Localization of the added electrons:

Starting from a neutrum $Fe_2O_3$(0001) iron-terminated surface, we created a surface iron vacancy and added 1 to 3 electrons. The figure S1a shows the charge density difference between the surface with 1 extra electron and its neutrum counterpart. Figures S1b and S1c follow the same idea and show the difference between the surface with 2/3 extra electrons and the one with 1/2 extra electrons. The extra charge is always located on the vacancy oxygen atoms first neighbor to the Fe vacancy.

\[
\Delta n = n(X) - n(X+1)
\]

Iron/oxygen atoms are in cyan/red, and the vacancy is illustrated with a dashed transparent circle.

**Figure S1** 2 side views of the charge density difference $\Delta n$ between the $Fe_2O_3$(0001) surface with a Fe vacancy charge with "X" electron and "X+1" electrons ($\Delta n = n(X) - n(X+1)$). Red/blue mean gain/loss of electrons. Iron/oxygen atoms are in cyan/red, and the vacancy is illustrated with a dashed transparent circle.
Fe chemical potential:

In Fig. S2 we show the unit cell used to evaluate the chemical potential of iron. For each structure we optimized the lattice parameter and found the most stable magnetic configuration, which is antiferromagnetic in the (111) planes for FeO, antiferromagnetic in the (0001) direction for Fe$_2$O$_3$ and ferromagnetic for FeO$_2$.

![Image of three structures](image)

<table>
<thead>
<tr>
<th>Composition</th>
<th>FeO</th>
<th>Fe$_2$O$_3$</th>
<th>FeO$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of atoms</td>
<td>216</td>
<td>270</td>
<td>324</td>
</tr>
<tr>
<td>Symmetry</td>
<td>Cubic</td>
<td>Hexagonal</td>
<td>Hexagonal</td>
</tr>
<tr>
<td>Dimensions (Å)</td>
<td>13.325 x 13.325 x 13.325</td>
<td>15.009 x 15.009 x 13.671</td>
<td>14.5 x 14.4 x 14.5</td>
</tr>
</tbody>
</table>

**Figure S2** An illustration of the three structures used to evaluate the iron chemical potential in rich and poor conditions. Iron/oxygen atoms are in cyan/red.
Finite size correction:

The $E_{corr}$ term in the formation energy equation shown in the main paper was estimated in the following way. First, all energies were aligned using as reference the Hartree potential value averaged in the center of 6 frozen layers of the slab, far from the vacancy defect where the extra charges are located. Then, following the scheme proposed by Dan Wang et. al. [D. Wang et. al., Phys. Rev. Lett. 114, 196801 (2015)] to correct the ionization energy in 2D/1D systems, we evaluated the ionization energy $IE$ at each charge state as a function of the vacuum size $L_z$, with $2 < L_z < 4$nm. The result of such analysis is shown in Fig. S3 for the 1- and 3- cases. We note that $IE$ changes almost linearly with $L_z$, but also shows a $1/L_z$ contribution, that is non vanishing even at significantly large $L_z$ values. This non-linear behavior is likely connected to the fact that we have a surface, and not a 2D system as in the work by Dan Wang. $IE$ was then fitted using the expression given in the reference [D. Wang et. al., Phys. Rev. Lett. 114, 196801 (2015)]:

$$IE = \beta_1 + \beta_2 L_z + \frac{\beta_3}{L_z},$$

where $\beta_i$ are parameters that depend on the charge state/unit cell size. $\beta_1$ is connected to the converged ionization energy $IE_0$ by the expression $\beta_1 = IE_0 + \alpha/\sqrt{S}$, where $S$ is the surface area; $\beta_2$ has a value close to $\frac{q^2}{24\epsilon_0 S}$, with some variation from this value given that the background charge is distributed in the whole cell, and not only in the vacuum region; $\beta_3$ was found to be nearly proportional to $q$ and $ln(S)$, having an almost constant value. The fitted curves obtained are shown in the same Fig. S3.

![Figure S3](image_url)

Figure S3 In (a) we show the ionization energies $\epsilon(-1/0)$ between the charge state 1- and neutral as a function of the vacuum length $L_z$. The lines connecting the points were fitted using the equation above. In (b) we show the same plot for the $\epsilon(-2/-1)$ case.

With the $\beta_i$ parameters determined, we evaluated the converged ionization energies $IE_0$ between oxidation states 1-/0, 2-/1- and 3-/2-, that were then used to give the correct formation energy as a function of the Fermi Energy, shown in Figure 3 of the main paper.
Projected Density of States (PDOS) for selected structures:

For each structures shown in figure 6 of the main paper we have evaluated the PDOS at the neutral charge state. The result is shown in figure S4, where we can see that, except for configuration (d), all structures have at least one unoccupied band near the top of the valence band maximum.

**Figure S4** In a/b/c/d, we show the PDOS at the neutral charge state for structures of figure 6a/6b/6c/6d of the main paper, respectively. The HOMO energy is set as zero. The upper and lower panels in each figure indicate spin up and down, respectively.
**Additional details of the phonon calculation:**

To get converged frequency and amplitude values, all structures were optimized with tight thresholds, equal to: \(10^{-3}\) bohr for the maximum geometry change and root mean square of the geometry change; \(10^{-4}\) bohr\(^{-1}\) * Hart for the maximum component and root mean square of the force. In the evaluation of the infrared (IR) spectrum, the finite difference method was used with an increment of 0.003 in each x/y/z direction for the construction of the Hessian. Since the total number of atoms is 278 and we are interested only in the surface modes, only the outermost 4 atomic layers and the water molecules were actually displaced.

To evaluate the error due to this approximation, we evaluated the IR spectrum considering different number of active layers, using a smaller 2x2 unit cell of a \(Fe_2O_3\)(0001) pristine surface (without an iron vacancy) interacting with a single partially dissociated water molecule, 18 of vacuum and a lower cutoff of 500Ryd. The number of active surface layers tested were 2, 4 and 6. In Fig. S5 we can see the IR curves for each case, where only the surface modes are plotted, i.e., the modes involving at least one Fe/O atom of the outermost 2 layers and/or an atom from water. From the figure we see that the surface modes are found to be in the \(720 \leq w_f \leq 850\) range, and the frequency/amplitude error considering 4 active layers is found to be less than \(10cm^{-1}/5\%\), respectively. This should be the approximate error that we have in the values reported in the main paper.

**Figure S5** Simulated infrared spectrum of a water molecule adsorbed in a pristine iron terminated \(Fe_2O_3\)(0001) surface. Each curve consists of a a complete phonon calculation with "n" active surface layers, with n equal to 2, 4 and 6. Only the modes involving the surface Fe/O/H atoms are drawn.