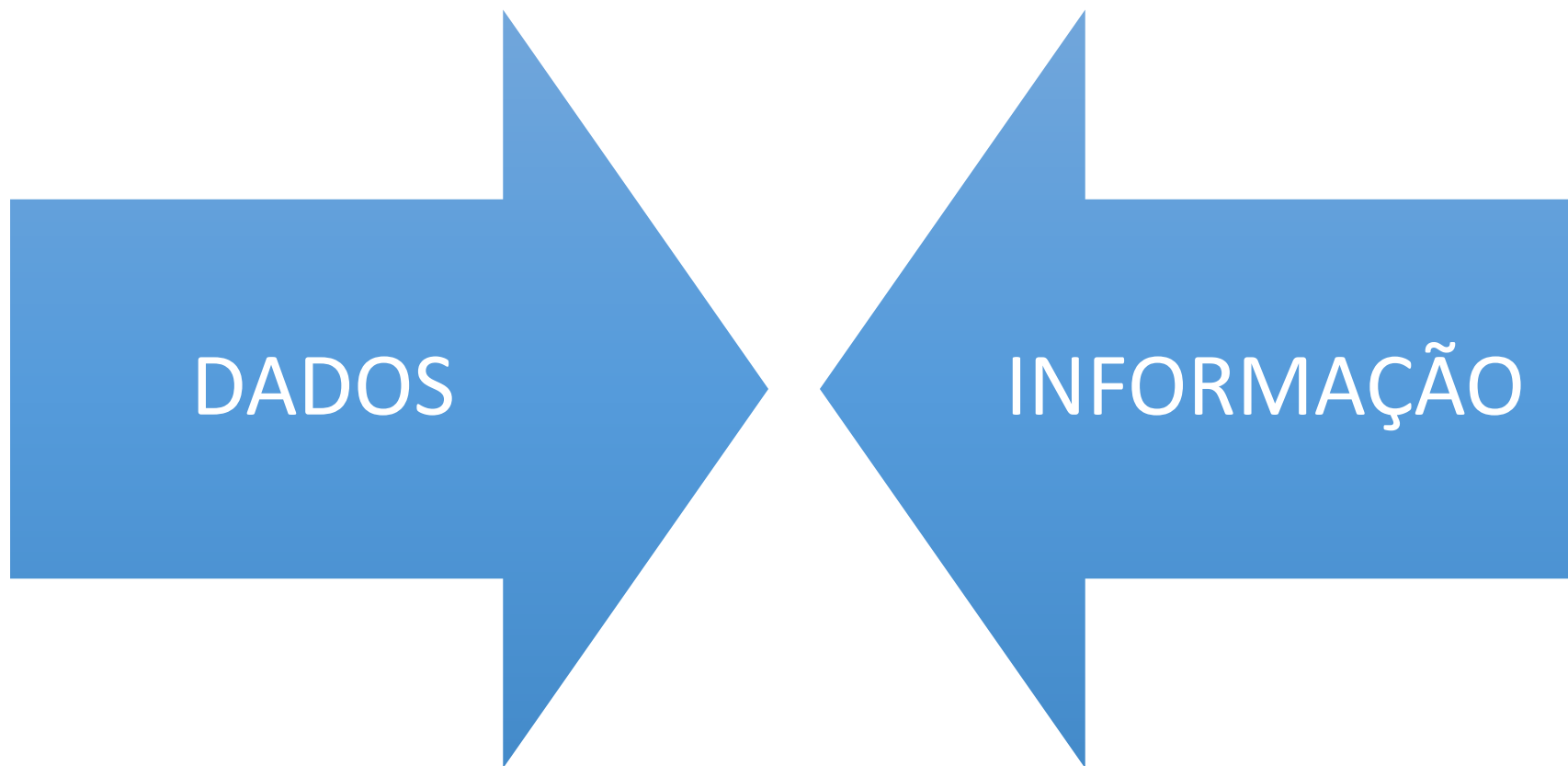
Wikipedia

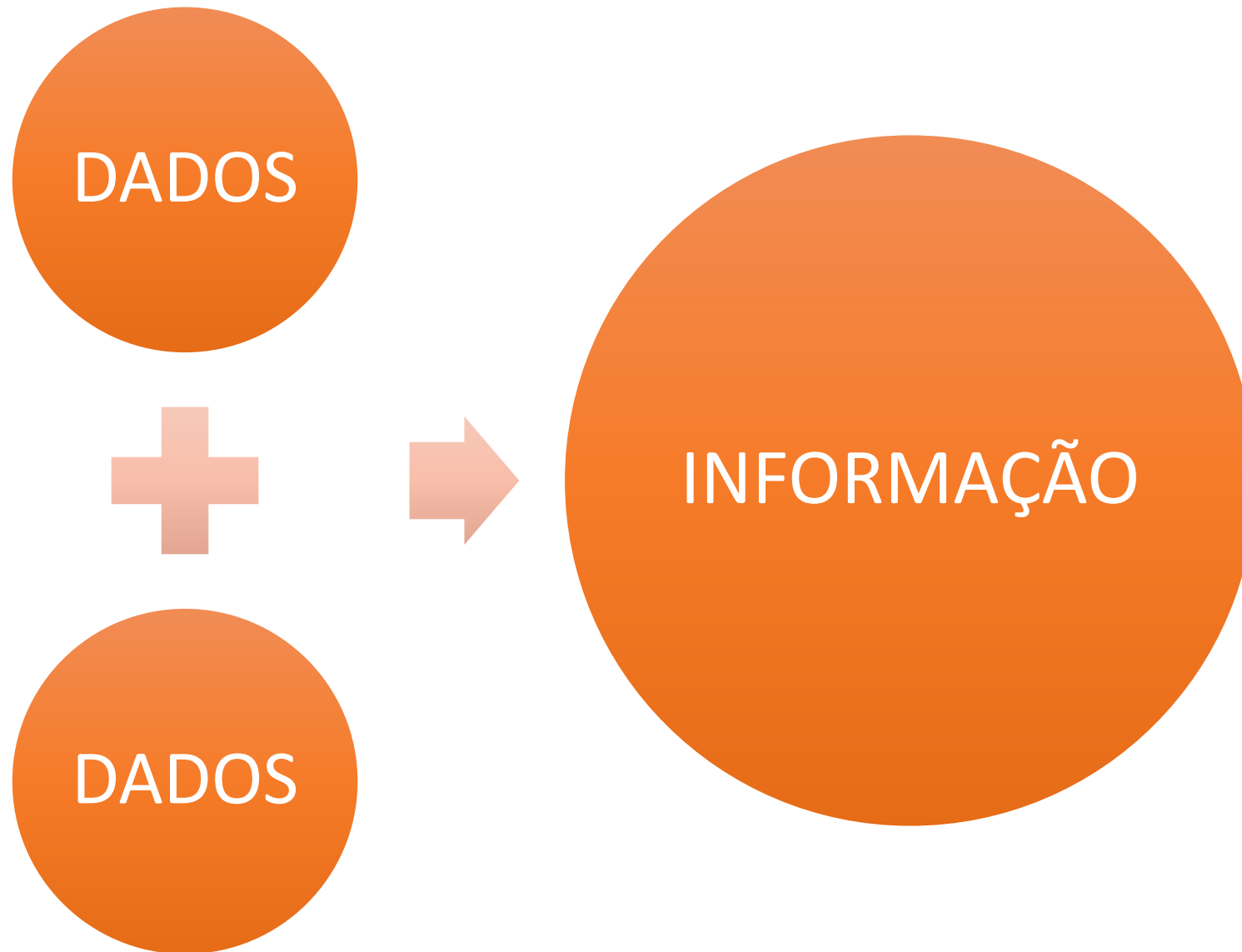


É inútil dizer «estamos a fazer o possível».
Precisamos de fazer o que é necessário.











QUANDO OS DADOS NÃO GERAM INFORMAÇÃO

- ✓ É PREFERIVEL NÃO TER NADA?
- ✓ É PREFERÍVEL TER DADOS?

Conclusões

1. A periodicidade das monitorizações das emissões gasosas deveria depender mais da variabilidade dos resultados obtidos, que dos caudais mássicos em causa
2. A periodicidade não deve ser fixa, mas ser alargada ou encurtada consoante a variabilidade dos resultados ao longo do tempo
3. A informação obtida por uma rede densa de sensores de qualidade do Ar, com incertezas conhecidas e maiores que as das Estações de Referência é muito maior e relevante que a obtida por 60 Estações de referência a cobrir todo o território.
4. O custo é muitíssimo mais baixo e teremos efetivamente informação.
5. Estações de Referência são importantes para aferir as medições indicativas mas o território tem que (para gerarmos informação) ser inundado por estações indicativas low cost que colmatam a falha de informação Espacial.
6. O modelo vigente neste momento em Portugal e suponho que na Europa trás implícito que os dados das emissões atmosféricas estão todos sub-estimados e os dados da Qualidade do Ar não podem ser representativos da Qualidade do Ar no País

Conclusões

7. Mesmo que todas as medições estejam muito bem feitas (incertezas muito baixas) o resultado do conjunto das medições não trás implícito uma confiança grande no resultado conjugado das medições porque;

7.1. Têm uma baixa representatividade temporal e/ou Espacial

7.2. Não temos garantia os equipamentos usados nas medições com baixa representatividade Espacial estejam devidamente calibrados e com manutenção adequada

7.3. Nas medições com baixa representatividade temporal, não temos informação de que esse espaço de tempo possa representar o normal funcionamento da fonte, e pior, nem sequer que durante o período e quem decorre a medição o valor médio obtido tenha uma incerteza baixa, i.e. que os valores que originaram o valor médio tenham um desvio padrão baixo.

Dito isto, está na altura de rever todo o modelo vigente e reinventar a forma de abordar a informação e não a confundir com ruído.

MUITO OBRIGADO