

# Bitmap

<https://quip.com/qgyxAEdi8MfL/Bitmap>

## 用户标签

如果用table column SQL join，很费时间

用户1：N标签 --> 1标签：N用户

Name	Sex	Age	Occupation	Phone	+	ID	Name	→	ID	Sex	Age	Occupation	Phone
小灰	男	90后	程序员	苹果		1	小灰		1	男	90后	程序员	苹果
大黄	男	90后	程序员	三星		2	大黄		2	男	90后	程序员	三星
小白	女	00后	学生	小米		3	小白		3	女	00后	学生	小米
Bitmap Conversion													
Sex	Bitmap		Age	Bitmap		Occupation	Bitmap		Phone	Bitmap			
男	1,2		90后	1,2		程序员	1,2		苹果	1			
女	3		00后	3		学生	3		三星	2			
									小米	3			

## Example

- 非90后，用全局(123)和XOR,(12) 而不是not;
- 90后且用苹果 &, 90后或用苹果 ||

## Advantage.

1. only 1 bit. HashMap each Id is int or long, which is 4/8 bytes.
2. Union/Intersection by bit manipulation

## Implementation

```
private long[] words;
private int size;
public boolean getBit(int bitIndex){
    // check boundary
    if(bitIndex < 0 || bitIndex > size-1) {
        throw new IndexOutOfBoundsException("out of bound error");
    }
    int wordIndex = getWordIndex(bitIndex);
    return (words[wordIndex] & (1L << bitIndex)) != 0; //0&1=0 false; -->1&1=1 true
}
public void setBit(int bitIndex) {
    // check boundary
    int wordIndex = getWordIndex(bitIndex);
    words[wordIndex] |= (1L << bitIndex);
}
private int getWordIndex(int bitIndex) {
    return bitIndex >> 6;
}
```